RAND

Consequences of Welfare Reform: A Research Synthesis

Jeffrey Grogger, Lynn A. Karoly, Jacob Alex Klerman

DRU-2676-DHHS

July 2002

Prepared for the Administration for Children and Families, U.S. Department of Health and Human Services

Labor and Population Program

PREFACE

The Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 replaced the Aid to Families with Dependent Children (AFDC) program with the Temporary Assistance for Needy Families (TANF) program. The TANF program implemented by PRWORA—and a series of state-level reforms instituted prior to 1996 known as state waivers—produced changes in the structure of welfare benefits, introduced time limits, strengthened requirements for mandatory participation in work-related activities, and changed various administrative procedures.

To inform public debate on issues relating to the reauthorization of the TANF program in 2002 and to help states in refining the designs of their TANF programs, the Administration for Children and Families (ACF) of the U.S. Department of Health and Human Services (USDHHS) contracted with RAND to synthesize the current state of knowledge about the effects of the TANF legislation and the TANF programs of individual states. To this end, this document—the final report for the synthesis project—considers a range of outcomes, including the welfare caseload, employment and earnings, use of other government programs, fertility and marriage, household income and poverty, food security and housing, and child development. The primary focus of the synthesis is on the *net effects* of TANF, taking into account the effects of other factors such as the economy and other policy changes that may have affected the outcomes of interest. Like the literature on which it is based, the synthesis considers both the effects of specific policies including benefit structures, time limits, work requirements, and sanction policies, as well as the effect of the TANF reforms as a whole.

This research was funded as Task Order Number 1 under contract number 282-00-0005 from the U.S. Department of Health and Human Services. Jeffrey Grogger is a RAND research associate and professor of public policy at the University of California, Los Angeles. Lynn Karoly and Jacob Klerman are senior economists at RAND. The opinions expressed and conclusions drawn in this report are the responsibility of the authors and do not represent the official views of the USDHHS, other Agencies, UCLA, or RAND.

CONTENTS

FIGURES

TABLES

SUMMARY

The Research Synthesis project was undertaken to inform public debate on issues relating to the reauthorization of the Temporary Assistance for Needy Families (TANF) program in 2002 and to help states in refining the designs of their TANF programs. To this end, this report—the final report of the project—synthesizes the current state of knowledge about the effects of the TANF legislation and the TANF programs of individual states. It considers a range of outcomes, including the welfare caseload, employment and earnings, use of other government programs, fertility and marriage, household income and poverty, food security and housing, and child well-being. The primary focus is on the *net effects* of TANF, taking into account the impact of other factors such as the economy and other policy changes that may have affected the outcomes of interest. Like the literature on which it is based, the synthesis considers both the effects of specific policies, such as benefit structures, time limits, work requirements, and sanction policies, as well as the effect of the TANF reforms as a whole.

We begin this summary by highlighting the key findings from the synthesis. We then provide some relevant background for the study, describe our approach to conducting the synthesis, and discuss the results in more detail, including the limitations of the current knowledge base. Finally, we consider directions for future research.

FINDINGS IN BRIEF

The following key findings emerge from our synthesis of the research literature:

- Many of the effects of welfare reform on welfare use have been well studied. Over a dozen econometric studies have attempted to estimate the effects of welfare reform taken as a whole, and all but a few report that reform had substantial effects on reducing the caseload.
- Most of the reforms that were introduced in the 1990's had positive effects on employment and earnings. Thus, it seems likely that welfare reform is responsible for a portion of the increase in work and earnings among single mothers during the last decade. Nearly all of the evidence, from both experimental and econometric studies, points in this direction.
- There is little information available about the effects of welfare reform on the use of other government programs. The research is generally consistent with the hypothesis that welfare reform has caused part of the recent decline in food stamp use, but it does not explain the mechanisms that underlie this linkage. Only a few studies analyze the effects of welfare waivers on the use of school nutrition programs; fewer still have considered housing subsidies. The existing evidence also provides too narrow a basis to draw general

conclusions about whether welfare reform caused part of the initial decline in Medicaid enrollment that followed the implementation of TANF.

- The evidence from both experimental and econometric studies is insufficient to draw any firm conclusions about the effects of welfare reform on marriage or fertility. Evidence from the Minnesota Family Investment Program (MFIP) suggests that providing generous financial work incentives, either alone or with work requirements, may increase marriage or keep existing marriages intact. The mixed results from the Canadian Self-Sufficiency Project (SSP), however, suggest caution in interpreting the MFIP results.
- Some welfare reform components can raise incomes and reduce poverty, although this result is not associated with all policy components and there is reason to believe that some of the initially favorable effects will not persist over time. Generous financial work incentives—high earned income disregards inside the welfare system or earnings supplements outside the welfare system—generate the strongest income gains and anti-poverty effects.
- There is evidence of both positive and negative effects on child well-being of various components of welfare reform. Positive and negative effects were observed for indicators that capture socio-emotional behavior, academic performance, and health. The most favorable effects are associated with financial work incentives, most likely because of the increase in family income that results from combining work and welfare. But even for these programs, there is some evidence of unfavorable impacts for some subgroups of participants, particularly for adolescent children and for younger children of parents who do not experience large income gains. Work requirements do not appear to have strong impacts on children, either favorable or unfavorable, although again there is evidence of unfavorable impacts for adolescents, especially in school performance.

BACKGROUND

The last decade has witnessed significant changes in welfare policy, beginning with state waivers under the Aid to Families with Dependent Children (AFDC) program and culminating in the Temporary Assistance for Needy Families (TANF) block grant implemented by the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA). These reforms produced changes in the structure of benefits, introduced time limits, strengthened requirements for mandatory participation in work-related activities, and changed various administrative procedures.

During the same period, welfare-related outcomes have also changed, with the welfare caseload falling to approximately 2.1 million families as of September 2001, less than half its all-time peak level of 5.0 million families in 1994; with the fraction of welfare recipients participating in welfare-to-work activities, or actually working, increasing rapidly; with employment rates and earnings of single mothers rising substantially; and with family income increasing and the poverty rate falling. These improvements in labor market and economic outcomes have been accompanied by a leveling off of the prior upward trend in nonmarital fertility.

Because these changes occurred as welfare reform took place, some observers have concluded that welfare reform caused them. However, this inference ignores the fact that other policy changes—such as increases in the Earned Income Tax Credit (EITC), expansion of subsidized health insurance de-linked from welfare receipt, and increases in the minimum wage—took place during the same time. Perhaps most important, there was a long and robust economic expansion. Thus, at least some of the improvements in welfare-related outcomes likely resulted from changes in other policies and the improving economy, rather than from changes in welfare programs.

Understanding the role welfare reform played in general—and the role that specific reform policies played in particular—is important, because policymakers at all levels are once again debating the direction of welfare policy. The fiscal provisions of PRWORA must be reauthorized by September 30, 2002, and it seems likely that that reauthorization will serve as an opportunity for Congress to review the decisions made in the 1996 legislation. Any significant changes at the federal level will, in turn, require subsequent changes at the state and local levels. In addition, it seems likely that many state legislatures will use the opportunity of federal reauthorization to revisit their choices about how to implement TANF in their states.

With funding from the Administration for Children and Families (ACF) of the U.S. Department of Health and Human Services (USDHHS), we have synthesized the current state of knowledge from the growing base of research literature on the effects of welfare policies, with the aim of informing the ensuing debates about TANF reauthorization and state implementation.

HOW WE CONDUCTED THE SYNTHESIS

Conceptually, we organized the synthesis using the two-dimensional matrix presented in Table S.1. Along the rows are reform policies, preceded by TANF as a whole (labeled "TANF as a bundle") and followed by its main components; along the columns are welfare-related outcomes. Table S.1 is only exemplary. The following is a complete list of the policies studied: (1) financial work incentives, including earnings disregards and benefit reduction rates; (2) requirements to work or participate in work-related activities (and sanctions for failing to satisfy those requirements); (3) time limits on the receipt of benefits; (4) family caps and minor residence requirements; and (5) parental responsibility requirements. The full set of outcomes includes the following: (1) welfare use and the caseload; (2) employment and earnings; (3) utilization of other government programs (e.g., food stamps and Medicaid); (4) family structure, specifically marriage and fertility; (5) income and poverty; (6) other measures of well-being (e.g., food security, housing security, and health insurance coverage); and (7) child well-being (e.g., child development and school progress).

Ideally, the goal of the synthesis would be to fill in each of the cells in the matrix, expressing in a common format how each policy affected each outcome. However, this goal is not realistic, because the literature we synthesize has not yet covered each policy-outcome pair. Moreover, some of the cells in the matrix are "populated" by studies that, for one reason or another, are too tentative or otherwise inconclusive to be relied on for policy purposes. Nevertheless, the empty cells provide useful diagnostic information: They tell us where additional research results would help in making policy decisions and where we do not have enough knowledge.

While the above paragraph makes the synthesis effort seem straightforward, filling in the cells in the matrix is not straightforward. The synthesis aims to answer the question: What is the effect of a given policy (e.g., a lower benefit reduction rate or a time limit) on a given welfare-related outcome (e.g., the caseload or child development), *holding all else equal*? If all else is *not* held equal, then confounding influences—or more simply, confounders—can yield misleading results. There are two general research strategies for dealing with confounding influences: random assignment and econometric methods using observational data. After an extensive review of available studies, we identified 34 random assignment studies and 33 econometric studies that form the basis for our synthesis across the policies and outcomes.[1]

Table S.1—Two-Dimensional Matrix of Policy Reforms and Outcomes

Policy Reforms (Impact of What)	Outcomes (Impact on What)				
	Caseload	Employment	Earnings	Income	. . .
TANF as a "bundle"					
Specific TANF policies					
Work-related activity requirements					
Time limits					
Financial work incentives					
. . .					

Each of these approaches has strengths and weaknesses and may be more or less successful in controlling for confounding factors. One strength of the random assignment studies is that, when implemented well, they control for all non-random confounding factors. One drawback of these studies is that they were largely conducted prior to the enactment of PRWORA and are not necessarily representative of the policies adopted by the states under TANF. In contrast, the econometric studies exploit the observed variation in policies as they were adopted over time and across states during the waiver period before PRWORA's passage and after the passage of PRWORA. In the absence of experimental conditions, these observational studies control for as many confounding factors as possible but they are subject to the critique that there may be other unobserved factors that are responsible for some of the changes in behavior rather than changes in welfare policies.

Another difference between the two approaches concerns entry effects. Random assignment studies estimate effects from the moment of randomization, usually for those already on (or at least applying for) welfare. They thus miss any effects of policies on the likelihood of applying for welfare or on behaviors such as education, marriage, and fertility that might make one less likely to apply for welfare. In contrast, observational studies are able to analyze the entire population and thus can capture entry effects.

---

[1]We also draw on the so-called "leaver" studies that examine post-exit outcomes under welfare waivers and PRWORA for former welfare recipients (e.g., USDHHS, 2001a). While these studies provide relevant context and are essential for monitoring the status of families that discontinue receiving aid, they do not purport to identify the causal impact of welfare reform on outcomes.

In filling in the matrix, it is not enough simply to tally all the findings across the available studies. Rather, we implement an approach for weighing the findings for each analysis and assessing the strength of the cumulative evidence for each policy-outcome pair. Our approach accounts for the number of studies in any given cell, the quality of the methodology employed, and the statistical significance of the estimated impacts. Using this approach, our synthesis assigns a qualitative summary of the direction of the effect of each policy on each outcome and an indicator of the depth of the knowledge base associated with that effect; that is, how much is known about the effect of that policy reform on that outcome.

WHAT THE SYNTHESIS SHOWS

Table S.2 shows the filled-in version of the matrix in Table S.1. The columns of the matrix correspond to the outcomes we listed above. In the case of child well-being, given the expectation that impacts may differ by the age of the child, we assess the impacts for three age groups defined at the time of follow-up: preschoolers, grade schoolers, and adolescents. The rows of the matrix correspond to the reform policies or groups of policies for which there is some evidence base in the literature. We differentiate between the impact of financial work incentives (e.g., higher earned income disregards, lower benefit reduction rates) alone (row (1)) or when tied to hours of work (row (2)) or when combined with mandatory work-related activities (rows (5) and (6)). In the latter case we differentiate between strong and weaker financial work incentives, where programs in the latter category involve implicit tax rates that may be higher than those under AFDC.[2] We also consider the impact of mandatory work-related activities alone (row (3)) and the sanctions that accompany them (row (4)). For time limits, we distinguish pre–time limit effects from post–time limit effects (row (7) and (8)). We also assess the impacts of family caps (row (9)) and parental responsibility requirements (row (10)), two of the other reform policies that have received some study. The effects of reform as a bundle, in row (11), pertain only to the pre–time limit period.

The cell entries qualitatively summarize the effect of each policy on each outcome. The words indicate the direction of the effect, while the shading indicates how much is known about the effect of that policy reform on that outcome (the knowledge base). Starting first with the effect, "increase" indicates that a majority of the studies that populate the policy-outcome pair in question show that the policy increases the outcome, "decrease" indicates the opposite, while "mixed" indicates that there are roughly as many results showing a decrease as an increase. "No change" indicates the estimated impacts are mixed in sign and nearly all are small and insignificant.

---

[2]The Vermont Welfare Restructuring Project (WRP) is an example of a program with a weaker financial work incentive which can be compared with a program like the Minnesota Family Investment Program (MFIP) which we classify as having a strong financial work incentive. (See Table 3.5 for details regarding the features of the financial work incentives for these programs.)

Table S.2—Summary of Synthesis Results

	Policy or Policy Bundle	Welfare Use (A)	Employment (B)	Earnings (C)	Use of Other Government Programs: Food Stamps (D)	Use of Other Government Programs: Medicaid (E)	Marriage (F)	Fertility (G)	Income (H)	Poverty (I)	Other Measures of Well-being: Food Security (J)	Other Measures of Well-being: Children's Health Coverage (K)	Other Measures of Well-being: Savings (L)
(1)	Financial Work Incentives	INCREASE	INCREASE	*	*	*	INCREASE		INCREASE	DECREASE	INCREASE	INCREASE	INCREASE
(2)	Financial Work Incentives Tied to Hours Worked	INCREASE §	INCREASE	INCREASE	INCREASE		*		INCREASE	DECREASE	INCREASE		INCREASE
(3)	Mandatory Work-related Activities	DECREASE	INCREASE	INCREASE	DECREASE	DECREASE	NO CHANGE	NO CHANGE	MIXED	DECREASE	DECREASE	DECREASE	
(4)	Sanctions for non-compliance	DECREASE											
(5)	Mandatory Work-Related Activities and Strong Financial Work Incentives	INCREASE	INCREASE	INCREASE	DECREASE	MIXED	*		INCREASE	DECREASE	*	INCREASE	*
(6)	Mandatory Work-Related Activities and Weak Financial Work Incentives	DECREASE	INCREASE	INCREASE	DECREASE	DECREASE			INCREASE		*	DECREASE	
(7)	Time Limits (Before Recipients Reach Limit)	DECREASE	INCREASE	*				*	*				
(8)	Time Limits (After Recipients Reach Limit)	DECREASE	MIXED	*	*				*				*
(9)	Family Cap	MIXED				*		MIXED					
(10)	Parental Responsibility							*					
(11)	Reform as a Bundle (Before Recipients ~~Reach Time Limits)~~	DECREASE	INCREASE	INCREASE	DECREASE	MIXED	MIXED	MIXED	INCREASE	DECREASE	*	*	*

	Policy or Policy Bundle	Child Well-being: Child Abuse and Neglect (all ages) (M)	Preschool Age at Follow-Up: Behavior Problems (N)	Preschool Age at Follow-Up: School Achievement Problems (O)	Preschool Age at Follow-Up: Health Problems (P)	Grade School Age at Follow-Up: Behavior Problems (Q)	Grade School Age at Follow-Up: School Achievement Problems (R)	Grade School Age at Follow-Up: Health Problems (S)	Adolescents at Follow-Up: Behavior Problems (T)	Adolescents at Follow-Up: School Achievement Problems (U)	Adolescents at Follow-Up: Health Problems (V)
(1)	Financial Work Incentives					DECREASE	MIXED	INCREASE	INCREASE	*	
(2)	Financial Work Incentives Tied to Hours Worked		*	*	*	*	DECREASE	DECREASE	INCREASE	MIXED	*
(3)	Mandatory Work-related Activities	*	MIXED	MIXED	MIXED	MIXED	MIXED	DECREASE	MIXED	INCREASE	MIXED
(4)	Sanctions for non-compliance	*									
(5)	Mandatory Work-Related Activities and Strong Financial Work Incentives					DECREASE	DECREASE	INCREASE	INCREASE	INCREASE	
(6)	Mandatory Work-Related Activities and Weak Financial Work Incentives	*							*	*	
(7)	Time Limits (Before Recipients Reach Limit)	*									
(8)	Time Limits (After Recipients Reach Limit)										
(9)	Family Cap	*									
(10)	Parental Responsibility				*			DECREASE			
(11)	Reform as a Bundle (Before Recipients ~~Reach Time Limits)~~	*				MIXED	*	DECREASE	MIXED	INCREASE	*

NOTES: * Cell has up to three moderate or high-quality studies with no significant impacts or a single moderate-quality study with a significant impact.

§ These programs increase the sum of welfare payments and the earnings supplement provided outside the welfare system, although welfare payments per se may decrease.

LEGEND

Knowledge base:	DIRECTION	DIRECTION	DIRECTION	
	Much evidence	Some evidence	Little evidence	No evidence

Turning to the knowledge base, cells populated by several high-quality studies, most of which yield similar and significant estimates, are indicated by the dark gray shading. At the other end of the spectrum, cells with a shallow knowledge base are indicated by no shading. These cells are populated by a single high-quality study that yielded a significant estimate, two moderate-quality studies that yield similar and significant estimates, or similar constellations of evidence. Cells between these two categories are indicated by intermediate gray shading. Cells populated by a single moderate-quality study, or one or more high-quality studies whose results were insignificant, are indicated by an asterisk, denoting a nearly empty knowledge base. Cells for which there are no studies are left blank.

The entries in column (B) indicate that most reforms or combinations of reforms considered in Table S.2 increase employment, although we assign varying degrees of confidence to this qualitative assessment. Column (C) shows that many of these policies also increase earnings. Beyond employment and earnings, however, the impacts of specific policies vary to a greater extent. Thus, we organize the remainder of our discussion of Table S.2 by the table rows.[3]

What Does the Synthesis Tell Us About the Effect of Financial Work Incentives?

Although all the recent reform policies are capable of increasing employment (column (B)), they involve different trade-offs between reducing welfare use (column (A)) and increasing income (column (H)). Programs with generous financial work incentives generally increase welfare use, as seen in the intersection of column (A) and row (1)—which we refer to hereafter as cell A1. This is also true for financial work incentives tied to hours of work through an earnings supplement outside the welfare system (cell A2), where transfer payments (welfare payments or the earnings supplement) increase, although welfare use per se may decrease. Welfare use also increases when generous financial incentives are combined with mandatory work-related activities (cell A5), but the opposite is true for programs with weaker incentives (cell A6).

We are unable to assign a direction for financial work incentives alone on the use of food stamps or Medicaid (cells D1 and E1). A shallow evidence base suggests financial work incentives tied to hours of work may increase food stamp use (cell D2). When combined with work requirements, it appears that both strong and weak financial work incentives may decrease food stamp use (cells D5 and D6), but the effect on Medicaid use depends on the strength of the financial work incentives. A very shallow knowledge base suggests that financial work incentives alone may increase marriage (cell F1), but we do not have enough evidence to say how marriage is affected when financial work incentives are tied to hours worked or combined with work requirements for single parents. There is some suggestive evidence from MFIP that programs that provide generous financial work incentives combined with work requirements may increase marriage or keep existing marriages intact. However, the mixed results for the Canadian SSP suggest caution in interpreting the MFIP results. There is no evidence base from which to assess the relationship between financial work incentives and fertility.

[3]In Chapters 4 to 10 of the report, we organize our synthesis around outcomes rather than policies, with chapters devoted to the various outcomes listed in the columns of Table S.2. We return to a discussion of the impact of specific reform policies and reform as a bundle in Chapter 11.

Since financial work incentives allow families to keep more of their benefits as their earnings rise, they also increase income and decrease poverty, as shown in columns (H) and (I). However, even the programs that do increase income do so by only modest amounts, even when the financial work incentives are quite generous—more generous than most state TANF plans. With one exception (cell K6), financial work incentives (alone, or tied to hours of work, or combined with work requirements) are also associated with improvements in other measures of well-being such as food security, children's health insurance coverage, and savings (the intersection of rows (1), (2), (5) and (6) with columns (J) to (K)). However, several of the relevant cells are empty, and those that indicate a favorable impact are derived from a shallow knowledge base.

The impact on child well-being is more uncertain, and, when we are able to assign a direction, it is almost always based on a shallow evidence base. For children who are school-aged at the time of follow-up, strong financial work incentives (alone, or tied to hours of work, or combined with work mandates) appear to decrease behavior problems and possibly also school achievement problems as well (the intersection of rows (1), (2), (5), and (6) with columns (Q) to (S)). The available measures of child health suggest potential unfavorable effects for financial work incentives alone (cell S1) or when combined with work requirements (cell S5), but the reverse is true for financial work incentives tied to hours of work (cell S2). Thus, it appears that for this age group, the increased income associated with reforms that incorporate strong financial work incentives may lead to some improvements in children's outcomes in certain domains.

In contrast, for adolescents at follow-up, the various policies that include financial work incentives consistently appear to increase behavior problems and school achievement problems (the intersection of rows (1), (2), (5), and (6) with columns (T) and (U)). The evidence base available to assess the impact of financial work incentives on outcomes for pre-school-aged children at the time of follow-up is almost nonexistent although some of the impacts recorded in Table S.2 for grade-school-aged children pertain to children who were preschoolers at the time of random assignment.

What Does the Synthesis Tell Us About the Effect of Mandated Work-Related Activities?

Mandated work-related activities have been studied more than any other reform. Consequently, most of the cells with the darkest shading are in row (3). A substantial body of evidence shows that they generally reduce welfare use (cell A3). When implemented in conjunction with the AFDC benefit structure, in which benefits fell nearly dollar-for-dollar as earnings rose, however, they have little effect on income (cell H3). Even so, the limited evidence available suggests that these programs decrease poverty somewhat (cell I3). This may indicate that such programs are able to raise incomes for families just below the poverty line. This is consistent with the evidence that such programs have greater effects on income among relatively advantaged recipients than among disadvantaged recipients.

However, viewed from a different perspective, row (5) of the table shows that it is possible to require work and raise income (and more substantially reduce poverty) at the same time. The key is to combine the work requirement with a strong financial incentive, so that earnings rise more rapidly than benefits fall. The price for raising incomes is higher welfare use, which illustrates a central trade-off facing efforts to reform welfare.

Turning to the other outcomes in Table S.2, the evidence base is deep in indicating that mandated work-related activities reduce food stamp use (cell D3), while more limited evidence suggests food security declines as well (cell J3). We also place less confidence in the negative impact on Medicaid use (cell E3) but somewhat more confidence on the corresponding decrease in children's health insurance coverage (cell K3). This policy has no effect on marriage or fertility (cells F3 and G3), a conclusion that is based on five years of follow-up data for seven programs and two years of follow-up data for five other programs (hence the dark shading). A somewhat less substantial evidence base provides a very mixed picture of the impact of these programs on child well-being for all three of the age groups shown in Table S.2 (the intersection of row (3) with columns (N) to (V)). The only favorable assessment is for health problems of grade schoolers (cell S3), while the one clear unfavorable impact is for school achievement problems of adolescents (U3).

What Does the Synthesis Tell Us About Sanctions?

Sanctions are another policy reform about which little is known. Many states have enacted sanctions substantially more stringent than those under AFDC. Consequently, many families have lost their aid, or at least part of their aid, because of sanctions. No experiments have been conducted to isolate the effects of sanctions, however; None of the experiments we consider involve any experimental variation in sanction policy, except in conjunction with other policy reforms. Some econometric studies of the caseload indicate that stricter sanctions have greater effects on welfare use, but evidence showing that substantial declines in welfare use preceded the imposition of such sanctions by several years clouds the interpretation of those findings (cell A4). Except for one econometric study of child maltreatment (cell M13), there are no other studies on the effects of sanctions on employment, earnings, income, or other measures of family well-being.

What Does the Synthesis Tell Us About Time Limits?

Although time limits have been analyzed in several econometric studies, the few experiments involving time limits also incorporated other major reforms, making it difficult to isolate the effects of time limits alone. Moreover, the econometric studies only speak to the impact of time limits before recipients begin to exhaust their benefits (row (7)). Nonexperimental estimates from two random assignment studies, along with one econometric study of employment, provide some insights into how behavior changes once recipients begin reaching the time limit. Such evidence supports only a few rather limited statements about time limits.

First, most of the econometric studies suggest that time limits reduce welfare use during the pre–time limit period (cell A7), indicating that recipients change their behavior even before their benefits are exhausted. Evidence from one set of econometric studies is consistent with the notion that recipients leave welfare before reaching the time limit in order to "bank" their months of eligibility for future use.

Second, only two studies suggest that time limits also increase employment during the pre–time limit period (cell B7), so we place less confidence on this cell entry. There is insufficient evidence or no evidence available for assigning the direction of impact of time limits before recipients reach the limit for any of the other outcomes shown in Table S.2, including child well-being.

Third, the knowledge base regarding the post–time limit effects of time limits is even shallower. Two studies show that welfare use falls sharply once recipients begin to exhaust their benefits. Effects on employment are mixed, but none of the evidence suggests that it changes much, either up or down, once recipients start reaching the limit. Clearly, the post–time limit consequences of time limits could increase substantially once a higher proportion of the caseload reaches the limit.

What Does the Synthesis Tell Us About Family Caps and Parental Responsibility?

The table documents that we know relatively little about how family caps and parental responsibility requirements affect key outcomes. Limited evidence points to a mixed impact of family caps on fertility (cell G9). An equally shallow evidence base also produces mixed evidence with respect to the impact of family caps on welfare use (cell A9). Parental responsibility requirements, specifically those related to well-baby and well-child services (e.g., vaccinations), have been assessed in terms of their direct impact on the behaviors they seek to change, with some evidence of favorable effects in terms of child health for young children (cell M18). However, how this policy affects other outcomes is unknown.

What Does the Synthesis Tell Us About Welfare Reform as a Bundle?

Looking beyond specific policy reforms, there is a modest base of evidence to assess the impact of welfare reform as a bundle—either as implemented under state waivers or TANF—on some key outcomes. The available evidence comes from econometric studies that evaluate the impact of reform as a whole under state waivers or TANF and also from a handful of random assignment studies that implemented reform bundles that resemble state TANF plans. These experimental evaluations do not, however, necessarily represent the range of policy bundles implemented across the states, especially at the less generous end of the spectrum (i.e., lower benefit levels, weaker financial work incentives, stricter work requirements and sanctions, and shorter time limits).

With this caveat in mind, welfare reform as a bundle appears to produce impacts similar to those seen for mandatory work-related activities with weak financial work incentives (compare rows (6) and (11)): a decline in welfare use and use of food stamps and an increase in employment, earnings, and income. This is plausible given that most states implemented weaker financial work incentives combined with mandatory work-related activities, and given that what is known about the behavioral impacts of time limits suggests that they operate in the same direction as financial work incentives and work mandates (compare cells A6 and A7, and B6 and B7).

However, this comparison is not available for all outcomes, and the evidence base for the impact of welfare reform as a bundle on the other outcomes in Table S.2 is shallower. The impact on Medicaid use, marriage, and fertility is mixed (cells E11, F11, and G11), while poverty appears to decrease (cell I11). There is too little evidence to assess the impact of reform as a bundle on other measures of well-being in columns (J) to (L). In the case of the child well-being outcomes in columns (N) to (V), the limited available evidence appears to show a mixed impact on behavior problems of young children and adolescents, and an increase in school achievement problems for adolescents. There is some indication of reduced health problems for grade schoolers. However, the cells that are signed in these columns are based exclusively on results from two random assignment studies. The bundle of reforms implemented in these two states is not very representative of the reforms implemented in other states in terms of the length of the time limit (two years or less in both cases) or the generosity of the financial work incentives (very generous in one case). Thus, the impacts in row (11) for these columns should be interpreted cautiously.

It is also important to note that, regardless of the depth of the knowledge base, the entries in row (11) represent the effects of reform as a bundle during the pre–time limit period. Post–time limit evidence is very limited, and most studies summarized in this row cover time periods prior to when recipients could have exhausted their benefits. Once recipients reach the limit in substantial numbers, these effects could change.

What Does the Synthesis Tell Us About Welfare Reform Effects on Subgroups?

The effects of reform do not appear to differ widely across subgroups. Many observers would view this as good news, since there was widespread concern when PRWORA was enacted that only relatively advantaged recipients would respond, leaving the most disadvantaged behind. The subgroup-specific analyses provide no compelling evidence of this. In some cases, subgroup-specific impacts are similar for persons of different levels of disadvantage. In other cases, different measures of disadvantage generate different patterns, some appearing to favor the relatively advantaged and some appearing to favor the relatively disadvantaged. In many cases, subgroup-specific estimates are insignificant, in part because subgroup-specific sample sizes are too small to generate precise results even when the program may have had a substantial effect.

THE STATE OF OUR KNOWLEDGE BASE

Another way to use Table S.2 is to look across the whole table and assess the state of the knowledge base (the types of shading or lack of shading in the various cells). The table reveals that the knowledge base is strongest for understanding the impact of various welfare reform policies on welfare use, employment, earnings, and income. The base is weakest for assessing the impact of policies on broader measures of well-being, especially child outcomes, most notably those for pre-school-age children. Among the policies, a solid base of research exists to evaluate the impacts of mandatory work-related activities on most outcomes, and it is nearly as strong for financial work incentives, either alone or when tied to hours worked or in combination with mandatory work-related activities. As we have already discussed, several reform policies have received less attention, most notably sanctions, family caps, and parental responsibility requirements. Overall, just under half the cells in our matrix (120 out of 242 cells) are empty, indicating no research base exists to assess the policy-outcome pair. Another 36 cells (those with an asterisk) are nearly empty.

Some of the gaps in the knowledge base are particularly relevant for policy. For example, there have been relatively few causal studies of how welfare reform has affected poor families' participation in the Medicaid program, as shown in column (E), or the health care coverage of children more generally, as shown in column (K). This omission is particularly important in light of the initial decreases in Medicaid enrollment that occurred following the implementation of TANF—despite 15 years of policy initiatives designed to increase the coverage of poor children. As seen in columns (F) and (G), less is known about the impact of individual welfare reform policies and reform as a whole on marriage and fertility despite continued interest among many policymakers in policies to promote the formation and maintenance of two-parent families and to reduce out-of-wedlock childbearing. With the increased emphasis on work for mothers of children as young as age one or even younger, it is unfortunate that so little is known about welfare reform and child development prior to school entry, as shown in columns (N) to (P). This is an issue that is particularly relevant for policies aimed at improving early care and education.

For the policy-outcome combinations where we have a more substantial knowledge base, a nearly universal limitation of our conclusions is that they apply mostly to the short run. Most of the studies present evidence from follow-up periods of roughly two years, although the 11 programs in the National Evaluation of Welfare-to-Work Strategies (NEWWS) and several others provide results based on four or five years of follow-up data. The limited available evidence suggests that some of the effects vary with time since the policy was implemented.

The short-run nature of the evidence limits our understanding of whether reform has accomplished its goals of reducing unwed childbearing, encouraging marriage, and maintaining two-parent families. Marriage and fertility involve substantially more inertia than other aspects of behavior. As a result, we would expect the effects of welfare reform on such outcomes to become apparent only over a longer horizon. With mostly a short-run follow-up period to draw on, it should come as little surprise that most of the evidence from high-quality studies is mixed and statistically insignificant.

The short-run nature of the data also poses a problem for assessing how welfare reform affects the well-being of children. Although some aspects of a child's well-being, such as behavior problems, may respond quickly in reaction to changes in his or her parent's behavior, other aspects, such as cognitive skills, are likely to take much longer to change. Furthermore, even effects apparent in the short-term may change as children are exposed to cumulatively lower levels of welfare use and higher levels of employment on the part of their parents. In the short run, depending on the reform policy or policies, there is evidence of both favorable and unfavorable impacts on various aspects of children's development. In the case of adolescents, there is more consistent evidence of unfavorable behavioral and school achievement impacts associated with the policies that have been evaluated, in some cases up to five years after reform. Whether these same patterns will continue in the longer run—or whether they will be attenuated or exacerbated—remains to be determined.

A more general omission is any understanding of how reform has affected families' decisions to enter welfare. Random assignment experiments are a powerful research design for revealing how policy reforms affect those entering or on welfare at a point in time. However, they provide no information at all on how families decide to join the rolls. Econometric studies of welfare use reflect the effects of entry decisions, but they do not distinguish them specifically. To date, there have been only a few econometric studies that focus specifically on welfare entry. This omission is significant because entry appears to be important. Theoretical considerations lead us to expect that most policy reforms are likely to affect both entry and exit. Recent empirical work indicates that as much as one-half of the recent decline in the caseload is attributable to declining rates of entry.

DIRECTIONS FOR FUTURE RESEARCH

Despite the gaps in our knowledge base, we draw a broader lesson for future research: Our knowledge base in 2002 is stronger because of research programs put in place in the late 1980s and early 1990s under the strong guidance of USDHHS. That increase in knowledge occurred only as a result of major expenditures on program development and evaluation. Likewise, the inclusion of research funding in the PRWORA legislation supported a continuation of the research and evaluation studies that were initiated prior to federal reform. Consequently, the available knowledge base associated with the welfare reforms implemented in the last decade is superior in many respects to that available for many other areas of social policy.

To add to that knowledge base, it is desirable to learn about current policies that are poorly understood and about reforms that may be proposed in the future. Since the research cycle is at least as long as the policy cycle, we need to continue to put research efforts in place now for what we will need to know when the nation next considers major welfare reform.

Several specific areas deserve priority. To begin, more long-run information on the effects of current policies is crucial. Current long-run studies should be continued and, where possible, extended. Long-term evaluations should include such outcomes as family structure and child well-being, where the impacts may become apparent only over time or where they may vary with the stage of child development. Further research is also needed to understand the effectiveness of alternative strategies for promoting the transition from welfare to work for subgroups of the welfare population, such as for recipients with substance abuse problems and those who experience domestic violence.

Other policies that are less well understood need further evaluation. Time limits represent an important example. The number of families exhausting their benefits may grow sharply in the near future as recipients reach the federal five-year time limit. Studies to assess how families respond are critical. Sanctions are among the most poorly understood of all of the policy reforms. This is an area where both econometric and experimental work would be useful. Econometric analyses that incorporate information on the likelihood of sanctioning and the monetary value of sanctions would provide a more complete understanding of this policy than the studies that are currently available. Experimentation could also help reveal how different levels of sanctions affect a broad range of outcomes. In either case, future research should continue on the path of expanding the range of outcomes examined, in addition to welfare use, employment, and earnings, which have been the focus of most studies to date.

Entry effects also need to be better understood, both to fully grasp how reform has affected welfare use and labor market behavior and to understand how it affects fertility and the utilization of important in-kind services. This is an area where experimentation has less to offer. What is needed are high-quality econometric studies that focus directly on entry decisions.

Initiatives sponsored by USDHHS and other agencies in several of these areas will add to the current knowledge base. For example, follow-up studies continue for a number of the experimental evaluations we examined in our synthesis, and evaluations are under way in a number of other states that implemented other bundles of reforms. Reports are expected soon with longer-term results for several of the experimental evaluations which will add to the small base of results currently available with follow-up periods as long as five years after randomization. Other studies are under way to understand issues regarding accessing Medicaid and the Food Stamp Program, to evaluate the effectiveness of programs serving particularly disadvantaged segments of the welfare population, and to evaluate alternative approaches to promoting job retention and advancement among TANF recipients.

As in the past, advancing such an ambitious research agenda will require substantial federal participation. Many of the experiments reviewed here were conducted to satisfy the requirement that waiver-era reforms be evaluated and because the federal government paid for a portion of the costs. Under PRWORA, states are no longer obligated to rigorously evaluate their reforms. If we are to increase our knowledge base between now and the next time the nation considers major welfare reform, federal funds need to be invested to continue the evaluation of state investments under TANF. Even given TANF's devolution of welfare policy to the states, a strong federal role in research and evaluation remains appropriate. As this study demonstrates, knowledge gained in one state may be broadly applicable in others. Because of these knowledge spillovers, the states cannot be expected to finance and carry out the needed amount of evaluation research without federal assistance.

ACKNOWLEDGMENTS

We benefited from detailed comments on initial draft chapters and the full draft report provided by the members of the technical advisory group to this project: Gordon Berlin (Manpower Demonstration Research Corporation), Rebecca Blank (University of Michigan), Jeanne Brooks-Gunn (Columbia University), Greg Duncan (Northwestern University), David Ellwood (Harvard University), David Fein (Abt Associates), Ron Haskins (Brookings Institution), Susan Mayer (University of Chicago), Lawrence Mead (New York University), Bruce Meyer (Northwestern University), Robert Moffitt (Johns Hopkins University), LaDonna Pavetti (Mathematica), Isabel Sawhill (Brookings Institution), John Karl Scholz (University of Wisconsin, Madison), Steve Trejo (University of Texas), and Sheila Zedlewski (Urban Institute). Discussion with other RAND and UCLA colleagues including John Adams, Steven Haider, Joe Hotz, and Robert Schoeni also yielded important insights. Elaine Reardon and Robert Schoeni also provided thorough technical reviews of the report.

Finally, we would like to thank several members of the RAND staff who contributed to this report. Glenn Daley, Caroline Danielson, and Dorothy Schmidt provided outstanding research assistance in assembling and abstracting the literature reviewed in this report. Boichi San developed an Access database for recording results from the literature review. We are grateful for the editorial contribution made by Paul Steinberg and Christina Pitcher and for the production assistance and administrative support provided by Mechelle Butler and Christopher Dirks.

ACRONYMS

ABAWD	Able-bodied adults without dependents
ABC	A Better Chance (Delaware)
ACF	Administration for Children and Families (USDHHS)
ADC	Aid to Dependent Children
AFDC	Aid to Families with Dependent Children
AFDC-UP	AFDC Unemployed Parent
AWWDP	Arkansas Welfare Waiver Demonstration Project
BPI	Behavioral Problems Index
CEA	Council of Economic Advisors
COS	Child Outcome Study (NEWWS)
CPS	Current Population Survey
CWEP	Community Work Experience Programs
CWPDP	California Work Pays Demonstration Project
DEFRA	Deficit Reduction Act of 1984
DoD	difference-of-differences
DoDoD	difference-of-difference-of-differences
EITC	Earned Income Tax Credit
EMPOWER	Employing and Moving People Off Welfare and Encouraging Responsibility (Arizona)
FDP	Family Development Program (New Jersey)
FIP	Family Investment Program (Iowa)
FNS	Food and Nutrition Service

FPL	Federal poverty line
FSA	Family Support Act of 1988
FSP	Food Stamp Program
FTP	Family Transition Program (Florida)
GAIN	Greater Avenues for Independence (California)
GAO	General Accounting Office
GED	General Educational Development (high-school equivalency certificate)
HCD	Human Capital Development
HCFA	Health Care Financing Administration (now Centers for Medicare and Medicaid Services)
IA	Income Assistance (Canada)
IDA	individual development account
IMPACT	Indiana Manpower Placement and Comprehensive Training Program
JOBS	Job Opportunities and Basic Skills Training Program
LEAP	Learning, Earning and Parenting (program) (Ohio)
LFA	Labor Force Attachment
MDRC	Manpower Demonstration Research Corporation
MFIP	Minnesota Family Investment Program
MFIP-IO	MFIP Incentives Only
NEWWS	National Evaluation of Welfare-to-Work Strategies
NLSY	National Longitudinal Survey of Youth
NSAF	National Survey of America's Families
NSLP	National School Lunch Program
OBRA	Omnibus Budget Reconciliation Act of 1981
PIP	Preschool Immunization Project (Georgia)
PPI	Primary Prevention Initiative (Maryland)
PPVT-R	Peabody Picture Vocabulary Test-Revised
PRWORA	Personal Responsibility and Work Opportunity Reconciliation Act

PSID	Panel Study of Income Dynamics
SCHIP	State Children's Health Insurance Program
SIPP	Survey of Income and Program Participation
SPD	Survey of Program Dynamics
SSP	Self-Sufficiency Project (Canada)
TANF	Temporary Assistance for Needy Families
TMA	Transitional Medical Assistance
TSMF	To Strengthen Michigan Families
UI	Unemployment Insurance
USBLS	U.S. Bureau of Labor Statistics
USDA	U.S. Department of Agriculture
USDHHS	U.S. Department of Health and Human Services
VIEW	Virginia Initiative for Employment not Welfare
VIP/VIEW	Virginia Independence Program/Virginia Initiative for Employment not Welfare
WIC	Special Supplemental Nutrition Program for Women, Infants, and Children
WIN	Work Incentive program
WRP	Welfare Restructuring Project (Vermont)
WRP-IO	WRP Incentives Only

Chapter One

INTRODUCTION

1.1. BACKGROUND

Over the years, the objectives of welfare reform have been to reduce dependency and promote work while still alleviating need. The Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) was enacted in 1996 in part to further those objectives. It also advanced three other goals: to reduce unwed childbearing, to promote marriage, and to maintain two-parent families.

PRWORA represents the culmination of the most recent era of welfare reform. That era began during the late 1980s, when many states sought and received approval to alter their state Aid to Families with Dependent Children (AFDC) programs through waivers under section 1115 of the Social Security Act. The waivers and the Temporary Assistance for Needy Families (TANF) block grant implemented by PRWORA produced changes in the structure of benefits, introduced time limits on the receipt of benefits, strengthened requirements for mandatory participation in work-related activities, and changed various administrative procedures.

Over the last decade, change was not limited to welfare programs: Welfare-related outcomes also changed. The welfare caseload fell by more than half, from its all-time peak of 5 million families in 1994 to approximately 2.1 million families in September 2001 (USDHHS, 2001b). The fraction of welfare recipients participating in welfare-to-work activities or actually working has increased rapidly. Employment rates of women leaving welfare range from 62 to 90 percent (USDHHS, 2001a). Among single women with children more broadly, the fraction employed increased from 69 percent in 1993 to 83 percent in 1999, a 20 percent increase (Grogger, forthcoming). Single mothers worked, on average, 7 more weeks in 1999 (for a total of 37 weeks) compared with 1993, and their earnings have increased by 35 percent over the same time period. Family income has also increased, and the poverty rate has fallen (Blank, 2000). These improvements in labor market outcomes have been accompanied by a leveling off of the prior upward trend in nonmarital fertility.

Because these changes in the caseload, in the labor market, and in other outcomes occurred as welfare reform took place, some observers have concluded that welfare reform caused these changes in behavior. However, this inference ignores the fact that other changes took place during the same time. Other policy changes, such as increases in the Earned Income Tax Credit (EITC), expansion of subsidized health insurance delinked from welfare receipt, and increases in the minimum wage, could have had equally important effects on the behavior of single mothers. The long and robust economic expansion of the 1990s also could have had important effects. Thus, at least some of the improvements in welfare-related outcomes likely resulted from changes in other policies and the improving economy rather than from changes in welfare programs.

Understanding the role welfare reform played in general—and the role specific reform policies played in particular—is important, because policymakers at the national, state, and local level are once again debating the direction of welfare policy. The fiscal provisions of PRWORA, which implemented welfare reform at the federal level through the TANF program, must be reauthorized by September 30, 2002. Reauthorization will serve as an opportunity for Congress to review the decisions made in the 1996 legislation, including various TANF program features such as financial work incentives, mandatory work-related activities, and time limits. Any significant changes at the federal level will, in turn, require subsequent changes at the state and local levels. In addition, it seems likely that over the next few years many state legislatures will revisit their choices about how to implement TANF in their state. Finally, as state budgets grow tighter, it is more important than ever to understand which policies are effective.

Ideally, these debates would be informed by research findings. Since the passage of PRWORA, the research literature on the effects of welfare policies has grown, both with the addition of new studies of relevant pre-TANF experience and with the emergence of a literature exploring post-TANF outcomes. However, incorporating the lessons from the research literature into the policy debate is not a simple matter. The literature is spread over several disciplines, each with its own jargon, and, owing to lags in the release of data and in the peer-review process, much of it has not yet been published. Moreover, the methodological challenges of distinguishing the effects of welfare reform from the effects of other policies and the economy are difficult. Studies sometimes come to different conclusions.

1.2. OBJECTIVE AND APPROACH

To allow policymakers to benefit from this large and diverse knowledge base, the Administration for Children and Families (ACF) of the U.S. Department of Health and Human Services (USDHHS) contracted with RAND to synthesize the current state of knowledge, with the aim of informing the ensuing debates about TANF reauthorization and state implementation. This study collects, summarizes, and synthesizes what is known about how specific welfare reforms and welfare reform as a whole have affected welfare-related outcomes.

In the remainder of this chapter, we discuss how we organize and implement the synthesis and the conceptual framework that guides our approach. Chapter 3 expands on some of the methodological challenges involved in implementing the framework.

1.2.1. How We Organized the Synthesis

We organized our synthesis using the two-dimensional matrix presented in Table 1.1. Along the rows are reform policies, preceded by TANF as a whole (labeled “TANF as a bundle”) and followed by the major components of state welfare reforms: requirements for work-related activities and welfare-to-work programs, time limits, financial work incentives implicit in the benefit structure, and so on. Along the columns are welfare-related outcomes, beginning with the welfare caseload, employment, earnings, and income, and continuing with other outcomes of interest. Using this framework, we seek to synthesize what is known about the causal effects of each policy and program (the rows of Table 1.1) on each outcome (the columns of Table 1.1). Ideally, the results of such a synthesis would allow us to answer the dual questions discussed above: (1) What caused the observed changes in outcome *Y*? and (2) What would be the effects (on each outcome) of changing policy or program *X*?

The four outcomes listed in Table 1.1 represent only a few of the welfare-related outcomes that we analyze. The full list includes the following: (1) welfare use and the caseload; (2) employment and earnings; (3) utilization of other government programs (e.g., food stamps and Medicaid); (4) family structure, specifically marriage and fertility; (5) income and poverty; (6) other measures of well-being (e.g., food security, housing security, and health insurance coverage); and (7) child well-being (e.g., child development and school progress).

Table 1.1—Two-Dimensional Matrix of Policy Reforms and Outcomes

Policy Reforms (Impact of What)	Outcomes (Impact on What)				
	Caseload	Employment	Earnings	Income	. . .
TANF as a "bundle"					
Specific TANF policies					
Work-related activity requirements					
Time limits					
Financial work incentives					
. . .					

Like the literature on which it is based, the synthesis considers both the effect of the TANF reforms as a bundle and the effects of specific policies. The policies listed in Table 1.1 represent a subset of the policies that we consider. The complete list is determined by the policies for which a research base exists. Those policies include the following: (1) financial work incentives, including earnings disregards and benefit reduction rates; (2) requirements to work or participate in work-related activities (and sanctions for failing to satisfy those requirements); (3) time limits on the receipt of benefits; (4) family caps and minor residence requirements; and (5) parental responsibility requirements. In the case of the second set of policies, requirements to work or participate in work-related activities, we also consider variation in program content or approach, such as the human capital development model (i.e., providing additional education and training before urging work) and work-first model (i.e., encouraging work immediately).[4]

[4]We also considered other reform policies, such as diversion and resource policies (e.g., individual development accounts, and auto and other resource limits that determine eligibility). However, we are not aware of any studies that can be used to identify the causal impact of these specific policy elements on the outcomes we consider. Rather, in the studies we synthesize, these elements are bundled together with other policies so that only their collective impact can be assessed.

Ideally, the goal of the synthesis would be to fill in each of the cells in the matrix, expressing in a common format how each policy affected each outcome. However, this goal is not realistic, because the literature that we synthesize here has not yet covered each policy-outcome pair. Moreover, some of the cells in the matrix are "populated" by studies that, for one reason or another, are too tentative or otherwise inconclusive to be relied on for policy purposes. Nevertheless, the empty cells provide useful diagnostic information: They tell us where additional research results would help in making policy decisions and where we do not have enough knowledge.

1.2.2. How We Completed the Matrix

The core of this synthesis is a review of the literature for each policy-outcome pair. We began by cataloging the existing literature. For each policy-outcome pair, we produced a table listing the studies, their key characteristics, and their findings. We considered the following questions: How did each study characterize the policy? What data were used to measure outcomes? What were the results?

The next step was to synthesize the literature. Where possible, we put the estimates on a common scale. We noted where different estimates resulted from different subpopulations or time periods, different data sources, and different methods. We then critiqued the studies, from a methodological perspective, asking the following questions: What methods were used to estimate causal effects? How did the analysis control for confounding factors? How confident should we be that the estimates recover the true causal effect of the program or policy on the outcome? In light of these various considerations, should some studies receive more weight in our synthesis than others? Given the literature as a whole, how confident should we be about our conclusions? Chapter 3 provides more discussion on these methodological issues.

Finally, we considered directions for future research, including the following questions: What research issues, that is, which policy-outcome pairs, are in particular need of attention? What methods should be used to address those issues? Is more time or different data needed to produce policy-relevant results?

1.3. ORGANIZATION OF THIS DOCUMENT

This document is organized into two parts: (1) two chapters that discuss broader issues relevant to each of the synthesis chapters; and (2) the core synthesis chapters themselves. In the first part, Chapter 2 considers the various welfare reform policies and programs, providing both historical context and basic information about specific policies and programs. It also presents the theoretical framework that predicts how the reforms should affect behavior and that serves as a guide to synthesizing the results. Chapter 3 considers the methodological challenges involved in implementing the framework described above, focusing in particular on methods for causal inference and issues with data sources.

The core of this document is found in the second part: Chapters 4 through 10, which are organized around the columns (outcomes) in Table 1.1. Specifically, Chapters 4 through 6 consider conventional welfare-related outcomes—welfare use, employment and earnings, and use of other government programs, respectively. Chapters 7 through 10 consider broader outcomes—family structure (fertility and marriage), household income and poverty, other measures of well-being, and child outcomes, respectively. Within each of these core chapters, we begin with a review of the facts: What is the level of the outcome? How has it varied over time? We then present a discussion of data and methodological issues specific to these outcomes. Having discussed these preliminaries, we then present tables with summaries of research results and provide a narrative synthesis. Finally, Chapter 11 provides a brief summary and discusses some directions for future research. Two appendixes are included as well, one providing results for different population subgroups analyzed in a subset of the studies, and another documenting our method for weighting studies in the synthesis.

Chapter Two

CONTEXT FOR UNDERSTANDING FEDERAL AND STATE WELFARE POLICY REFORMS

In this chapter, we provide some context for understanding federal and state welfare policy reforms, beginning with a very brief historical overview of federal welfare policy. We then introduce the economic model that provides a framework for our synthesis of the literature. Finally, we discuss four broad categories of welfare reforms and use the economic model to discuss their expected effects on behavior.

2.1. A BRIEF HISTORY OF FEDERAL WELFARE POLICY[5]

Title IV-A of the 1935 Social Security Act established the Aid to Dependent Children (ADC) program, renamed Aid to Families with Dependent Children (AFDC) in a 1962 amendment to the Social Security Act. Where previously there had been no federal program of poor support, ADC/AFDC was a joint state-federal program:

> designed to release from the wage earning role the person whose natural function is to give her children the physical and affectionate guardianship necessary . . . [to] rear them into citizens capable of contributing to society.[6]

Initially, the program was designed to serve needy children under 16 with only one able-bodied parent at home. Over time, other family members—e.g., older children (1940), the mother or other caretaker (1950), the child of an unemployed parent and that parent (1961), an unborn child in the last trimester of pregnancy (1981)—became eligible for benefits.

For whatever eligibility rules were in effect, the federal government defined the basic program as an entitlement. All eligible individuals who applied had to be enrolled, sufficient funding had to be found to pay for the benefits, and recipients were entitled to certain due-process protections. The states set the standard of "need" and the payment level and operated the program with state (or sometimes county) employees. Funding was provided by both the state and federal governments in proportions that varied across the specific programs and across the states.

---

[5]This section draws on various editions of the *Green Book* prepared by the U.S. House of Representatives Committee on Ways and Means.

[6]As quoted in Garfinkel and McLanahan (1986), Chapter 4.

In 1960, the official program goals were revised to include work for recipients, and beginning in 1968, the Work Incentive (WIN) program was established to make work or training programs available for certain AFDC recipients. Those programs applied to mothers whose youngest child was age 6 or older and were for the most part small and ineffectual. However, in the early 1980s when welfare agencies received additional authority to operate their own work and training programs, these programs became the object of considerable policy innovation and careful evaluation. Those innovations and evaluations served as a research base for the 1988 Family Support Act (FSA).

The FSA reinforced this emphasis on self-sufficiency through the mutual efforts of the recipient and the government. On the work dimension, the FSA replaced WIN and other welfare-to-work programs with the Job Opportunities and Basic Skills Training (JOBS) program. JOBS increased funding and strengthened the participation mandate for welfare-to-work programs and included financial penalties for recipients who failed to comply. The age-of-youngest-child exemption was lowered from 6 to 3 (or as low as 1, at state option), reflecting, in part, increase in labor force participation among nonwelfare mothers with pre-school-age children. During this period, with the increased emphasis on self-sufficiency, the program rules also changed to provide some financial incentives for work by allowing working recipients to keep more of their earnings. In practice, implementation of the FSA and JOBS was crippled by the recession of the early 1990s. The combination of a worsened job market for welfare recipients and a cash crunch in many state governments resulted in an undersized and underfunded program.

Beginning in the late 1980s, the waiver authority granted to the Secretary of Health and Human Services under section 1115 of the Social Security Act was utilized to grant waivers to particular rules and regulations governing state implementation of AFDC. States could petition USDHHS to implement experimental, pilot, or demonstration projects they believed would result in a more effective welfare program. These experiments were required to be cost-neutral and to include a rigorous evaluation (usually random assignment). By the time of PRWORA's passage, USDHHS had approved waivers for more than 40 states, many of them for statewide reforms (USDHHS, 2000, p. 259).

In August 1996, PRWORA passed and was signed into law, replacing AFDC with TANF. Beyond providing aid so that needy children could be cared for in their homes, the key objectives of TANF included ending the dependence of needy parents on government benefits by promoting job preparation, work, and marriage; preventing and reducing the incidence of out-of-wedlock pregnancies; and encouraging the formation and maintenance of two-parent families. In implementing these objectives, PRWORA eliminated many federal requirements on state welfare programs and allowed states nearly complete discretion in setting eligibility standards. The entitlement status of welfare was abolished, and federal funding of the AFDC and JOBS programs (as well as Emergency Assistance to Needy Families) was consolidated into a single TANF Block Grant. The TANF Block Grant was funded at 1994-95 spending levels. Thus, as caseloads and cash assistance declined, more funds were available for services that were to be provided to fewer cases.

During the year following passage of PRWORA, almost all states started the process of replacing their AFDC programs with their TANF program. Implementing those programs often stretched well into 1998, and refinements continue to be made. In most states, the changes have been substantial. Exploiting their new discretion and block-grant funding, most states changed their welfare programs beyond what was required by TANF (Gais et al., 2001). As a result, welfare policies that a decade earlier had varied across states primarily by the size of the welfare benefit now vary along dozens of dimensions.

2.2. KEY POLICY REFORMS UNDER WAIVERS AND TANF

Policy reforms under waivers, and even more so under TANF, change many different aspects of the basic AFDC/JOBS model. Here, we review the four changes that have received the most attention in the research literature: (1) the financial work incentives implicit in the benefit structure, (2) mandates to participate in work-related activities, (3) time limits, and (4) restrictions on living arrangements for minors and family caps. Many of these reforms were initially instituted under section 1115 waivers during the mid-1990s; some represent extensions of policies implemented under prior reform legislation.[7]

Our discussion begins with a brief review of the standard economic model of welfare programs. We then use this theory in our discussion of the likely effect of each of these reforms. Specifically, for each of these reforms, we briefly describe the policies that were in place at the beginning of the reform era. We then describe the range of policies that states adopted under waivers and TANF

2.2.1. The Standard Economic Model

The economic model of the consumer's choice of labor supply is useful for integrating the results of the studies. The model begins by noting that the typical consumer (potential welfare recipient) would prefer both more leisure and more income.[8] More work would yield greater earnings, but at the expense of less leisure. As a result, the consumer faces a trade-off. Her decisions about whether to go on welfare and whether and how much to work will depend on her wage and the structure of the welfare system. [9]

This simple model predicted behavior fairly well in the pre-reform environment. Under AFDC, the key policy parameters were the size of the welfare grant and the benefit reduction rate, which is the implicit tax rate by which the recipient's grant is reduced as her earnings rise. The standard economic model predicts that a larger grant should increase welfare use while reducing employment and hours of work. It also predicts that a lower benefit reduction rate should increase employment. These predictions have been largely borne out by a substantial body of systematic empirical evidence (Moffitt, 1992).

The model also predicts how consumers should respond to some of the recent reforms; the effects of other reforms can be predicted by extending the basic model. As we discuss the

---

[7]A variety of other reforms were approved under waivers and have been incorporated into state TANF programs. These include removing restrictions on eligibility for two-parent families, increasing asset limits that determine eligibility, instituting parental responsibility requirements (e.g., child immunizations and other preventative health care, and child school attendance), and providing extended transitional child care and health insurance after welfare exit.

[8] "Leisure" here is used to mean all activities other than work. It includes activities referred to in the vernacular as leisure, but also such activities as household work. In terms of the model, what is important is that the activities referred to as leisure be preferable to work, all else equal, and not be paid.

[9]Throughout, we use the feminine pronoun in referring to welfare recipients. The overwhelming fraction of adult welfare recipients are female.

recent reforms, we note how the model predicts that they should affect behavior, or how the model can be modified to incorporate their effects. The model serves an important function by providing us with a framework for synthesizing the literature—that is, for distilling a whole from the parts represented by several dozen independently conducted studies.

2.2.2. Financial Work Incentives

The debate about financial work incentives dates back to the 1967 Amendments to the Social Security Act. Prior to 1967, welfare benefits were reduced by one dollar for each dollar that the recipient earned, so increased work did not result in increased income. This benefit structure discouraged work. The 1967 Amendments sought to encourage work by allowing recipients to keep the first $30 of their monthly earnings (referred to as a $30 “earnings disregard”) and by reducing the benefit reduction rate for additional earnings to 67 percent (i.e., the recipient retained one-third of each additional dollar of earnings). Subsequent legislation—1981 Omnibus Budget Reconciliation Act (OBRA), 1984 Deficit Reduction Act (DEFRA), and 1988 FSA—further modified these formulas, but the incentives remained approximately constant.[10] Beyond some minimal threshold, a welfare recipient who worked kept at most a third of her earnings. From OBRA 1981 until the eve of PRWORA, after the fourth month of work, the benefit reduction rate was 100 percent.

States experimented with financial work incentives under waivers and incorporated them into their TANF plans. Although a few states maintain the AFDC incentive schedule, most now have more generous schemes in place. At least one state (Connecticut) now allows a recipient to keep all of her earnings up to the federal poverty level without losing any benefits.

The economic model discussed above predicts that introducing a financial incentive may have complex effects on behavior. Moreover, the effects may vary between consumers receiving welfare and consumers who were originally income-ineligible prior to the introduction of the incentive. The effect of the financial incentive on welfare recipients is relevant for interpreting the results from the random assignment experiments. Its effect on both groups is relevant for interpreting the results from econometric studies. We discuss these two effects in turn.

The financial incentive increases the return to an hour of work. As a result, it should encourage welfare recipients to begin working. Its effect on the labor supply and earnings of recipients who were already working prior to the policy change is ambiguous. In addition to increasing the return to an hour of work, introducing the financial incentive raises the recipient’s take-home income for a given number of hours of work. The effect on the return to work, known as the “substitution effect“ in the economics literature, provides an incentive to increase work hours. The effect on take-home income, known as the “income effect,“ provides an incentive to decrease work hours. Because the substitution and income effects work in opposite directions, the net effect of a financial incentive on labor supply and earnings of welfare recipients is ambiguous (Blank, Card, and Robins, 2000). Most observers would expect the substitution

[10]Among the changes with OBRA 1981 was a limit to work expenses of $75 per month plus actual child care costs up to $160 per child. In addition, the $30 plus one-third formula applied only to the first four months of work. DEFRA 1984 extended the $30 disregard to 12 months, but kept the one-third disregard at only four months. The 1988 FSA increased the limit on work expenses to $90 and the limit on child care to $175 per child over age two and $200 per child under age two.

effect to dominate for low-income families, which would imply that stronger financial work incentives should increase labor supply and earnings among welfare recipients. However, we find some cases where earnings appear to decrease, suggesting that the income effect may dominate in some cases.

Beyond its effects on work, the financial incentive should increase the welfare use of welfare recipients. Because the financial incentive allows the recipient to keep more earnings while remaining on aid, it provides an incentive to remain on aid longer. It is important to note that the incentive to remain on welfare arises from an incentive to combine work and welfare; financial work incentives should decrease the fraction of the caseload that receives aid without working.

The effects that the financial incentive may have on originally ineligible workers will depend on how the incentive is implemented. Expanding the financial incentive increases the level of earnings at which a working family becomes ineligible for aid. If this new eligibility threshold is applied to both new welfare applicants and ongoing recipients, then the increase in the eligibility threshold will make some previously ineligible families newly eligible to receive welfare. Such families may reduce their labor supply and start receiving benefits.[11] If the new eligibility threshold is applied only to ongoing recipients, as is the case in many states, then previously ineligible workers would remain ineligible.

In summary, the model predicts that, among welfare recipients, a financial incentive should increase employment and increase welfare use by raising the fraction of recipients who combine work and welfare. Its predicted effects on hours of work, earnings, and income are ambiguous because of opposing income and substitution effects. If the new eligibility threshold is applied to new applicants, then the increase in welfare use should be greater than it would be if the new eligibility threshold applied only to ongoing recipients. Applying the new threshold to new applicants would also make the program more likely to reduce labor supply and earnings.

Finally, there may also be a reporting effect, sometimes referred to as “smoke-out.” Financial work incentives may cause some recipients to report some previously unreported earnings in order to reduce the risk of being caught cheating. Work requirements, which are discussed in the next subsection, may also have such effects. However, smoke-out does not reflect a change in work effort, but merely a change in reporting. This implies that some of the recent improvement in employment and earnings may be apparent rather than real. Despite its potential importance in interpreting recent labor market trends, we are unaware of any evidence on the magnitude of the smoke-out effect.

2.2.3. Mandatory Work-Related Activities and Sanctions for Noncompliance

While work requirements date back to the WIN program implemented in 1968 and the Community Work Experience Programs (CWEP) created under OBRA 1981, the baseline for recent work-related activity mandates was the 1988 FSA’s JOBS program. JOBS was intended to

[11]Indeed, some families that remain just above the new eligibility threshold may reduce their hours (and earnings) in order to qualify for benefits.

be a mandatory program, exempting primarily single parents with children under age three.[12] Nonexempt welfare recipients were required to participate in work-related activities, which often involved basic skills programs. Those who failed to participate were subject to sanctions, which involved forfeiting the adult's portion of the AFDC benefit. However, less than half of the welfare caseload was required to participate in JOBS, and actual participation rates were much lower.

Because JOBS' work mandates were widely perceived as too weak, many states strengthened their work-related activity mandates under waivers. Common modifications include higher hours requirements, more restrictive definitions of permissible work-related activities, and lower age-of-youngest-child-exemptions. Many states reoriented their welfare to work programs as well, emphasizing job search and employment (so-called "work-first" programs) over basic skills and education.

TANF accelerated and in some cases required such changes. Under TANF, states were required to increase the fraction of their caseload participating in work-related activities. Furthermore, TANF limited the extent to which education and training could be used to satisfy this requirement. In practice, the sharp caseload decline and TANF's caseload reduction credit—which implicitly treated any fall in the caseload not related to stricter eligibility rules as participation—rendered these aggregate participation rate requirements largely meaningless.

The effects of a perfectly enforced work-related activity mandate can be predicted by extending the standard economic model to incorporate the requirement that welfare recipients spend the prescribed amount of time engaging in permissible work-related activities. The mandate is obviously predicted to increase participation in such activities, and to the extent that the permissible activities include working, they may increase employment as well. Work-related activity mandates should also decrease welfare use. Given most states' benefit structures, an increase in work will raise earnings. Especially in low break-even states, work will yield income high enough to make a family income ineligible for welfare. Furthermore, earnings near that level may induce a family to leave welfare—the small payment is not worth the effort and stigma of remaining on welfare. In addition, by reducing leisure—and the utility derived from that leisure, work-related activity mandates may induce some recipients to find a job and others to leave welfare for other sources of support (e.g., family or friends).

In the more realistic case of imperfectly enforced work mandates, these effects may be muted. The extent to which they are muted will depend on the extent of enforcement. The less strictly the mandate is enforced, the less pressure recipients will feel to participate in the work-related activities and the less likely they will be to leave the welfare rolls. In the real world, enforcement is likely to be important in determining the effects of mandated work-related activities.

Enforcement can take various forms. Monitoring is one method that has been used to achieve compliance with the mandates. Sanctions are another. Most states have raised sanctions for noncompliance, both for first and subsequent violations. Some states now have full-family sanctions for initial violations, and, in some states, third violations can lead to lifetime case closures.

[12]Beyond mothers of children under age three and those working at least 30 hours a week, the other exemptions included the ill, the disabled, those over 60, those living in remote areas, those needed in the home, those in their last trimester of pregnancy, and those for whom guaranteed child care was not available.

Sanctions can also be incorporated into the standard model using insights from deterrence theory, which has long been used to study street crime, insider trading, and lesser forms of rule-breaking such as shirking on the job (Becker 1968; DeMarzo, Fishman, and Hagerty, 1998; Dickens et al., 1989). In deterrence theory, the consumer decides whether to comply with a rule as a function of two things: the likelihood of being detected if she fails to comply, and the penalty associated with detection. The model predicts that compliance will rise as the detection risk rises and as the penalty for noncompliance increases. In the welfare context, we would expect higher detection risk and higher penalties for noncompliance to increase participation in work-related activities, increase employment, and decrease welfare use. Of course, some families who fail to comply will be detected and sanctioned, which will further decrease the caseload.

2.2.4. Time Limits

Time limits, which reduce or eliminate a recipient's benefits after a specified amount of time on aid, are among the most radical of the reforms instituted under TANF.[13] AFDC was an entitlement under which families were to receive aid for as long as they remained otherwise eligible. Prior to TANF, however, about a dozen states implemented waivers to time limit cash assistance. Consistent with this trend, with a few exceptions, TANF mandated a 60-month maximum time limit for adult receipt of federally funded benefits.

Although about half of the states have imposed the federal 60-month limit, many have adopted shorter limits. Some specify an intermittent limit, for example, allowing for only 24 months on aid within any 60-month period. Some have both intermittent and lifetime limits. Michigan and Vermont effectively have no time limit, having chosen to fund aid receipt in excess of the federal limit from state-only funds.[14] States also vary as to their exemption and extension policies.

Time limits may have complex effects on both welfare use and other dimensions of behavior. Like financial work incentives, they have mechanical effects, reducing welfare use and caseloads as families reach the time limit. They may have behavioral effects as well, leading families to exit the welfare rolls before they actually exhaust their benefits to preserve ("bank") their months of eligibility for future use.

Grogger and Michalopoulos (1999) have extended the standard economic model to include time limits. Their model predicts that time limits should have the greatest behavioral effects among the families with the youngest children. The reason is that these families have the longest time before their youngest child turns eighteen, making them ineligible for aid. Thus, they have the longest time over which to spread their benefits, giving them the greatest incentive to save some for the future.

---

[13]Although the term "time limits" also has been used to refer to the deadlines by which recipients must satisfy work-related activity mandates, we reserve the term for what has also been referred to as "termination" or "benefit-reduction" time limits.

[14]There is some ambiguity about whether and how New York will continue benefits after the time limit.

2.2.5. Minor Residence Requirements and Family Caps

PRWORA did not focus only on work. PRWORA specifically called attention to family structure, with explicit goals of reducing the incidence of out-of-wedlock pregnancies and encouraging the formation and maintenance of two-parent families. It sought to solve the problem created by the perception that AFDC encouraged young women to have children out of wedlock and discouraged them from marrying.

Beginning in the waiver period and continuing with increased intensity into the TANF period, states took several actions to address these concerns. First, the FSA gave states the option of requiring minors to live with an adult (usually a parent, but also with a guardian, or in a supervised setting) in order to received benefits. These provisions are commonly referred to as "minor residence rules." By the time of PRWORA's passage, 13 states had adopted such provisions as part of their state TANF plans and another 11 states had adopted these provisions under waivers. PRWORA mandated such provisions for each state's TANF program.

Similarly, under AFDC, welfare payments increased with family size. It was claimed that this provision gave welfare mothers an incentive to have more children. Under a "family cap," the welfare benefit does not increase with the birth of a new child who is conceived while the mother is receiving cash assistance. Beginning with New Jersey in 1992, 14 states received waivers for such a provision. Under TANF, 16 states provided no increases in benefits with an additional child, 2 states provided only a partial increase, 3 states provided the increase in the form of a voucher, and 2 states had flat grants independent of the number of children for all families (regardless of whether a child was born while the mother was receiving welfare).

Although family caps and minor residence rules were the most common reforms to address concerns about family structure, they were not the only ones. Other reforms included changing the eligibility rules for two-parent families (what had been the AFDC-Unemployed Parent program) to lessen disincentives to marriage and closer links to family planning programs. However, these other reforms have received little study and, therefore, receive less attention in the synthesis that follows.

While the primary objective of these policies is to reduce teen childbearing and subsequent pregnancies among welfare participants, these policies may reduce welfare receipt by reducing nonmarital births (especially first births) and more generally by making welfare less attractive. Such policies may therefore reduce welfare entry, as well as increase welfare exits.

Chapter Three

METHODOLOGICAL ISSUES

Synthesizing the literature on the causal effects of the reform policies described in the last chapter is the goal of this report. To do so, we first must assess the quality of each individual study addressing a particular topic (i.e., a policy-outcome combination). We then must assess the quality and quantity of the entire body of evidence available on that topic. Three methodological issues influence our assessment of individual studies. The first issue concerns the methods that a study uses to draw causal inferences. The second issue concerns the nature and characterization of the policy variation that the studies use to estimate the effects of the policy. The third issue concerns the data used to measure the outcomes of interest.

In this chapter, we first discuss these factors and how they contribute to our assessment of individual studies. We then provide summary information about the studies we include in our synthesis. We conclude by discussing how we weigh the results from multiple studies to draw conclusions about the effects of a particular policy on a particular outcome.

3.1. METHODS OF CAUSAL INFERENCE IN INDIVIDUAL STUDIES

This synthesis aims to answer the question: What is the effect of a given policy (e.g., a lower benefit reduction rate or a time limit) on a given welfare-related outcome (e.g., the caseload or child development), *holding all else equal*? To do this, we review studies that attempt to assess the causal effects of welfare reform. Although we restrict our attention to causal studies, it is important to note that not all welfare reform studies attempt to assess causation. Nor are causal studies the only types of analyses that are useful for assessing the success of welfare reform.

Examples of some important noncausal studies are the leavers studies that track the behavior of families that have left the welfare rolls. The results from 15 such studies, most of which were funded by USDHHS, are summarized by USDHHS (2001a). Those results show that 17 to 38 percent of leavers return to the welfare rolls within one year after their exit, and between 62 and 90 percent work at some time during that same year. On average, leaver families have post-exit incomes that are similar to the incomes that they had while on aid.[15]

[15]In addition to information about leavers' welfare use, employment, and income, USDHHS (2001a) summarizes results regarding leavers' earnings, use of other government programs, and other forms of material hardship. Other leaver studies include Ahn et al. (2000), Coulton and Verma (2000), Du et al. (2000), Fogarty and Kranley (2000), Foster (1999), Julnes et al. (2000), Loprest (2000), Loprest and Acs (2000), Midwest Research Institute (2000), Moses and Macuso (1999), Rockefeller Institute (1999), Ryan et al. (1999), Verma (2000), Westra and Routley (1999, 2000). For

Such studies provide information that is essential for monitoring the status of families that leave aid. However, they do not provide estimates of the causal effects of welfare reform. They cannot compare the behavior we *do observe* to the behavior we *would have observed* if reform had not taken place. In the language of the evaluation literature, they provide no "counterfactual" against which to assess the effects of reform.

We focus this synthesis on the causal effects of welfare reform because this is what a policymaker needs to know to make informed policy choices. For example, suppose Congress considers eliminating the federal time limit. In considering such a policy change, members of Congress would want to know how welfare-related outcomes would differ between two different policy scenarios: (1) a baseline scenario that leaves the time limit in place and (2) an alternative scenario that eliminates the time limit. When making the comparison, the legislator would want to hold everything else constant, such as the effect of the economy.

This thought experiment defines what we mean by the effect of a policy, but it does not tell us how to measure it. The reason is that we do not observe the outcome under the counterfactual scenario. Rather, we observe only the outcome corresponding to the policy that was actually chosen. The challenge facing the analyst is to devise a research design that predicts what outcomes would have been under the counterfactual policy.

This requires the researcher to design and implement a research strategy that holds all else constant. If the researcher fails to hold constant other factors that could independently influence the outcome, such as the economy, then the resulting estimates of the effects of the policy may be misleading. In the evaluation literature, such estimates are termed biased or inconsistent. They may reflect not only the effect of the policy of interest, but also the effects of the other factors. These other factors that could yield misleading results are referred to as confounding influences, or simply confounders.

The literature on the effects of welfare policies has adopted two general research strategies for dealing with confounding influences: random assignment and econometric analyses of observational data. Both of these methods have strengths and weaknesses. Because both approaches contribute to our understanding of the effects of welfare reform, we discuss each of them in some detail. In the core synthesis chapters, we consider evidence from both types of studies.

3.1.1. Random Assignment

One attractive approach to the problem of confounding factors is random assignment.[16] Rather than relying on existing variation in policies or programs, the analyst induces random variation. To test a new program, a study population is chosen, such as all persons receiving aid at a particular time. Then, each member of the study population is assigned either to the control group, which is subject to the baseline policy environment, or the treatment group, which is

summaries of these and other leaver studies, see GAO (1999c), Committee on Ways and Means (2000), Acs and Loprest (2000), Isaacs and Lyon (2000), and Cancian et al. (1999a, 1999b, 2000).

[16]For a discussion of random assignment in the social science context, see Burtless (1995) and Heckman and Smith (1995).

subject to the new policy environment. The assignment is determined by the logical equivalent of a coin toss.

In principle, this approach holds everything constant except the policy whose effect the analyst seeks to estimate. Since families assigned to the new program differ from those assigned to the baseline program only by a flip of a coin, confounding influences, such as the economy, should be identical for the two groups. If randomization is implemented properly, there should be no systematic differences across the two groups other than those attributable to the different policy environments. Thus, the average effect of the policy, which is referred to as the "treatment effect" or the "impact" of the policy, can be estimated by the difference in mean welfare-related outcomes between the two groups.[17]

Such random assignment experiments can be a powerful evaluation tool. The crucial importance of controlling for confounding factors and the potential of random assignment for doing so led ACF-USDHHS to require random assignment evaluations as a component of section 1115 waivers.[18] We review many of the studies emerging from such waiver evaluations in the chapters that follow.

Despite their advantages, however, random assignment studies have a number of disadvantages. First, random assignment evaluations can be conducted only when random assignment was performed at implementation. Random assignment is not always feasible. Even when it would be feasible, it is expensive and difficult to implement. As a result, random assignment is not always built into a program's implementation

Second, random assignment evaluations of welfare reform capture the effect of the new policy only from the point of randomization, almost always for women who are on welfare. Some reforms, however, such as work requirements and policies designed to affect fertility and family formation, are expected to deter people from ever using welfare. Since individuals deterred from entering the program will never be included in the study population, conventional random assignment evaluations will not capture the effects of reform on welfare entries. This is important because recent evidence suggests that more than half of the decline in the welfare caseload results from changes in entry rates rather than from changes in exit rates (Haider, Klerman, and Roth, 2001).

Third, random assignment experiments may not reproduce the environment of a universally implemented program. Broader implementations may affect labor markets or service providers (e.g., the capacity of educational institutions). Experiments may be more likely to be implemented in locations with above-average management capability and may attract the best managers. Thus, implementation in other sites or more broadly in a given site may yield smaller effects.

Fourth, random assignment studies are not immune to problems that can bias their findings. For example, experimental contamination may result when treatment group members "cross-over," by moving to a location that is not part of the evaluation or when control group members become eligible to receive the program services. Sample attrition for subsequent follow-up

[17]In practice, studies usually report more efficient results that use regression methods partially to control for the remaining (random) differences between the two groups.

[18]On the ACF-USDHHS experience with waivers, see Harvey, Camasso, and Jagannathan (2000).

data collection may be nonrandom and therefore bias the estimated treatment effects (Heckman, Smith, and Taber, 1998).

A final important problem involves participants' perceptions of the rules that apply to them. In several of the random assignment studies we summarize below, members of the treatment group were confused about which of the new policies applied to them. In others, members of the control group incorrectly believed that they were subject to the new reforms. This latter form of confusion is of particular concern in exceptional-control evaluation designs, where almost all the population is subject to the new "treatment" rules and only a small fraction of the study population is held back under the old "control" rules. In an environment where most recipients are subject to welfare reform and welfare reform receives considerable public and media attention, it may be difficult to persuade the controls that they are not subject to the new rules themselves.[19] Consequently, the control group may behave more like the treatment group, thereby biasing the estimated program impacts toward zero.

3.1.2. Econometric Methods for Observational Data

An alternative to random assignment is to analyze observational data, that is, to compare outcomes across different policy regimes (time-place combinations), typically using administrative data or national survey data. This is the primary method available to evaluate reforms that were not incorporated into the random assignment experiments. Unlike conventional random assignment, analyses of observational data can capture the effects of reform on welfare entries.

The key methodological problem with the analysis of observational data is that, while the policy environment will vary across observations, many confounding factors will vary as well. For policy analysis, we want to measure the effect of the policies, holding all else equal. To estimate that effect, we need to control for these confounding influences.

Regression analysis is the standard approach to this problem. To control for the effect of the economy, for example, econometric studies usually include the local unemployment rate as an independent variable in a linear regression model. In this way, standard regression methods control for *observable* confounding factors under the implicit assumptions that their effects are linear and additive. If these assumptions are correct and the observed variable (i.e., the unemployment rate) adequately represents the potentially confounding factor (i.e., the economy), then regression techniques eliminate any bias that could otherwise arise as a result of such a confounding factor.

Because many of the factors influencing welfare-related outcomes are difficult or even impossible to measure, the principal challenge facing econometric studies is controlling for *unobserved*, or *unmeasureable*, confounding influences. Now even more than before PRWORA, states determine their welfare programs. However, states differ in many ways besides their

[19]See Meyers, Glaser, and MacDonald (1998) on financial incentive changes as part of California's Work Pays Experiment. Harvey, Camasso, and Jagannathan (2000) note that at least some control group subjects in the section 1115 waiver demonstration studies believed they were subject to time limits, a family cap, or one of the other state welfare waiver provisions being evaluated. Similarly, Miller et al. (2000, Table B.1) report that many members of the Minnesota Family Investment Program (MFIP) treatment and control groups thought they were subject to time limits, even though MFIP did not involve time limits.

welfare programs, and some of those differences—in general attitudes or political sentiment, for example—may affect both welfare-related outcomes and welfare policy in the state. If such unobservable differences are not somehow controlled for, then the analyst may erroneously attribute changes in welfare-related outcomes to changes in welfare policy that are, in fact, attributable to unobserved factors. This problem of unobservable influences that may confound the relationship between welfare policy and welfare-related outcomes goes by many names in the research literature, including "unobserved heterogeneity," "policy endogeneity," "omitted variable bias," and "spurious correlation." It is the central threat to the validity of econometric studies.

In the literature on welfare reform, the standard approach to this problem is known as the difference-of-differences (DoD) method. DoD controls for two types of unobservable influences: those that vary between states, but are constant within a state over time; and those that vary over time, but similarly for all states. An example of a state-specific, time-invariant unobservable might be the state's general political leaning. An example of a uniform national trend might be the macroeconomic policy environment.

To illustrate how the DoD approach works, we consider a very simple example in which there are two states and two time periods. Both state 1 and state 2 have the baseline policy in the first period. State 1 adopts the new policy in the second period, but state 2 does not. The DoD estimator of the effects of the new policy compares the before-and-after change in welfare-related outcomes in state 1 to the before-and-after change in state 2. Because both before-and-after changes are computed within the respective states, both control implicitly for time-invariant, state-specific factors that could confound the relationship between the policy change and the welfare-related outcome. Because both changes are computed between the same two time periods, their difference (that is, the difference of differences) nets out the influence of any nationwide trends.

Most of the econometric literature evaluating the effects of welfare reform uses a generalization of this DoD approach. [20] In that generalization, this basic DoD insight is implemented in a multiple regression framework using dummy variables for each state and for each year. Because it controls for such state fixed-effects, the DoD estimator is often referred to as the "state fixed-effects" estimator. The multiple regression framework allows for more than two states and more than two periods. It also allows for the fact that states adopt new policies at different times. It can allow for state-specific linear (or quadratic) time trends to capture smoothly trending changes in a state over time. Finally, it allows the analyst to control explicitly for observable factors such as the unemployment rate and benefit levels.[21]

3.2. MEASURING THE POLICY ENVIRONMENT

Ideally, we would like to learn the effect of each individual policy embedded in the TANF reforms. Furthermore, we would like to learn the extent to which policies interact with each other, such that the effect of adopting two policies together is greater than (or less than) the

[20]On DoD, see Meyer (1995). See also the discussion in Moffitt and Ver Ploeg (1999).

[21]Because data from multiple states are necessary to implement the DoD procedure, we omit from the synthesis analyses based on data from single states, including Figlio and Ziliak (2000), Henry et al. (2000), and Klerman and Haider (2000).

sum of the effects of adopting each policy alone.[22] To do so, we need to observe outcomes when a single policy or a given bundle of policies is implemented. Several issues arise in measuring the policy environment that are relevant for both random assignment and econometric studies. In this section, we focus on issues associated with the nature and characterization of the policy environment in random assignment and econometric studies in turn.

3.2.1. Characterizing the Policy Environment in Random Assignment Studies

Random assignment studies are designed to measure the impact of the "treatment," that is, the program features that differ between the experimental and control groups. The experimenter thus controls the policy environment being evaluated through the design of the program. In the case of welfare reform evaluations, these program features include financial incentives to encourage work, requirements to participate in work-related activities, sanctions policies, parental responsibility requirements, and so on.

For a few of the random assignment studies we consider, the treatment consists of a single policy reform, such as a family cap or a parental responsibility requirement. Two studies employ dual-treatment designs. These involve two experimental groups (in addition to the control group), both of which experience financial work incentives and one of which additionally is subject to mandated work-related activities. The dual-treatment design provides information about the effect of financial work incentives and the incremental effect of the work-related activity mandate. However, most of the random assignment studies were implemented to evaluate multifaceted state waiver programs rather than specific reform policies. Therefore most studies involve a single treatment group that is subject to multiple policy reforms. Such designs shed light on the effects of reform as a bundle, although they generally cannot be used to estimate the impact of specific reforms.

While the policy or policies being evaluated in a random assignment study may be known in principle, another issue that affects experimental studies is program implementation. Programs with the same name are often implemented very differently in one place than they are in another. Insight into implementation issues is often provided by a process study involving analyses of program records, a caseworker survey, or a recipient survey. These process analyses might reveal that some locations have the management capacity to train and motivate employees to implement a new program, while other places appear to be less successful. One might argue that we should be interested in the "average" effect of implementation. There is concern, however, that the sites participating in demonstration programs have more management capacity (or use it more effectively in the demonstration sites) than would be the case for the average site trying to implement the reforms in the context of a statewide program. In the individual synthesis chapters, we sometimes note external

---

[22]We are also interested in how the effects of specific policies vary along other dimensions. Does a policy's effect vary across subgroups (e.g., whites versus blacks)? Does a policy's effect vary when the economy is good versus when the economy is bad? These interactions are potentially important (e.g., perhaps a work-based strategy will work when the economy is robust—because jobs are available—but would not be as effective when the economy experiences a downturn), but they are much harder to estimate. The synthesis chapters that follow find few studies that address these important issues.

evidence that suggests that some of the variation across sites may result from differences in the quality of implementation.

Finally, while our synthesis focuses on a set of random assignment studies implemented during the 1990s, primarily as part of implementing welfare waivers, the individual policy components or even the bundle of components evaluated in the experiments do not necessarily equate with the state-level reforms implemented under TANF. Thus, while the impact estimates from the evaluations may capture the effect of implementing a given reform or set of reforms in a given setting during a given time period, they may not generalize to the impact that we would expect to see for the TANF program as implemented in an entire state in the post-TANF era. The findings from the random assignment studies are most useful in demonstrating the direction and magnitude of the effects for a particular reform or group of reforms, but we must be more circumspect in the inferences we draw from these studies about the impacts of policies as actually implemented across the states.

3.2.2. Characterizing the Policy Environment in Econometric Studies

While random assignment studies generate their own policy variation to estimate the effects of the policy, econometric studies make use of the variation in policies that exists between states and over time. This requires the analyst to characterize the welfare policies that are in place in each state in each year. This has proven to be a difficult undertaking, in part because there are so many policy components to characterize and in part because a single policy component can vary along several dimensions. Moreover, those dimensions may be difficult to quantify and even more difficult to measure as actually implemented at the state or local level.[23]

The approach taken by most analysts has been to specify a policy change in terms of the date on which the policy (or policy bundle) was adopted. The analyst constructs a dummy variable that is equal to one after the policy is in place and equal to zero before the policy is in place.[24] The analyst then includes that dummy variable in her regression model. In the context of a DoD regression, the coefficient attached to such a variable indicates how much the welfare-related outcome changed in the state that adopted the policy once the policy was adopted, implicitly using states that did not adopt the policy change at the same time to control for unobservable confounding factors.

There are some ambiguities associated with this approach. For example, some analysts assume that the policy was in place once it was passed into law, whereas others assume that it was not in place until it was officially implemented.[25] In practice, the appropriate date is probably even later, when the program is rolled out, staff are trained, and news of the new environment

---

[23]Some states initially implemented their policy changes in only a portion of the state, which further complicates efforts to characterize the states' policies.

[24]When the new policy was introduced within a calendar year, the dummy variable is usually replaced by a variable measuring the fraction of the year during which the new policy was in place.

[25]Even when analysts agree on the appropriate way to conceptualize the policy environment, the devolution of welfare policy to the state—and even local—level has created challenges for researchers who want to assemble the required information about the timing and nature of specific reforms as adopted and implemented. Some of the information is recorded in official documents (e.g., waiver applications), and some has been collected in a coordinated fashion (e.g., the Urban Institute's Assessing the New Federalism project, and the State Policy Documentation Project of the Center for Law and Social Policy and the Center on Budget and Policy Priorities), but the characterization of programs as implemented remains incomplete.

reaches recipients. Such in-the-field implementation dates are difficult to operationalize, which explains why they have never been used in observational studies. As a result, estimates of program effects are likely to be too conservative (i.e., too small).

The main virtue of the dummy-variable approach is that it is simple and transparent. Analysts may disagree about whether the legal adoption date or the official implementation date best characterizes the date on which the policy was put in place, but such disagreements are narrow. Moreover, one can test whether the difference is important empirically.

However, an important drawback of using adoption dates alone to characterize reform policies is that they capture only one dimension of policy variation. They provide no information about other dimensions of variation that might have important effects on behavior. For example, many states implemented financial work incentives during the waiver period. Some states cut the benefit reduction rate from essentially 100 percent to 75 percent, whereas others cut them to 50 percent or less. Economic theory predicts that bigger incentives should have stronger effects on employment, but the dummy variable approach treats all financial work incentives as being equal. Thus, it misses an important dimension of policy variation.

Furthermore, the dummy variable approach may be more susceptible to confounding factors than policy variables that capture additional dimensions of variation. Dummy welfare reform variables tend to equal zero in the early part of the sample period and to equal one at the end of the sample period. Thus, they are correlated with trends, or more precisely, in the context of DoD regressions, with state-specific trends that deviate from national trends. If the analyst fails to control for such trends, the results may be biased estimates of the effects of reform, since the reform dummy is correlated with the trends.

One approach to this problem is to characterize the reform more completely. Rather than simply including a dummy variable, it is sometimes possible to include a variable describing the intensity of the reform (e.g., the size of a financial work incentive). In this case, we would expect to find not only an effect when the policy is adopted, but also a larger effect when a stronger form of the policy is adopted. The variation across states allows more precise estimates. In addition, the additional implication that the effect should vary with the strength of the reform can be tested.

Of course, defining other dimensions of variation is not always simple. An example is illustrated by several analysts' characterization of sanction policies. Starting in the waiver era and continuing into the post-PRWORA period, many states stiffened the sanctions they impose on recipients who violate their work requirements or personal responsibility mandates. In some states, an initial violation results in the adult's portion of the grant being deleted until the recipient comes into compliance with the requirement. At the other end of the spectrum, some states cancel the family's entire grant until the family comes into compliance (and in some states for a minimum duration). In many states, sanctions become more stringent for repeat offenders. Seven states impose lifetime, full-family sanctions for repeated noncompliance with work requirements, even if the family comes back into compliance.[26]

[26]One study suggests that as many as 540,000 recipients may have received full-family sanctions between 1997 and 1999 (Goldberg and Schott, 2000). However, that study fails to account for families that would have left welfare anyway. Thus it probably overstates the net effect of sanctions on the welfare caseload.

Four different sets of analysts have offered characterizations of states' sanction policies, coding them as lenient, stringent, or in between. The characterizations are shown in Table 3.1. The aspects of state policies used to rate the different states vary between analysts; as a result, the summary ratings vary to a considerable extent. Pennsylvania is a noteworthy case in point; its sanction polices are rated as lenient by two sets of analysts, moderate by another, and severe by another. Indeed, the four sets of ratings are in agreement for only 25 of the 51 states. This poses a problem for comparing results across studies. If analysts cannot agree on what a strict sanction policy is, the effects of a "strict" sanction policy may vary across studies for reasons that have more to do with measurement than with real behavior. Moreover, none of the rankings incorporates information about the monetary value of the sanctions. This is important because a partial-family sanction in a high-benefit state may result in the same financial penalty as a full-family sanction in a low benefit state. Likewise, none of the rankings incorporates information about the rate at which sanctions are imposed, which is shown in the last two columns of the table. As mentioned in Chapter 2, standard deterrence theory predicts that both the severity of the sanction and the probability of detection should affect behavior. This suggests that any characterization of sanctions that omits information about the likelihood that they are applied is incomplete.

Another obstacle to characterizing states' welfare reform policies is associated with implementation issues. Policy dummies, or even more detailed measures of policy characteristics, usually only capture variation in official statutes and regulations. However, the de facto variation in implementing a statute may be as important as the de jure variation in the actual statutes. For example, states with the same full-family sanction policy have varying numbers of people who have actually been sanctioned; states with similar time-limit policies vary in the fraction of people who receive extensions when they reach the time limit.

Finally, even assuming that the policy environment could be accurately captured for econometric analysis, the fact that most states implemented policies in bundles rather than individually poses an additional hurdle for statistical inference. This is similar to the problem of policy bundling in the random assignment studies. If policies are adopted together, there is less variation along each policy dimension to separately measure the effect of an individual policy. The limited number of post-reform state-year cells available for study, resulting both from the currency of the reforms and the lags in data release, makes it particularly difficult to distinguish the effect of one policy from that of another. As a result, econometric studies have typically been more successful in estimating the effects of reform-as-a-bundle than in estimating the separate effects of individual reforms.

Table 3.1—Four Characterizations of States' Initial Sanction Policies

State	Study: CEA (1999)	Study: GAO (2000)	Study: Burke and Gish (1998), as cited by Rector and Youssef (1999)	Study: Pavetti and Bloom (2001)	All measures agree?	Percentage under full-family sanction	Percentage under any sanction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alabama	Intermed.	Intermed.	Intermed.	Stringent		4.5	9.8
Alaska	Lenient	Lenient	Lenient	Lenient	Yes	0.0	4.3
Arizona	Intermed.	Intermed.	Intermed.	Intermed.	Yes	5.1	5.1
Arkansas	Stringent	Lenient	Stringent	Lenient		0.3	3.2
California	Lenient	Lenient	Lenient	Lenient	Yes	0.1	2.3
Colorado	Intermed.	Intermed.	Intermed.	Intermed.	Yes	0.0	5.4
Connecticut	Intermed.	Stringent	Intermed.	Intermed.		2.7	3.1
Delaware	Stringent	Stringent	Intermed.	Stringent		0.8	15.3
DC	Lenient	Lenient	Lenient	Lenient	Yes		7.4
Florida	Stringent	Stringent	Stringent	Stringent	Yes	0.3	1.8
Georgia	Stringent	Stringent	Stringent	Stringent	Yes	2.3	2.3
Hawaii	Lenient	Stringent	Lenient	Stringent		0.0	0.0
Idaho	Stringent	Stringent	Stringent	Stringent	Yes	1.1	1.1
Illinois	Intermed.	Intermed.	Intermed.	Intermed.	Yes	1.2	6.0
Indiana	Intermed.	Lenient	Lenient	Lenient		0.9	5.5
Iowa	Intermed.	Lenient	Intermed.	Stringent		0.6	3.3
Kansas	Stringent	Stringent	Stringent	Stringent	Yes	0.0	1.0
Kentucky	Intermed.	Lenient	Lenient	Intermed.		1.6	5.7
Louisiana	Intermed.	Stringent	Intermed.	Stringent		0.4	3.2
Maine	Lenient	Lenient	Lenient	Lenient	Yes	0.1	5.3
Maryland	Stringent	Stringent	Intermed.	Stringent		0.0	11.3
Massachusetts	Intermed.	Lenient	Intermed.	Stringent		4.7	4.7
Michigan	Intermed.	Intermed.	Intermed.	Stringent		2.4	4.5
Minnesota	Lenient	Lenient	Lenient	Lenient	Yes	0.3	7.6
Mississippi	Stringent	Stringent	Stringent	Stringent	Yes	0.0	0.9
Missouri	Lenient	Lenient	Lenient	Lenient	Yes	1.4	12.3
Montana	Lenient	Lenient	Lenient	Lenient	Yes	1.0	8.0
Nebraska	Stringent	Lenient	Stringent	Stringent		2.2	2.2
Nevada	Stringent	Intermed.	Intermed.	Intermed.		0.8	3.2
New Hampshire	Lenient	Lenient	Intermed.	Lenient		0.0	4.8
New Jersey	Intermed.	Stringent	Intermed.	Stringent		2.8	8.0
New Mexico	Intermed.	Intermed.	Intermed.	Intermed.	Yes	0.0	0.0
New York	Lenient	Lenient	Lenient	Lenient	Yes	0.0	0.0
North Carolina	Lenient	Lenient	Lenient	Intermed.		0.5	29.1
North Dakota	Intermed.	Intermed.	Intermed.	Stringent		7.0	7.0
Ohio	Stringent	Stringent	Stringent	Stringent	Yes	0.0	0.9
Oklahoma	Stringent	Lenient	Stringent	Stringent		0.0	2.2
Oregon	Intermed.	Intermed.	Intermed.	Intermed.	Yes	0.5	0.6
Pennsylvania	Stringent	Lenient	Lenient	Intermed.		0.0	6.3
Rhode Island	Lenient	Lenient	Lenient	Lenient	Yes	0.2	3.0
South Carolina	Stringent	Stringent	Stringent	Stringent	Yes	0.2	5.7
South Dakota	Intermed.	Stringent	Intermed.	Stringent		0.9	0.9
Tennessee	Stringent	Stringent	Stringent	Stringent	Yes	0.3	0.3
Texas	Lenient	Lenient	Intermed.	Intermed.		0.0	15.5
Utah	Intermed.	Intermed.	Intermed.	Stringent		0.0	4.0

Table 3.1—Continued

State	Study: CEA (1999)	GAO (2000)	Burke and Gish (1998), as cited by Rector and Youssef (1999)	Pavetti and Bloom (2001)	All measures agree?	Percentage under full-family sanction	Percentage under any sanction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Vermont	Intermed.	Lenient	Lenient	Intermed.			0.0
Virginia	Stringent	Intermed.	Stringent	Stringent		0.7	0.7
Washington	Lenient	Lenient	Lenient	Lenient	Yes	0.0	5.6
West Virginia	Stringent	Intermed.	Intermed.	Intermed.			0.0
Wisconsin	Stringent	Stringent	Stringent	Stringent	Yes	4.6	22.8
Wyoming	Stringent	Stringent	Stringent	Stringent	Yes	2.0	2.0
Total/Average					25	1.1	6.1

NOTES: Columns (6) and (7) are from GAO (2000) and pertain to 1998. The terminology used to describe the severity of sanctions differs among authors. Our "lenient" category corresponds to the categories described as "partial/partial" by CEA (1999); "partial" by GAO (2000); "weak" by Rector and Youssef (1999); and "lenient" by Pavetti and Bloom (2001). Our "intermediate" category corresponds to the categories described as "partial/full" by CEA (1999); "graduated" by GAO (2000); "moderate" or "delayed full-check" by Rector and Youssef (1999); and "moderate" by Pavetti and Bloom (2001). Our "stringent" category corresponds to the categories described as "full/full" by CEA (1999); "full-family" by GAO (2000); "initial full-check" by Rector and Youssef (1999); and "stringent" by Pavetti and Bloom (2001).

3.3. DATA SOURCES FOR WELFARE OUTCOMES

To use either randomized trials or econometric methods to estimate the effects of a policy or program, the analyst requires data on welfare-related outcomes. In this section, we review the most commonly used data sources and discuss their utility. We also discuss sample sizes and statistical power, issues that are relevant for both econometric and experimental analyses of welfare-related outcomes.

3.3.1. Data Sources

Tables 3.2 and 3.3 summarize the major sources of administrative and survey data, respectively. Randomized trials often abstract their own data from these administrative records and from their own surveys. Econometric studies usually analyze existing sources of administrative and survey data.

As seen in Table 3.2, administrative data cover many of the outcomes of interest, and, for welfare program participants, they cover the entire caseload at a given time. The drawbacks of administrative data include issues of data quality, lack of coverage of non-TANF participants (e.g., leavers) and eligible nonparticipants (e.g., those choosing not to enter), and limited information on individual socioeconomic characteristics. Certain outcomes, such as measures of child well-being, are not typically available in administrative data sources. There is also considerable variation in state-level data systems and in their suitability for research, as well as in the extent of historical information and cross-state comparability of data systems.

Table 3.2—Sources of Administrative Data for Analysis of Welfare Reform

Data Source	Outcomes	Coverage	Notes
State reports to ACF-USDHHS	Caseload	All states, pre- and post-TANF	Aggregate program counts; some aggregate information on distribution of caseload by demographic group
State reports to ACF-USDHHS	Work activities; participation rate	All states, post-TANF	Aggregate data only on total work activities and participation rates, and numbers in specific program components; some JOBS data available
State-specific welfare program data	Caseload; aid payments; sanctions; program activities	Within a single state; availability of historical data varies widely	Issues of data quality; systems are not consistent across states, making cross-state comparisons difficult
Unemployment insurance data	Employment; earnings	Within a single state; availability of historical data varies widely	Gaps in coverage; data relatively comparable across states and some cross-state efforts have been mounted; limited numbers of covariates (difficult to identify at-risk population)
Other social welfare program administrative data	Participation in Food Stamp Program, Medicaid, subsidized housing, child care subsidies, foster care, child support, etc.	Within a single state; availability of historical data varies widely	
Other administrative data (e.g., birth certificates)	Births, etc.	Nationwide (e.g., births) or state-specific; historical data varies	Welfare recipients not identified

Table 3.3—Sources of Survey Data for Analysis of Welfare Reform

Data Source	Outcomes	Coverage	Notes
Current Population Survey (CPS)	Program participation and income/poverty status in previous calendar year; employment and earnings at interview and in previous calendar year; family structure	Nationwide, relatively consistent survey content back to 1968	Sample size: 55,000 households; increasing in March 2001 and beyond
Survey of Income and Program Participation (SIPP) and Survey of Program Dynamics (SPD)	Monthly data for same outcomes as CPS, plus (monthly) program entry and exit	Nationwide, relatively consistent survey content back to 1984	Sample size: varies, about 30,000 households at any point in time; survey is a panel, following respondents for about two-and-a-half years
Panel Study of Income Dynamics (PSID)	Same as CPS, plus program entry and exit and measures of child well-being	Nationwide sample of families followed since 1968	Sample size: 4,800 families in original cohort augmented by split-offs to 8,000 families in 2001
National Longitudinal Survey of Youth (NLSY)	Same as CPS, plus program entry and exit and measures of child well-being	Nationwide cohort of youth followed since 1979	Sample size: about 11,500 youth in original cohort plus the children of the original cohort of young women followed since 1986
National Survey of America's Families (NSAF)	Similar to CPS plus hardship, housing, health status and health care use, attitudes, knowledge of service availability	Representative population in 13 states interviewed in 1997 and 1999	Sample size: about 44,000 households in repeated cross-sections
State surveys	Vary	Current and recent recipients	Details vary across states; issues with locating former recipients; limited cross-state comparability
Program evaluation surveys	Vary	Treatment and control groups	Details vary across evaluations

The major sources of survey data shown in Table 3.3 also cover many of the welfare-related outcomes of interest, often for large nationally representative samples observed both before and after welfare reform. These databases are typically rich in the socioeconomic information they contain, and they usually cover both program participants and nonparticipants. Some surveys track respondents over time so that the dynamics of behavior can be studied over both short and long horizons. The limitations of these databases include relatively small samples of welfare program participants, which is an acute problem for most longitudinal surveys; nonrandom survey nonresponse in cross-section surveys; attrition in panel surveys; and reporting errors for participation in many welfare programs. State-level survey data based on samples drawn from administrative records of welfare recipients can be hampered by problems with locating and tracking those no longer on aid.

3.3.2. Statistical Power

Whether the available data are sufficient to estimate policy effects precisely using econometric methods is the subject of some discussion in the literature. (See, in particular, Adams and Hotz, 2001.) The econometric studies discussed in this synthesis almost all use DoD methods. As such, they require that the outcomes be consistently measured across time, both before and after reform, and across states, and that there be enough observations in each state-year cell to construct at least a rough estimate of the outcomes of interest in the population of interest.[27]

These seemingly simple requirements make most conventional data sources unusable for econometric analyses. The requirement of consistent data across states rules out state-specific administrative data files. The requirement of consistent data across years rules out most single interview studies or studies that began after reform was under way. The requirement that it be possible to construct a rough estimate of the outcomes of interest in the population of interest rules out almost all other data sources. To understand this, note that welfare participation is a relatively rare behavior. The peak national welfare caseload (in persons) was about 15 million, and in early 2001 the figure was 5.5 million; in percentage terms, this is 5.5 and 2.1 percent of the population, respectively. Whether these numbers are “large” from a social perspective is part of the policy debate. However, from a survey research perspective, these are quite small numbers. Even a moderate-sized random sample of 10,000 households is likely to yield only a few hundred households with any welfare recipients. The resulting state-specific estimates of both the rate of welfare receipt and the rate of change of welfare receipt will be noisy (i.e., it will differ considerably from the true value because of sampling variability). Much of the variation will not be a result of variation in the true number of people receiving welfare, but instead of variation in who happens to be sampled, that is, to classical sampling variability.

Thus, the requirements for consistent national data, across multiple years and for large samples, appear to rule out all survey data except the CPS, the SIPP, and perhaps the NLSY and the PSID.

Finally, it is worth noting that power issues may also arise in the experimental evaluations, especially in the analysis of effects for population subgroups. While power calculations are

[27]Only a “rough estimate” is needed because, to some extent, statistical methods can be used to smooth over the sampling variability in individual cells.

used to ensure that the total sample in treatment and control groups is sufficiently large to detect an impact of a given size with a high probability, the likelihood of detecting impacts of the same size on smaller subgroup samples may be much smaller. Thus, detecting impacts in subgroups will require larger (often much larger) samples and thus a much higher cost of evaluation. Furthermore, some sites are simply not large enough to support the required samples. One approach for addressing this issue is to pool results across studies and then consider subgroup differences. This is the strategy adopted by Michalopoulos and Schwartz (2000).

3.4. SUMMARY OF STUDIES INCLUDED IN THE SYNTHESIS

As noted above, our synthesis draws on both random assignment evaluations and econometric studies. To better understand these studies, we review their key features. We begin with the details of the random assignment studies.[28]

3.4.1. Features of Random Assignment Studies

Tables 3.4 and 3.5 summarize the features of the experimental evaluations that we draw on for the synthesis and provide a useful reference for the discussion of these evaluations in the chapters that follow. We include only those studies that reasonably approximate the types of policies implemented under TANF. As a result, we exclude some evaluations that consider very specialized reforms such as those that focus on service delivery for teen parents on welfare or at risk of welfare participation (e.g., the Learning, Earning, and Parenting (LEAP) program, the New Chance program, and the Teen Parent Demonstration program), child support policies (e.g., New York Child Assistance Program), and specialized service delivery (e.g., the Postemployment Services Demonstration program). Furthermore, we exclude some of the earlier welfare-to-work experiments that predate the 1988 Family Support Act (e.g., San Diego's Saturation Work Initiative Model and the early GAIN experiments in several California counties). We also exclude Project Independence, which was Florida's initial JOBS program. It was the precursor to Florida's Family Transition Program (FTP), which we do include.

Table 3.4 describes basic features of the experiments such as the location of the demonstration, whether it was part of a statewide reform, the population served, the period of randomization and length of the follow-up, the sample sizes in treatment and control groups, the policy environment for the controls (typically AFDC/JOBS), and contextual information in the form of the unemployment rate and welfare benefit level. Table 3.5 provides details on the policy reforms applicable to the treatment group. This includes information on the central reform components of financial work incentives, mandatory work-related activities, and time limits, as well as other reforms such as sanctions, family caps, parental responsibility requirements, transitional child care and health insurance, changes in eligibility for two-parent families, and various other features (e.g., changes in asset limits and use of personal responsibility agreements).

---

[28]Since the econometric studies typically focus on one outcome, we defer a discussion of the methods for these studies to the synthesis chapters that follow. In contrast, the random assignment studies typically consider multiple outcomes so that the summary provided in this section serves as a reference for all the synthesis chapters.

Table 3.4—Selected Design Features of Random Assignment Studies Included in Synthesis

Name	State	Sites	Demo part of statewide reform?	Cases served	RA period start	RA period end	F/U length	Sample sizes: Total	T	C	Controls	Initial conditions in state (for RA start year): U rate (%)	Max $ grant (a)	Cite(s)
A. Programs that focus on financial work incentives														
California Work Pays Demonstration Project (CWPDP)	CA	3 counties	No	Single-parent recipients (b)	Oct 92	Dec 92	42 months	7,841	5,211	2,630	AFDC/ JOBS	9.3	663	Becerra et al. (1998) Hu (2000)
Welfare Restructuring Project Incentives Only (WRP-IO)	VT	6 welfare districts	Yes	Single-parent recipients and applicants (b)	Jul 94	Jun 95	42 months	2,196	1,087	1,109	AFDC/ JOBS	4.7	638	Bloom et al. (1998) Hendra and Michalopoulos (1999) Bloom, Hendra and Michalopoulos (2000)
Minnesota Family Investment Program Incentives Only (MFIP-IO)	MN	3 urban counties	No	Urban single-parent long-term (> 24 mos. in last 36 mos.) recipients	Apr 94	Mar 96	through 6/98	1,769	835	934	AFDC/ JOBS	4.0	532	Miller et al. (1997) Miller et al. (2000) Gennetian and Miller (2000)
				Urban single-parent recent applicants	Apr 94	Mar 96	through 6/98	3,113	980	2,133				
B. Programs that focus on financial work incentives tied to hours of work														
New Hope	WI	2 areas of Milwaukee	No	Poor families employed FT at RA	Jul 94	Dec 95	through 12/98	418	218	200	No New Hope benefits	4.7	518	Bos et al. (1999) Bos and Varga (2001)
				Poor families not employed FT at RA	Jul 94	Dec 95	through 12/99	935	459	476				
Self-Sufficiency Project (SSP) (c)	Canada (BC, NB)	Province-wide	No	Single-parent recipients	Nov 92	Mar 95	36 months	5,729	2,880	2,849	Traditional Income Assistance	10.5 (BC) 12.8 (NB)	1,131 (BC) 747 (NB)	Michalopoulos et al. (2000) Morris and Michalopoulos (2000)
Self-Sufficiency Project Plus (SSP-Plus) (c)	NB, Canada	lower NB	No	Single-parent recipients	Nov 94	Mar 95	18 months	596	293	303	Traditional Income Assistance	N.A.	N.A.	Quets et al. (1999)
Self-Sufficiency Project Applicants (SSP-A) (c)	BC, Canada	Vancouver and lower British Columbia	No	Single-parent applicants (no IA for at least 6 months prior to RA)	Feb 94	Feb 95	30 months	2,852	1,422	1,430	Traditional Income Assistance	N.A.	N.A.	Michalopoulos, Robins and Card (1999)

Table 3.4—Continued

Name	State	Sites	Demo part of statewide reform?	Cases served	RA period start	RA period end	F/U length	Sample sizes: Total	Sample sizes: T	Sample sizes: C	Controls	U rate (%)	Max $ grant (a)	Cite(s)
C. Programs that focus on mandatory work-related activities														
LA Jobs-1st GAIN	CA	Los Angeles County	No	Single parent recipients and applicants (b)	Apr 96	Sep 96	24 months	15,683	11,521	4,162	AFDC/ JOBS plus Work Pays (d)	7.2	607	Freedman et al. (2000b)
Atlanta Labor Force Attachment (LFA)	GA	Atlanta	No	Recipients and applicants	Jan 92	Jan 94	5 years (e)	2,938	1,441	1,497	AFDC/ JOBS plus "fill-the-gap" budgeting (f)	7.0	280	Freedman et al. (2000a) McGroder et al. (2000) Hamilton et al. (2001)
Grand Rapids Labor Force Attachment (LFA)	MI	Grand Rapids	No	Recipients and applicants	Sep 91	Jan 94	5 years (g)	3,012	1,557	1,455	AFDC/ JOBS	9.3	474	Same as above
Riverside Labor Force Attachment (LFA)	CA	Riverside	No	Recipients and applicants	Jun 91	Jun 93	5 years	6,726	3,384	3,342	AFDC/ JOBS plus Work Pays after late 1993 (d)	7.7	694	Same as above
Portland	OR	Portland	No	Recipients and applicants; no cases with substantial barriers	Feb 93	Dec 94	5 years	4,028	3,529	499	AFDC/ JOBS	7.3	460	Same as above
Atlanta Human Capital Development (HCD)	GA	Atlanta	No	Recipients and applicants	Jan 92	Jan 94	5 years (e)	2,992	1,495	1,497	AFDC/ JOBS	7.0	280	Same as above
Grand Rapids Human Capital Development (HCD)	MI	Grand Rapids	No	Recipients and applicants	Sep 91	Jan 94	5 years (g)	2,997	1,542	1,455	AFDC/ JOBS	9.3	474	Same as above
Riverside Human Capital Development (HCD)	CA	Riverside	No	Recipients and applicants, low education	Jun 91	Jun 93	5 years	4,938	1,596	3,342	AFDC/ JOBS plus Work Pays after late 1993 (d)	7.7	694	Same as above
Columbus Integrated	OH	Columbus	No	Recipients and applicants	Sep 92	Jul 94	5 years (h)	4,672	2,513	2,159	AFDC/ JOBS	7.3	334	Same as above
Columbus Traditional	OH	Columbus	No	Recipients and applicants	Sep 92	Jul 94	5 years (h)	4,729	2,570	2,159	AFDC/ JOBS	7.3	334	Same as above
Detroit	MI	Detroit	No	Recipients and applicants	May 92	Jun 94	5 years (i)	4,459	2,226	2,233	AFDC/ JOBS	8.9	459	Same as above
Oklahoma City	OK	Oklahoma City	No	Applicants	Sep 91	May 93	5 years (e)	8,677	4,309	4,368	AFDC/ JOBS	6.7	341	Same as above
Indiana Manpower Placement and Comprehensive Training Program (IMPACT) Basic Track	IN	Statewide	Yes	Recipients and applicants, less job ready	May 95	Dec 95 (j)	2 years (k)	3,856	3,090	766	AFDC/ JOBS	4.7	288	Fein et al. (1998)

Table 3.4—Continued

Name	State	Sites	Demo part of statewide reform?	Cases served	RA period start	RA period end	F/U length	Sample sizes			Controls	Sample sizes		Cite(s)
								Total	T	C		U rate (%)	Max $ grant (a)	
D. Programs that focus on financial work incentives and mandatory work-related activities														
Welfare Restructuring Project (WRP)	VT	6 welfare districts	Yes	Single-parent recipients and applicants (b)	Jul 94	Jun 95	42 months	4,376	3,267	1,109	AFDC/ JOBS	4.7	638	Bloom et al. (1998) Hendra and Michalopoulos (1999) Bloom, Hendra and Michalopoulos (2000)
Minnesota Family Investment Program (MFIP)	MN	3 urban counties (l)	No	Urban single parent long-term (> 24 mos. in last 36 mos.) recipients (b)	Apr 94	Mar 96	through 6/98	1,780	846	934	AFDC/ JOBS	4.0	532	Miller et al. (1997) Miller et al. (2000) Gennetian and Miller (2000)
				Urban single parent recent applicants (b)	Apr 94	Mar 96	through 6/98	4,049	1,916	2,133				
To Strengthen Michigan Families (TSMF)	MI	4 local service offices	Yes	Single-parent recipients (b)	Oct 92	Oct 92	4 years	8,739 (m)	4,462	4,277	Until 10/94: AFDC/ JOBS After 10/94: Modified AFDC/ JOBS	8.9	459	Werner and Kornfeld (1997)
				Single-parent applicants (b)	Oct 92	Sept 95	1 to 2 years	6,042 (m)	3,017	3,025				
Family Investment Program (FIP)	IA	9 counties	Yes	Recipients	Sept 93	Sept 93	14 quarters	6,684	4,461	2,223	AFDC/ JOBS	4.0	426	Fraker and Jacobson (2000)
				Applicants	Oct 93	Mar 95	8 quarters	6,009	3,973	2,036				
E. Programs that focus on other individual reforms														
Arkansas Welfare Waiver Demonstration Project (AWWDP)	AR	10 counties	N.A.	Recipients and applicants	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	Turturro et al. (1997)
Family Development Program (FDP) (n)	NJ	10 counties (o)	Yes	Recipients	.	Oct 92	through 12/96	4,875	3,268	1,607	AFDC/ JOBS	8.5	424	Camasso, Harvey and Jagannathan (1996) Camasso et al. (1998, 1999)
				Applicants	Oct 92	Dec 94	through 12/96	3,518	2,233	1,285				
Primary Prevention Initiative (PPI)	MD	6 welfare offices (4 urban, 2 rural)	Yes	Recipients and applicants	Jun 92	Aug 95	1 to 2 years	1,775 (p)	911	864	AFDC/ JOBS	6.7	377	Minkovitz et al (1999)
Preschool Immunization Project (PIP)	GA	Muscogee County	Yes	Recipients (q)	Nov 92	Nov 92	4 years	2,801 (r)	1,076	1,725	AFDC/ JOBS	7.0	280	Kerpelman, Connell, and Gunn (2000)

Table 3.4—Continued

Name	State	Sites	Demo part of statewide reform?	Cases served	RA period start	RA period end	F/U length	Sample sizes			Controls	Sample sizes		Cite(s)
								Total	T	C		U rate (%)	Max $ grant (a)	
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)														
Employing and Moving People Off Welfare and Encouraging Responsibility (EMPOWER)	AZ	3 in Phoenix 1 on Navajo reservation	Yes	Recipients (including those receiving TMA)	Oct 95	Oct 95 (s)	36 months (t)	2,934	1,476	1,458	AFDC/ JOBS	5.1	347	Kornfeld et al. (1999)
Indiana Manpower Placement and Comprehensive Training Program (IMPACT) Placement Track	IN	Statewide	Yes	Recipients and applicants, more job ready	May 95	Dec 95 (j)	2 years (k)	5,595	4,537	1,058	AFDC/ JOBS	4.7	288	Fein et al. (1998)
Virginia Independence Program (VIP) and Virginia Initiative for Employment not Welfare (VIEW) (u)	VA	3 cities: Lynchburg, Petersburg, and Portsmouth (v)	Yes	Recipients (w)	Jul 95	Jul 95	27 months	7,568	3,784	3,784	AFDC/ JOBS	4.5	354	Gordon and Agodini (1999)
A Better Chance (ABC)	DE	5 pilot offices	Yes	Single parent recipients and applicants (x)	Oct 95	Sept 96 (y)	max. 18 mos. (z)	3,959	2,138	1,821	AFDC/ JOBS	4.3	338	Fein and Karweit (1997) Fein (1999) Fein and Lee (2000) Fein et al. (2001)
Family Transition Program (FTP)	FL	Escambia County	No	Recipients and applicants	May 94	Feb 95	4 years	2,815	1,405	1,410	AFDC/ JOBS	6.6	303	Bloom et al. (1999) Bloom et al. (2000a)
Jobs First	CT	Manchester New Haven	Yes	Recipients and applicants	Jan 96	Feb 97	4 years	4,803	2,396	2,407	AFDC/ JOBS	5.7	636	Bloom et al. (2000b) Hendra, Michalopoulos and Bloom (2001) Bloom et al. (2002)

NOTES: Abbreviations: T=treatment; C=control; U=unemployment; BC=British Columbia; NB=New Brunswick; N.A.=not available; IA=Candian Income Assistance; FT=full-time; RA=random assignment; TMA=transitional medical assistance.

(a) For one adult and two children.
(b) Evaluation also includes sample of two-parent families with results reported separately.
(c) All monetary values in Canadian dollars.
(d) Under Work Pays, the earnings disregard was $120 and the BRR was 67 percent, and a higher needs standard was used for "fill-the-gap" budgeting.
(e) Controls became subject to treatment conditions beginning in the fourth quarter of 1996.
(f) A higher needs standard (equal to $424 in 1993 for a family of three) was used for "fill-the-gap" budgeting.
(g) Controls assigned before January 1993 became subject to treatment conditions three years after RA.
(h) Controls became subject to treatment conditions beginning in the fourth quarter of 1997.
(i) Controls became subject to treatment conditions three years after RA.
(j) Randomization scheduled to end in December 1999; evaluation includes participants in first 8 months.
(k) Those entering after June 1995 observed for up to 6 months after Basic and Placement track distinction was eliminated in June 1997.
(l) Demonstration also implemented in 4 rural counties with results reported separately.
(m) Sample sizes are for combined one- and two-parent families. 87% of recipient cases and 80% of applicant cases are one-parent families.
(n) FDP provisions phased in between October 1992 and October 1993.
(o) Implementation of the FDP provisions was delayed in two counties until January 1995. Some results only pertain to 8 counties with implementation by October 1993.
(p) Sample sizes refer to number of children age 3 to 24 months with complete medical records abstraction for analysis of vaccination status at 1- and 2-year followups; a larger sample of families were in the experiment.
(q) Reforms applied to recipients and applicants but only former group included in evaluation.
(r) Sample sizes refer to number of children up to age 6 with complete medical records abstraction; 2,500 families were in the treatment and control groups.
(s) Randomization continued for new applicants from November 1995 to July 1997; evaluation includes only recipients as of October 1995.
(t) EMPOWER REDESIGN implemented in August 1997 under TANF applied to both treatment and controls.
(u) VIP provisions implemented in July 1995 but VIEW provisions phased in at five demonstration sites between July 1995 and October 1997.
(v) Evaluation also includes 2 counties but staggered implementation means no exposure to new rules at time of follow-up.
(w) Evaluation also includes sample of applicants between July 1995 and September 1996 but staggered implementation means little exposure to new rules at time of follow-up.
(x) Evaluation also includes sample of two-parent families but sample sizes were too small for separate analysis.
(y) Randomization continued through February 1997; evaluation includes participants in first 12 months.
(z) Controls became subject to treatment conditions beginning March 1997.

Table 3.5—Key Reforms (Treatment) of Random Assignment Studies Included in Synthesis

Name	Financial work incentives	Mandatory work-related activity	Sanctions	Time limits	Family cap	Parental responsibility	Transitional		Two parent families	Other features
							Child care	Health insur.		
A. Programs that focus on financial work incentives										
CPWDP	After 9/93: Eliminated 4-month time limit on AFDC disregard (first $30 and 33% of remaining earnings)								100-hour rule eliminated	• Reduction in AFDC grant (8.5% from 10/92 to 9/93) • Asset limit (including vehicle value) increased • Restricted account for education, house or business
WRP-IO	Enhanced disregards (first $150 and 25% of any remaining earnings)		"Vendor payment sanction" (state takes control of grant); noncompliance leads to loss of grant				Y	Y	100-hour rule and work history requirement eliminated	• Asset limit (including vehicle value) increased
MFIP-IO	Enhanced disregards (38% of guarantee level and 38% of any remaining earnings up to 140% of PL)									• Streamlined administrative procedures • Direct reimbursement of childcare providers • Food Stamp cash-out • Asset limit (including vehicle value) increased
B. Programs that focus on financial work incentives tied to hours of work										
New Hope	Earnings supplement for minimum of 30 hours/week at unsubsidized or community service job									• Subsidized health insurance and child care
SSP (a)	Earnings supplement (half difference between gross earnings and benchmark income set to $30K in NB and $37K in BC to start) for minimum of 30 hours/week		Monthly supplement withheld for third or higher episode of less-than-full-time employment	3 years						• Must take up program w/in 1 yr. • Can not simultaneously receive traditional income assistance (IA) • Unearned income and income from other family members disregarded
SSP-Plus (a)	Earnings supplement (half difference between gross earnings and benchmark income set to $30.6K to start) for minimum of 30 hours/week		Monthly supplement withheld for third or higher episode of less-than-full-time employment	3 years						• Must take up program w/in 1 yr. • Can not simultaneously receive traditional income assistance (IA) • Unearned income and income from other family members disregarded • Voluntary employment-related services (resume service, Job Club, job leads, job coaching, self-esteem workshop) • Blueprint for self-sufficiency developed
SSP - A (a)	Earnings supplement (half difference between gross earnings and benchmark income set to $37K to start) for minimum of 30 hours/week		Monthly supplement withheld for third or higher episode of less-than-full-time employment	3 years						• Had to remain on IA for 1 yr. and then take up program w/in 1 yr. • Can not simultaneously receive traditional income assistance (IA) • Unearned income and income from other family members disregarded

Table 3.5—Continued

Name	Financial work incentives	Mandatory work-related activity	Sanctions	Time limits	Family cap	Parental responsibility	Transitional: Child care	Transitional: Health insur.	Two parent families	Other features
C. Programs that focus on mandatory work-related activities										
LA Jobs-1st GAIN		• Job search first • Strong "work first" message and Job Clubs for supervised job search	Adults-only sanction for non-compliance with program activities; high enforcement							
Atlanta LFA		• Job search first • Exemption if youngest child under age 3 • Case managers indicated that education and training services were available as a second step after initial job search	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Grand Rapids LFA		• Job search first • Exemption if youngest child under age 1 • Clients encouraged to enroll in education programs in addition to working • Caseworkers believed clients might be justified in turning down temporary or part-time jobs	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Riverside LFA		• Job search first • Exemption if youngest child under age 3 • Clients encouraged to take low-paying or part-time jobs as a first step	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Portland		• Job search or training/education first depending on disadvantage • Exemption if youngest child under age 1 • Job search clients encouraged to find good jobs (i.e., with benefits, higher-paying)	Adults-only sanction for non-compliance with program activities; high enforcement				Y			• Integrated case management • Excluded those with serious barriers to participation
Atlanta HCD		• Education/training first • Exemption if youngest child under age 3 • Clients given choice in type of education activity	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Grand Rapids HCD		• Education/training first • Exemption if youngest child under age 1 • Clients given choice in type of education activity	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff)

Table 3.5—Continued

Name	Financial work incentives	Mandatory work-related activity	Sanctions	Time limits	Family cap	Parental responsibility	Transitional		Two parent families	Other features
							Child care	Health insur.		
Riverside HCD		• Education/training first • Exemption if youngest child under age 3 • Short stay in basic education stressed • Clients moved into active job search once literacy target achieved	Adults-only sanction for non-compliance with program activities; high enforcement							• Traditional case management structure (separate income maintenance and welfare-to-work program staff) • Limited to those with/o diploma or GED, or low reading/math score, or limited English proficiency
Columbus Integrated		• Education/training first • Exemption if youngest child under age 3 • Clients given choice in type of education activity	Adults-only sanction for non-compliance with program activities; high enforcement				Y			• Integrated case management
Columbus Traditional		• Education/training first • Exemption if youngest child under age 3 • Clients given choice in type of education activity	Adults-only sanction for non-compliance with program activities; high enforcement				Y			• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Detroit		• Education/training first • Exemption if youngest child under age 1	Adults-only sanction for non-compliance with program activities; low enforcement				Y			• Traditional case management structure (separate income maintenance and welfare-to-work program staff)
Oklahoma City		• Education/training first • Exemption if youngest child under age 1 • Importance of education as a way of increasing job skills stressed for all clients	Adults-only sanction for non-compliance with program activities; low enforcement				Y			• Integrated case management (but limited resources weakened this feature)
IMPACT Basic Track		• Work first • Mandatory 20 hours a week in E&T or work-related activities			No increase with added children born to current recipients	• Immunizations and school attendance • Pregnant/parenting minors must live with parents • Cooperate with CSE				• Personal responsibility agreement required • Assignment to tier determined by caseworker administered assessment

Table 3.5—Continued

Name	Financial work incentives	Mandatory work-related activity	Sanctions	Time limits	Family cap	Parental responsibility	Transitional: Child care	Transitional: Health insur.	Two parent families	Other features
D. Programs that focus on financial work incentives and mandatory work-related activities										
WRP	Enhanced disregards (first $150 and 25% of any remaining earnings)	Half-time (single parents with youngest child under 13) or full-time (single parents with no child under 13 or 2-parent families) paid or comm. service job after 30 months on aid (15 months for 2-parent able-bodied primary wage earner)	"Vendor payment sanction" (state takes control of grant); noncompliance leads to loss of grant				Y	Y	100-hour rule and work history requirement eliminated	• Asset limit (including vehicle value) increased
MFIP	Enhanced disregards (38% of guarantee level and 38% of any remaining earnings up to 140% of PL)	Mandatory E&T for recipients on aid > 24 in past 36 mos. (single parents) or > 6 in past 12 mos. (two-parents) and no child<1 year old (single parents only) and working < 30 hours per week	10% of grant. May be lower than sanction for controls, which was reduction of grant by adult's portion						100-hour rule and work history requirement eliminated	• Streamlined administrative procedures • Direct reimbursement of childcare providers • Food Stamp cash-out • Asset limit (including vehicle value) increased
TSMF	Enhanced disregards (first $200 and 20% of any remaining earnings)	• Until 10/94: 20 hours/week in work, E&T, self-improvement, or community service • After 10/94: participation in Work First (applied to controls, too)	After 10/94: 25% of grant, plus FSP sanction			• After 10/94: Immunizations			100-hour rule and work history requirement eliminated	• Personal responsibility agreement required • All earnings and savings of dependent children disregarded • After 10/94: Vehicles exclude from asset tests; allowed deductions for investments in self-employment
FIP	Enhanced disregards (first $200 and 50% of any remaining earnings)	Mandatory E&T participation at levels specified in individual agreement (including possible unpaid work experience or community service)	Assignment to Limited Benefit Plan (cash grant reduced for 3 months and then eliminated for succeeding 6 months)				Y		100-hour rule and work history requirement eliminated	• Personal responsibility agreement required • Asset limit (including vehicle value) increased • Balance in Individual Investment Account and interest/dividend income disregarded
E. Programs that focus on other individual reforms										
AWWDP					No increase with added children born to current recipients (b)	• Family planning information and services				
FDP (c)	Earned disregard up to 50% of grant level	• More extensive case management and supportive services • Exemption if youngest child under age 2	Strengthened		No increase with added children born to current recipients (d)			Y		
PPI			$25 monthly penalty for failure to verify preventative care at 6-month intervals			• Semi-annual verification of preventative health care services including vaccinations for pre-school age children				
PIP			Loss of portion of grant for non-immunized child			• Verify at time of eligibility and semi-annually or annually thereafter that pre-school age children receive vaccinations				

Table 3.5—Continued

Name	Financial work incentives	Mandatory work-related activity	Sanctions	Time limits	Family cap	Parental responsibility	Transitional: Child care	Transitional: Health insur.	Two parent families	Other features
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
EMPOWER			Automatic min. 1-, 3- and 6-month sanction of adult portion of grant for 1st, 2nd, or 3rd noncompliance with JOBS requirements	24 months in 60 months (adult only) (e)	No increase with added children born to current recipients	• Pregnant/parenting minors must live with parents	Y	Y	100-hour rule eliminated	• Contributions to Individual Development Account for training and education disregarded • Teens age 13 and over not exempted from JOBS
IMPACT Placement Track	Fixed grant up to PL	Mandatory 20 hours a week in work-related activities	Noncompliance leads to loss of adult portion of grant for 2 mos. min., increasing to 36 mos. for third penalty	24 months in 60 months (adult only) (f)	No increase with added children born to current recipients	• Immunizations and school attendance • Pregnant/parenting minors must live with parents • Cooperate with CSE	Y	Y	100-hour rule eliminated	• Personal responsibility agreement required • Increased asset limit if working • Assignment to tier determined by caseworker administered assessment
VIP/VIEW (g)	Families receive full TANF grant as long as net earnings plus TANF put them below the federal poverty line (h)	• Job search required for 90 days (h) • If regular employment not found, participation in Community Work Experience Program required in exchange for benefits (h) • Exemption allowed for parents with child under 18 months, and medical exemptions tightened (h)	• Case closed if personal responsibility agreement not signed (h) • Full family sanction for non-compliance with job search or work requirements after signing agreement (h) • Other sanctions for failure to comply with CSE, parental responsibility provisions • Months in sanction count towards time limit (h)	24 months; ineligible for 36 months (h)	No increase with added children born to current recipients (but any child support received for capped child is disregarded)	• Stronger cooperation with CSE • Immunizations and school attendance • Minors with children must live with parents	Y (h)	Y (h)	Eligibility determination same as for one-parent families	• Transportation assistance while on TANF and for one year after case closes (h) • Allowed to accumulate $5,000 in savings for use towards own business, education, or home ownership • Personal responsibility agreement required (h) • Diversion payments for forgoing welfare for 160 days
ABC	• More generous disregards • Fill the gap budgeting	• Mandatory work activities in first 24 months • Work required in second 24 months • Pay-for-performance community service job if not able to find work	Progressive sanctions leading to case closure after 5 months of continuous noncompliance with work and parenting requirements	48 months; ineligible for 96 months (i)	No increase with added children born to current recipients	• Parenting class • Immunizations and school attendance • Pregnant/parenting minors must live with parents • Substance abuse treatment and family planning • Cooperate with CSE	Y	Y	Eligibility determination same as for one-parent families	• Personal responsibility agreement required
FTP	Enhanced disregards (first $200 and 50% of any remaining earnings)	• Mandatory participation in job search and placement activities (j) • Intensive case management and enhanced services • Exemption if youngest child under 6 months	Sanctions for noncompliance with work and parenting requirements	24 in 60 months or 36 in 72 months depending on recipient characteristics		• School attendance and parental contact with teachers • Immunizations	Y			• Asset limit (including vehicle value) increased
Jobs First	All earned income disregarded as long as earnings below PL	• Mandatory work first employment services; E&T as last resort • Exemption if youngest child is under age 1 (and child not conceived while on welfare)	• First instance: 20% grant reduction for 3 months • Second: 35% reduction for 3 months • Third: cancelled for 3 months	21 months (with possible extensions)	$50 increase for additional child conceived while mother on aid (half the increase under AFDC)		Y	Y	Non-financial eligibility requirements made similar to one-parent family requirements	• Asset limit (including vehicle value) increased • All child support passed through and first $100 disregarded

NOTES: For full program names and citations, see Table 3.4. Abbreviations: PL=federal poverty line; CSE=child support enforcement; E&T=employment and training services; NB=New Brunswick; BC=British Columbia; IA=Canadian Income Assistance.

(a) All monetary values in Canadian dollars.

(b) 37 percent of treatment group thought subject to family cap compared with 20 percent of controls. 52 percent and 46 percent of treatment and control groups respectively did not know how much more money they would get with an added child.

(c) Implementation of the FDP provisions was delayed in two counties until January 1995. Some results only pertain to 8 counties with implementation by October 1993.

(d) 36 percent of treatment group thought no additional benefit with added child compared with 35 percent of controls.

(e) 61 percent of treatment group believed state had time limits compared with 56 percent of control group.

(f) 69 percent of treatment group thought subject to time limits compared with 43 percent of control group.

(g) VIP provisions implemented in July 1995 but VIEW provisions phased in at five demonstration sites between July 1995 and October 1997.

(h) VIEW provision phased in between July 1995 and October 1997.

(i) 84 percent of treatment group thought subject to time limits compared with 66 percent of control group.

(j) Many control group members were subject to Project Independence provisions requiring participation in mandatory work-related activity.

The studies are grouped in both tables according to their central reform or reforms. Our categorization contains six groups. The first group, shown in Panel A of Table 3.4, consists of three experiments that focus on financial work incentives: California's Work Pays Demonstration Program (CWPDP) and the Incentives-Only components of the Vermont Welfare Restructuring Project (WRP-IO) and the Minnesota Family Investment Program (MFIP-IO). MFIP-IO and WRP-IO were parts of dual-treatment experiments where the Incentives-Only groups experienced financial work incentives and the other experimental groups were subject to work-related activity mandates as well.

The programs listed in Panel B also provide financial work incentives, but they are implemented as earnings supplements outside the welfare system. Moreover, the earnings supplements were available only to participants who worked at least a minimum number of hours. New Hope was conducted in Wisconsin; the Self-Sufficiency Project (SSP) programs were carried out in Canada.

The third group of studies, shown in Panel C, consists of programs that imposed or strengthened requirements for mandatory work-related activities. These studies include Los Angeles Jobs-First GAIN (Greater Avenues for Independence), the 11 sites included in the National Evaluation of Welfare-to-Work Strategies (NEWWS), and the Basic Track of the Indiana Manpower Placement and Comprehensive Training Program (IMPACT).

The studies listed in Panel D combine mandatory work-related activities and a financial work incentive. In addition to the full WRP and MFIP programs, programs in Michigan (To Strengthen Michigan Families, TSMF) and Iowa (Family Investment Program, FIP) are included in this group. As Table 3.5 shows, MFIP and FIP provided more generous financial work incentives than WRP or TSMF.

Category E consists of four programs that focus on various other reforms. The Arkansas Welfare Waiver Demonstration Project (AWWDP) and New Jersey Family Development Program (FDP) each evaluate a family cap provision alone or with other reforms. The Maryland Primary Prevention Initiative (PPI) and Georgia Preschool Immunization Project (PIP) evaluate parental responsibility requirements focused on immunizations or preventative health care for children more generally.

The six evaluations in category F add time limits to a program with financial work incentives and/or mandatory work-related activities. Arizona's EMPOWER (Employing and Moving People Off Welfare and Encouraging Responsibility) program combines a time limit with somewhat stricter JOBS sanctions. The other five programs—Indiana's IMPACT Placement Track program, the Virginia Independence Program/Virginia Initiative for Employment not Welfare (VIP/VIEW), Delaware's A Better Chance (ABC) program, Florida's FTP, and Connecticut's Jobs First program—combine a time limit with financial work incentives and a work requirement. Of all the programs we consider, these incorporate the most TANF-like bundles of reforms. They provide information about the effects of reform as a bundle.

Table 3.4 reveals that the evaluations we draw on were implemented in the 1990s under state waivers prior to the passage of PRWORA, with randomization periods that range from mid-1991 to late 1996. Thus, the reforms implemented under the studies listed in Table 3.4 are not necessarily representative of the range of individual reforms or range of policy bundles implemented across the states under PRWORA, especially at the less generous end of the spectrum (i.e., lower benefit levels, weaker financial work incentives, stricter work requirements and sanctions, and shorter time limits). Furthermore, for most settings, the economy was steadily improving during the period of randomization and follow-up. Even so, there is considerable variation across the evaluations in the initial state of the economy and the generosity of the state welfare program in terms of benefit levels (see Table 3.4).

In terms of measuring outcomes, Table 3.4 shows that about two years of follow-up data are typically available, although some programs have observed participants for up to five years post-randomization. Unless otherwise noted, the sample sizes shown are the maximum number of study participants available for analysis. In some cases, results discussed in subsequent chapters are based on smaller samples, especially when outcomes derive from survey data where the samples are often a subset of the full study population.

Most programs served both longer-term recipients and new applicants, with both single-parent and two-parent families eligible. When results are available separately for single parents, we show sample sizes specific to that group and report results in the synthesis chapters that exclude two-parent families. (When results are only available for a combined sample, the single-parent families usually dominate the sample.) Likewise, when available, we separately report sample sizes and results for recipients (those on welfare at the time of randomization) and applicants (those randomized at the time of application to welfare). In addition to stratifying results for one- and two-parent families and for recipients and applicants, many of the random assignment studies we analyze also report results for other subgroups of the study population, for example defined by educational attainment, employment history, or various composite measures of disadvantage. Appendix A discusses the key results from these subgroup analyses for each of the outcomes we consider in the synthesis. These findings are referenced in the individual chapters as well.

Finally, it is worth noting at this stage that a number of the methodological issues summarized in Section 3.1.1 above apply to the random assignment studies summarized in Tables 3.4 and 3.5. These methodological concerns will affect the weight we place on these particular studies throughout the chapters that follow. In particular, experiments will only yield valid estimates of the effect of ongoing (nonexperimental) implementation if they mimic the conditions of ongoing (nonexperimental) implementation. Failure of recipients to understand the policies that affected them—at a level similar to the level that would be expected in an ongoing (nonexperimental) program—violates this condition. As a result, the resulting estimates may be too small.

Confusion about program rules was a problem in a number of evaluations. In Arizona, 56 percent of the control group versus 61 percent of the treatment group thought they were affected by time limits (Kornfeld et al., 1999). In Delaware, a similar problem occurred with 66 percent of the controls reporting that they thought they had a time limit, compared with 84 percent of the treatment group (Fein and Karweit, 1997). Indiana is another example of this type of control group contamination. This confusion over key policy provisions leads us to place less weight on these programs in the chapters that follow.

In addition, in VIP/VIEW in Virginia and FDP in New Jersey, implementation of the treatment reforms was staggered, so that the "exposure" to the treatment varies across the study population. Another issue is that there are several cases where the treatment changed during the period of randomization (e.g., Michigan's TSMF) or during the period of follow-up (Indiana's IMPACT). This complicates the interpretation of the treatment impacts, which become a mixture of the two regimes. In several other studies (e.g., Arizona's EMPOWER and

Delaware's ABC), the experiment was terminated and the reforms (or modified set of reforms) were applied to all study participants. Some of the long-term results from NEWWS may also be affected by control-group crossover, because at many of the sites, the control groups became eligible for program services during the fourth or fifth year of the follow-up. This control-group crossover limits the period during which "pure" treatment effects can be measured. This type of crossover was most likely to occur when a state implemented its TANF program in the late 1990s.

3.4.2. Outcomes Covered by Synthesis Studies

Table 3.6 summarizes the outcomes covered by the econometric and random assignment studies we include in our synthesis. The columns of the table pertain to the outcome chapters that follow: welfare caseload, employment and earnings, use of other government programs, and so on. In the case of the econometric studies, we tally the number of studies that examine a given outcome and note which studies analyze the CPS, which studies analyze administrative data (the two primary sources of data for econometric studies that meet our quality criteria), and which studies analyze other data sources. For random assignment studies, we simply indicate when the impact analysis includes one or more measures in each outcome category.[29]

This tabulation reveals that, with the exception of the random assignment studies in category E (other reforms), all the random assignment studies and the bulk of the econometric studies cover welfare utilization. All of the experimental studies also cover employment and earnings, use of government programs (with the exception of CWPDP and the Canadian SSP), and income and poverty, but far fewer econometric analyses examine these outcomes. A smaller number of demonstration studies examine family structure, other measures of well-being, and child well-being. For the last two outcome areas, virtually all the evidence comes from random assignment studies, since survey data with the other required characteristics (i.e., panel data and large, national samples) and administrative data generally do not cover these outcomes.

3.5. ASSESSING RESULTS FROM MULTIPLE STUDIES

The previous section illustrates the range of studies available for the synthesis, as well as the outcomes covered by those studies. However, it is not enough to simply tally all the findings across the available studies. Rather, we need an approach for weighing the findings from each analysis and assessing the strength of the cumulative evidence for each policy-outcome pair. In this section, we discuss how we take the methodological issues raised in this chapter into account when synthesizing findings across studies.

---

[29]For the Arkansas and New Jersey family cap demonstrations listed in Panel E of Table 3.6, we note that although the impact analyses provide some results for welfare utilization, employment and earnings, or the use of other government programs, we do not discuss these findings in Chapters 4, 5, and 6, respectively. Instead, since the main focus of these demonstrations is the impact on fertility, we only discuss these experiments in terms of their impact on this outcome in Chapter 7. Likewise, the impact analyses for the two parental responsibility demonstrations (PPI and PIP) are really relevant only for Chapter 10 on child outcomes.

Table 3.6—Outcomes Analyzed by Econometric and Random Assignment Studies Included in Synthesis

Study	State	Welfare use	Employment and earnings	Use of other government programs	Fertility and marriage	Income and poverty	Other measures of well-being	Child outcomes
I. Econometric Studies								
Econometric – Administrative Data	—	13		4				1
Econometric – CPS Data	—	6	5	1	3	4		
Econometric – Other Data	—	3			3			
II. Experimental studies (random assignment)								
A. Programs that focus on financial work incentives								
CWPDP	CA	X	X		X	X		
WRP-IO	VT	X	X	X		X	X	X
MFIP-IO	MN	X	X	X	X	X	X	X
B. Programs that focus on financial work incentives tied to hours of work								
New Hope	WI	X	X	X		X	X	X
SSP	Canada	X	X		X	X	X	X
SSP Plus	Canada	X	X			X		
SSP Applicants	Canada	X	X			X		
C. Programs that focus on mandatory work-related activities								
LA Jobs-1st GAIN	CA	X	X	X	X	X	X	X
Atlanta LFA	GA	X	X	X	X	X	X	X
Grand Rapids LFA	MI	X	X	X	X	X	X	X
Riverside LFA	CA	X	X	X	X	X	X	X
Portland	OR	X	X	X	X	X	X	X
Atlanta HCD	GA	X	X	X	X	X	X	X
Grand Rapids HCD	MI	X	X	X	X	X	X	X
Riverside HCD	CA	X	X	X	X	X	X	X
Columbus Integrated	OH	X	X	X	X	X	X	X
Columbus Traditional	OH	X	X	X	X	X	X	X
Detroit	MI	X	X	X	X	X	X	X
Oklahoma City	OK	X	X	X	X	X	X	X
IMPACT Basic Track	IN	X	X	X		X		
D. Programs that focus on financial work incentives and mandatory work-related activities								
WRP	VT	X	X	X		X	X	X
MFIP	MN	X	X	X	X	X	X	X
TSMF	MI	X	X	X		X		X
FIP	IA	X	X	X		X		
E. Programs that focus on other individual reforms								
AWWDP	AR	(X)	(X)	(X)	X			
FDP	NJ	(X)	(X)		X			
PPI	MD							X
PIP	GA							X
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)								
EMPOWER	AZ	X	X	X	X	X	X	
IMPACT Placement Track	IN	X	X	X		X		
VIP/VIEW	VA	X	X	X		X		
ABC	DE	X	X	X	X	X		X
FTP	FL	X	X	X	X	X	X	X
Jobs First	CT	X	X	X	X	X	X	X

NOTES: For full program names and citations, see Table 3.4. X=results discussed in relevant synthesis chapter; (X)=results not discussed in synthesis chapter.

In general, because even the best studies have limitations, the results from a single study provide a weak basis for making policy decisions. One's confidence would increase if the results were consistent when different data were analyzed and when different, valid methods were employed to deal with the problem of confounding influences. When different studies yield similar results, it is less likely that the results stem from either data problems or inadequate controls for confounding factors. Such robust results are more likely to represent the true effects of the policy in question.

Ideally, in assessing the effects of a particular policy on a particular outcome, we would have a large number of studies on which to draw, where the studies were based on different data and employed different methodological approaches. Of course, this ideal is not entirely realistic, since only ten years have passed since states first began experimenting with statewide waivers to the AFDC program, and only six years have passed since the enactment of PRWORA. As seen in Table 3.6, there are a few policy-outcome combinations for which many studies are available, but many more for which only a few exist.

Moreover, it is not only the number of studies that matter; rather, it is the number of studies using different data and a mix of methods in drawing the same conclusions. In some cases, there appear to be several studies on a particular topic, but the studies are based on the same (or nearly the same) underlying data and use similar methods. In other cases, we have studies using different methods and data to explore the same topic. Confluent results from studies based on different data provide stronger evidence than confluent results from studies based on largely similar data.

Finally, as the discussion from the preceding section suggests, quality matters at least as much as quantity. Some random assignment studies carried out randomization properly, used data from multiple sources, and succeeded in communicating to the study participants which set of rules applied to them. These are the highest-quality random assignment studies, to which we give considerable weight. We give less weight to lower-quality studies, particularly those where the study participants (i.e., treatment and control group members) were unclear on the rules that applied to them.

Similarly, econometric studies differ in terms of their quality. Some provide rigorous controls for unobservable confounding factors and employ policy measures that capture multiple dimensions of policy variation. These represent the highest-quality econometric studies, to which we also give considerable weight. We give somewhat less weight to studies that employ controls for unobservable confounding factors but use only dummy variables to represent reform policies. We give little weight to studies that provide no controls for unobservables.

Since random assignment and econometric methods represent quite different approaches to the problem of confounding influences, we consider confluent results from high-quality random assignment studies and high-quality econometric studies to provide the strongest type of evidence on the effects of a particular policy. Of course, a multiplicity of high-quality studies of both types that point in the same direction yields stronger evidence still.

In the chapters that follow, we will see that there are very few policy-outcome combinations for which evidence of this quality and quantity is available. Thus, there are relatively few cases, albeit some important ones, where the research literature allows us to draw definitive conclusions about the effects of welfare reform. There are many policy-outcome pairs for which only one or two high-quality studies exist, and still more for which a few lower-quality studies are available. When the results from such studies point in the same direction, they provide suggestive evidence about the effects of the policy, particularly when the results are consistent with theoretical predictions. Nevertheless, evidence of this type is necessarily less conclusive than that from higher-quality studies. For a large number of policy-outcome pairs, little if any evidence is available.

Chapter Four

THE CASELOAD AND WELFARE USE

4.1. BACKGROUND

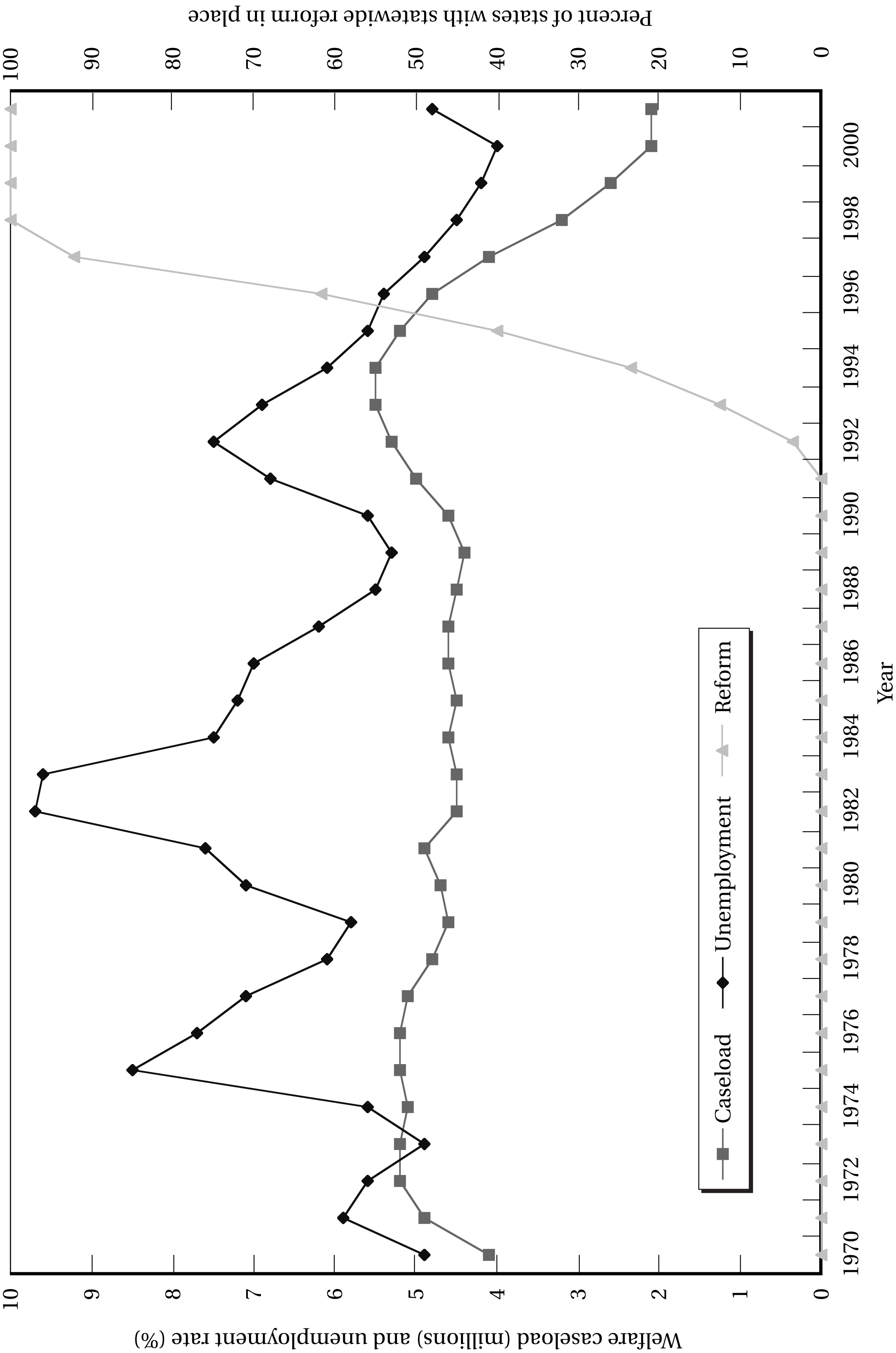
We begin the core synthesis chapters by analyzing the effects of welfare reform on welfare use. Of all the welfare-related outcomes we consider, welfare use may be most directly affected by reform. It has also received the most research attention. Much of this attention stems from the dramatic changes in welfare caseloads that took place during the 1990s.

Figure 4.1 presents data for the period 1970 to 2001 on the welfare caseload, defined as the fraction of the U.S. population receiving cash aid (either under AFDC or TANF). For most of that period, the caseload was fairly stable. However, beginning in the late 1980s, it increased substantially, rising from 4.5 percent to 5.5 percent between 1988 and 1993. After 1993, caseloads started falling. By 2000, they had reached 2.1 percent, a 35-year low. As of June 2001, they remained at that level.

Many studies have attempted to explain this precipitous decline. One suggested explanation is the economy. A useful measure of economic conditions is the unemployment rate, which is also plotted in Figure 4.1. Prior to 1990, changes in the welfare caseload were weakly associated with changes in the unemployment rate, as evidenced by the small increases in the caseload during the 1975 and 1980 recessions. The eligibility restrictions in OBRA 1981 have been offered to explain why caseloads did not rise further as unemployment approached 10 percent, in the early 1980s, but the caseload also remained roughly constant during the earlier recession. Only since 1990 have changes in the caseload closely tracked changes in the unemployment rate. Both increased sharply during the early 1990s and decreased sharply thereafter. In 2001, due to the softening economy, the unemployment rate rose to 4.8 percent.

Another suggested explanation for the drop in welfare caseloads is welfare reform. The caseload decline coincides with sharp increases in the number of states reforming their welfare programs, first under waivers and then under PRWORA. As shown in Figure 4.1, the first statewide waivers were implemented in 1992; by 1998, all states had implemented their TANF plans.

The role played by welfare reform is the topic of what follows. Estimates of the effects of welfare reform from a number of random assignment studies are the subject of the next section. That section also includes a brief summary of the results discussed in Appendix A regarding subgroup differences in the impacts of reform policies on the welfare caseload. Following that, we discuss the results from a number of econometric studies. In section 4.4, we synthesize the



SOURCE: Caseload: USDHHS (2001b); unemployment rate: USBLS (2002); and statewide reform: authors' tabulations of data from CEA (1999).

Figure 4.1—The Welfare Caseload, Unemployment, and Statewide Reform: 1970–2000

studies to convey what is known about the effects of welfare reform on welfare use. We conclude with a summary of our findings.

4.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECTS OF WELFARE REFORM ON WELFARE USE

The estimates that we report for the random assignment studies are typically referred to as "impact" estimates. They represent the difference between the average welfare-related outcome among the treatment group and the average welfare-related outcome among the control group. The precise outcome measures, and the follow-up period over which they are calculated, vary from study to study and are reported in column (4) of Table 4.1. The impact estimates themselves appear in column (6). Column (7) reports percentage impacts, obtained by dividing the impact estimates by the control group means in column (5).

For several studies, we report more than one estimate. The reason is that most studies report estimates that are disaggregated in various ways. The most common disaggregations involve ongoing recipients versus new applicants. Where possible, we present separate results for these groups. In many cases, however, the original study presents the results only in aggregated form.[30]

For some programs, we present results for different time periods following random assignment. We do this when the program impacts seem to change over time. We also report multiple results when the program's effect on other outcomes, such as employment, changes over time. This facilitates comparisons across outcomes in later chapters. Finally, we present multiple results for the few studies that report impacts for periods both before and after their time limits become binding.

Where possible, we report in Table 4.1 estimates based on single-parent families, which are the largest group to receive welfare. In a few studies, both single-parent and two-parent families participated, but results were not reported separately. In those cases, we present the overall estimates. Since the single-parent group is so much larger than the two-parent group, that group tends to dominate those results.

4.2.1. Programs That Focus on Financial Work Incentives

As indicated in Chapter 3, three studies provide information on the effects of financial work incentives: CWPDP, WRP, and MFIP. Extended financial work incentives were one of the main reforms included in CWPDP, although the program impacts also reflect the effects of the program's reduced benefit level. WRP and MFIP were dual-treatment experiments. The treatment groups for the full WRP and full MFIP programs were subject to both work-related activity mandates and financial work incentives, whereas the Incentives Only treatment groups, WRP-IO and MFIP-IO, were subject only to the financial work incentives. [31]

[30]More detailed subgroup-specific impacts are presented in Appendix A.

[31]Members of the treatment groups were subject to some other policy changes as well, such as extended transitional child care or a food stamp cash-out. Because the MFIP treatment group received its food stamp benefits in the form of cash, the MFIP welfare use measure is an indicator of whether the participant received cash aid (welfare plus cashed-out food stamp benefits) in the case of the treatment group, or whether the participant received cash aid (welfare) or food stamps benefits in the case of the control group. For both groups, the welfare use indicator also reflects receipt of General Assistance.

Table 4.1—Estimated Impact of Welfare Reform on Welfare Use: Random Assignment Studies

Name	Cases served	Data	Measure	Welfare use: Control mean	Welfare use: Impact	Welfare use: %
A. Programs that focus on financial work incentives						
CWPDP	Single parent recipients	A	Avg. welfare receipt, year 3	67.0	1.0	1.5%
WRP-IO	Single-parent recipients and applicants	A	Ever received welfare, last 3 mos. of FU	37.4	0.3	0.8%
MFIP-IO	Urban single parents recipients	A	Avg. quarterly welfare receipt, year 1	90.7	2.8 ***	3.1%
		A	Avg. quarterly welfare receipt, year 3	63.6	10.5 ***	16.5%
	Urban single parents applicants	A	Avg. quarterly welfare receipt, year 1	65.8	8.4 ***	12.8%
		A	Avg. quarterly welfare receipt, year 3	36.6	10.3 ***	28.1%
B. Programs that focus on financial work incentives tied to hours of work						
New Hope	Poor families employed FT at RA	A	Months receiving welfare, year 1 of 2-yr FU	3.4	-0.1	-2.9%
		A	Months receiving welfare, year 2 of 2-yr FU	2.6	-0.8 **	-30.8%
	Poor families not employed FT at RA	A	Months receiving welfare, year 1 of 2-yr FU	5.9	0.0	0.0%
		A	Months receiving welfare, year 2 of 2-yr FU	3.6	0.3	8.3%
SSP	Single-parent recipients	A	Monthly receipt of IA or SSP year 2	78.9	7.6 ***	9.6%
		A	Monthly receipt of IA or SSP year 3	70.7	9.8 ***	13.9%
SSP Plus	Single-parent recipients	A	Receipt of IA or SSP, Q5	81.1	4.3	5.3%
SSP Applicants	Single-parent applicants	A	Receipt of IA or SSP, Q5	61.5	3.7 **	6.0%
		A	Receipt of IA or SSP, Q9	49.6	6.4 ***	12.9%
C. Programs that focus on mandatory work-related activities						
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Received welfare, Q8	66.2	-4.6 ***	-6.9%
Atlanta LFA	Recipients and applicants	A	Received welfare, Q8	67.0	-5.7 ***	-8.5%
Grand Rapids LFA	Recipients and applicants	A	Received welfare, Q8	60.9	-7.4 ***	-12.2%
Riverside LFA	Recipients and applicants	A	Received welfare, Q8	56.4	-6.4 ***	-11.3%
Portland	Recipients and applicants; no cases with substantial barriers	A	Received welfare, Q8	53.0	-11.7 ***	-22.1%

Table 4.1—Continued

Name	Cases served	Data	Measure	Welfare use: Control mean	Welfare use: Impact	Welfare use: %
Atlanta HCD	Recipients and applicants	A	Received welfare, Q8	67.0	-3.5 **	-5.2%
Grand Rapids HCD	Recipients and applicants	A	Received welfare, Q8	60.9	-6.5 ***	-10.7%
Riverside HCD	Recipients and applicants	A	Received welfare, Q8	60.0	-4.1 **	-6.8%
Columbus Integrated	Recipients and applicants	A	Received welfare, Q8	53.0	-6.8 ***	-12.8%
Columbus Traditional	Recipients and applicants	A	Received welfare, Q8	53.8	-4.6 ***	-8.6%
Detroit	Recipients and applicants	A	Received welfare, Q8	73.7	-3.6 ***	-4.9%
Oklahoma City	Applicants	A	Received welfare, Q8	40.8	-2.5 **	-6.1%
IMPACT Basic Track	Recipients and applicants-basic track	A	Received welfare, Q4	52.4	2.2	4.2%
D. Programs that focus on financial work incentives and mandatory work-related activities						
WRP	Single-parent recipients and applicants	A	Ever received welfare, last 3 mos. of FU	37.4	-2.1	-5.6%
MFIP	Urban single-parent recipients	A	Avg. quarterly welfare receipt, year 1	90.7	1.7 *	1.9%
		A	Avg. quarterly welfare receipt, year 3	63.6	7.6 ***	11.9%
	Urban single-parent applicants	A	Avg. quarterly welfare receipt, year 1	65.8	8.4 ***	12.8%
		A	Avg. quarterly welfare receipt, year 3	36.6	6.4 ***	17.5%
TSMF	Recipients	A	Monthly welfare receipt over 4-yr FU	60.4	-1.5 ***	-2.5%
	Applicants	A	Monthly welfare receipt over 1-yr FU	64.1	-2.1	-3.3%
		A	Monthly welfare receipt over 2-yr FU	54.7	-1.9 **	-3.5%
FIP	Recipients	A	Welfare receipt, Q4	76.2	3.3 ***	4.3%
		A	Welfare receipt, Q8	57.3	1.3	2.3%
	Applicants	A	Welfare receipt, Q4	34.8	2.2	6.3%
		A	Welfare receipt, Q8	23.8	1.2	5.0%

Table 4.1—Continued

Name	Cases served	Data	Measure	Welfare use: Control mean	Welfare use: Impact	Welfare use: %
E. Programs that focus on other individual reforms						
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)						
EMPOWER (a)	Recipients	A	Monthly welfare receipt, months 1-36	41.1	-1.0	-2.4%
IMPACT Placement Track	Recipients and applicants-placement track	A	Received welfare, Q4	52.6	-9.3 ***	-17.7%
		A	Received welfare, Q8	29.3	-3.9	-13.3%
VIP/VIEW	Recipients	A	Welfare receipt in Q8	53.3	-1.2	-2.3%
ABC	Recipients and applicants	A	Months on welfare, Q1-Q4	9.1	0.0	0.0%
FTP	Recipients and applicants	A	Avg. percent receiving aid, year 2	44.4	-0.8 *	-1.8%
		A	Avg. percent receiving aid, year 3	32.0	-6.9 ***	-21.6%
		A	Avg. percent receiving aid, year 4	20.7	-8.8 ***	-42.5%
JOBS First	Recipients and applicants	A	Ever received aid, Q7	53.9	6.8 ***	12.6%
		A	Ever received aid, Q8	51.0	-5.7 ***	-11.2%
		A	Ever received aid, Q16	28.0	-9.3 ***	-33.2%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; Q=quarter; RA=random assignment; FT=full-time.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) Phoenix site only, cash assistance.

According to the economic model discussed in Chapter 3, financial work incentives should increase employment. Both CWPDP and WRP-IO increased welfare use slightly, but neither impact was significant. In contrast, the results from MFIP-IO show sizeable and significant increases in welfare use among both ongoing recipients and new applicants.

The different results generated by these different programs potentially could be explained by a number of factors. However, one particularly important difference is in the generosity of the programs' financial work incentives, as seen in Table 3.5. Both CWPDP and WRP involved fairly weak financial work incentives. In CWPDP, the treatment and control groups were subject to the same earnings disregards during the first four months of employment; the treatment group experienced more generous financial work incentives only after working for four months. The WRP treatment group actually faced a higher benefit reduction rate than the control group during the first four months of work. Moreover, the differential incentive remained fairly small during the fifth through twelfth months of work. In contrast, the MFIP incentive was fairly generous, which may explain why it had a relatively strong effect.

4.2.2 Programs That Focus on Financial Work Incentives Tied to Hours of Work

The programs listed in Panel B of Table 4.1 involve financial work incentives in the form of earnings supplements that are conditioned on full-time work. In all cases, the earnings supplement is paid outside the welfare system. As a result, these programs may be thought of as alternatives to traditional welfare.

The New Hope program had little effect on AFDC use among families not working full-time when randomly assigned, but it did significantly decrease second-year AFDC receipt among families initially satisfying the full-time work requirement. Unfortunately, Bos et al. (1999) do not report how New Hope affected the rate of transfer receipt, that is, the rate at which the treatment group received support from either AFDC or the earnings supplement. They report that 74 percent of the treatment group received the supplement at some point over the 24-month follow-up period, making it likely that the program raised the rate of transfer receipt. However, they do not report supplement receipt in a way that would allow us to eliminate possible double counting of persons receiving both types of aid. Thus, we cannot say for certain whether the results from New Hope accord with the standard economic model, which predicts that the total transfer rate should rise.

The SSP programs all raised the total transfer rate, that is, the rate at which recipients received either traditional welfare (Income Assistance, or IA) or the SSP supplement. Only in the SSP Plus program, where the sample size was small (596), was the effect insignificant. SSP decreased IA receipt (not shown), but the total transfer rate increased by virtue of the number of participants willing to work full time in exchange for the supplement.

4.2.3. Programs That Focus on Mandatory Work-Related Activities

Panel B of Table 4.1 reports on 13 welfare-to-work programs. Eleven are part of NEWWS; the others are L.A. Jobs-First GAIN and Indiana's IMPACT program Basic Track.

The welfare-to-work programs in all but one of the sites resulted in lower levels of welfare use. This is largely consistent with the predictions that work requirements should make welfare less attractive from the standard economic model discussed in Chapter 2. The average reduction in welfare use is 5.1 percentage points. Relative to the control-group mean, the average reduction is 8.7 percent.

Across the programs, there is evidence that the job-search-oriented programs generated somewhat greater reductions in welfare use than the skills-oriented programs during the first two years of the follow-up. The job-search-oriented programs—L.A. Jobs-First GAIN, Atlanta Labor Force Attachment (LFA), Grand Rapids LFA, and Riverside LFA—reduced welfare use by an average of 6 percentage points, whereas the skills-oriented programs—Atlanta Human Capital Development (HCD), Grand Rapids HCD, Riverside HCD, and the programs in Columbus, Detroit, Oklahoma City, and Indiana—averaged 3.9 percentage-point reductions. Moreover, in the three NEWWS sites that ran both an LFA and an HCD program, the LFA programs had larger effects on welfare use. The Portland program had the largest effects of all, which may bode well for its hybrid model. Then again, Portland's larger effects may be attributable to the fact that, unlike the other sites, the Portland program excluded recipients with substantial barriers to employment from participating in the demonstration (Freedman et al., 2000a, p. ES-21). The Detroit, Oklahoma City, and Indiana programs yielded the smallest reductions in welfare use, which may be attributable to lower levels of enforcement. Both Columbus programs yielded similar effects, providing no clear evidence that alternative case management approaches matter.

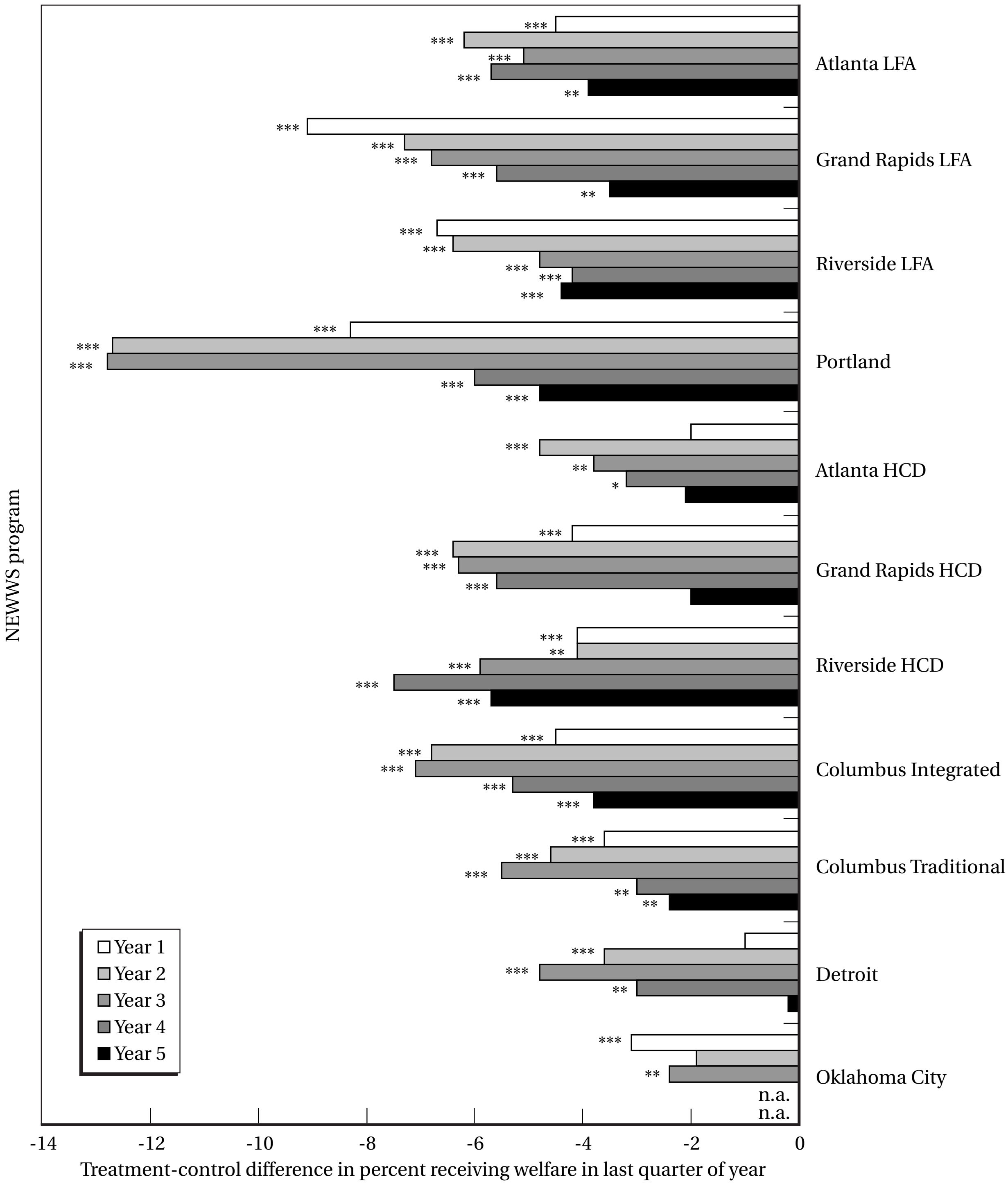
Recent data from NEWWS provide information on the longer-term effects of mandatory work-related activities. Program impacts by year after random assignment are presented in Figure 4.2. In all cases but one, the effects of the program fade over time: The longer-term impacts are smaller than the shorter-term impacts.[32]

4.2.4. Programs That Focus on Financial Work Incentives and Mandatory Work-Related Activities

Panel D of Table 4.1 presents results from four programs that combine financial work incentives with mandatory work-related activities. Whether these programs raise or lower welfare use cannot be predicted from the theoretical framework discussed in Chapter 2. By themselves, work mandates should decrease welfare use, whereas financial work incentives should increase it. Thus, the net effect will depend on the relative strength of the two opposing influences.

Of the four programs, WRP and TSMF reduced welfare use (albeit insignificantly in the case of WRP), whereas MFIP and FIP increased it. As can be seen from Table 3.5, the two programs that increased welfare use also had relatively generous financial work incentives; the other two programs had less generous financial work incentives. Although other factors could contribute to the differences as well, the impacts are generally consistent with the notion that the effects of financial work incentives are most likely to dominate the effects of work mandates when the financial incentives are strong.

[32]In principle, this could be because of the control-group crossover that occurred during years 4 and 5, when control group members were allowed to participate in the welfare-to-work programs that were previously available only to members of the treatment group. However, Hamilton et al. (2001) provide evidence that there was little actual crossover, suggesting that crossover had little to do with the program fade-out.



SOURCE: Hamilton et al. (2001), Table D.2.

NOTE: Treatment-control difference is statistically significant at the *=10%, **=5%, ***=1% level.

Figure 4.2—Impact Estimates for Welfare Receipt in 11 NEWWS Programs, Years 1 to 5

4.2.5. Programs That Focus on TANF-Like Bundles of Reforms

The six programs listed in Panel F of Table 4.1 combine time limits with work-related activity mandates (EMPOWER and the IMPACT Placement Track), financial work incentives (FTP), or both (VIP/VIEW, ABC, and Jobs First). Because they include a number of major reforms, these programs may be the most similar to the TANF plans adopted by the states after the passage of PRWORA. These programs provide some insights into the effects of reform as a bundle, although their broad focus makes it difficult to isolate the effects of any specific reform from the general program impacts. However, two of them do shed some light on what happens when families begin to reach the time limit.

Implementation issues bear on the interpretation of the impact estimates from a number of these studies. For example, in Virginia, although the implementation of the VIEW reforms took place at different times in different counties, the data on which the analysis is based pertain to the same time period for all of the study sites. Thus, the sample period includes both pre-reform and post-reform data. In fact, in two of the sites, it includes only pre-reform data, since VIEW was implemented there in the last month of the sample period. The presence of pre-reform data would tend to mask the effects of the program, since the pre-reform behavior of the treatment and control groups should be the same if randomization is carried out properly. As a result, we pay the VIP/VIEW results relatively little attention, both here and in later chapters.[33]

There are further questions about the extent to which these programs reflect the effects of their time limits. Most of these studies cover only the pre–time limit period, that is, the period prior to when any of the participants could have exhausted their benefits. Thus, with the exceptions of the FTP and Jobs First evaluations, these studies provide no information on the mechanical effects of time limits.

Furthermore, because of implementation issues, it is doubtful that the time limits included in these programs could have had much effect on behavior. As discussed in Chapter 3, there was substantial confusion about time limits among the study participants in the EMPOWER, IMPACT, and ABC programs. As a result, the impact estimates for these programs may reflect only the effects of their other policy reforms.

The other policy reform in EMPOWER involved changes to JOBS work-related activity mandates. However, the changes were fairly minor, amounting to a slight stiffening of sanctions for noncompliance without any changes in required activities or exemptions. This may explain why the program had essentially no effect on welfare use.

IMPACT's Placement Track component also included mandatory work-related activities. Unlike EMPOWER, however, IMPACT imposed substantially more rigorous mandates and a search-oriented welfare-to-work program. Its impacts on welfare use are roughly comparable to those of the programs that focus solely on search-oriented work-related activities. However, only the first-year effect is significant.

---

[33]In principle, we could use the site- and quarter-specific estimates provided in Gordon and Agodini (1999) to compute impact estimates over the post-VIEW period for the three sites that implemented VIEW prior to the end of the sample period. For a number of reasons, we do not take this approach. First, the follow-up periods are short for two of the three sites. Second, it is not possible for us to construct standard errors for such estimates. Third, the site with the largest post-VIEW impacts also had large pre-VIEW differences between the treatment and control groups, raising questions of whether randomization was properly conducted at that site.

Besides its poorly understood time limits, ABC involved mandatory work-related activities and a financial work incentive. The program as a whole had no effect on welfare use. It is possible that the opposing incentives of the program's two operative reforms offset each other.

FTP also involved policy reforms with conflicting incentives for welfare use. Its 24-month time limit was relatively well understood, with 88 percent of the treatment group and 29 percent of the control group reporting that they were subject to time limits.[34] The program also involved a fairly generous financial incentive. During year two of the follow-up period, FTP reduced welfare use by 0.8 percentage points. This suggests that the opposing incentives of the financial incentive and the time limit nearly offset each other, at least during the pre–time limit period.

Jobs First also provided conflicting incentives. Its 21-month time limit was well understood; 89 percent of the treatment group and 23 percent of the control group reported that they were subject to time limits. Like the time limit, its strengthened work-related activity mandate should have decreased welfare use. However, the program also included a very generous financial incentive: Members of the program group could earn up to the federal poverty line without having their benefit reduced. The strength of this financial incentive may explain why Jobs First actually increased welfare use by 6.8 percentage points during the last quarter of the pre–time limit period. Apparently, the effect of the extraordinarily generous financial incentive outweighed the effects of the work-related activities mandate and the time limit.

FTP and Jobs First are the only programs to provide insights into how TANF-like reform programs affect welfare use once the time limit begins to become binding. In FTP, the post–time limit period begins with year three; in Jobs First, it begins with quarter eight. Both programs had a sizeable reduction welfare use during the post–time limit period. Moreover, the negative impact grew over time. On the one hand, this may indicate merely that families that exhaust their benefits are indeed dropped from the rolls. On the other hand, given the substantial uncertainty surrounding the question of whether states would indeed enforce time limits, the finding that at least two states have done so is an important observation (Blank, forthcoming).

Finally, the change in welfare impacts between the pre– and post–time limit periods may shed some light on the mechanical effects of time limits. In FTP, none of the recipients could have exhausted their benefits prior to the end of year two. In Jobs First, none of the recipients could have exhausted their benefits prior to the end of quarter seven. To construct an estimate of what happens when recipients begin to reach the time limit, we subtract the pre–time limit impact from the post–time limit impact

This is not an experimental estimate of the mechanical effects of time limits, because program participants were not randomized with respect to the time at which they reached the time limit. Rather, it can be interpreted as a DoD estimate. The difference between the treatment and control groups estimates the impact of the program, and the difference between the pre– and post–time limit impacts estimates the mechanical effect of the time limit. This DoD approach will yield a valid estimate only if the effects of the programs' other policy reforms do not change between the two periods. This condition is more likely to be satisfied the closer the time periods used to construct the estimate. For this reason, we focus on the last pre–time limit period and the first post–time limit period. However, the estimates only indicate what happens

[34]Treatment group members who were deemed to be particularly disadvantaged received a 36-month time limit.

as a specific fraction of recipients reaches the time limit. The mechanical effects of time limits could become larger as more recipients exhaust their benefits.

In both cases, the impact of the program falls sharply as recipients begin to reach the limit. In FTP, the program impacts fall from –0.8 to –6.9 between years two and three. This amounts to 14 percent of the year-two control-group mean. In Jobs First, the program impact falls from 6.8 to –5.7, a relative decline of 23 percent. These are substantial changes.

4.2.6. Subgroup Differences

In Appendix A, we discuss what is known about the effects of various reforms on the welfare caseload for different segments of the welfare population. For the most part, the evidence is limited: There is little clear evidence on how the effects of the various policy reforms vary across subgroups. As for programs involving financial work incentives, there are no obvious patterns. Of the three studies involving TANF-like bundles of reform that provide subgroup estimates, there is no clear tendency for the reforms to have greater or lesser effects among the more disadvantaged. There is better evidence about programs that focus on mandatory work-related activities. This evidence suggests that such policies are similarly effective for most subgroups of the recipient population. At the same time, however, it suggests that search-oriented programs decrease welfare use among more disadvantaged groups by a somewhat greater amount than do skills-oriented programs.

4.3. ECONOMETRIC STUDIES OF THE EFFECTS OF WELFARE REFORM ON WELFARE USE

In addition to the random assignment studies, several econometric studies have attempted to estimate the effects of welfare reform. Although most of these studies focus on the effects of reforms as a bundle, several attempt to estimate the effects of specific reforms. We survey the estimates of specific reforms in Section 4.3.2 and focus on the effects of reform as a bundle in Section 4.3.3. However, we start with an overview of the similarities and differences of the econometric studies evaluated.

4.3.1. Similarities and Differences of Econometric Studies

A central challenge facing these studies is to disentangle the effects of reform from the effects of the economy. As seen in Figure 4.1, both were trending in ways that should have reduced the caseload. In the language of the research literature, such simultaneous trends are referred to as "collinear." Solving the collinearity problem, that is, distinguishing the effects of reform from the effect of the economy, has been a concern in all of the econometric analyses of welfare reform. It is an even greater problem in estimating the effects of specific reforms, since the effect of each reform must be distinguished not only from that of the economy, but also from those of the other reforms.

Most of the econometric studies are based on several years of annual state-level administrative data, most of which focus on annual state-level caseloads. Three studies directly analyze percent changes in caseloads between two points in time. Four studies are based on individual-level survey data. Two others use survey data aggregated by state, year, and various demographic measures. One reanalyzes data from the FTP demonstration.

Although these studies differ in many ways, they share some similarities. They involve regression models in which a measure of either the aggregate caseload or individual-level welfare use is to be explained by some or all the following factors: one or more measures of welfare reform, a measure of the generosity of the state's welfare program, and one or more measures of the economy. Typically, the analysis includes the current value and possibly lagged values of the annual state-level unemployment rate to control for the economy and distinguish the effects of the economy from the effects of the reform. Studies based on individual-level data typically control for a number of individual-level characteristics known to predict welfare use, such as the mother's age, education, race, and family size. Some of the aggregate studies also include state-level averages of such characteristics as control variables. Most of the analyses also include state-fixed effects and state-specific time trends to deal with unobservable confounding factors. A few include lagged dependent variables, that is, past values of the caseload.

Although the studies vary in the control variables they include, which in turn affects the quality of their results, they also differ in smaller ways. Some studies include measures of economic conditions beyond the unemployment rate. A few are based on monthly rather than annual data. Some of the studies based on aggregate data define the caseload as the number of persons on aid divided by the population, whereas others use the number of cases divided by the population. Some use the entire state's population as the denominator, whereas others use the population of women within certain age ranges. Estimates from the studies of reform as a bundle, which we cover in Section 4.3.3 below, suggest that such differences in detail have little impact on the estimated effects of reform.

Six studies depart from this pattern to an extent sufficient to warrant separate attention. Three focus on welfare transition rates. Hofferth, Stanhope, and Harris (2000a, 2000b) use individual-level longitudinal data from the PSID to estimate the effects of a number of specific reforms on rates of entry to and exit from welfare. Mueser et al. (2000) use administrative data to estimate the effect of reform as a bundle on entry and exit rates in five cities.

While most of the other studies analyze the level of welfare use over a period of several years, two studies analyze the change in the welfare caseload over a single time interval. Rector and Youssef (1999) use administrative data to analyze the percent decline in state-level caseloads (recorded as a positive number) between January 1997 and January 1998. MaCurdy, Mancuso, and O'Brien-Strain (2000) also use administrative data, focusing on the percent change in state-level caseloads between August 1996 and March 1999. Mead (2001) focuses on the percent change in state-level caseloads between 1994 and 1998. Unlike most of the other studies we review, none of these studies includes explicit controls for unobservable confounding factors. However, by focusing on changes in the caseload, rather than levels, they may partially control for such factors implicitly.[35]

Most of the aggregate studies use the logarithm of the annual state-level caseload as their dependent variable. For these specifications, the coefficients from the regression models are interpreted as the percent change in the caseload associated with a one-unit change in the explanatory variable. Thus, the coefficient on a welfare-reform dummy is interpreted as the

[35]Differencing the data within each state provides an alternative to the state-fixed effects approach for controlling for state-specific unobservables. However, the approach requires that both the dependent and independent variables be differenced, whereas these studies difference only the dependent variable and, in some cases, a few of the independent variables.

percent change in the caseload associated with the reform. Estimates from the studies that analyze percent changes in the caseload are interpreted the same way. In the survey-based studies, the dependent variable indicates whether the family was on welfare over some time period. The coefficients from these models are interpreted as the percentage-point change in welfare use associated with a unit change in the regressors, similar to the impact estimates from the random assignment studies. To aid in comparing the two types of estimates, we report the corresponding percent changes for all estimates in column (12) of Tables 4.2 and 4.3.[36]

Finally, since most econometric studies present more than one estimate of the effects of reform, we attempt to include in Table 4.2 the results from the authors' preferred specifications. In a few cases, we have included estimates from other specifications as well, when those additional estimates add important insights into the effects of a particular reform or reform bundle. In a couple of cases, we were unable to determine which specification the author(s) preferred. In those cases, we included estimates from what we considered to be the highest-quality specification.

4.3.2. Effects of Specific Reforms

We first consider econometric estimates of the effects of specific reforms. By the criteria described in Chapter 3, some of these estimates are of low quality, since they are based on regression models that fail to control for unobservable confounding factors. Most fall into the moderate quality category. These studies include controls for unobservable confounding factors but rely solely on dummy (or modified dummy) variables to capture the effects of specific reform policies. Thus, they utilize only temporal variation in policy to estimate the effects of the policy. Only a few studies employ both controls for unobservables and policy measures that capture additional dimensions of policy variation, thus providing high-quality evidence on the effects of specific reforms. However, even high-quality studies may suffer from power problems of the type discussed in Chapter 3.

Financial Work Incentives

Panel A of Table 4.2 presents estimates of the effects of financial work incentives. Of the six estimates, four accord with the prediction from the standard economic model described in Chapter 2. Estimates from the Council of Economic Advisors (CEA) (1997), Ziliak et al. (2000), and CEA (1999) show that incentives increase welfare use. The estimate from CEA (1999), which is based on a measure that captures the generosity of each state's financial incentive, and thereby counts as high-quality evidence, is significant. Hofferth, Stanhope, and Harris (2000a, 2000b) estimate financial work incentives to decrease the exit rate from welfare, but to have no effect on reentry rates.

[36]Unfortunately, Hofferth, Stanhope, and Harris (2000a, 2000b) do not provide enough information for us to transform their estimates in this way.

Table 4.2—Estimated Impact of Specific Welfare Reforms on Welfare Use: Econometric Studies

											Other Controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. Var.	Specific policy measure	Includes LDV's?	Coeff. (s.e.)	% effect	Economy	Demogr.	Fixed Effects	Policy
A. Financial work incentives														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if FWI's more generous than AFDC	no	0.11 (2.16)	0.11	U, U-1		S, Y, State time trends	B, WRA-A, WRA-D, WRA-S, TL, FC
Ziliak et al. (2000)	Monthly state-level caseloads	total population	87	96	AFDC cases/women 15-44	Log	Dummy=1 if FWI's more generous than AFDC (sum of current and lag coeffs.)	yes	1.79 a, b	2.01	U, U-1,...,U-6		S, month dummies	WRA-D, TL, FC
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if FWI's more generous than AFDC	no	-4.569 (4.318)	-4.57	U, U-1		S, Y, State time trends	B, WRA-A, WRA-D, WRA-S, TL, FC
CEA (1999)	Annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	Log of value of earnings disregard for family earning $750/month	no	5.38 (2.24)	3.44	U, U-1, U-2		S, Y, State time trends	B, MW, WRA-A, WRA-S, TL, FC
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Modified dummy>0 if FWI's more generous than AFDC	no	-0.584(0.190)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, WRA-A, WRA-D, WRA-S, TL, FC
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Modified dummy>0 if FWI's more generous than AFDC	no	0.050 (0.306)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, WRA-A, WRA-D, WRA-S, TL, FC
B1. Work-Related Activities - Age Exemptions														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if exemptions more stringent than JOBS	no	2.64 (7.09)	2.64	U, U-1		S, Y, State time trends	B, FWI, WRA-D, WRA-S, TL, FC
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if exemptions more stringent than JOBS	no	5.733 (4.695)	5.73	U, U-1		S, Y, State time trends	B, FWI, WRA-D, WRA-S, TL, FC
CEA (1999)	Annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	No age exemptions	no	4.86 (6.31)	4.86	U, U-1, U-2		S, Y, State time trends	B, MW, FWI, WRA-S, TL, FC
							Age exemption for child< 6 months	no	11.56 (7.56)	11.56				
							Age exemption for child 6 months to 3 years	no	12.37 (5.03)	12.27				
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Age exemption stricter than JOBS (3 years old)	no	0.767 (0.209)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-D, WRA-S, TL, FC
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Age exemption stricter than JOBS (3 years old)	no	0.905 (0.328)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-D, WRA-S, TL, FC

Table 4.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. Var.	Specific policy measure	Includes LDV's?	Coeff. (s.e.)	% effect	Other Controls: Economy	Other Controls: Demogr.	Other Controls: Fixed Effects	Other Controls: Policy
B2. Work-Related Activities - Deadlines														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if work requirement deadline	no	2.86 (2.83)	2.86	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-S, TL, FC
Ziliak et al. (2000)	Monthly state-level caseloads	total population	87	96	AFDC cases/women 15-44	Log	Dummy=1 if work requirements (sum of current value and lags)	yes	-0.283 a	-0.32	U, U-1,...,U-6		S, month dummies	FWI, TL, FC
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if work requirement deadline	no	-9.211 (5.600)	-9.21	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-S, TL, FC
Rector and Youssef (1999)	Percent reduction in monthly state-level caseloads	total population	97	98	Reduction in caseload, January 1997-June 1998	Percent	Dummy variable=1 if immediate work requirement	no	10.96 (5.45)	-10.96	U			WRA-S
MaCurdy, Mancuso, and O'Brien -Strain (2000)	Percent change in monthly state-level caseloads	total population	96	99	Reduction in caseload, August 1996-March 1999	Percent	Dummy variable=1 if immediate work requirement	no	0.0004 (0.040)	0.4	Change in U; change in 20th-percentile wage	A, E, R, No. kids, No. young kids, %immigrants, unwed birth rate, %non-citizens		B, WRA-S, TL
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Dummy=1 if work requirement deadline	no	0.114 (0.293)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-S, TL, FC
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Dummy=1 if work requirement deadline	no	-0.398 (0.489)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-S, TL, FC

Table 4.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. Var.	Specific policy measure	Includes LDV's?	Coeff. (s.e.)	% effect	Other Controls: Economy	Demogr.	Fixed Effects	Policy
B3. Work-Related Activities - Sanctions														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if sanctions more stringent than JOBS	no	-9.69 (3.00)	-9.69	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, TL, FC
Levine and Whitmore (1998)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if sanctions more stringent than JOBS	no	-8.05 (2.60)	-8.05	U, U-1		S, Y, State time trends	B, Any waiver
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if sanctions more stringent than JOBS	no	-2.043 (5.641)	-2.04	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, TL, FC
CEA (1999)	Annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	Partial sanctions that are more severe than JOBS	no	-9.71 (3.85)	-9.71	U, U-1, U-2		S, Y, State time trends	B, MW, FWI, WRA-A, TL, FC
							Graduated sanctions	no	-18.14 (4.82)	-18.14				
							Full-family sanctions	no	-39.36 (7.07)	-39.36				
Rector and Youssef (1999)	Percent reduction in monthly state-level caseloads	total population	97	98	Reduction in caseload, January 1997-June 1998	Percent	Dummy variable=1 if moderate sanctions	no	11.34 (10.38)	-11.34	U			WRA-D
							Dummy variable=1 if delayed full-family sanctions		13.66 (4.73)	-13.66				
							Dummy variable=1 if initial full-family sanctions		24.81 (5.14)	-24.81				
MaCurdy, Mancuso, and O'Brien -Strain (2000)	Percent change in monthly state-level caseloads	total population	96	99	Reduction in caseload, August 1996-March 1999	Percent	Dummy variable=1 if delayed full-family sanctions	no	-0.091 (0.061)	-9.1	Change in U; change in 20th-percentile wage	A, E, R, No. kids, No. young kids, %immigrants, unwed birth rate, %non-citizens		B, WRA-D, TL
							Dummy variable=1 if initial full-family sanctions		-0.196 (0.073)	-19.6				
Mead (2001)	Percent change in annual state-level caseloads	Total population	94	98	Change in caseload, 1994-1998	Percent	Dummy variable=1 if delayed full-family sanctions		-9.87 (3.33)	-9.87		AFDC caseload in 1994		WRA-O, CSE, Legislative effectiveness, state individualism
							Dummy variable=1 if initial full-family sanctions		-15.9 (4.07)	-15.9				
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Dummy variable =1 if sanctions stricter than JOBS	no	0.221 (0.192)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, TL, FC
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Dummy variable =1 if sanctions stricter than JOBS	no	0.030 (0.277)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, TL, FC

Table 4.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. Var.	Specific policy measure	Includes LDV's?	Coeff. (s.e.)	% effect	Other Controls: Economy	Other Controls: Demogr.	Other Controls: Fixed Effects	Other Controls: Policy
B3. Work-Related Activities - Other														
Mead (2001)	Percent change in annual state-level caseloads	Total population	94	98	Change in caseload, 1994-1998	Percent	Change in percent of welfare adults active in JOBS, 1994-1996		-0.838 (0.294)	-0.84		AFDC caseload in 1994		WRA-S, CSE, Legislative effectiveness, state individualism
							Change in JOBS clients assigned to post-secondary education		0.369 (0.149)	0.37				
C. Time limits														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if time limit in place	no	-6.37 (3.74)	-6.37	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, WRA-S, FC
Ziliak et al. (2000)	Monthly state-level caseloads	total population	87	96	AFDC cases/women 15-44	Log	Dummy=1 if time limit in place	yes	-1.268 a, b	-1.27	U, U-1,...,U-6		S, month dummies	FWI, WRA-D, FC
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if time limit in place	no	-6.790 (7.000)	-6.79	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, WRA-S, FC
CEA (1999)	Annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	Modified dummy>0 if time limit or work req. deadline in place	no	-3.75 (4.93)	-3.75	U, U-1, U-2		S, Y, State time trends	B, MW, FWI, WRA-A, WRA-S, FC
MaCurdy, Mancuso, and O'Brien -Strain (2000)	Percent change in monthly state-level caseloads	total population	96	99	Reduction in caseload, August 1996-March 1999	Percent	Dummy variable=1 if 5-year time limit	no	-0.049 (0.080)	-4.9	Change in U; change in 20th-percentile wage	A, E, R, No. kids, No. young kids, %immigrants, unwed birth rate, %non-citizens		B, WRA-D, WRA-S
							Dummy variable=1 if shorter time limit		-0.105 (0.090)	-10.5				
Grogger and Michalopoulos (forthcoming)	Micro-level administrative data from FTP demonstration	Families in Escambia County, FL with children < 14/15	94	97	Monthly welfare use indicator	Level	Treatment group dummy * (age*)	no	0.007 (0.002)	-1.6		A, E, R, No. kids, age youngest kid, past welfare use, past employment	Y, random assignment to treatment	
Grogger (2000)	CPS micro data	single mothers 16 - 54	78	98	equals 1 if welfare in last year	Level	Modified dummy>0 if time limit in place	no	0.018 (0.017)	5.8	U	A, E, R, No. kids, age youngest kid	S, Y	B, MW, Any reform
							Time limit dummy * (age**)		0.007 (0.002)	-2.3				
Grogger (forthcoming)	CPS micro data	single mothers 16 - 54	78	99	equals 1 if welfare in last year	Level	Modified dummy>0 if time limit in place	no	0.0236 (0.0157)	7.9	U	A, E, R, No. kids, age youngest kid	S, Y	B, MW, Any Reform, EITC
							Time limit dummy * (age**)		0.0066 (0.0015)	-2.2				
Grogger (2002)	SIPP micro data	single mothers 16 - 54	90	99	equals 1 if welfare in last month	Level	Time limit dummy	no	0.015 (0.016)	6.0	U	A, E, R, No. kids, age youngest kid	S-Y fixed effects	B, Any Reform
							Time limit dummy * (age**)		0.012 (0.001)	-4.8				
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Dummy variable =1 if time limit in place	no	-0.509 (0.368)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, WRA-S, FC
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Dummy variable =1 if time limit in place	no	-0.484 (0.427)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, WRA-S, FC

Table 4.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. Var.	Specific policy measure	Includes LDV's?	Coeff. (s.e.)	% effect	Other Controls: Economy	Demogr.	Fixed Effects	Policy
D. Family caps														
CEA (1997)	Annual state-level caseloads	total population	76	96	AFDC recipients/pop	Log	Modified dummy>0 if family cap in place	no	-0.49 (2.76)	-0.49	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, WRA-S, TL
Ziliak et al. (2000)	Monthly state-level caseloads	total population	87	96	AFDC cases/women 15-44	Log	Dummy=1 if parental responsibility encouraged	yes	-0.551 a, b	-0.62	U, U-1,...,U-6		S, month dummies	FWI, WRA-D, TL
Moffitt (1999)	Annual state-level caseloads	total population	77	95	AFDC cases/pop	Log	Modified dummy>0 if family cap in place	no	-10.580 (4.751)	-10.58	U, U-1		S, Y, State time trends	B, FWI, WRA-A, WRA-D, WRA-S, TL
CEA (1999)	Annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	Modified dummy>0 if family cap in place	no	6.71 (3.06)	6.71	U, U-1, U-2		S, Y, State time trends	B, MW, FWI, WRA-A, WRA-S, TL
Hofferth, Stanhope, and Harris (2000b)	PSID micro data	welfare recipients	89	96	Equals 1 if family leaves welfare	Logit	Dummy variable =1 if family cap in place	no	-0.191 (0.213)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, WRA-S, TL
		past welfare recipients	89	96	Equals 1 if family re-enters welfare	Logit	Dummy variable =1 if family cap in place	no	0.237 (0.283)	n.a.	U	A, E, R, No. kids, age youngest kid, work experience, disability	S, post-93 dummy	B, FWI, WRA-A, WRA-D, WRA-S, TL
E. Child support enforcement														
Huang et al. (2000)	Annual state-level caseloads	Women 15-44	76	96	AFDC Basic cases/females 15-44	Log	Average child support payment to welfare families	no	-0,272 (0.045)	-27.2	U, U-1, U-2, median wage, 10th% wage	A, E, R, Immigration	S, Y	B, Any waiver, UP program, party control, Medicaid spending
Mead (2001)	Percent change in annual state-level caseloads	Total population	94	98	Change in caseload, 1994-1998	Percent	Percent of AFDC families receiving child support	no	-0.357 (0.132)	-0.36		AFDC caseload in 1994		WRA-S, WRA-O, Legislative effectiveness, state individualism

NOTES: Abbreviations: LDV=lagged dependent variable; s.e.=standard error; U=unemployment rate;U-m=mth lag of unemployment rate; A=age, E=education, R=race, B=maximum welfare benefit, MW=minimum wage; EITC=Earned Income Tax Credit; S=state; Y=year; FWI=financial work incentives; WRA-A= age exemptions from work-related activities; WRA-D= deadlines for satisfying work-related activity mandates; WRA-S=sanctions for non-compliance with work-related activities; WRA-O=other features of work-related activities; TL=time limits; FC=family cap; CSE=child support enforcement; n.a. = cannot be calculated from available information.
age* = (age of youngest child -14) for families with 3-year time limit; = -(age of youngest child -15) for families with 2-year time limit.
age** = (age of youngest child -13).
(a) Standard error cannot be computed from information provided.
(b) Most of underlying coefficients were insignificant.

Work-Related Activities

Five studies attempt to estimate the effects of more stringent age exemptions from states' mandates to engage in work-related activities. The results are presented in Panel B1 of Table 4.2. Only the coefficients from Rector and Youssef (1999) and the exit study by Hofferth, Stanhope, and Harris (2000b) have the expected sign. The other estimates are insignificant and indicate that stricter exemptions work to increase caseloads rather than to decrease them. As for the effects of more stringent deadlines for satisfying work-related activity mandates, presented in Panel B2, only one estimate is significant.

The studies are more consistent about the effects of increased sanctions. Six of the nine studies listed in Panel B3 report significant estimates that indicate that stiffer sanctions reduce the caseload. CEA (1999) provides high-quality evidence, employing a specification that both provides explicit controls for unobservables and allows the effects of sanctions to vary with their severity. Its estimates are negative and significant and indicate that the stiffest sanctions have the greatest effects on welfare use. Indeed, the estimated effects of full-family sanctions are very large, indicating that they reduce welfare use by 39 percent.

Rector and Youssef (1999) and MaCurdy, Mancuso, and O'Brien-Strain (2000) employ a similar set of policy measures. Their results are qualitatively similar to those from CEA (1999), although the magnitudes of their estimates are smaller.

Interpreting the magnitudes of these estimates warrants some caution. In other analyses not shown here, MaCurdy, Mancuso, and O'Brien-Strain (2000) regress changes in state-level caseloads between 1989 and 1992 on policy changes implemented between 1992 and 1996. Since policy changes made after 1992 logically cannot affect behavior prior to 1992, these regressions shed some light on the policy endogeneity problem, that is, on the extent to which behavior influenced policy, rather than the other way around. The coefficient on the full-family sanctions dummy is statistically significant and, interpreted at face value, suggests that sanctions reduced pre-reform caseloads by 18 percent. Since no states implemented waivers involving full-family sanctions until late in 1994 (CEA, 1999), this effect clearly cannot be attributed to sanction policy. Rather, it may be evidence of policy endogeneity, indicating that states with large (percentage) reductions in their caseload during the pre-waiver period were more likely to seek waivers for full-family sanctions. Alternatively, it may reflect the effects of some other policy change that typically preceded the sanctions in states that eventually received sanction waivers.

Time Limits

Eleven studies address the effects of time limits. All can be interpreted as at least implicitly estimating the behavioral effects of time limits, rather than their mechanical effects, because the sample periods analyzed generally ended before recipients began exhausting their benefits. All but one of the estimates suggest that time limits reduce welfare use. However, of the studies based on aggregate data, only the estimate from CEA (1997) is significant, and then only at the 10 percent level. The estimates from Hofferth, Stanhope, and Harris (2000a, 2000b) are insignificant as well.

Grogger and Michalopoulos (forthcoming) provide explicit estimates of the behavioral effects of time limits based on a reanalysis of data from Florida's FTP program. Their analysis is structured around a theoretical model that predicts that families with the youngest children should reduce their welfare use the most once time limits are imposed. The reason is that such families have the longest period over which to spread their limited benefits, and thus the greatest incentive to save their benefits for future use.

The estimate reported in Panel C of Table 4.2 is the coefficient on an interaction term between the FTP treatment group dummy and a function of the age of the youngest child in the family.[37] The coefficient is statistically significant and suggests that the time limit component of FTP indeed reduced welfare use in a manner that was greatest for families with the youngest children.

In three complementary studies, Grogger (2000, 2002, forthcoming) estimates the effects of time limits using family-level data from the CPS and the SIPP. He reports that, for families whose youngest child is less than 13, time limits reduce welfare use by the most among the families with the youngest children.[38] For families whose youngest child is over 13, many of which will become ineligible before they could reach the federal five-year limit, time limits have no significant effect.

Family Caps

Six studies estimate the effects of family caps using modified dummy variables. Two of the coefficients in Panel D are significant. They are mixed in sign, however, providing little reliable guidance as to the effects of this important policy reform on welfare use.

Child Support Enforcement

Two studies consider the effects of child support enforcement using variables that reflect the extent of child support payments to families on welfare. Huang et al. (2000) provide high-quality evidence, including using state and year dummies in the regression model and measuring the effects of policy changes via the average payment to welfare families in each state and year. Their estimate is highly significant and suggests that higher child support payments substantially reduce welfare use. Mead (2000) employs a similar policy measure and obtains similar results.

4.3.3. Effects of Reform as a Bundle

Over a dozen studies have attempted to estimate the effects of welfare reform as a bundle. All of these studies characterize reform using modified dummy variables. Some distinguish the effects of waivers from the effects of TANF. The results are presented in Table 4.3.

---

[37]The function is given by age* = (Age of the youngest child – 14) for families with three-year time limits and youngest children less than 15 and by age* = (Age of the youngest child – 15) for families with two-year time limits and youngest children less than 15. For families with older youngest children, age* = 0, since such families will become ineligible when their youngest child turns 18, which will take place before they could possibly exhaust their benefits.

[38]In these studies, age** = (Age of the youngest child – 13) for families with youngest children less than 13. For families with older youngest children, age** = 0.

Table 4.3—Estimated Impact of Waivers or TANF Reforms on Welfare Use: Econometric Studies

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Includes LDV's?	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr.	Other controls: Fixed Effects	Other controls: Policy
A. Studies of the caseload based on administrative data														
CEA (1997)	annual state-level caseloads	total population	76	96	AFDC recipients/population	Log	Any waiver	no	-5.78 (1.94)	-5.78	U, U-1		S, Y, State time trends	B
Levine and Whitmore (1998)	annual state-level caseloads	total population	76	96	AFDC recipients/population	Log	Any wavier	no	-1.52 (2.05)	-1.52	U, U-1		S, Y, State time trends	B, WRA-S
Moffitt (1999)	annual state-level caseloads	total population	77	95	AFDC cases/population	Log	Any waiver	no	-5.751 (2.600)	-5.75	U, U-1		S, Y, State time trends	B
Blank (2000)	annual state-level caseloads	total population	77	96	AFDC cases/population	Log	Any waiver	no	-0.064 (0.018)	-6.4	U, U-1, U-2, median wage, 20th percentile wage	A, E, R, Immigration, Non-martial births	S, Y	B, UP program, party control, Medicaid spending
Wallace and Blank (1999)	annual state-level caseloads	total population	80	96	AFDC cases/population	Log	Any waiver	no	-0.072 (0.020)	-7.2	U, U-1, U-2, median wage, 20th percentile wage	A, E, R, Immigration, Non-martial births	S, Y	B, UP program, party control, Medicaid spending
	monthly state-level caseloads	total population	80	98	total AFDC cases	Log	Any waiver	no	-0.138 **	-13.8	U, U-1,..., U-12		S, state-specific month dummies	
							TANF		-0.347 **	-34.7				
CEA (1999)	annual state-level caseloads	total population	76	98	AFDC recipients/population	Log	Any waiver	no	-9.4***	-9.4	U, U-1, U-2		S, Y, State time trends	B, MW
							TANF		-18.84***	-18.84				
Huang et al. (2000)	Annual state-level caseloads	Women 15-44	76	96	AFDC Basic cases / females 15-44	Log	Any waiver	no	-0.080 (0.022)	-8.0	U, U-1, U-2, median wage, 10th% wage	A, E, R, Immigration	S, Y	B, CSE, UP program, party control, Medicaid spending
Figlio and Ziliak (1999)	annual state-level caseloads	total population	76	96	AFDC recipients/population	Log	Any waiver	yes	0.505 (1.274)	1.3 (a)	U, U-1,...,U-4		S, Y, State time trends	B
Bartik and Eberts (1999)	Annual state-level caseloads	total population	84	96	AFDC cases/population	Log	Any waiver	yes	0.007 (0.007)	7.6 (a)	U, U-1, U-2; EG, EG-1, EG-2; WP, WP-1, WP-2; HS, HS-1, HS-2		S, Y	B

Table 4.3—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Includes LDV's?	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr.	Other controls: Fixed Effects	Other controls: Policy
B. Studies of welfare use based on survey data														
1. Models of welfare use														
Moffitt (1999)	CPS aggregated	women 16-54	77	95	welfare users/population	Level	Any waiver	no	-0.010 (0.004)	-20.4	U, U-1	A, E, ExY, ExU, ExU-1	S, Y, State time trends	B
Grogger (2000)	CPS micro data	single mothers 16 - 54	78	98	Any welfare during year	Level	Any reform (waiver or TANF)	no	-0.021 (0.008)	-6.6	U	A, E, R, No. kids, age youngest kid	S, Y	B, MW
O'Neill and Hill (2001)	CPS micro data	single mothers 18 - 44	82	99	Any welfare during year	Level	Any waiver	no	-0.51	-1.6	U, W, college wage premium	A, E, R, No. kids, age youngest kid, ever married, urban-rural	S, trend and trend squared, state-year trends	B
							TANF		-6.48***	-20.9				
2. Models of welfare use disaggregated by maternal education														
Moffitt (1999)	CPS aggregated	women 16-54, educ<12	77	95	welfare users/population	Level	Any waiver	no	-0.017 (0.006)	-34.7	U, U-1	A, E, ExY, ExU, ExU-1	S, Y, State time trends	B
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	welfare users/population	Level	Any waiver	no	-0.0086 (0.0038)	-8.2	U, U-1, EG, each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	no	0.0020 (0.0031)	4.3				
		women 16-54, educ>12					Any waiver	no	0.0003 (0.0025)	1.7				
		women 16-54, educ<12					TANF	no	-0.019 (0.0088)	-18.1				
		women 16-54, educ=12					TANF	no	-0.0115 (0.0077)	-25.0				
		women 16-54, educ>12					TANF	no	0.0058 (0.0059)	32.2				
C. Study of welfare dynamics using administrative data														
Mueser et al (2000)	Quarterly county-level caseloads	Central counties of five urban areas	90 to 93	97	AFDC entrants/population	Log	Any waiver	no	-0.022 (0.055)	-2.2	U		County dummies, county-specific trends, county-specific quarter dummies	
							TANF		-0.130 (0.051)	-13.0				
					AFDC leavers/AFDC caseload	Log	Any waiver	no	0.215 (0.049)	21.5	U		County dummies, county-specific trends, county-specific quarter dummies	
							TANF		0.110 (0.052)	11.0				

NOTES: Abbreviations: LDV=lagged dependent variable; s.e.=standard error; U=unemployment rate; U-m=mth lag of unemployment rate; EG=employment growth; A=age, E=education, R=race, B=maximum welfare benefit, MW=minimum wage; EITC=Earned Income Tax Credit; S=state; Y=year; WP=wage premium predicted from state industrial mix; HS = fraction of employees with high school diplomas or better, calculated from state industrial mix; CSE=child support enforcement; WRA-S=sanctions for non-compliance with work-related activities.

** = statistically significant at the 5 percent level; no standard error reported.

*** = statistically significant at the 1 percent level; no standard error reported.

(a) Long-run effect.

Of the nine sets of estimates that are based on administrative data and reported in Panel A of Table 4.3, seven indicate that the introduction of any (statewide) waiver reduced the caseload. These estimates range from –1.5 percent to –13.8 percent. The lowest estimate is from Levine and Whitmore (1998) and comes from a model that also controls for stricter sanctions. It should be interpreted as the effect of reform policies other than stricter sanctions and, thus, is not directly comparable to the estimates from the other studies. The two studies that estimate the effects of TANF report even larger effects, ranging from –18.8 to –34.7 percent. Most of these estimates are significant and all suggest that reform, under both waivers and TANF, has reduced welfare caseloads. [39]

Four sets of authors provide results from the March CPS. Grogger (2000) and O'Neill and Hill (2001) analyze data from single mothers, who are the primary recipients of cash aid. They analyze individual-level data on welfare use, using a dummy dependent variable that is equal to one for women who report welfare use in the previous year.

Moffitt (1999) and Schoeni and Blank (2000) focus on women between the ages of 15 and 54. They first aggregate the data into cells defined by state of residence, year, age, and education. These cells constitute their units of observation. Their dependent variable is the rate of welfare use within each of the cells (not the logarithm of that rate). These authors estimate separate models by different levels of education. This allows us to see whether the estimated effects of reform are concentrated primarily among the poorly educated, who make disproportionate use of the welfare system. If instead the estimated effects of welfare reform were similar across all levels of education, we would be concerned that the estimates did not truly reflect the effects of reform, but rather of some unobservable confounding factor.

Moffitt analyzes the effect of waivers on welfare use; Grogger analyzes the effect of reform, defined by the presence of either a statewide waiver or a TANF plan; and Schoeni and Blank and O'Neill and Hill consider the effects of waivers and TANF separately. All use the unemployment rate to control for the state of the economy, each states' maximum welfare benefit, and a set of variables to control for maternal education and age. O'Neill and Hill control for wages as well. Grogger and O'Neill and Hill also control for family size, race, and the age of the youngest child in the family. All the models control for state fixed-effects. Most use year dummies to control for general trends in welfare use. O'Neill and Hill are the exceptions, using more restrictive linear and quadratic terms in time, instead.

Of the estimates from Moffitt, Grogger, and O'Neill and Hill, all but one are negative and significant. In relative terms, they indicate that welfare reform reduced welfare use by 2 to 20 percent, which is within the range of estimates from the studies based on administrative caseloads. O'Neill and Hill report that TANF has larger effects than waivers, echoing the results from CEA (1999) and Wallace and Blank (1999).

As noted above, Moffitt and Schoeni and Blank report separate results for groups defined by their level of education, which are presented in Panel B2 of Table 4.3. Regarding the effects of waivers, both sets of authors find that welfare reform has its largest effects on high school dropouts and smaller effects on women who attended college. In both studies, only the effects

[39] Wallace and Blank (1999) note that many of these models fail to track pre-PRWORA caseload trends very well.

for dropouts are statistically significant. Schoeni and Blank find TANF to have larger effects than waivers, which is consistent with the evidence from other authors. Their estimates for dropouts and high school graduates are significant, whereas their estimate for women with higher education is positive though insignificant.

Panel C of Table 4.3 reports results from a study of welfare dynamics, that is, of entries and exits from the welfare rolls. Mueser et al. (2000) find that reform has affected both types of welfare transitions, although the effects vary by the type of reform. Waivers have small and insignificant effects on welfare entries, whereas TANF reduced welfare entries significantly. Both types of reforms increase welfare exits significantly, but the effects of waivers are stronger.

Although the majority of the evidence presented in Table 4.3 is consistent with the notion that reform as a bundle reduced welfare use, two studies suggest that it may have actually raised welfare use, albeit with an insignificant positive effect (Figlio and Ziliak, 1999; Bartik and Eberts, 1999). Beyond their conclusions, these two outliers differ from the other studies by including lagged dependent variables (that is, past values of the welfare caseload) as explanatory variables in their regression models. In the jargon of econometrics, models that include lagged dependent variables are referred to as dynamic, whereas models without them are referred to as static. The authors argue that such dynamic models are necessary to capture potentially sluggish adjustment of caseloads to changes in policy and economic conditions. According to Figlio and Ziliak (1999), it is the presence of the lagged dependent variables, more than any other reason, that explains the difference in results between the static and dynamic models.

For understanding the effects of welfare reform, this raises two important questions. First, why do the lagged dependent variables make such a difference? Second, which set of estimates is correct, if either?

Adding lagged dependent variables to a model raises a number of technical issues that do not arise in the context of static models. First, state fixed-effects models are inconsistent in the presence of lagged dependent variables. Although Ziliak et al. (2000) choose an alternative approach to deal with the unobserved heterogeneity problem, that method (known as first-differencing) may also yield inconsistent estimates in the presence of lagged dependent variables (Nickell, 1981).[40]

Furthermore, adding lagged dependent variables to the model exacerbates an already difficult collinearity problem. From inspecting Figure 4.1, it is easy to see that lagged values of the caseload are highly correlated with both the unemployment rate and the adoption of state welfare reforms, both of which are already highly correlated with each other. Adding lagged dependent variables to the model thus turns the difficult problem of distinguishing the effects of welfare reform from the effects of the economy into the even more difficult problem of distinguishing the effects of welfare reform from the effects of the economy and from recent trends in the caseload itself.

Such an undertaking might nevertheless be worthwhile if, in the end, we were left with a model that provided a deeper understanding of welfare dynamics and the effects of policy on the behavior of welfare recipients. However, recent work suggests that adding lagged dependent

---

[40]Ziliak et al. (2000) claim that their sample period is long enough to avoid these problems, but they provide no direct evidence to support their claim.

variables to an otherwise static regression model is unlikely to yield such an understanding. This insight stems from a recent study that analyzes the relationship between the nominally dynamic regression models that appear in the caseload literature and true welfare dynamics (Klerman and Haider, 2000).

In the welfare setting, dynamics refer to “flows,” that is, to transitions on and off the welfare rolls by welfare entrants and welfare leavers. With information on those flows, we can compute the “stocks,” that is, the caseload, at any point in time. With information on how policy affects those flows, we can compute how policy affects the stocks, both instantaneously and over time. Clearly, such information would serve a number of important purposes.

In the sense of Figlio and Ziliak (1999) and Bartik and Eberts (1999), however, the notion of dynamics refers not to welfare transitions and how welfare flows affect welfare stocks, but rather to how past welfare stocks are correlated with current welfare stocks. A key question is thus whether correlations among past and current stocks provide even indirect information about welfare transitions. Klerman and Haider (2000) address this question and provide a number of conditions under which nominally dynamic regression models, that is, those that include lagged values of the caseload as regressors, provide information on welfare dynamics. The conditions are highly restrictive. One condition requires welfare exit rates to be independent of the amount of time that the recipient has already spent on welfare. A substantial body of empirical evidence points to the contrary, indicating that exit rates fall as the spell length increases (Blank, 1989; Bane and Ellwood, 1994).

As a result, the nominally dynamic models of Figlio and Ziliak (1999) and Bartik and Eberts (1999) are unlikely to provide even indirect information about the effects of welfare reform on welfare dynamics. Of course, the same can be said about the static models: None of them addresses the issue of welfare dynamics at all. At the same time, however, the technical issues associated with consistently estimating static models are less difficult than those associated with the nominally dynamic models, and, in addition, the collinearity issues that confront the static models, while considerable, are less daunting than those confronting the nominally dynamic models.

Moreover, the results from the more disaggregated studies of welfare reform also suggest that reform has affected welfare use. The studies of Moffitt and Schoeni and Blank find that the effects of reform are concentrated among women with low levels of education, which increases our confidence that the results are “real,” and not merely the result of unobserved confounding factors. Put differently, if they had found that welfare reform affected college women to the same extent that it affects dropouts, they would have cast doubt on all the prior studies. However, because they found the effects to be strongest among dropouts, they increase our confidence that the estimates reflect the effects of reform, rather than unobservable confounding influences. Perhaps more importantly, direct evidence on welfare dynamics indicates that welfare reform affects entries and exits in plausible ways (Mueser et al., 2000; Hofferth, Stanhope, and Harris, 2000a, 2000b).

A further piece of evidence also suggests that the lagged dependent variable models understate the effects of reform. Ziliak et al. (2000) use nominally dynamic models to estimate the effects of a number of specific policy reforms, including work-related activity mandates. Their estimate, which appears in Panel B2 of Table 4.2, indicates that work-related activity mandates have a very small effect on the caseload.[41] This contrasts with the results from the 13 random assignment studies that focused on work-related activity mandates, 12 of which produced significant decreases in welfare use.

Finally, beyond the narrow question of whether nominally dynamic econometric models understate the effects of welfare reform, there is the more general question of whether welfare reform contributed to the unprecedented decline in welfare use that took place during the 1990s. On this more general question, further evidence can be brought to bear. In the next chapter, we show that almost all of the random assignment experiments increased employment. The few econometric studies on the topic concur that reform increased work among welfare-prone populations.

This is not altogether surprising, since all the major policy reforms would be expected to increase employment. Although different reforms have different incentives for welfare use, all the major reforms provide positive incentives for work. Nevertheless, this observation makes an important point: The major welfare reforms implemented during the 1990s affected behavior. Given that they affected behavior, the only logical argument that we could make to support the contention that they did not affect welfare use is that, on a nationwide basis, the conflicting incentives resulting from the different reforms exactly cancelled each other out.

The results from Table 4.1 show that, in individual cases, such “policy canceling” can occur. Nevertheless, it seems unlikely to have occurred on a nationwide basis. The reason is that the financial work incentives implemented in MFIP, FTP, and Jobs First, which provide perhaps the strongest evidence of policy canceling, were among the most generous in the country. Nationwide, although 36 states had implemented some sort of financial incentive by 1997, 37 had implemented work-related activity mandates that were more demanding than AFDC/JOBS, all but two had implemented more stringent sanctions for noncompliance, and all but three had implemented time limits.[42] Of course, most states implemented several such polices in combination. From a purely numerical perspective, it seems unlikely that the typical financial incentive could have completely offset the combined effects of work-related activity mandates, sanctions, and time limits, each of which appears to have reduced welfare use.

For all these reasons, as well as the direct evidence provided by most of the econometric studies, we conclude that welfare reform played an important role in reducing the welfare caseload during the late 1990s. This is not meant to deny the importance of other factors. The economy played an important role, and as several analysts have suggested, it may have been the single most important explanation for why caseloads fell. However, despite a few studies that make claims to the contrary, the bulk of the evidence suggests that welfare reform made an important contribution.

---

[41]Moreover, only two of the five coefficients that contribute to the estimate in Table 4.2 were both negative and significant.

[42]These figures are based on the authors' tabulations of data from CEA (1999).

4.4. EVALUATING THE EFFECTS OF WELFARE REFORM ON WELFARE USE

Having presented the results of several studies, in this section we attempt to synthesize them to convey what is known about the effects of welfare reform on welfare use. We consider the random assignment and econometric studies together, weighing both the quantity and quality of the evidence. We begin by discussing the effects of specific reforms and then turn to discussing their effects as a bundle.

4.4.1. Effects of Specific Reforms

Financial Work Incentives

The effects of financial work incentives are the primary focus of three high-quality random assignment studies and six econometric studies, only one of which rates as high-quality by the criteria discussed in Chapter 3. Estimates from MFIP-IO and the high-quality econometric study (CEA, 1999) indicate that stronger financial work incentives are associated with higher rates of welfare use, as the standard economic model would predict. The CWPDP and WRP-IO programs had no significant effect on welfare use, but their financial work incentives were fairly weak. Estimates from all but one of the lower-quality econometric analyses are insignificant.

The programs that combine financial work incentives and mandated work-related activities also shed some light on the effects of the incentives. Since the financial incentive should increase welfare use, all else equal, whereas the work-related activity mandate should decrease it, the net effect of such programs is ambiguous. Of the four random assignment studies that combine these two reforms, those with stronger financial work incentives tend to generate positive net effects on welfare use, while those generating insignificant or negative effects involved relatively weak financial incentives.

Evidence from the three SSP programs, which tied financial incentives to hours of work, is consistent with this pattern as well. Those programs provided strong financial work incentives in the form of earnings supplements for consumers who satisfied the programs' work requirements. Those programs substantially increased the rate at which consumers received transfer payments.[43]

Mandatory Work-Related Activities

Mandatory work-related activities have received a substantial amount of study, both from random assignment and econometric analyses. As a result, certain conclusions about these policies can be drawn fairly strongly. Of the 13 random assignment studies that focus on work mandates, 12 generated significant declines in welfare use during the first two years after random assignment. Programs with stronger enforcement generally had larger effects. The programs stressing job search generally yielded greater decreases than the programs stressing

---

[43]New Hope may appear to be an exception to this general rule, since it provided a strong financial incentive but did not raise AFDC use. However, for programs such as New Hope, the economic model predicts an increase in the rate of transfer payments, that is, in receipt of the earnings supplement. Unfortunately, Bos et al. (1999) do not provide impact estimates for receipt of the earnings supplement that are comparable to the impact estimates for AFDC.

skills development, but the mean difference was relatively small, and the contrast involved only four search-oriented programs. The impacts of both types of programs faded over time.

Much less can be said about other aspects of states' work-related activity mandates. Among the several econometric studies that analyze age-exemption thresholds or shorter deadlines for satisfying the mandates, the conclusions are quite mixed. Only one of these studies satisfies our criteria for providing high-quality evidence, and it yields the perverse (though generally insignificant) result that more stringent exemption criteria increase the caseload.

Nine studies estimate the effect of sanctions for noncompliance with the work mandates. Seven report that sanctions significantly reduce welfare use. Three studies report that stricter sanctions have greater effects than weaker sanctions. However, the interpretation of those estimates is clouded by results suggesting that sanction policies implemented after 1994 reduced caseloads between 1989 and 1992. Moreover, none of the studies estimates the effects of the monetary value of sanctions, and none incorporates any information about the frequency with which sanctions are actually imposed.

Time Limits

Of the seven moderate-quality econometric studies that implicitly estimate the behavioral effects of time limits, most suggest that time limits reduce welfare use, although only one is even marginally significant. Four high-quality econometric studies, two of which are very similar, find that time limits reduce welfare use the most among families with the youngest children. This suggests that time limits have behavioral effects, because the families that reduce their current welfare use the most are those with the most to lose by prematurely exhausting their benefits. However, all these studies rely on a number of assumptions to isolate the effects of time limits from the effects of other reforms.

Only two studies provide evidence on the mechanical effects of time limits. Evidence from both FTP and Jobs First suggests that welfare use falls considerably as families begin to exhaust their benefits.

Family Caps

Family caps have been the subject of six econometric studies, none of which provide high-quality evidence on their effects. Their results are mixed, providing little insight into the question of whether family caps have any effect on current welfare use.

Child Support Enforcement

Two econometric studies, including one that provides high-quality evidence, estimate the effects of child support enforcement on welfare caseloads. Both estimate that child support enforcement has substantially reduced the rolls.

4.4.2. The Effects of Reform as a Bundle

Six random assignment studies involve TANF-like bundles of reforms. Of those, four suggest that reform as a bundle reduced welfare use. Three of the programs showing the smallest reductions had time limits that were poorly understood by participants. If they had been better understood, the programs probably would have reduced welfare use more. The only program to raise welfare use prior to time limits becoming binding was Jobs First, which included an extraordinarily generous financial incentive.

In addition, over a dozen econometric studies have attempted to estimate the effects of reform as a bundle on aggregate welfare use. Except for a few analyses that use lagged values of the caseload as controls for current caseloads, these studies generally find that reform has reduced the welfare rolls. For the reasons we detail above, we place relatively little weight on the studies that employ lagged caseloads.

Because we view the lagged-caseload studies less favorably than Bell (2001), who recently reviewed much of the econometric literature on the caseload decline, it is worthwhile to explain why our conclusions differ. First, Bell is more optimistic than we are that the lagged-caseload models provide a reasonable approximation to the process of welfare entry and exit that must underlie any change in the caseload. Although we agree with Bell about the importance of analyzing such welfare transitions directly, our reading of recent research suggests that lagged-caseload models are unlikely to provide insights into the process driving entries and exits (Klerman and Haider, 2000).

Beyond this technical difference of opinions, however, we think there is a further reason why our conclusions differ: We consider a broader range of evidence. Bell did not review either of the studies that analyze welfare transitions directly. Moreover, he limited his review to econometric studies. By considering the results from Mueser et al. (2000) and Hofferth, Stanhope, and Harris (2000a, 2000b), we see that the indirect evidence on welfare dynamics from the lagged-caseload studies often contradicts the direct evidence from the welfare-transitions studies. By considering the random assignment studies, we see that the lagged-caseload results regarding the effects of specific policy reforms often contradict the results from high-quality experiments. In our view, this additional evidence is persuasive.

Estimates from the other studies of reform as a bundle generally suggest that reform has played an important role in reducing the caseload. However, precise estimates vary widely. The estimated reductions attributable to waivers range from 12 to 31 percent. The three studies that attempt to distinguish the effects of TANF from the effects of waivers generally find TANF to have even larger effects.

At the same time, nearly all the econometric studies agree that the economy played an important role in reducing the caseload.[44] Most of the analyses suggest that the economy accounted for one-fourth to one-half of the 1993–1996 decline in the welfare rolls. The few analysts who consider the post-PRWORA period explicitly generally credit the economy with a smaller fraction of the caseload decline over that period, which is consistent with the fact that the decline in the unemployment rate slowed during the late 1990s. Other social policy changes

[44]The exception is Rector and Youssef (1999).

had important effects as well, with one estimate suggesting that the changes to the EITC explained 16 percent of the decline in welfare use (Grogger, forthcoming). Welfare reform appears to have played an important role in reducing the caseload, but it was hardly the only factor underlying the unprecedented declines of the mid- to late-1990s.

4.5. CONCLUSIONS

Many of the effects of welfare reform on welfare use have been well-studied. Over a dozen econometric studies have attempted to estimate the effects of reform as a bundle, and all but a few report that reform had substantial effects on the caseload. Moreover, as we explain above, the contradictory evidence comes from a small number of studies that employ a technique that poses considerable technical challenges. All but one of the econometric studies concur that the economy played an important role in reducing caseloads during the 1990s. Regardless of how effective welfare reform might have been, in the absence of the booming labor market, the decline in welfare use would have been substantially smaller.

In terms of the effects of specific reforms, over a dozen experimental studies have focused on mandatory work-related activities, and most find that such policies reduce welfare use by a significant and substantial amount. The few that find otherwise generally involved weakly enforced mandates or included a generous financial incentive along with the work mandate.

Beyond work-related activity mandates, however, evidence on the effects of specific reforms becomes thinner. Evidence on the effects of financial work incentives is consistent with the notion that substantially increasing the generosity of financial work incentives increases welfare use. Much of the evidence for this conclusion is inferred from programs that combined financial work incentives with mandatory work-related requirements. Only four high-quality studies focus directly on financial work incentives.

Sanctions for noncompliance with mandated activities have been the focus of substantial policy interest. Several studies find that stricter sanctions lead to a greater reduction in the caseload, but some of that reduction may have preceded the implementation of the sanctions. None of the studies to date attempts to monetize the effects of sanctions, which may be important since a full-family sanction in a low-benefit state may actually cost the family less than a partial sanction in a high-benefit state. Even more troubling is the fact that no study has estimated how the frequency of sanctions affects welfare use. States vary widely in the extent to which they actually impose sanctions, and deterrence theory (Becker, 1968) suggests that moderate sanctions imposed with a high frequency may be as effective as severe sanctions that are seldom imposed.

Similarly, there have been relatively few high-quality studies of the effects of time limits. Four econometric studies (two of which are quite similar) show that time limits have greater effects on families with younger children. This suggests that time limits have behavioral effects, since it suggests that families that have more to lose by prematurely reaching the limit are more likely to reduce their current welfare use. Two random assignment studies show sharp declines in welfare use starting at the time when families begin to exhaust their benefits. The importance of such mechanical effects is likely to grow in the near future, as increasing numbers of recipients reach the five-year limit on federal funding.

After time limits, the number of studies providing impacts for specific policy reforms gets even smaller. There is one high-quality study on the effects of age exemptions for mandatory work-related activities and one on the effects of child support enforcement. There are none on the effects of family caps.

In summary, if we think about the welfare column of our policy-outcome matrix, a few cells are well-filled. There are a few more that contain evidence from a few high-quality studies. The fact that these results generally accord with predictions from the standard economic model gives us more confidence in them than we would have based on their numbers alone. However, many of the rows are nearly empty, including some that involve important policy reforms.

Chapter Five

EMPLOYMENT AND EARNINGS

5.1. BACKGROUND

Beyond reducing welfare dependence, one of the key goals of PRWORA is to increase work. By all accounts, work has risen in recent years. The fraction of welfare recipients working rose from 7 percent in 1992 to 33 percent in 1999. In 1998, the fraction of welfare recipients starting jobs exceeded 50 percent in 10 states. The fraction of job entrants still employed after three months exceeded 75 percent in 29 states (USDHHS, 2000).

Among single mothers more generally, recent trends in employment and earnings are the mirror image of recent trends in welfare use. Employment among single mothers rose from 69 percent in 1993 to 83 percent in 1999, a gain of 20 percent. Weeks worked during the year, a broader measure of work effort that we will refer to as "labor supply," rose from an average of 29.5 to 36.7 over the same period. Mean trends in earnings have also been favorable. Measured in 1998 dollars, average annual earnings among female family heads (including those without earnings) stood at roughly $12,300 in 1993. By 1999, they had risen to nearly $16,600, a gain of 35 percent (Grogger, forthcoming).

Although these average trends are favorable, some low-income families may have lost economic ground in recent years. Earnings among the lowest 20 percent of female-headed families fell by roughly 10 percent between 1995 and 1997, but appear to have rebounded in the meantime (Primus et al., 1999; Haskins, 2001). Fifteen welfare-leavers studies sponsored by USDHHS show that roughly 50–65 percent of persons leaving welfare are employed in their first quarter off aid (USDHHS, 2001a). Approximately 60–70 percent find employment at some point within the first year. However, many of these jobs are fairly unstable, since only about 40 percent of welfare leavers on average are employed in each of their first four quarters off aid.

These studies also show that earnings among employed welfare leavers average roughly $1,800 to $3,400 in their first quarter off the rolls. There is some evidence of earnings growth, but it is small. By their fourth quarter off welfare, leavers with earnings typically make between $2,100 and $3,900 (USDHHS, 2001a). Such relatively flat earnings trajectories could stem either from low wage growth or from intermittent work. Recent research shows that the wages of low-income women rise with work experience in a manner that is comparable to other workers when experience is measured as actual hours of previous work (Gladden and Taber, 2000; Loeb and Corcoran, 2001). This suggests that low earnings growth is due to sporadic employment.

Although this descriptive evidence provides an invaluable context for interpreting the effects of recent policy changes, our interest lies not in merely describing the experience of low-income families in the post-reform era, but rather in assessing how welfare reform in general, and specific reform measures in particular, have affected employment and earnings. The economic model discussed in Chapter 2 predicts that nearly all the recent reforms should increase the employment of single parents. Specifically, because financial work incentives increase the amount of earnings a working recipient may keep, they promote employment. Other reforms encourage recipients to seek work as well, either as a condition for receiving aid or in anticipation of reaching their time limit. The one possible exception involves work requirements that mandate participation in an education-focused welfare-to-work program. Recipients participating in such a program may actually work less than they would have otherwise, at least while they are taking part in the program.

However, the earnings effects of some of the reforms are ambiguous, because of the countervailing incentives discussed in Chapter 2. As noted there, financial work incentives increase the rate of return from work, but they also raise income for a given level of employment. The effect on returns, known as the "substitution effect" in the economics literature, acts to increase labor supply, but the effect on income, known as the "income effect," may induce the consumer to reduce her labor supply. The net effect of the financial work incentive on labor supply and earnings is thus ambiguous, depending on the relative magnitudes of the substitution and income effects. All else equal, we would expect smaller income effects among families with lower levels of income, simply because such families are less able to "afford" to reduce their labor supply in response to a financial incentive.

To illustrate this point, consider two women, each of whom earns $6/hour. The first works 40 hours per week, the second works 10 hours per week. A decrease in the benefit reduction rate from 67 to 50 percent raises each worker's net wage by $1/hour. If the workers do not change their hours, this would result in a monthly earnings gain of $160 for the first worker and $40 for the second worker. Alternatively, the first worker could reduce her labor supply by 10 hours per week and still enjoy greater earnings than she made before the benefit reduction rate reduction. However, if the second worker reduced her hours by the same amount, her earnings would fall to zero.

The logic of this example is sometimes used to argue not only that substitution effects should be larger among families with lower income, but that, among welfare families, the substitution effect should dominate the income effect, causing reductions in benefit reduction rates to increase labor supply and earnings as well as employment. However, some of the evidence below suggests that, even at the levels of income received by some welfare recipients, the income effect may outweigh the substitution effect.

Other polices with complex effects on labor supply and earnings include work requirements and time limits. Both policies encourage employment, either directly or indirectly. However, both policies may hasten job search, which could lead recipients to accept jobs with lower wages and perhaps other characteristics that might result in a less durable match between the worker and the employer. Thus, it is conceivable that these policies could result in fewer hours of work and lower earnings.

As in Chapter 4, we consider evidence on the effects of reform from both random assignment and econometric studies. The two types of studies typically employ different types of employment and earnings data. Although some random assignment studies collect survey data, most use administrative data from the states' unemployment insurance (UI) systems. Every quarter, all employers covered by the UI system are required to report the earnings of all their employees who earn $50 or more. This information constitutes the earnings data used in most random assignment studies, although the studies sometimes analyze annual earnings or even earnings over a multiyear follow-up period instead. Because the units of measurement vary among studies, we report not only the impacts from the studies themselves, but also a normalized monthly impact measure. Typically, a study participant is considered to be employed if she has any reported earnings in a given quarter, although some studies adopt an annual measure of employment instead.

The main problem with such administrative data is that they miss some earnings. The UI system covers about 90 percent of all jobs in the United States; uncovered sectors include self-employment, federal government employment, some state and local government employment, some domestic jobs, and some agricultural jobs (U.S. Bureau of Labor Statistics, 1989). Of course, the informal sector is uncovered as well, which means the UI data miss income from people who provide informal child care, take in laundry, and otherwise work for cash. A comparison of employment data from administrative and survey sources suggests that administrative data underestimate self-reported employment among welfare leavers by about 10–20 percent (Isaacs and Lyon, 2000, Table 2C).

All the econometric studies are based on data from the March CPS. They therefore use fairly similar measures of employment and earnings. Employment is typically measured as a dummy variable that equals one if the survey respondent worked for pay in the year preceding the survey and equals zero otherwise. Respondents also indicate the number of weeks they worked for pay in the previous year, which provides a useful measure of labor supply. In addition, respondents report their income from earnings in the previous year. As in the econometric literature on welfare use, some researchers analyze the individual-level data directly, whereas others first aggregate the data into cells defined by the respondent's education, age, state, and year.

Many of the analytical issues that arise in econometric studies of the caseload also arise in studies of employment and earnings. In particular, welfare reform is but one of several factors that contributed to recent changes in the labor market behavior of single mothers. The strong economy played an important role, as did the EITC (Meyer and Rosenbaum, 2001; Grogger, forthcoming; Hotz, Mullin, and Scholz, 2001). Other, largely unobservable factors may have played a role as well. As in the caseload literature, researchers have attempted to account for such confounding influences by including controls in their regression models for the economy, characteristics of survey respondents, other policy variables, and state-specific fixed-effects and trends.

In this chapter, we review the evidence on how much welfare reform has increased employment among the welfare-eligible population. We also examine the effects of welfare reform on earnings, which represent the primary means of support for low-income families not receiving welfare. The effects of welfare reform on earnings bear importantly on its consequences for income and other broad measures of family well-being that we address in later chapters.

As in Chapter 4, we begin with a discussion of the random assignment studies, including a brief summary of the results from Appendix A with analyses by subgroups. Following that, we discuss the results from a number of econometric studies. We then synthesize the studies to convey what is known about the effects of welfare reform on employment and earnings. We conclude with a summary of our findings.

5.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECTS OF WELFARE REFORM ON EMPLOYMENT AND EARNINGS

Most of the evidence on the effects of welfare reform on employment and earnings comes from random assignment studies. Table 5.1 summarizes their findings. As in Chapter 4, when the estimated effects of a program seem to change with time since random assignment, or when time limits begin to bind, we report multiple estimates from the same program. Otherwise, we present a single estimate.

5.2.1. Programs That Focus on Financial Work Incentives

Results from the three programs that focus on financial work incentives are presented in Panel A of Table 5.1. According to Becerra et al. (1998), CWPDP reduced employment by 2 percentage points. Although the program impact is insignificant, it appears to contradict the prediction from the economic model discussed in Chapter 2. However, results from a reanalysis of the CWPDP data show that the program increased employment (by 3.1 percentage points in the third year of the follow-up). Moreover, the employment impacts from that reanalysis are statistically significant (Hotz, Mullin, and Scholz, 2002). However, the authors provide no reconciliation of their results with those of Becerra et al. (1998). Since it is beyond the scope of this synthesis to provide a reconciliation of these contradictory findings, we omit the results of CWPDP from the discussion below.

WRP Incentives-Only had positive but small and insignificant effects on employment and earnings. This seems consistent with the weak incentive that was offered by the program. The MFIP Incentives-Only program, with its stronger financial incentive, significantly increased employment among both recipients and applicants in the first year after random assignment. Although the effects fade over time, they remain significant among recipients, for whom the first-year effects were stronger as well.

As for earnings, MFIP Incentives-Only had no significant effect, and most of the insignificant estimates are negative rather than positive. This may happen because the income effect, which provides an incentive to decrease hours of work, outweighs the substitution effect, which provides an incentive to increase hours of work. [45] As noted in Chapter 2, we would expect the income effect to be larger at higher levels of income. MFIP has a relatively high benefit level, providing $9,228 per year to a family of three with no other income. Compared to other welfare programs, this represents a relatively high level of income.

---

[45]Evidence on hours of work would enable us to draw this conclusion more directly.

Table 5.1—Estimated Impact of Welfare Reform on Employment and Earnings: Random Assignment Studies

Name	Cases served	Data	Measure	Employment: Control mean	Employment: Impact	Employment: %	Earnings: Control mean	Earnings: Impact	Earnings: %	Earnings: Normalize to monthly
A. Programs that focus on financial work incentives										
CWPDP	Single parent recipients	A	Avg. employment, earnings, year 3	37.0	-2.0	-5.4%	$2,372	-$160	-6.7%	-$13
WRP-IO	Single-parent recipients and applicants	A	Ever emp, avg. quarterly earnings, last 3 mos. of FU	48.4	2.8	5.8%	$1,502	$14	0.9%	$5
MFIP-IO	Urban single parents recipients	A	Avg. quarterly employment, earnings, year 1	32.8	7.0 ***	21.3%	$537	$50	9.3%	$17
		A	Avg. quarterly employment, earnings, year 3	44.7	3.6 *	8.1%	$1,298	-$48	-3.7%	-$16
	Urban single parents applicants	A	Avg. quarterly employment, earnings, year 1	48.8	2.7 *	5.5%	$1,216	-$66	-5.4%	-$22
		A	Avg. quarterly employment, earnings, year 3	55.3	0.0	0.0%	$2,017	-$136	-6.7%	-$45
B. Programs that focus on financial work incentives tied to hours of work										
New Hope	Poor families employed FT at RA	A	Ever employed, total earnings, year 1 of 2-yr FU	94.7	2.5	2.6%	$10,480	-$253	-2.4%	-$21
		A	Ever employed, total earnings, year 2 of 2-yr FU	91.8	2.6	2.8%	$11,550	-$889	-7.7%	-$74
	Poor families not employed FT at RA	A	Ever employed, total earnings, year 1 of 2-yr FU	77.9	9.9 ***	12.7%	$4,380	$916 ***	20.9%	$76
		A	Ever employed, total earnings, year 2 of 2-yr FU	76.7	6.6 ***	8.6%	$6,129	$473	7.7%	$39
SSP (a)	Single-parent recipients	A	Monthly emp. and annual earnings, year 2	30.4	9.8 ***	32.2%	$3,198	$1,254 ***`	39.2%	$105
		A	Monthly emp. and annual earnings, year 3	32.5	7.2 **	22.2%	$3,852	$865 ***	22.5%	$72
SSP Plus (a)	Single-parent recipients	A	Employment, earnings, Q5	31.1	16.2 ***	52.1%	$221	$120 ***	54.3%	$120
SSP Applicants (a)	Single-parent applicants	A	Employment, earnings, Q5	38.1	4.1 **	10.8%	$552	$78 **	14.1%	$78
		A	Employment, earnings, Q9	42.8	12.1 ***	28.3%	$610	$242 ***	39.7%	$242

Table 5.1—Continued

Name	Cases served	Data	Measure	Employment: Control mean	Employment: Impact	Employment: %	Earnings: Control mean	Earnings: Impact	Earnings: %	Earnings: Normalize to monthly
C. Programs that focus on mandatory work-related activities										
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Ever emp., avg. total earnings, 2-year FU	57.6	9.6 ***	16.7%	$6,385	$1,627 ***	25.5%	$68
Atlanta LFA	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	61.6	4.5 ***	7.3%	$5,006	$813 ***	16.2%	$34
Grand Rapids LFA	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	70.1	7.6 ***	10.8%	$4,639	$1,035 ***	22.3%	$43
Riverside LFA	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	45.0	15.1 ***	33.6%	$4,213	$1,276 ***	30.3%	$53
Portland	Recipients and applicants; no cases with substantial barriers	A	Ever emp., avg. total earnings, years 1 and 2	60.9	11.2 ***	18.4%	$5,291	$1,842 ***	34.8%	$77
Atlanta HCD	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	61.6	2.8 **	4.5%	$5,006	$496 **	9.9%	$21
Grand Rapids HCD	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	70.1	5.3 ***	7.6%	$4,639	$580 **	12.5%	$24
Riverside HCD	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	38.9	9.3 ***	23.9%	$3,133	$317	10.1%	$13
Columbus Integrated	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	72.2	1.7	2.4%	$6,892	$673 **	9.8%	$28
Columbus Traditional	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	72.2	1.3	1.8%	$6,882	$677 **	9.8%	$28
Detroit	Recipients and applicants	A	Ever emp., avg. total earnings, years 1 and 2	58.2	4.1 ***	7.0%	$4,001	$367 *	9.2%	$15
Oklahoma City	Applicants	A	Ever emp., avg. total earnings, years 1 and 2	65.0	-0.9	-1.4%	$3,514	$5	0.1%	$0
IMPACT Basic Track	Recipients and applicants-Basic Track	A	Employed in Q4; earnings in year 1	44.6	0.7	1.6%	$2,345	-$146	-6.2%	-$12

Table 5.1—Continued

Name	Cases served	Data	Measure	Employment: Control mean	Employment: Impact	Employment: %	Earnings: Control mean	Earnings: Impact	Earnings: %	Earnings: Normalize to monthly
D. Programs that focus on financial work incentives and mandatory work-related activities										
WRP	Single-parent recipients and applicants	A	Ever emp, avg. quarterly earnings, last 3 mos. of FU	48.4	8.7 ***	18.0%	$1,503	$135 *	9.0%	$45
MFIP	Urban single-parent recipients	A	Avg. quarterly employment, earnings, year 1	32.8	13.3 ***	40.5%	$537	$163 ***	30.4%	$54
		A	Avg. quarterly employment, earnings, year 3	44.7	11.5 ***	25.7%	$1,298	$143 *	11.0%	$48
	Urban single-parent applicants	A	Avg. quarterly employment, earnings, year 1	48.8	3.0 **	6.1%	$1,216	-$70	-5.8%	-$23
		A	Avg. quarterly employment, earnings, year 3	55.3	2.8 **	5.1%	$2,017	$15	0.7%	$5
TSMF	Recipients	A	Avg. annual emp., earnings over 4-yr FU	36.2	1.6 ***	4.4%	$3,120	$223 ***	7.1%	$19
	Applicants	A	Avg. annual emp., earnings over 1-yr FU	41.5	-0.6	-1.4%	$3,109	$134	4.3%	$11
		A	Avg. annual emp., earnings over 2-yr FU	39.0	1.0	2.6%	$3,426	$12	0.4%	$1
FIP	Recipients	A	Avg. quarterly emp., earnings in Q4 of 2-yr FU	48.2	1.3	2.7%	$916	$94 **	10.3%	$31
		A	Avg. quarterly emp., earnings in Q8 of 2-yr FU	56.1	1.2	2.1%	$1,334	$57	4.3%	$19
	Applicants	A	Avg. quarterly emp., earnings in Q4 of 2-yr FU	56.8	4.8 ***	8.5%	$1,757	$152 **	8.7%	$51
		A	Avg. quarterly emp., earnings in Q8 of 2-yr FU	58.4	3.1 *	5.3%	$2,097	$121	5.8%	$40

Table 5.1—Continued

Name	Cases served	Data	Measure	Employment: Control mean	Employment: Impact	Employment: %	Earnings: Control mean	Earnings: Impact	Earnings: %	Earnings: Normalize to monthly
E. Programs that focus on other individual reforms										
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
EMPOWER (b)	Recipients	A	Quarterly emp., earnings, Q1-Q10	39.1	-1.1	-2.8%	$886	-$29	-3.3%	-$10
		S	Employment, monthly earnings at mid-1998 interview	51.3	3.4	6.6%	$597	$60	10.1%	$60
IMPACT Placement Track	Recipients and applicants-placement track	A	Employed in Q4; earnings in year 1	50	7.6 ***	15.2%	$3,139	$815 ***	26.0%	$68
		A	Employed in Q8; earnings in year 2	54.4	2.8	5.1%	$4,944	$559 ***	11.3%	$47
VIP/VIEW	Recipients	A	Avg. emp., earnings in year 2 of 2-yr FU	51.3	2.9 ***	5.7%	$2,777	$193 *	6.9%	$16
ABC	Recipients and applicants	A	Any emp., total earnings, Q1-Q4	48.0	9.4 ***	19.6%	$3,378	$446	13.2%	$37
FTP	Recipients and applicants	A	Avg. annual emp, earnings in year 2	43.2	6.5 ***	15.0%	$3,278	$661 ***	20.2%	$55
		A	Avg. annual emp., earnings in year 3	44.6	6.7 ***	15.0%	$3,852	$910 ***	23.6%	$76
		A	Avg. annual emp., earnings in year 4	48.0	1.8	3.8%	$4,640	$567 ***	12.2%	$47
JOBS First	Recipients and applicants	A	Ever emp., avg. earnings in Q7	48.6	8.6 ***	17.7%	$1,424	$149 ***	10.5%	$50
		A	Ever emp., avg. earnings in Q8	50.3	7.7 ***	15.3%	$1,531	$198 ***	12.9%	$66
		A	Ever emp., avg. earnings in Q16	53.1	7.6 ***	14.3%	$2,149	$129 *	6.0%	$43

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; E=earnings; W=cash welfare payments; FS=Food Stamp payments; EITC=Earned Income Tax Credit; CC=out-of-pocket child care expenses; FU=follow-up; HH=household; Q=quarter.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.
(a) Results in Canadian dollars.
(b) Phoenix site only, cash assistance.

5.2.2. Programs That Focus on Financial Work Incentives Tied to Hours of Work

The results of the New Hope program, presented in Panel B, differ between families who were working full time at the beginning of the program and those who were not. Among families working full time at random assignment, New Hope had no effect on employment over the follow-up period. It reduced earnings in both years, albeit insignificantly. Among families not working full time at random assignment, New Hope increased employment and earnings substantially.

The difference between these two groups may have to do with the relative strength of the income and substitution effects arising from New Hope's financial incentive. The average control-group earnings among families working full-time at random assignment was about $11,000. This is comparable to the income provided by MFIP. Moreover, these participants satisfied the work requirement already when the program began and had employment rates over 90 percent throughout the follow-up period. Bos et al. (1999) show that the program led these participants to reduce their annual hours of work, which suggests that the income effect arising from New Hope's financial incentive may have dominated the substitution effect. In other words, the opportunity to enjoy greater income from the earnings supplement, while at the same time spending fewer hours away from home, outweighed the opportunity to enjoy even greater income gains by working more. For families who were not working full-time at random assignment, who worked and earned substantially less, the opposite may have been true: The opportunity to enjoy additional income may have outweighed the cost of more hours away from home. The requirement to work full time to receive the supplement would have reinforced the positive substitution effect among these participants as well.

The SSP and SSP Plus programs required recipients to be working full time within one year of random assignment to qualify for the programs' earnings supplement. Both programs increased employment and earnings substantially in the period following the one-year mark, although SSP's effects faded a bit in the following year. The SSP Applicant study required recipients to be working full-time within two years after random assignment to qualify for the supplement. It increased employment and earnings substantially in the quarter following the two-year mark, but also raised employment and earnings in the interim period.

5.2.3. Programs That Focus on Mandatory Work-Related Activities

The employment and earnings effects of programs that focus on mandatory work-related activities are presented in Panel C of Table 5.1. Of the 13 programs, 12 had positive effects on employment over a two-year follow-up period, and the one negative effect is quite small. Moreover, nine of the 12 positive estimates are significant.

On average, these programs increased employment by 5.6 percentage points during the first two years, which amounts to an average 10 percent gain over the control groups. The LFA programs, which emphasize job search, resulted in larger average employment gains than the HCD programs, which emphasize skill-building and generally require the recipient to participate in classroom activities. The average employment increase among the search-oriented programs was 9.2 percentage points, compared to 3 percentage points among the skills-oriented programs.

The earnings results from these programs were similarly positive. Twelve of thirteen programs produced positive effects on earnings, nine of which were significant at least at the 5 percent level. The one negative effect was insignificant. The average earnings gain over the first two years of the follow-up exceeded $700; only four of the programs failed to increase earnings by at least $400. Again the gains were greater for the search-oriented programs than the skills-oriented programs. Among the four work-first programs, two-year earnings gains averaged $1,188. Among the human-capital programs, they averaged $371.

Five-year employment impacts from NEWWS are presented in Figure 5.1; five-year earnings impacts are shown in Figure 5.2. As in Chapter 4, there is evidence of program fade-out. Annual employment impacts in years four and five averaged 2.0 percentage points, compared to 4.8 percentage points in years one and two. Long-term annual earnings impacts averaged $324, compared to short-term impacts of $378.

In the three sites that ran LFA and HCD programs simultaneously, the gap between the LFA impacts and the HCD impacts faded over time as well. For the first two years of the program, the average LFA impacts on annual employment and earnings were 8.7 percentage points and $561, respectively. The average HCD impacts were 4.7 percentage points and $267. For years three through five, the average LFA impacts on annual employment and earnings were 2.8 percentage points and $355, and the average HCD impacts were 2.2 percentage points and $291. The longer-term LFA-HCD differentials are much smaller than the differentials that appeared during the first two years of the program, suggesting that some of the early differential was due to the fact that LFA participants were looking for work during the initial period of the program rather than taking part in training activities.[46]

5.2.4. Programs That Combine Financial Work Incentives and Mandatory Work-Related Activities

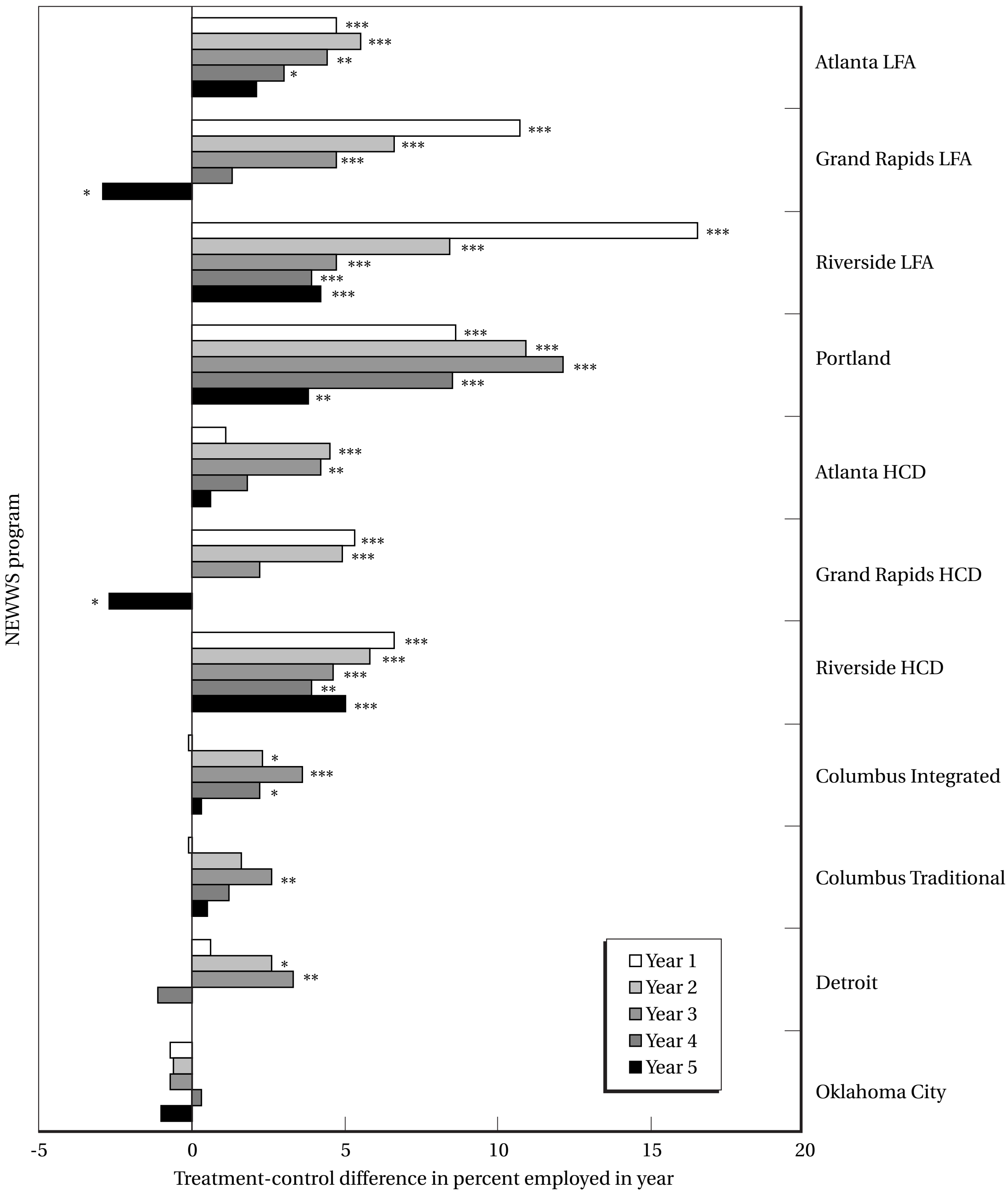
The results from four programs that combine financial work incentives and mandatory work-related activities are summarized in Panel D of Table 5.1. All these programs report at least some significant and positive employment effects. However, the effects vary across studies.

Similarly, all four programs produced at least some significant earnings gains. Here too, however, the effects are heterogeneous. In some cases, the effects change with time since random assignment; in others, they appear to differ between ongoing recipients and recent applicants.

Although these programs exhibit some consistent patterns, some of their effects seem attributable to their specific designs. The nature of the various mandated work-related activities seems likely to play an important role. For this reason, we discuss the results from each program in some detail.

Both the MFIP and WRP programs have stronger effects on employment and earnings than their incentives-only counterparts. The results from WRP suggest that a requirement to work in exchange for welfare may be effective in increasing employment and earnings, even when it has

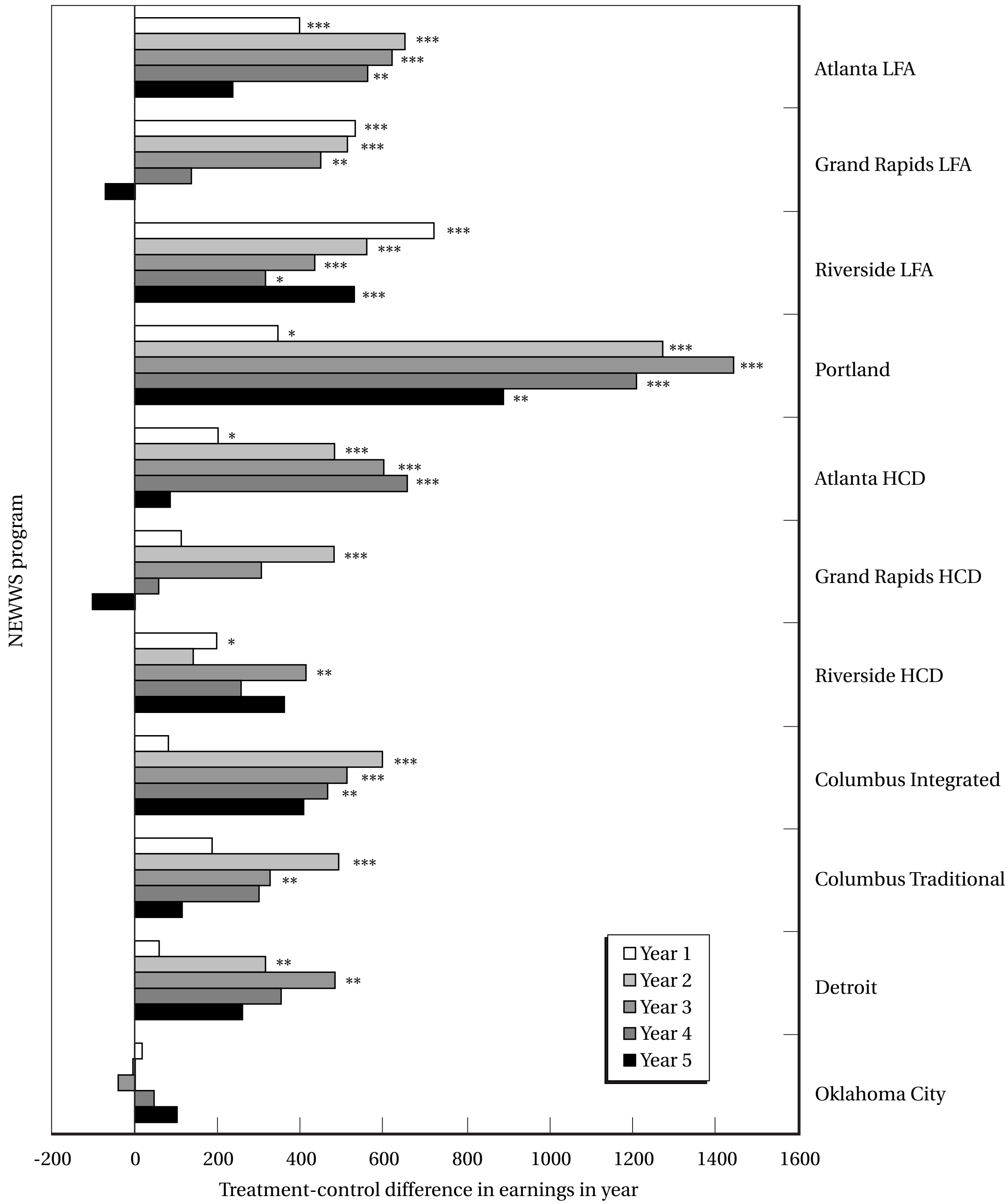
[46]This is consistent with evidence from Hotz, Imbens, and Klerman (2000). They analyze longer-term data from the GAIN program and find that the gap between work-oriented programs and skills-oriented programs fades over time.



SOURCE: Hamilton et al. (2001), Table C.1.

NOTE: Treatment-control difference is statistically significant at the *=10%, **=5%, ***=1% level.

Figure 5.1—Impact Estimates for Employment in 11 NEWWS Programs, Years 1 to 5



SOURCE: Hamilton et al. (2001), Table C.2.

NOTE: Treatment-control difference is statistically significant at the *=10%, **=5%, ***=1% level.

Figure 5.2—Impact Estimates for Earnings in 11 NEWWS Programs, Years 1 to 5

little effect on welfare use, as was seen earlier in Table 4.1. MFIP's work requirement seems to decrease the extent to which the program's effects fade over time. The earnings results from MFIP also suggest that work requirements, which applied to ongoing recipients throughout the follow-up period, may provide a means of overcoming the negative income effects that arise from the combination of a strong financial incentive and relatively high benefit level.

The FIP program combined a fairly strong financial incentive with an education-focused welfare-to-work program, which may explain its results. The FIP evaluation showed that the program increased the fraction of participants who combined work and welfare, precisely as the standard economic model would predict. Moreover, this effect was similar in magnitude and significant for both recipients and applicants. However, the program had quite different effects on the rate at which it moved participants into the workforce and off welfare entirely, which is the other means by which it could have increased employment. Whereas the program had essentially no such effect on recent applicants, it significantly decreased the fraction of ongoing recipients who abandoned welfare for work. As a result, the program increased overall employment for both ongoing recipients and recent applicants, but the effect was significant only for recent applicants. One possible explanation for this pattern is that applicants were more likely to satisfy the work-related activity mandate by working, while ongoing recipients were more likely to satisfy it by taking part in the state's welfare-to-work program. However, since the evaluation provides no information on welfare-to-work participation rates, this explanation is speculative.

Michigan's TSMF program had small positive effects on employment, mirroring its small negative effects on welfare use. These effects may result from features of the program's design. TSMF had a weak financial incentive, and its welfare-to-work program was not particularly work-focused during its first two years. When the program did adopt a work-first approach, the work-related activity mandate was applied to *both* the treatment and the control group (although the control group was subject to lesser sanctions for violating the work mandate). With little effective difference in the conditions applying to the treatment and control groups, we might expect to find small effects.

5.2.5. Programs That Focus on TANF-Like Bundles of Reforms

Most of the programs that included TANF-like bundles of reforms had positive effects on employment and earnings. The exception is Arizona's EMPOWER program. Administrative data suggest that the program may have caused a slight decline in employment and earnings, whereas survey data show small increases. However, none of the estimates from EMPOWER are statistically significant. This could be the result of substantial confusion regarding the program's time limit, which was its principal policy reform. Indiana's Placement Track program significantly raised employment in the first year, but the effect faded over time. Its effects on earnings were larger and somewhat more persistent.

VIP/VIEW, ABC, FTP, and Jobs First generated increases in employment and earnings, most of which were significant. Most of the gains were fairly sizeable as well, particularly for earnings. The smallest effects stem from the VIP/VIEW program, which may be attributable to its phased-in evaluation design.

The pre– and post–time limit impacts of FTP and Jobs First are fairly similar. Both programs had positive and sizeable impacts on employment during the pre–time limit period. In FTP, the impacts were slightly larger in year three, whereas in Jobs First, they were slightly smaller. Their earnings impacts rose slightly in the immediate post–time limit period, but then fell by the fourth year. Neither employment nor earnings changed a great deal as recipients began to exhaust their benefits.

5.2.6. Evidence on Other Employment-Related Outcomes

Beyond providing evidence on the primary employment and earnings outcomes, a subset of the studies listed in Table 5.1 provide estimates on how the various programs have affected other employment-related outcomes, such as weekly hours, wages, job characteristics, and wage growth. However, many of those estimates are nonexperimental in nature, since the outcomes are functions not only of the recipient's treatment status, but also of her post-randomization employment status. To illustrate this point, note that wage growth can be calculated only for persons working at two specified time points (unless we wish to attribute a zero wage to persons not working). However, persons employed at a point in time do not represent random subsamples of the treatment or control groups. Thus, simple comparisons of wage growth between the control and treatment groups does not in general provide estimates of the causal effect of the treatment. Moreover, some of the estimates are based on small survey samples, rather than the larger samples for which administrative records are available.

Partly because of these reasons, the results are not very conclusive. Although there are individual exceptions, there is no clear evidence from the random assignment studies that welfare reform increases wages, wage growth, or the fraction of workers who receive employer-provided health benefits, paid vacation, or paid sick leave. There is even less basis for discriminating between the effects of different types of reforms.

5.2.7. Subgroup Differences

Although only a subset of the random assignment studies provides subgroup analyses, the subset is larger in the case of employment and earnings than it was in the case of welfare use. These results from the available studies are discussed in Appendix A. These studies provide no clear evidence that any of the reforms consistently works to the benefit or detriment of relatively disadvantaged groups in terms of employment and earnings. Some of the results are mixed, however, and the number of studies on which this conclusion is based is generally small.

5.3. ECONOMETRIC STUDIES OF THE EFFECTS OF WELFARE REFORM ON EMPLOYMENT AND EARNINGS

Only five econometric studies estimate the impact of reform on employment and labor supply. As in the caseload literature, these studies typically use "modified dummy variables" to represent welfare reform policies. Therefore they provide moderate-quality evidence on the effects of those policies. Some studies represent some reform policies using measures that capture additional dimensions of policy variation, thus providing high-quality evidence on the effects of such policies.

The results of these studies, reported in Table 5.2, are fairly consistent. Four studies estimate the effect of reform as a bundle. O'Neill and Hill (2001) find reform to have positive and significant effects on the employment of single mothers, with the effects of TANF being stronger than the effects of waivers. Both Moffitt (1999) and Schoeni and Blank (2000) find waivers to have their largest effects among the least-educated women, increasing both their employment and their labor supply by about 4 to 5 percent. Schoeni and Blank report that TANF has similar effects on weeks worked but not on employment. Grogger (forthcoming) finds that reform increased both employment and labor supply among single mothers with infants, by about 4 and 8 percent, respectively, although those effects decrease with the age of the youngest child.

Two studies estimate the effects of time limits. Grogger (forthcoming) finds them to have the greatest effects among families with the youngest children, increasing employment among single mothers with infants by about 2.5 percentage points, or about 3.5 percent. This is consistent with the theory of Grogger and Michalopoulos (1999), but the magnitude and significance of the estimates are smaller for employment than for welfare use. Meyer and Rosenbaum (2001) estimate the effect of a hybrid time limit measure, defined to include both termination time limits and time limits for work. They find these policies to increase employment by about 3 percent. They also find that terminating benefits substantially increases employment.

Meyer and Rosenbaum also provide evidence on the effects of financial work incentives. Their model includes a variable that measures the expected benefit payable to a welfare recipient who works, which varies with the generosity of the incentive. Meyer and Rosenbaum's estimate indicates that increasing financial incentives so as to raise annual benefit payments to working recipients by $1,000, which amounts to roughly a two-thirds increase, increases employment by 12.8 percent.

Only three econometric studies estimate the effects of welfare reform on earnings. These estimates are mixed. Schoeni and Blank report that both waivers and TANF increase earnings among the least educated by about 5 to 7 percent, although only the waiver effect is significant. Moffitt, with his shorter sample period, reports only one significant earnings effect, for women with a high school diploma. This estimate indicates that waivers increased earnings among that group by $560, or about 6 percent. Grogger finds that reform increased earnings among single mothers with infants by about 10 percent when he estimates the relationship in logarithms, but when he estimates the model in levels, which allows him to include families with no earnings, the effect is insignificant. He finds no significant evidence that time limits affect earnings, either in logs or levels.

5.4. EVALUATING THE EFFECTS OF WELFARE REFORM ON EMPLOYMENT AND EARNINGS

Having presented the results of several studies, in this section we attempt to synthesize them to convey what is known about the effects of welfare reform on employment and earnings. We consider the random assignment and econometric studies together, considering both the quantity and quality of the evidence. We begin by discussing the effects of specific reforms and then turn to their effects as a bundle.

Table 5.2—Estimated Impact of Welfare Reform on Employment and Earnings: Econometric Studies

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr.	Fixed Effects	Policy
A. Employment													
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	Fraction working in previous year	Level	Any waiver	0.020 (0.007)	3.7	U, U-1, EG, & each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	-0.005 (0.006)	-0.6				
		women 16-54, educ>12					Any waiver	-0.002 (0.005)	0.2				
		women 16-54, educ<12					TANF	0.015 (0.017)	2.9				
		women 16-54, educ=12					TANF	0.004 (0.015)	0.5				
		women 16-54, educ>12					TANF	-0.024 (0.011)	-2.9				
Grogger (forthcoming)	CPS micro data	female headed families 16-54	78	99	Employed last year	Level	Any reform (waiver or TANF)	0.026 (0.012)	3.7	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-0.002 (0.001)	-0.3				
							Modified time limit dummy	-0.014 (0.017)	-2.0				
							Time limit dummy * (age*)	-0.003 (0.002)	-0.5				
O'Neill and Hill (2001)	CPS micro data	single mothers 16 - 54	83	2000	Employed last week	Level	Any waiver	2.34 ***	3.9	U, W, college wage premium	A, E, R, No. kids, age youngest kid, ever married, urban-rural	S, trend and trend squared, state-year trends	B
							TANF	6.59 ***	11.0				
Meyer and Rosenbaum (2001)	CPS micro data	female family heads, 19-44	84	96	Employed last year	Level	Modified dummy for any time limit waiver	0.014 (0.007)	2.3		A, E, R # kids, any kids< 6,3,2,1	S, Y, month	B, MW, EITC, Medicaid, training, child care
							Dummy for any time limit terminations	0.022 (0.011)	3.7				
							Welfare benefit for working recipients (in $1,000s)	0.077 (0.007)	12.8				

Table 5.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr.	Other controls: Fixed Effects	Other controls: Policy
B. Labor Supply													
Moffitt (1999)	CPS aggregated	women 16-54	77	95	Annual weeks worked	Level	Any waiver	0.3 (0.3)	1.0	U, U-1		S, Y, State time trends	B
	CPS aggregated	women 16-54, educ<12	77	95	Annual weeks worked	Level	Any waiver	1.5 (0.6)	5.0	U, U-1	A, E	S, Y, State time trends	B
		women 16-54, educ=12					Any waiver	0.5 (0.6)	1.8				
		women 16-54, educ=13-15					Any waiver	0.5 (0.6)	1.6				
		women 16-54, educ>16					Any waiver	0.2 (0.3)	0.6				
Schoeni and Blank (2000)		women 16-54, educ<12			Annual weeks worked		Any waiver	0.732 (0.355)	4.0	U, U-1, EG, & each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	-0.175 (0.291)	-0.5				
		women 16-54, educ>12					Any waiver	0.031 (0.239)	0.1				
		women 16-54, educ<12					TANF	-0.090 (0.832)	-0.5				
		women 16-54, educ=12					TANF	0.300 (0.728)	0.9				
		women 16-54, educ>12					TANF	-0.377 (0.553)	-1.0				
Grogger (forthcoming)	CPS micro data	female headed families 16-54	78	99	Annual weeks worked	Level	Any reform (waiver or TANF)	2.425 (0.600)	8.0	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-0.216 (0.073)	-0.7				
							Modified time limit dummy	-0.830 (0.923)	-2.7				
							Time limit dummy * (age*)	-0.116 (0.098)	-0.4				

Table 5.2—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr.	Other controls: Fixed Effects	Other controls: Policy
C. Earnings													
Moffitt (1999)	CPS aggregated	women 16-54	77	95	Annual earnings	Level	Any waiver	274.3 (161.8)	2.8	U, U-1		S, Y, State time trends	B
	CPS aggregated	women 16-54, educ<12	77	95	Annual earnings	Level	Any waiver	87.8 (318.6)	0.9	U, U-1	A, E	S, Y, State time trends	B
		women 16-54, educ=12					Any waiver	560.0 (318.6)	5.7				
		women 16-54, educ=13-15					Any waiver	441.4 (318.5)	4.5				
		women 16-54, educ>16					Any waiver	154.7 (318.7)	1.6				
Schoeni and Blank (2000)		women 16-54, educ<12			Family head's pre-tax annual earnings	log	Any waiver	0.050 (0.023)	5.0	U, U-1, EG, & each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	-0.006 (0.018)	-0.6				
		women 16-54, educ>12					Any waiver	-0.013 (0.015)	-1.3				
		women 16-54, educ<12					TANF	0.065 (0.053)	6.5				
		women 16-54, educ=12					TANF	0.028 (0.026)	2.8				
		women 16-54, educ>12					TANF	-0.016 (0.035)	-1.6				

Table 5.2—Continued

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr.	Fixed Effects	Policy
Grogger (forthcoming)	CPS micro data	female headed families 16-54	78	99	Annual pre-tax earnings ($1000s)	Level	Any reform (waiver or TANF)	830 (501)	6.5	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-21 (68)	-0.2				
							Modified time limit dummy	-705 (862)	-5.5				
							Time limit dummy * (age*)	39 (73)	0.3				
			78	99	Annual pre-tax earnings ($1000s)	Log	Any reform (waiver or TANF)	0.102 (0.046)	10.2	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-0.008 (0.004)	-0.8				
							Modified time limit dummy	-0.079 (0.054)	-7.9				
							Time limit dummy * (age*)	-0.002 (0.005)	-0.2				

NOTES: Abbreviations: s.e.=standard error; U=unemployment rate; U-1=lagged unemployment rate; EG=employment growth; A=age, E=education, R=race, B=maximum welfare benefit, MW=minimum wage; EITC=Earned Income Tax Credit; S=state; Y=year.

age* = (age of youngest child -13).

5.4.1. The Effects of Specific Reforms

We draw evidence on the effects of financial work incentives from three studies: the WRP-IO program, the MFIP-IO program, and the Meyer-Rosenbaum econometric analysis. By the standards discussed in Chapter 3, all three studies provide high-quality evidence. Meyer-Rosenbaum suggests that increasing financial work incentives increases employment. The random assignment studies qualify that conclusion. They suggest that strong incentives, such as those in MFIP-IO, increase employment, whereas weak incentives, such as those in WRP-IO, have no effect.

The earnings effects of these two programs are more complicated, which should come as little surprise given the conflicting income and substitution effects that arise from their financial incentive. WRP-IO had no effect on earnings, just as it had no effect on employment. However, MFIP-IO had no significant effect on earnings, despite its positive effects on employment. Indeed, most of its insignificant earnings effects were actually negative rather than positive. One explanation that is consistent with these results is that, at sufficiently high income levels, such as that provided by MFIP, the income effect that arises from the financial incentive dominates the substitution effect. Put differently, the opportunity that the incentive presents to enjoy higher income while spending less time away from home may outweigh the countervailing incentive to earn even more by taking greater advantage of the earnings subsidy.

These general patterns also appear among the programs that involved financial work incentives tied to hours of work. Both New Hope and SSP entailed strong financial incentives in the form of earnings supplements that were available to consumers who met the programs' work requirements. With one exception, these programs increased employment and earnings. The exception involved New Hope families who were working full time at the beginning of the experiment. There was no change in employment among those participants, and their earnings decreased (albeit insignificantly) rather than increased. Like the MFIP group, these New Hope recipients already enjoyed relatively high incomes. They also worked a great deal, leaving little room for the program to increase their employment. Thus the program's negative earnings impacts on this group may have been the result of income effects that led these workers to prefer additional time at home over additional income from work.

Over a dozen high-quality random-assignment studies provide evidence on the effects of mandatory work-related activities. The results indicate that this type of policy generally increases both employment and earnings. In the short run, work programs that focus on job search generally yielded greater effects than programs that focus on skills building. After the first few years, the impacts of the program faded, although most remained positive. The impact differential between the search-oriented programs and the skills-oriented programs also faded with time.

Adding mandatory work-related activities to programs involving financial work incentives generally increases their effects on both employment and earnings, at least in the two programs that allow such a comparison. The two policies have reinforcing effects, as the economic model discussed in Chapter 2 would predict. The results from WRP suggest that requiring work in exchange for welfare can increase employment, even though it may not decrease welfare use. The results from MFIP suggest that work requirements may overcome the negative income effect on earnings that arises when a strong financial incentive is added to a fairly high level of benefits. Adding the mandatory activities also reduces the extent to which the effects of these programs fade out over time. Two other programs that combined financial work incentives with mandatory work-related activities also increased employment and earnings, although it is unclear why TSMF mostly benefited ongoing recipients, whereas Iowa's FIP mostly benefited new applicants.

Evidence on the behavioral effects of time limits comes from two econometric studies. Meyer and Rosenbaum (2001) find that time limits increase employment among single mothers. Grogger (2000) finds that such increases are greatest among families with the youngest children, although he finds little such evidence for earnings.

Three studies provide evidence on whether time limits have any mechanical effects on employment or earnings. Meyer and Rosenbaum find that employment rises when recipients reach the limit. However, the evidence from FTP and Jobs First suggests that neither employment nor earnings change much as recipients begin to exhaust their benefits.

5.4.2 The Effects of Reform as a Bundle

The four econometric studies that analyze the effects of reform as a bundle all report that welfare reform increased employment among welfare-prone groups such as single mothers and women with low levels of education. This is consistent with the evidence that many of the specific policies included in the typical bundle increased employment. Indeed Meyer and Rosenbaum attribute 14–20 percent of the 1992–1996 increase in employment among single mothers to the effects of time limit waivers. Grogger attributes 7 percent of the 1993–1999 increase to time limits and another 6 percent to other reforms. O'Neill and Hill attribute larger fractions of the increase to reform as a bundle, crediting waivers for 22 percent of the increase between 1992 and 1996 and TANF for 62 percent of the increase between 1996 and 1999.

At the same time, the econometric studies indicate that the economy played an important role in raising the work effort of welfare-prone groups. Grogger (forthcoming) attributes 21 percent of the 1993–1999 increase in employment among single mothers to the improving economy. O'Neill and Hill attribute 35 percent of the 1992–1996 increase, and 17 percent of the 1996–1999 increase, to the economy.

Other policy changes had important effects as well. Meyer and Rosenbaum credit tax changes, mostly involving expansions to the EITC, for 27–35 percent of the increase in single mothers' employment between 1992 and 1996. Grogger estimates that the EITC explains 21 percent of their 1993–1999 increase in earnings.

Although the evidence on most specific reforms is limited, the collective results from the random assignment studies, together with those from the econometric studies, seem consistent with the notion that reform played an important role in improving the labor market outcomes, and particularly the employment, of single mothers. However, welfare reform was hardly the only factor at play. Both the economy and other policy changes played important roles in increasing poor families' employment and earnings.

5.5. CONCLUSIONS

The studies that we reviewed in this chapter suggest most of the reforms that were introduced during the 1990s had positive effects on employment and earnings. Thus, it seems likely that reform as a bundle is responsible for a portion of the increase in work and earnings among single mothers. Nearly all of the evidence, from both the experiments and the econometric studies, points in this direction.

The quantity of evidence on specific reform policies varies greatly across the specific reforms. Fourteen experimental studies involve financial work incentives, either as their primary focus or in conjunction with other major reforms. Most of those programs yielded increases in earnings. Mandated work-related activities have been the subject of thirteen high-quality random assignment studies. Nearly all of those programs increased employment and earnings in the short run. However, their long-run impacts were smaller.

For other reforms, the number of relevant studies is lower. There have been no studies of the effects of welfare sanctions on employment or earnings. Two have provided estimates suggesting that time limits raise employment in the period before participants begin exhausting their benefits. Three have provided mixed results regarding the mechanical effects of time limits.

Of the few econometric studies, all are in agreement that the expanding economy of the 1990s had important effects on the employment and earnings of single mothers. Studies that analyze the effects of recent EITC expansions find that they played an important role as well.

Nevertheless, despite welfare reform, the strong economy of the 1990s, and the EITC, the earnings of program participants generally remain low. Average annual earnings among NEWWS participants never exceeded $8,000. Even the average earnings of single mothers as a whole, at $16,600, just equals the federal poverty standard for a family of four.

Chapter Six

USE OF OTHER GOVERNMENT PROGRAMS

6.1. BACKGROUND

Many families affected by welfare reform are likely to remain eligible for other U.S. government programs in the "safety net." The largest such programs are the Food Stamp Program (FSP) and Medicaid. Low-income families also often qualify for housing assistance and various other nutrition programs, such as the School Lunch Program, the School Breakfast Program, and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

The quantity of research analyzing the effects of welfare reform on the use of these government programs varies widely. Most of the random assignment studies estimate food stamp impacts. This is probably because, in many states, the same administrative data system is used to track both welfare receipt and food stamp use. Thus, administrative data on food stamp receipt are nearly as readily available to researchers as administrative data on welfare use. Aggregate caseload data and survey-based information on food stamps are also available, but they have been less utilized for econometric studies than the corresponding welfare data.

Fewer random assignment studies, and only one econometric study, have considered the effects of welfare reform on Medicaid enrollment. Moreover, most of the random assignment studies that address the question do so using survey data, which generally include fewer observations than the administrative data used to analyze other outcomes. Only a few studies consider participation in other government programs, and those that do are based entirely on survey data.

Before discussing the results from studies that attempt to estimate the causal effects of welfare reform on the use of these various government programs, we provide some background on the programs, recent trends in participation, and, where relevant, provisions of PRWORA that affect those programs directly. The amount of detail we provide on each program is roughly proportional to the number of research studies available to synthesize.

6.1.1. The Food Stamp Program

The FSP provides low-income households with coupons—or more recently, electronic benefits accessed by means akin to a debit card—that can be used to purchase food. Households must have incomes less than 130 percent of the federal poverty line (FPL) to qualify. In addition, the household's net income—equal to its income less certain deductions—must fall below the FPL. The household must also satisfy asset tests.

Benefit levels are uniform in the continental United States. A family of three with no other source of income is eligible for $341 in monthly benefits. For families with other income, including income from TANF, benefits are reduced by 30 cents for every dollar of net income. As a result, a three-person household with income equal to half the FPL ($610/month) would be eligible for at least $235/month in food stamps. A three-person household with income equal to the FPL would be eligible for somewhere between $89 and $201 in food stamps (Zedlewski and Gruber, 2001).

Unlike cash aid programs, eligibility for food stamps is based on household composition rather than family structure. As a result, the FSP serves a larger client base than the TANF program. Of the 7.3 million households (which included 17.1 million persons) served by the FSP in 2000, 26 percent also received assistance from TANF (USDA, 2001a).

Although TANF households generally are eligible for food stamps, in 1999 only 81 percent of TANF households received FSP benefits (USDHHS, 2000). Non-TANF households must go through an eligibility determination process that requires substantial documentation of income sources. When families leave welfare, they may lose their eligibility for food stamps. If they become ineligible because of income gains, they stop receiving food stamp benefits. If they are unaware that they may remain eligible for food stamps, or if they are unable or unwilling to take time off from work or to document their sources of income, they may also stop receiving food stamps.[47]

Beyond these potential indirect effects on FSP use, PRWORA directly affected FSP eligibility for two groups of recipients: legal immigrants and able-bodied adults without dependents (ABAWDs). Legal immigrants who entered the United States after the passage of PRWORA were barred from receiving most kinds of aid, including food stamps. Legal immigrants present prior to the passage of PRWORA were to be dropped from the FSP rolls within a year, although the 1997 Balanced Budget Act repealed the latter provision. PRWORA limited ABAWDs to three months of FSP receipt in any 12-month period, unless they were working or participating in an approved work program. States can apply for waivers to exempt from the time limit ABAWDs who live in areas with high unemployment rates or insufficient jobs.

Recent trends in the food stamp caseload resemble recent trends in the welfare caseload. During the late 1980s, the program served about 19 million people per month. By 1994, the number of persons served by the FSP grew to 27.5 million; by 2000, it had fallen to 17.2 million, the lowest level since 1978. In relative terms, the FSP caseload fell 37 percent between 1994 and 2000.

Although participation rates for immigrants and ABAWDs fell substantially after 1996, they account for a small proportion of the overall caseload decline because they account for a relatively small share of all food stamp households. The group accounting for the largest share of the decline was single-parent families, which includes a number of families leaving cash aid (USDA, 1999).

Leavers studies show that former welfare recipients leave the FSP at a high rate. Among 13 cohorts tracked in twelve USDHHS-funded leaver studies, the fraction of leavers receiving food

[47]TANF leavers participate in the FSP at a lower rate than TANF recipients, but at a higher rate than comparable families who never received welfare (Zedlewski and Brauner, 1999).

stamp benefits in the first quarter after leaving welfare ranged from 23 to 78 percent (USDHHS, 2001a). Six of the studies had rates ranging from 35 to 54 percent, whereas two were lower and four were higher. Data from the National Study of America's Families are similar, showing that only about one-third of recent welfare leavers continued to receive food stamps (Zedlewski and Gruber, 2001). Acs et al. (2001) suggest that, among long-term welfare recipients, these rates did not change much during the 1990s.

Although much of the decline in FSP use results from a decrease in eligibility arising from gains in income, over half of the decrease stems from a decrease in take-up rates on the part of income-eligible households (USDA, 2001b). National Study of America's Families data show that about one-half of all leaver families have post-welfare incomes less than the FPL (Loprest, 2001), which implies that more than half of all leaver families remain income-eligible for food stamps. However, among leavers with incomes between 50 and 100 percent of the FPL, only 45 percent continued receiving food stamps. Among leavers with incomes below 50 percent of the FPL, only 51 percent received food stamps (Zedlewski and Gruber, 2001). Low take-up among the former group may stem from the relatively low benefit payment for which they would be eligible (Blank and Ruggles 1996). However, low take-up among the latter group represents nonparticipation by families who would be eligible for a relatively large payment.

Despite the usefulness of such contextual information, trends and patterns among welfare leavers do not establish whether welfare reform has caused FSP use to fall. As with other welfare-related outcomes, changes in the economy and other antipoverty policies, such as the EITC, may also underlie the recent declines in the FSP caseload. In Section 6.2 below, we review the evidence from a number of random assignment and econometric studies on the extent to which welfare reform has caused the decline in the FSP.

6.1.2. Medicaid and the State Children's Health Insurance Program

The Medicaid program, which enrolls roughly 40 million Americans at an annual cost of about $176 billion, serves three largely distinct populations: the elderly, the blind or disabled, and poor families with children. Poor families with children account for about two-thirds of Medicaid beneficiaries but only about 25 percent of annual Medicaid expenditures (Health Care Financing Administration, 2000). Since this is the group most relevant to welfare reform, we restrict our attention to it below.

Originally, poor families with children received coverage under Medicaid primarily by qualifying for cash welfare. However, beginning in the mid-1980s, expansions of coverage to poor pregnant women and poor children not receiving welfare weakened the link between the two programs. The largest expansions provide eligibility to children depending on their age and their family's income. By 1990, the states were required to provide coverage for all children under age six who lived in families with incomes less than 133 percent of the FPL. They were also required to cover all children born after September 30, 1983, in families with incomes below 100 percent of the FPL (Congressional Research Service, 1993). As a result, by 1997, all children under age 15 in families with incomes below the poverty line were eligible for Medicaid, regardless of whether their families received (or even qualified for) welfare. Many states exercised the option to extend coverage to older children and to children in families with higher incomes.

The Family Support Act of 1988 extended Medicaid to families that left welfare because of employment or earnings via Transitional Medical Assistance (TMA). Under FSA, TMA was available for one year after leaving welfare. Many states have extended the eligibility period under their TANF programs.

PRWORA officially ended the link between welfare and Medicaid by eliminating automatic Medicaid eligibility for welfare families and establishing a new eligibility category. Under section 1931 of the Social Security Act, states are required to extend Medicaid coverage to families who would have qualified for welfare under the AFDC eligibility rules that the state had in place as of July 1996. Section 1931 also provides states with the flexibility to expand Medicaid to cover more low-income families, including two-parent working families, at the state's option.

Health care coverage was further expanded by the 1997 Balanced Budget Act, which established the State Children's Health Insurance Program (SCHIP). SCHIP is designed to provide coverage for children in families with incomes up to 200 percent of the FPL who otherwise are ineligible for Medicaid (HCFA, 2000). About 30 states have chosen to cover children in families with incomes up to at least this level.

When Medicaid was tied to welfare, enrollments were fairly stable. Between 1975 and 1985, the number of poor children and adults covered by Medicaid grew from just over 14 million to just over 15 million. Enrollment grew to 17.2 million by 1990, then to 24.8 million by 1995 (Gruber, 2000). The number of covered persons in poor families fell between 1995 and 1997, which is the most recent year for which consistently defined data are available.[48] Coverage of poor children fell by 12.4 percent between 1995 and 1997; coverage of poor adults fell by 11.2 percent (Committee on Ways and Means, 2000, Table 15-14).

The nationwide trend in enrollments masks substantially different trends among the states. While the overall decline in enrollment among persons in poor families was 5.3 percent between 1995 and 1997, ten states experienced increases in enrollment. At the same time, nine states experienced decreases of 10 percent or more (Ku and Bruen, 1999).

Leavers studies generally show low rates of Medicaid coverage among persons leaving welfare, but like the nationwide trend data, there is a fair amount of heterogeneity across different locales. Nine state-specific studies funded by USDHHS provide coverage rates for family heads in their first quarter after leaving welfare. Of these, eight report coverage rates ranging from 42 to 69 percent. Wisconsin showed higher coverage, at 80 percent (USDHHS, 2001a, Appendix B). Data from the National Study of America's Families show that 56 percent of leavers are covered by Medicaid in the first six months after leaving aid (Garrett and Holahan, 2000).

Children are more likely than adults to retain eligibility, but the differential varies greatly by site. In Wisconsin, 86 percent of children were covered by Medicaid in the first quarter after their families left welfare, compared with 80 percent of adults. In Missouri, 42 percent of adults were covered, but 85 percent of children retained eligibility (USDHHS, 2001a). The National Survey of America's Families shows that 70 percent of children in families that had been off welfare for six months or less retained their Medicaid coverage (Garrett and Holahan, 2000).

[48]In 1998, for the first time, capitation payments were counted as services for reporting purposes (HCFA, 2000, p. 14).

Since Congress specifically sought to ensure that families would not lose health care coverage as a result of welfare reform, there has been substantial interest in explaining recent declines in Medicaid enrollments. Some analysts have suggested that the strong economy has played a role, whereas others have pointed to confusion on the part of recipients about the new relationship between welfare and Medicaid programs (Kenney and Haley, 2001). Others have suggested that states' TANF diversion programs may have deterred some families from completing their Medicaid applications (GAO, 1999c; Dion and Pavetti, 2000; Guyer, 2000). In Section 6.2, we synthesize the evidence on the extent to which welfare reform has caused the recent declines in Medicaid enrollments.

6.1.3. Other Programs

Although there has been relatively little research into the effects of welfare reform on participation in other nutrition programs or subsidized housing programs, it is useful to summarize some of the most important programs here. After the FSP, the largest nutrition programs are the National School Lunch Program (NSLP), the School Breakfast Program, and WIC. The school nutrition programs provide meals to primary and secondary school students; students from families with incomes below 130 percent of the FPL receive their meals for free and students from families with incomes between 130 and 185 percent of the FPL receive their meals at a reduced price. The WIC program provides vouchers for pregnant women, new mothers, and children up to age 5 that can be redeemed for specific food items that contain nutrients often lacking from the diets of low-income families.

Unlike the trend in food stamp participation, which was similar to trends in welfare use, participation in these three smaller nutrition programs grew through most of the 1990s. The number of children receiving free or reduced-price school lunches rose from 11.6 million per month (during the school year) in 1990 to 15.5 million in 2000. The number of children receiving free or reduced-price school breakfasts rose from 3.5 million per month in 1990 to 6.3 million in 2000. The number of persons served by the WIC program rose from 4.5 million in 1990 to 7.4 million in 1997, then fell to 7.2 million in 2000. The costs of operating the school nutrition programs in 2000 (including the tiny Special Milk Program) amounted to $7.6 billion; the cost of the WIC program was $3.97 billion (USDA, 2001d). The General Accounting Office (GAO) (1999) interprets increases in school nutrition programs as evidence of program substitution, that is, as evidence that families are turning to programs other than the FSP in order to satisfy their demand for nutrition assistance.

Housing assistance of various kinds is provided through a number of different programs that are run by several different agencies (Olsen, 2001). However, most rental assistance funds targeted to very-low-income households are administered by the Department of Housing and Urban Development. Such aid is provided in one of two forms: project-based assistance or household-based rental subsidies (Committee on Ways and Means, 2000). The number of renters assisted by these programs grew more or less steadily during the last decade, from 1.6 million households in 1990 to 2.1 million households in 2000.

6.2. RANDOM ASSIGNMENT AND ECONOMETRIC STUDIES OF THE EFFECTS OF WELFARE REFORM ON USE OF OTHER GOVERNMENT PROGRAMS

As noted above, there are relatively fewer random assignment and econometric studies examining the effects of welfare reform on the use of other government programs. Thus, unlike in the previous chapters, we combine the discussion of random assignment and econometric studies, breaking the discussion up by the three categories of other government programs presented above. The limited amount of information from the random assignment studies regarding subgroup differences in the outcomes covered in this chapter is summarized in Appendix A.

6.2.1. Food Stamps

Of the welfare-related outcomes considered in this chapter, the Food Stamp Program is the most studied. Most of the random assignment studies analyze the effects of welfare reform on food stamp use. In addition, four econometric studies address the question using national data from either administrative sources or the CPS.

Random Assignment Studies

Table 6.1 presents results from the random assignment studies. The only study focusing on financial work incentives that provides information about food stamp usage is the WRP Incentives-Only program, since MFIP involved a food stamp cashout. The WRP Incentives-Only program had no significant effect on food stamp use, just as it had no significant effect on welfare use.

Of the programs focusing on financial incentives tied to hours of work, food stamp impacts are available only for New Hope, because the SSP programs were conducted in Canada. The New Hope program resulted in increased food stamp use at the end of the second year among families employed at the time of random assignment. Since the same group was less likely to use welfare at that time (see Table 4.1), this suggests that these New Hope participants partially replaced their reduced welfare payments with increased food stamp benefits. However, among families not initially employed full time, New Hope had no significant effect on food stamp use, just as it had no effect on welfare use.

Most of the programs that focused on mandatory work-related activities resulted in significant declines in food stamp use. The four exceptions were the Atlanta LFA, Atlanta HCD, Oklahoma City, and IMPACT Basic Track programs. The average reduction in FSP use across the thirteen programs was 3.5 percentage points, which amounts to 62.5 percent of their average 5.6 percentage point reduction in welfare use. This suggests that about 65 percent of the families leaving welfare discontinued their food stamp benefits. [49]

The programs combining financial work incentives and mandatory work-related activities had effects on food stamp use that are similar to their effects on welfare use. WRP, which involved both a weak financial incentive and a delayed work requirement, had no significant effect on

[49]The five-year NEWWS report does not provide year-by-year impacts for food stamp use.

either. Reductions in food stamp use among TSMF recipients are similar to their reductions in welfare use. For the most part, the reductions in food stamp use are slightly smaller, although the earlier cohort of TSMF applicants actually reduced their food stamp use by a slightly greater amount than their welfare use. An exception to the general rule is FIP, where food stamp use fell despite increases in welfare use.[50]

Among programs that focused on TANF-like reform bundles, Indiana's Placement Track recipients had lower food stamp use than the controls. The reduction in food stamp use was less than the reduction in welfare use in the fourth quarter of the program but was slightly greater by the eighth quarter. Arizona's EMPOWER program and Virginia's VIP/VIEW program also had somewhat larger effects on food stamps than on welfare use. The food stamp results for Delaware's ABC program cannot be compared to the corresponding welfare-use results because the measures differ across the two outcomes.

In the pre–time limit period, FTP and Jobs First had effects on food stamp use that were similar to their effects on welfare use, although the food stamp impacts are insignificant. In the post–time limit period, food stamp use was lower in the treatment group compared with the control group for both programs, but only the Jobs First impact was significant, and then only in quarter 16. All the post–time limit impacts were substantially smaller than the corresponding impacts on welfare use.

The changes in program impacts between the pre– and post–time limit periods suggest that reaching the welfare time limit had relatively little effect on food stamp use. In FTP, food stamp use actually rose by 0.5 percentage points between years two and three. In Jobs First, food stamp use fell by 3.7 percentage points between quarters seven and eight, which amounts to about 30 percent of the corresponding decline in welfare use. Apparently, most of the families who exhausted their welfare benefits continued their food stamp use, at least initially. Although the impact of FTP stayed roughly constant through year four, the impact of Jobs First became more negative.

Econometric Studies

The four econometric studies of the effects of welfare reform on food stamp use are summarized in Table 6.2. The studies by Wallace and Blank (1999) and Currie and Grogger (2001) employ models that are comparable to the static models of welfare caseloads summarized in Chapter 4. They also employ similar policy measures to capture the effects of welfare reform waivers and TANF. Figlio, Gunderson, and Ziliak (2000) and Wilde et al. (2000) employ lagged dependent variable specifications similar to the welfare caseload model of Figlio and Ziliak (1999).

The estimates of Wallace and Blank and Currie and Grogger are fairly similar, despite the fact that they are based on different data. Wallace and Blank estimate that waivers reduced food stamp use by 2.5–3.2 percent, depending on whether they use monthly or annual food stamp caseload data. They estimate that welfare reform reduced food stamp caseloads by about 14

[50]Information on the use of food stamps and Medicaid among FIP applicants is provided only for the first four quarters of the follow-up period.

Table 6.1—Estimated Impact of Welfare Reform on Use of Food Stamp Program: Random Assignment Studies

Name	Cases served	Data	Measure	Food Stamp use: Control mean	Food Stamp use: Impact	Food Stamp use: %
A. Programs that focus on financial work incentives						
WRP-IO	Single-parent recipients and applicants	A	Ever received FS, last 3 mos. of FU	50.1	1.0	2.0%
B. Programs that focus on financial work incentives tied to hours of work						
New Hope	Poor families employed FT at RA	A	Months receiving FS, year 1 of 2-yr FU	5.3	-0.3	-5.7%
		A	Months receiving FS, year 2 of 2-yr FU	4.5	1.0 **	22.2%
	Poor families not employed FT at RA	A	Months receiving FS, year 1 of 2-yr FU	7.5	-0.1	-1.3%
		A	Months receiving FS, year 2 of 2-yr FU	5.2	0.4	7.7%
C. Programs that focus on mandatory work-related activities						
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Received FS, Q8	64.5	-4.2 ***	-6.5%
Atlanta LFA	Recipients and applicants	A	Received FS, Q8	76.9	-1.2	-1.6%
Grand Rapids LFA	Recipients and applicants	A	Received FS, Q8	67.3	-5.8 ***	-8.6%
Riverside LFA	Recipients and applicants	A	Received FS, Q8	54.4	-7.6 ***	-14.0%
Portland	Recipients and applicants; no cases with substantial barriers	A	Received FS, Q8	63.3	-4.6 ***	-7.3%
Atlanta HCD	Recipients and applicants	A	Received FS, Q8	76.9	-1.0	-1.3%
Grand Rapids HCD	Recipients and applicants	A	Received FS, Q8	67.3	-3.8 **	-5.6%
Riverside HCD	Recipients and applicants	A	Received FS, Q8	57.6	-5.5 ***	-9.5%
Columbus Integrated	Recipients and applicants	A	Received FS, Q8	64.0	-6.0 ***	-9.4%
Columbus Traditional	Recipients and applicants	A	Received FS, Q8	64	-4.0 ***	-6.3%
Detroit	Recipients and applicants	A	Received FS, Q8	81.7	-3.5 ***	-4.3%
Oklahoma City	Applicants	A	Received FS, Q8	56.0	0.4	0.7%
IMPACT Basic Track	Recipients and applicants-basic track	A	Received FS, Q4	61.8	1.0	1.6%

Table 6.1—Continued

Name	Cases served	Data	Measure	Food Stamp use: Control mean	Food Stamp use: Impact	Food Stamp use: %
D. Programs that focus on financial work incentives and mandatory work-related activities						
WRP	Single-parent recipients and applicants	A	Ever received FS, last 3 mos. of FU	50.1	-1.6	-3.2%
TSMF	Recipients	A	Monthly FS receipt over 4-yr FU	69.1	-0.9 ***	-1.3%
	Applicants	A	Monthly FS receipt over 1-yr FU	67.6	-0.8	-1.2%
		A	Monthly FS receipt over 2-yr FU	61.2	-2.3 ***	-3.8%
FIP	Recipients	A	FS receipt, Q4	79.3	-1.9 *	-2.4%
		A	FS receipt, Q8	66	-2.1 *	-3.2%
	Applicants	A	FS receipt, Q4	44.9	-0.2	-0.4%
		A	FS receipt, Q8	n/a		
E. Programs that focus on other individual reforms						
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)						
EMPOWER (a)	Recipients	A	Monthly FS receipt, months 1-36	47.2	-1.3 *	-2.8%
IMPACT Placement Track	Recipients and applicants-placement track	A	Received FS, Q4	66.2	-7.0 ***	-10.6%
VIP/VIEW	Recipients	A	Welfare receipt in Q8	70.1	-2.6 ***	-3.7%
ABC	Recipients and applicants	S	Percent receiving FS at spring 1997 interview	69.5	-4.2 *	-6.0%
FTP	Recipients and applicants	A	Avg. percent receiving FS, year 2	60.6	-0.9	-1.5%
		A	Avg. percent receiving FS, year 3	48.8	-0.4	-0.8%
		A	Avg. percent receiving FS, year 4	40.7	-0.7	-1.7%
JOBS First	Recipients and applicants	A	Ever received FS, Q7	61.6	3.1 **	5.0%
		A	Ever received FS, Q8	58.8	-0.6	-1.0%
		A	Ever received FS, Q16	42.5	-3.3 **	-7.8%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; Q=quarter; FS=Food Stamps; FT=full-time.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level;
*** = statistically significant at the 1 percent level.

(a) Phoenix site only, cash assistance.

Table 6.2—Estimated Impact of Welfare Reform on Use of Food Stamp Program: Econometric Studies

											Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Includes LDV's?	Economy	Demogr.	Fixed Effects	Policy
Wallace and Blank (1999)	FS caseload	Total population	80	96	annual FS caseload/total population	Log	Any waiver	-0.032 (0.017)	-3.2	no	U, U-1, U-2, median wage, 20th percentile wage	%elderly, %immigrants-1, %immigrants-2, %female heads, %non-marital births	S, Y, State time trends	B, party control
			80	98	monthly FS caseload	Log	Any waiver (12 lags)	-0.025 (0.000)	-2.5		U (12 lags)		S, state-specific month dummies	
							TANF (12 lags)	-0.137 (0.000)	-13.7					
Currie and Grogger (2001)	CPS micro data	HH w/ incomes < 300% of FPL	80	98	Any FS last year	Level	Any waiver	-7.800 (3.500)	-5.4	no	U	A, E, R, # kids, MSA	S, Y, State time trends	B, EBT, mean recertification interval
							TANF	-16.200 (5.900)	-11.2					
Figlio, Gunderson, and Ziliak (2000)	FS caseload	Total population	80	98	annual FS caseload / population	Log	Any waiver	1.014 (0.747)	1.0	yes (4 lags)	U (4 lags), EG (4 lags)		S, Y, State time trends	B, EBT, ABAWD waiver
Wilde, et al. (2000)	FS caseload	Total population	80	98	annual FS caseload / population	Log	Any waiver	0.621 (0.757)	0.6	yes (4 lags)	U (4 lags), EG (4 lags)		S, Y, State time trends	B, EBT, ABAWD waiver, political climate

NOTES: Abbreviations: LDV=lagged dependent variable; s.e.=standard error; U=unemployment rate;U-1=lagged unemployment rate; EG=employment growth; A=age, E=education, R=race, B=maximum welfare benefit, MW=minimum wage; ABAWD=able-bodied adults without dependents; EBT=electronic benefit transfer system; EITC=Earned Income Tax Credit; S=state; Y=year.

percent. Currie and Grogger's estimates, which are based on CPS data, indicate that waivers reduced food stamp use by 5.4 percent and that TANF reduced food stamp use by 11.2 percent. Both studies suggest that welfare reform accounts for a substantial fraction of the decline in the FSP caseload.

Figlio et al. and Wilde et al. both include four lags of food stamp caseloads in their model and report that waivers had no significant effect on food stamp use. These models are subject to the same technical and interpretational problems as the lagged dependent variable models of welfare caseloads, as discussed in Chapter 4. As in the studies of welfare caseloads, the lagged dependent variable models yield results that run contrary not only to the results from the other econometric studies, but also to the results from most of the random assignment studies.

6.2.2. Medicaid

Random Assignment Studies

Only seven of the random assignment studies report information on Medicaid receipt.[51] Their results are summarized in Table 6.3. Much of the information about Medicaid use comes from surveys rather than administrative data. As a result, sample sizes are smaller and significance levels are correspondingly lower.[52]

For the most part, these programs had effects on Medicaid qualitatively similar to their effects on welfare use. Both MFIP-IO and MFIP increased Medicaid coverage at the 36-month mark, although only the estimate for regular MFIP is significant. This is consistent with these programs' effects on welfare use. Both programs increased welfare use significantly in year three, and since welfare recipients were automatically entitled to receive Medicaid at the time, Medicaid eligibility should have risen as well.

Similarly, L.A. Jobs-First GAIN decreased Medicaid use, much as it decreased welfare use. The program decreased welfare use by 1.12 months and Medicaid use by 0.89 months, over the 24-month follow-up period. (Freedman et al., 2000b, Table 4.1, p. 74). This suggests that roughly 80 percent of those who left welfare as a result of the program also discontinued their Medicaid receipt.

In the Michigan TSMF program, Medicaid eligibility fell for ongoing recipients by about the same amount as welfare use. For the applicant groups, however, the effects of the program on Medicaid were insignificant, even though the program reduced welfare use among the early cohort of applicants by a significant amount. FIP is again an exception to the general rule, resulting in decreased Medicaid use despite increases in welfare use. Delaware's ABC program had no significant effect on Medicaid, just as it had no significant effect on welfare use. FTP

---

[51]The NEWWS programs provide impact estimates for TMA, but not for Medicaid coverage more generally. Since TMA is available only to persons leaving welfare because of employment or an earnings gain, it is a function of both Medicaid-related policy changes and the programs' impacts on employment. Moreover, TMA is only one means by which study participants may receive Medicaid coverage. For these reasons, TMA impacts are not comparable to the more general Medicaid impacts presented in Table 6.3 and are excluded from our analysis.

[52]None of the studies provides information about SCHIP, since SCHIP was not funded until 1998, after the follow-up periods of the studies that included questions about Medicaid participation.

Table 6.3—Estimated Impact of Welfare Reform on Medicaid Coverage: Random Assignment Studies

				Medicaid coverage		
Name	Cases served	Data	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives						
MFIP-IO	Urban single parents recipients	S	Adult covered by Medicaid 36 mos. after RA	66.2	4.2	6.3%
B. Programs that focus on financial work incentives tied to hours of work						
C. Programs that focus on mandatory work-related activities						
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Months of Medicaid coverage, years 1 and 2	21.2	-0.9 ***	-4.2%
D. Programs that focus on financial work incentives and mandatory work-related activities						
MFIP	Urban single-parent recipients	S	Adult covered by Medicaid at 36-month interview	66.2	6.4 *	9.7%
TSMF	Recipients	A	Adult average monthly Medicaid eligibility over 4-yr FU	67.6	-1.7 ***	-2.5%
	Applicants	A	Adult average monthly Medicaid eligibility over 1-yr FU	70.2	0.6	0.9%
		A	Adult average monthly Medicaid eligibility over 2-yr FU	61.2	-1.0	-1.6%
FIP	Recipients	A	Any paid claims, Q4	89.5	-3.1 ***	-3.5%
		A	Any paid claims, Q8	76.5	-1.7	-2.2%
	Applicants	A	Any Medicaid action, Q4	58.1	2.4	4.1%
		A	Any Medicaid action, Q8	n/a		
E. Programs that focus on other individual reforms						
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)						
ABC	Recipients and applicants	S	Percent receiving Medicaid at interview	76.7	-2.0	-2.6%
FTP	Recipients and applicants	S	Respondent covered by Medicaid at 4-year survey	36.8	-2.6	-7.1%
JOBS First	Recipients and applicants	S	Respondent covered by Medicaid at 3-year survey	60.4	9.1 ***	15.1%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; Q=quarter.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

had little effect on Medicaid at the time of the survey, despite reducing welfare use. Jobs First substantially increased Medicaid coverage at the time of the three-year follow-up at the same time that it substantially decreased welfare use. Jobs First extended TMA eligibility one additional year, which may account for this pattern, although the impact of this particular policy component can not be isolated in the Jobs First experimental design. The only other program that extended TMA was ABC, and that program produced no change in welfare use or Medicaid use.

Econometric Studies

The single econometric study to examine the effects of welfare reform on Medicaid caseloads is summarized in Table 6.4. That study, by Ku and Garrett (2000), uses state-level administrative data on nonelderly, nondisabled adults and children receiving Medicaid. The authors fit separate models for adults and children and include an extensive set of policy variables and controls in both specifications. They also include state and year dummies to control for unobservables.

The models provide some evidence that welfare waivers affect Medicaid use. Although the authors find that earnings disregards and family caps had no significant effects on Medicaid use, they find that fewer exemptions from JOBS work requirements had a negative and marginally significant effect. Since waivers generally reduced the fraction of the welfare caseload exempt from work requirements, this suggests that welfare waivers may have reduced Medicaid receipt among both adults and children.

6.2.3. Other Programs

Only four sets of random assignment studies have analyzed the effects of welfare reform on participants use of other nutrition programs: the NEWWS programs, L.A. Jobs-First GAIN, Indiana's IMPACT evaluation, and Delaware's ABC evaluation. The results of the effects of these programs on participation in the various nutrition programs are summarized in Table 6.5. All these results are based on survey data.[53] To our knowledge, there are no econometric studies that analyze the effects of welfare reform on school nutrition programs or the WIC program.

Of the 18 estimates included in Table 6.5, all but one are fairly small and insignificant. The one exception involves the Atlanta HCD program, where School Lunch use rose by a statistically significant 3.4 percentage points. Since we would expect roughly 1 in 20 statistical estimates to be significant at the 5 percent level even if the programs truly had no effect, these results suggest that welfare reform has had little if any effect on participation on these other nutrition programs. They suggest that, in these programs, at least, families do not use school nutrition programs as a substitute for the FSP.

The only random assignment studies to include information about housing subsidies are L.A. Jobs-First GAIN, the NEWWS programs, EMPOWER, FTP, and Jobs First. All information comes

[53]Estimates for IMPACT are not available separately for participants in the Basic and Placement Tracks.

Table 6.4—Estimated Impact of Welfare Reform on Medicaid Coverage: Econometric Studies

										Other controls				
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Includes LDV's?	Economy	Demogr.	Fixed Effects	Policy
Ku and Garrett (2000)	State-level Medicaid caseload	Adults in 44 states	84	96	annual non-disabled adult recipients/annual adult population	Log	Any waiver	-0.008 (0.022)	-0.8	no	U, earnings	%men insured, %FHHH, %Hispanic	S, Y	B, size of Medicaid expansion, medically needy program, UP program
							Earnings disregards	0.110 (0.082)	11.0					
							Family cap	-0.042 (0.032)	-4.2					
							Percent exempt from JOBS	0.114 (0.083)	11.4					
Ku and Garrett (2000)	State-level Medicaid caseload	Children in 44 states	84	96	annual non-disabled child recipients/annual child population	Log	Any waiver	-0.002 (0.021)	-0.2	no	U, earnings	%men insured, %FHHH, %Hispanic	S, Y	B, size of Medicaid expansion, medically needy program, UP program
							Earnings disregards	-0.031 (0.081)	-3.1					
							Family cap	-0.024 (0.031)	-2.4					
							Percent exempt from JOBS	0.153 (0.080)	15.3					

NOTES: Abbreviations: LDV=lagged dependent variable; s.e.=standard error; FHHH=female-headed households; U=unemployment rate; B=maximum welfare benefit, S=state; Y=year.

Table 6.5—Estimated Impact of Welfare Reform on Use of Nutrition Programs: Random Assignment Studies

Name	Cases served	Data	Measure	Nutrition program participation: Control mean	Impact	%
A. Programs that focus on financial work incentives						
B. Programs that focus on financial work incentives tied to hours of work						
C. Programs that focus on mandatory work-related activities						
LA Jobs-1st GAIN	Single-parent recipients and applicants	S	Ever participate in school meal program during FU	66.3	2.4	3.6%
Atlanta LFA	Recipients and applicants	S	Ever participate in school meal program during FU	86.2	1.8	2.1%
Grand Rapids LFA	Recipients and applicants	S	Ever participate in school meal program during FU	67.1	1.2	1.8%
Riverside LFA	Recipients and applicants	S	Ever participate in school meal program during FU	78.1	-1.8	-2.3%
Portland	Recipients and applicants; no cases with substantial barriers	S	Ever participate in school meal program during FU	66.1	-1.6	-2.4%
Atlanta HCD	Recipients and applicants	S	Ever participate in school meal program during FU	86.2	3.4 **	3.9%
Grand Rapids HCD	Recipients and applicants	S	Ever participate in school meal program during FU	67.1	-1.5	-2.2%
Riverside HCD	Recipients and applicants	S	Ever participate in school meal program during FU	81.4	0.4	0.5%
Columbus Integrated	Recipients and applicants	S	Ever participate in school meal program during FU	75.6	-1.4	-1.9%
Columbus Traditional	Recipients and applicants	S	Ever participate in school meal program during FU	75.6	-0.9	-1.2%
Detroit	Recipients and applicants	S	Ever participate in school meal program during FU	60.2	1.2	2.0%
Oklahoma City	Applicants	S	Ever participate in school meal program during FU	59.6	-2.1	-3.5%
D. Programs that focus on financial work incentives and mandatory work-related activities						
E. Programs that focus on other individual reforms						
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)						
IMPACT	Recipients and applicants-basic and placement tracks	S	Participating in school lunch at time of two-year survey	60.9	2.7	4.4%
		S	Participating in school breakfast at time of two-year survey	47.3	-2.5	-5.3%
		S	Participating in WIC at time of two-year survey	23.2	-1.7	-7.3%
ABC	Recipients and applicants	S	Participating in school lunch at time of Spring 1997 survey	52.3	-0.9	-1.7%
		S	Participating in school breakfast at time of Spring 1997 survey	45.0	-0.1	-0.2%
		S	Participating in WIC at time of Spring 1997 survey	35.4	-2.3	-6.5%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; Q=quarter.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

from surveys in which participants were asked whether they lived in public housing projects or received rent subsidies. The results are summarized in Table 6.6.

All but two of the estimates are insignificant.[54] Of the two significant estimates, one is negative and the other is positive. These estimates provide little reason to think that these types of reforms affect whether welfare recipients live in public or subsidized housing.

6.3. EVALUATING THE EFFECTS OF WELFARE REFORM ON USE OF OTHER GOVERNMENT PROGRAMS

In the random assignment studies, food stamp use generally follows welfare use. In programs where welfare use rose, such as MFIP, food stamp use rose as well. In studies where welfare use fell, food stamp use tended to fall as well. FIP and New Hope represent exceptions to this general pattern. In FIP, food stamp use fell despite increases in welfare use. In New Hope, families employed full time at random assignment increased their food stamp use while decreasing their welfare use.

In programs where welfare use fell, the corresponding decrease in food stamp use was generally smaller. This is consistent with the notion that welfare reform explains part, but not all, of the observed decline in the food stamp caseload that has taken place since the early 1990s. Evidence from two of the three econometric studies that address this issue is consistent with this view as well.

None of these studies explains why a decrease in welfare use results in a decrease in food stamp use. Most families leaving welfare, and most poor families generally, are eligible for food stamps, regardless of whether they receive cash aid. Thus, there is no automatic or legal link between welfare use and food stamp receipt. Various analysts have suggested that food stamp use has fallen in response to welfare reform because of administrative practices on the part of some welfare agencies, because the transactions costs of maintaining eligibility are higher for working families, and because poor families fail to understand that they may be eligible for food stamps even if they are ineligible for, or do not elect to receive, cash welfare.

There is substantially less evidence about the effects of welfare reform on Medicaid coverage. The single econometric study on the issue finds that less lenient age exemptions for work requirements reduce Medicaid use, but those estimates are only marginally significant. The random assignment studies suggest that programs that increase welfare use increase Medicaid receipt, whereas programs that decrease welfare use also decrease Medicaid receipt. Again, FIP is an exception to this rule, as is Jobs First. Moreover, there are only seven such studies that analyze Medicaid, which is a small number on which to base any general conclusions. It is also important to keep in mind that some recipients leaving welfare for work may replace their Medicaid coverage with employer-provided health insurance. We consider the effects of welfare reform on health coverage more generally in Chapter 9 below.

The effects of welfare reform on other government programs have been studied even less. Only four sets of studies analyzed participation in school nutrition programs. One of the NEWWS

[54]Impacts from the Jobs First four-year survey were also insignificant.

programs yielded a significant increase in program use; all the other estimates were rather small and insignificant. Only the NEWWS programs and L.A. Jobs-First GAIN consider housing subsidies; of the two significant effects, one was positive and one was negative. Regarding the effects of welfare reform on the use of these other government programs, there is simply too little information available to draw any firm conclusions.

6.4. CONCLUSIONS

The results synthesized in this chapter are generally consistent with the hypothesis that welfare reform has caused part of the recent decline in food stamp use. However, they tell us nothing about the mechanisms that underlie this linkage. As noted above, the research literature has suggested that food stamp use has fallen in response to welfare reform because of administrative practices on the part of some welfare agencies, because the transactions costs of maintaining eligibility are higher for working families, and because poor families fail to understand that they may be eligible for food stamps even if they are ineligible for, or do not elect to receive, cash welfare. To ensure that poor families are more effectively covered by this important safety net program, it is important to understand the reasons why welfare reform has caused food stamp use to fall.

However, perhaps the main message from this chapter is that there is very little information available from which to draw conclusions about the effects of welfare reform on the use of other government programs. Only a few studies analyze the effects of welfare waivers on the use of school nutrition programs. Fewer still have considered housing subsidies.

In the case of the Medicaid program, this limited research base is unfortunate in light of the policy objectives of recent reforms. During the decade prior to PRWORA, the federal government expanded the eligibility criteria for Medicaid with the dual objectives of weakening the links between welfare and health coverage and providing coverage for larger numbers of poor children. In the context of PRWORA, the legislation explicitly sought to ensure that reduced eligibility for welfare would not entail reduced eligibility for Medicaid. Nevertheless, Medicaid coverage of persons in poor families fell during the late 1990s at the same time that rates of uninsurance rose. Unfortunately, the evidence that exists provides too narrow a basis from which to draw general conclusions about whether welfare reform has caused part of the decline in Medicaid.

Table 6.6—Estimated Impact of Welfare Reform on Use of Public or Subsidized Housing: Random Assignment Studies

Name	Cases served	Data	Measure	Public housing			Subsidized housing		
				Control mean	Impact	%	Control mean	Impact	%
A. Programs that focus on financial work incentives									
B. Programs that focus on financial work incentives tied to hours of work									
C. Programs that focus on mandatory work-related activities									
LA Jobs-1st GAIN	Single-parent recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	16.1	0.5	3.1%	8	1.8	23.4%
Atlanta LFA	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	33.1	1.3	3.9%	19.4	1.1	5.7%
Grand Rapids LFA	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	9.4	-0.7	-7.4%	11.5	-1.6	-13.9%
Riverside LFA	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	6.3	-0.4	-6.3%	7.9	-0.5	-6.3%
Portland	Recipients and applicants; no cases with substantial barriers	S	Lived in public/subsidized housing at end of 2-year FU	18.6	1.1	5.9%	17.0	-6.5 **	-38.2%
Atlanta HCD	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	33.1	-1.0	-3.0%	19.4	4.4 **	22.7%
Grand Rapids HCD	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	9.4	0.3	3.2%	11.5	-0.5	-4.3%
Riverside HCD	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	7.3	-2.1	-28.8%	8.3	-1.0	-12.0%
Columbus Integrated	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	27.6	0.5	1.8%	8.2	-1.1	-13.4%
Columbus Traditional	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	27.6	-0.5	-1.8%	8.2	-2.5	-30.5%
Detroit	Recipients and applicants	S	Lived in public/subsidized housing at end of 2-year FU	9.4	0.0	0.0%	1.8	0.2	11.1%
Oklahoma City	Applicants	S	Lived in public/subsidized housing at end of 2-year FU	12.6	0.3	2.4%	9.5	-0.3	-3.2%

Table 6.6—Continued

				Public housing			Subsidized housing		
Name	Cases served	Data	Measure	Control mean	Impact	%	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities									
E. Programs that focus on other individual reforms									
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)									
EMPOWER (a)	Recipients	S	Lived in public or subsidized housing at wave 1 interview	29.1	0.6	2.1%			
FTP	Recipients and applicants	S	Lived in public or subsidized housing at 4-year interview	22.1	-1.3	-5.9%			
JOBS First	Recipients and applicants	S	Lived in public or subsidized housing at interim interview	48.4	-4.4	-9.1%			

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; Q=quarter.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) Phoenix site only, cash assistance.

FAMILY STRUCTURE

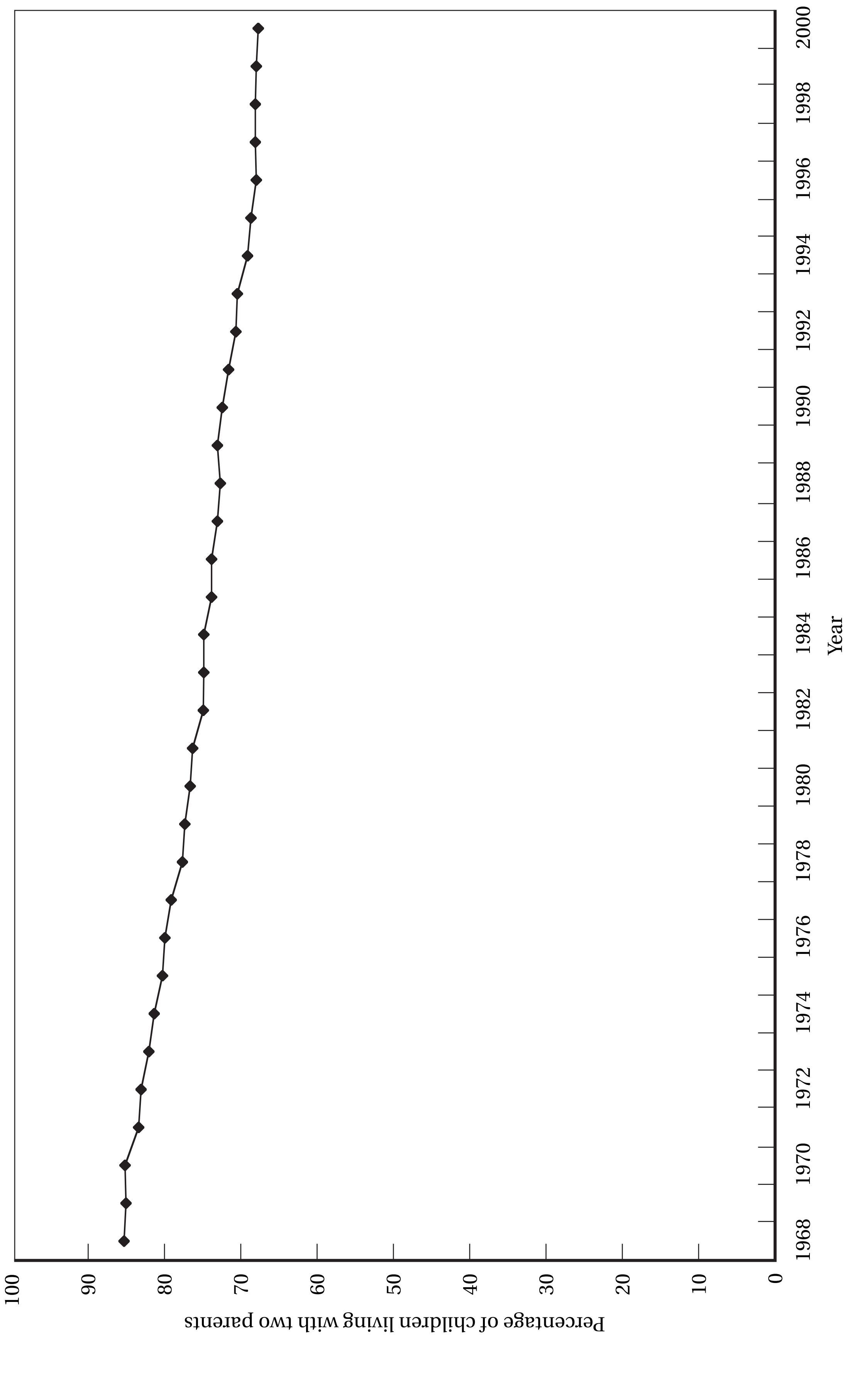
7.1. BACKGROUND

As noted in Chapter 1, in addition to promoting work and reducing dependency, PRWORA aimed to reduce unwed childbearing, to promote marriage, and to maintain two-parent families. In this chapter, we turn to the impact of welfare reform on family structure, considering both marriage and childbearing.

PRWORA's focus on reducing unwed childbearing, promoting marriage, and maintaining two-parent families was partially motivated by concern about trends in those outcomes. Up until about 1970, more than 85 percent of American children were being raised in two-parent families. Over the succeeding three decades, that figure fell to under 70 percent (see Figure 7.1) because of increases in nonmarital childbearing and, to a lesser extent, increases in divorce. Figure 7.2 shows that while in the 1950s less than 5 percent of births were to unmarried women, beginning in the early 1960s, this percentage began to increase sharply. By the early 1990s, one-third of births were to unmarried women. This rise in nonmarital childbearing was an important cause of the decrease in the share of children being raised by two parents.

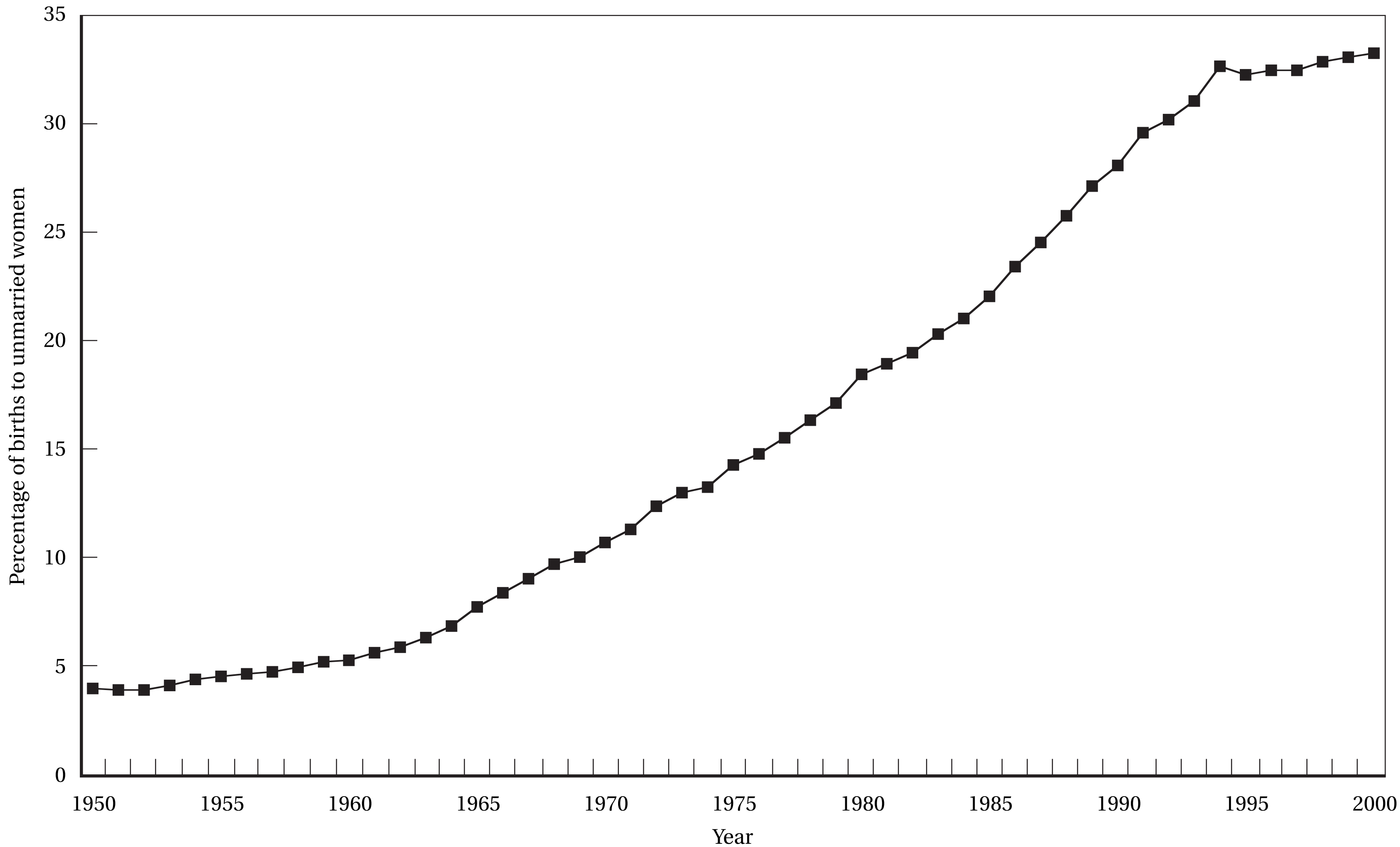
As seen in both Figures 7.1 and 7.2, some of the trends in family structure and fertility appear to have slowed or stabilized in the latter part of the 1990s, about the time welfare reform was under way. Both the percentage of children living in two-parent families and the percentage of births to unmarried women has been approximately constant since 1994. The overall trend evident in Figure 7.1 is consistent with other recent analyses of family structure, with some evidence that the relative changes in one- versus two-parent families is more pronounced for families with lower income or less education, precisely the groups that are more likely to be affected by welfare reform (Acs and Nelson, 2001; Dupree and Primus, 2001).

In the case of fertility, the leveling-off of the trend for nonmarital childbearing seen in Figure 7.2 has been accompanied by a decline in teen fertility rates during the 1990s (Martin et al., 2001). For example, across all race and ethnic groups, the drop in teen fertility from 1991 to 2000 is 28.9 percent for 15–17-year-olds and 15.8 percent for 18–19-year-olds. Furthermore, the drop is particularly large for blacks (40.3 percent and 23.6 percent for 15–17-year-olds and 17–19-year-olds, respectively). However, some of the decline occurred in the early 1990s, before widespread welfare reform efforts, raising questions about the role that reform played in reducing teen fertility.



SOURCE U.S. Bureau of the Census (2001).

Figure 7.1—Percentage of Children Living with Two Parents: 1968–2000



SOURCE: U.S. Department of Health and Human Services (1985), Table I-3, and U.S. National Center for Health Statistics (2002), Table D.

Figure 7.2—Percentage of Births to Unmarried Women: 1950–2000

These trends are suggestive that welfare reform may have had some impact on fertility and family structure, and a number of provisions implemented by the states initially under section 1115 waivers and then TANF were designed to directly affect these outcomes. As noted in Chapter 2, a number of states instituted family caps with the objective of reducing additional childbearing for mothers already on welfare. Minor residency requirements are another feature designed to make unwed teen childbearing less attractive. In addition, by eliminating differences in eligibility for two-parent versus one-parent families (e.g., the "100-hour rule" and work history requirement), states aimed to diminish any disincentive toward marriage associated with welfare eligibility rules.[55]

PRWORA's emphasis on family structure outcomes was partially motivated by an extension of the economic model of the effect of welfare programs developed in the earlier chapters. That extension views women as considering the structure of welfare programs when making choices not only about welfare and work, but also when making choices about family structure—whether to have children, whether to marry the father, and whether to subsequently divorce.

The theory's implications follow from noting that welfare has primarily been paid to single mothers, but not to childless women, nor (under most circumstances) to married women.[56] Welfare therefore lowers the price of raising a child when unmarried relative both to not having a child and relative to having a child and marrying (or not divorcing). Therefore, this model suggests that any policy change that makes welfare relatively more attractive (e.g., higher benefit levels or financial work incentives) will raise fertility (and especially nonmarital fertility) and decrease marriage. Conversely, any policy change that makes welfare relatively less attractive (e.g., a family cap, mandatory work-related activities, or time limits) will lower fertility (and nonmarital fertility) and increase marriage. However, when such reforms are enacted together, the combined effect on marriage and fertility is ambiguous.

These implications of economic theory assume that welfare is not available to married couples. However, welfare was potentially available to married couples under the AFDC Unemployed Parent (AFDC-UP) program and continues to be available under TANF. Making welfare payments to married couples increases the incentive to have children, but lowers the disincentive to marriage (Hu, 2000). As noted above, to further reduce the disincentives to be married, most states have reduced or eliminated the differential treatment of two-parent families under their TANF programs. For two reasons, however, the effects of the provisions of such welfare programs for married women are likely to be small. First, most married couples have income sufficiently high to make them income ineligible for welfare. Second, and perhaps as a consequence, the AFDC-UP program (under TANF, two-parent programs) are quite small in most states.

---

[55]The "100-hour rule" under AFDC required that, in addition to being financially eligible for benefits, the primary wage-earner could work no more than 100 hours per month. To meet the work history requirement, the family also had to show that the primary wage-earner had earned at least $50 in at least 6 of the last 13 calendar quarters or had been eligible for unemployment compensation during the past year.

[56]Under AFDC-UP, welfare was potentially available to two-parent families only when there was substantial previous labor market experience (six of the last thirteen quarters, but less than 100 hours of work in the current month). This condition made it unlikely that teens or young two-parent families would qualify.

There are other mechanisms by which welfare reform may affect family structure. For example, if welfare-to-work programs succeed in raising earnings and income, they might make women more attractive spouses and, thus, raise the propensity to marry. At the same time, increased work may limit the time available for searching for a marital partner; then again, interactions at the workplace may ease marital search. As yet another example, low household income may increase the emotional and financial strain on a marriage, so that welfare reforms that raise total income might be expected to increase marriage and, in particular, to help those currently married to stay married.

Although welfare reform was motivated in part by trends in marriage and fertility, these outcomes are less well studied in both the experimental and econometric literatures. Of the random assignment studies we review in this report, WRP, IMPACT, TSMF, FIP, New Hope, SSP Plus, SSP Applicants, VIP/VIEW, PPI, and PIP do not analyze either marriage or fertility. CWPDP, MFIP, and SSP examine only marriage, while AWWDP and FDP consider only fertility. The remaining programs—L.A. Jobs-First GAIN, the 11 NEWWS programs, EMPOWER, ABC, FTP, and Jobs First—analyze both outcomes. Two econometric studies consider either marriage or living arrangements, while there are four econometric analyses of fertility.

Compared with the outcomes examined in Chapters 4, 5, and later in 8, the more limited research on marriage and fertility can be attributed to several factors. First, although PRWORA motivated reform in part by goals related to marriage and childbearing, many of the state programs evaluated under waivers were designed more to influence work and welfare use. Even so, a few of the programs that included family caps, minor residency requirements, and changes in two-parent eligibility requirements do not evaluate either marriage or fertility (e.g., IMPACT and VIP/VIEW).

Second, unlike welfare use, employment and earnings, and some measures of income, marriage and fertility behavior are harder to measure using administrative data (although this is the source of information on fertility for FDP and AWWDP). Thus, those demonstration studies that do not have participant surveys are less likely to consider these outcomes. Third, even when resources are devoted to measuring these outcomes, changes in marital status and additional childbearing while on welfare are relatively rare events and changes in behavior may not be immediate, whether for the recipient generation or for the next generation of daughters of the recipients. As a result, studies with short follow-up periods may be less likely to detect significant changes in these outcomes. In addition, survey data often have smaller samples and are subject to measurement error (e.g., recall bias and differential non-response), leading these analyses to have lower power.[57] Consequently, these outcomes may not be included in impact analyses, and when they are, there may be limited statistical power to detect significant changes in behavior.

Fourth, the influence of welfare reform on marriage and fertility behavior is likely to affect women who are not on welfare just as much, if not more, than those who are on welfare. While

---

[57]This use of survey data is in contrast to most of the analyses of the previous chapters (considering welfare use, employment and earnings, and use of other government programs), but like most of the analyses in later chapters (considering income, other measures of well-being, and child development). As discussed below, population level analyses of fertility have access to Vital Statistics/birth certificate data. Like administrative data, birth certificate data are available for the entire population (not just a sample), in every time period, and without recall bias. However, only aggregate birth certificate data are available. No study has matched experimental data to individual-level birth certificate data, so random assignment analyses cannot use these data.

welfare reform may affect the likelihood that a woman on welfare has additional children or gets married or stays married, it should also affect these decisions for women who are *at risk* of welfare participation. For these women, welfare reform may affect their likelihood of entering welfare. However, as noted in Chapter 3, conventional demonstration studies are not designed to capture welfare entry effects, so they will miss this pathway by which reforms may affect family structure. This is a significant limitation of the demonstration studies and stresses the need for high-quality econometric studies.

The remainder of this chapter proceeds by considering first the random assignment studies and then the econometric studies of family structure and its two primary components: marriage and fertility. Since there is only one demonstration study with subgroup analyses of marriage and fertility, we discuss these results along with the main results rather than in a separate section (or in Appendix A). After discussing the random assignment and econometric studies in turn, we proceed to a synthesis of the experimental and econometric evidence. The final section offers our conclusions.

7.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECTS OF WELFARE REFORM ON FAMILY STRUCTURE

In this section, we consider the effects of random assignment studies on family structure, namely marriage, household size, and fertility. As noted above, most of the demonstration studies that consider these outcomes use survey data to assess whether a participant in the treatment or control group has had an additional child since random assignment or the participant's marital status at the time of the follow-up survey. The follow-up interval ranges from 18 months to five years.

In assessing current marital status, studies differ in whether they differentiate between those who are married versus those who are married and living with their spouse. Some studies also report impacts for cohabitation, separate from being married, or combined with those who are married. A few studies also measure whether there was any change in marital status since random assignment, given that the respondent may have married and subsequently become separated, divorced, or widowed by the time of the follow-up. Finally, two studies measure household composition in terms of household size and the number of adults and children. Changes in household size may result from changes in marital status or additional childbearing, but also for other reasons such as "doubling up" with other relatives or nonrelatives, or departures of older children who move out of the household. Where possible, our discussion focuses on marriage with a spouse present, the concept that most closely aligns with PRWORA's goals, but often only results for other outcomes (e.g., any cohabitation, marital or nonmarital) are available.

Fertility is typically measured for the survey respondent, and the measure is whether the respondent has had any children since random assignment. In one study, EMPOWER, childbearing while on welfare is measured both for case heads and unwed minors in the welfare case unit. That study and ABC also differentiate births since random assignment from conceptions since random assignment (defined as births more than 10 months since random assignment). For the other studies, some of the measured births may have been conceived prior to the time when the program rules becoming effective.

Finally, two studies use information from welfare data systems to measure children born to study participants. However, recording of births in welfare data systems is incomplete. Current welfare recipients have an incentive to report births. A reported birth will enable the child to be enrolled in Medicaid, and, in the absence of a family cap, the family's welfare payment will increase. Births to mothers not receiving welfare are not recorded in any welfare data system. This is an important limitation because PRWORA's interest in reducing out-of-wedlock childbearing is not limited to births among welfare recipients.

In the remainder of this section, we focus first on the results for marriage and household size, followed by the results for fertility. In both cases, we organize our discussion by the major reform or reforms considered by the demonstrations.

7.2.1. Marriage and Household Size

Table 7.1 records the results for the random assignment studies that examine marriage and household size. With the exception of Panel E, at least one study in each of the other policies or groups of policies in Panels A to F examines a measure of marriage.

Programs That Focus on Financial Work Incentives

As seen in Panel A of Table 7.1, results for two programs that focus on financial work incentives provide some evidence for an increase in marriage. Hu (2000) estimates the effects of the CWPDP on marital status. While he finds no effect for AFDC-Basic (i.e., single parent) cases, he finds that the experiment increased marriage for AFDC-UP cases. The effect appears to be the result of less divorce. The interpretation of these results is, however, complicated by policy bundling. The CWPDP waiver included a financial incentive; it also included a cut in the AFDC benefit level (at zero earnings) and removed some of the restrictions on eligibility of two-parent families (similar to those in MFIP discussed below). It is not clear which of the components of the bundle caused the marital status effect.

MFIP-IO is a pure financial incentive program, and its financial work incentives were deliberately designed to encourage marriage. Some restrictions on eligibility for two-parent families were eliminated, and the treatment of stepparent earnings was liberalized. Consistent with this intention, the experimental evaluation of the financial work incentives alone (i.e., MFIP-IO) suggests that marriage increases. For single parent recipients, the fraction married at the time of the 36-month follow-up interview is 11.0 percent in the treatment group versus 5.8 percent in the control group, a statistically significant difference of 5.2 percent. The impact is also positive on the combined status of married or cohabiting, but the difference is not significant. For single parent applicants, treatment group members are less likely to be married—or married or cohabiting—but again the difference is not significant.

Programs That Focus on Financial Work Incentives Tied to Hours of Work

Among the programs that tied financial work incentives to hours of work, SSP is the only one that assesses the impact on marriage behavior (Panel B of Table 7.1). The structure of SSP's incentives was specifically designed to lower disincentives to marry. Canada's Income Assistance program counts a husband's income when calculating the welfare benefit. If the

Table 7.1—Estimated Impact of Welfare Reform on Marital Status and Household Size: Random Assignment Studies

			Marital Status				Change in Marital Status				Household Size			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives														
CWPDP	Single-parent recipients	S	R is married at 29-41 mo FU (%)	13.7	2.1	15.3%								
	Two-parent recipients	S	R is married at 29-41 mo FU (%)	71.2	7.6 **	10.7%								
MFIP-IO	Urban single-parent recipients	S	R is married at the 36-mo FU (%)	5.8	5.2 **	89.7%								
		S	R is married or living with partner at the 36-moFU (%)	20.8	2.7	13.0%								
	Urban single-parent applicants	S	R is married at the 36-mo FU (%)	15.1	-2.2	-14.6%								
		S	R is married or living with partner at the 36-moFU (%)	29.6	-2.6	-8.8%								
B. Programs that focus on financial work incentives tied to hours of work														
SSP (a)	Single-parent recipients	S	R is married at 36-mo FU (%)	9.5	-0.6	-6.3%								
		S	R is married or in common law relationship at 36-mo FU (%)	17.3	0.1	0.6%	R ever married or in common law relationship as of 36-mo FU (%)	19.2	0.3	1.6%				

Table 7.1—Continued

			Marital Status				Change in Marital Status				Household Size			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
C. Programs that focus on mandatory work-related activities														
LA Jobs-1st GAIN	Single-parent recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	6.9	2.2	31.9%								
		S	R is living with partner at 2-yr FU (%)	8.5	-1.1	-12.9%								
Atlanta LFA	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	4.0	-0.3	-7.5%								
		S	R is married and living with spouse at 5-yr FU (%)	8.4	1.3	15.5%								
Grand Rapids LFA	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	11.8	1.3	11.0%								
		S	R is married and living with spouse at 5-yr FU (%)	20.5	2.3	11.2%								
Riverside LFA	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	13.4	-2.7 *	-20.1%								
		S	R is married and living with spouse at 5-yr FU (%)	22.0	-1.4	-6.4%								
Portland	Recipients and applicants; no cases with substantial barriers	S	R is married and living with spouse at 2-yr FU (%)	9.0	-0.2	-2.2%								
		S	R is married and living with spouse at 5-yr FU (%)	23.6	-6.2	-26.3%								
Atlanta HCD	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	4.0	-1.2	-30.0%								
		S	R is married and living with spouse at 5-yr FU (%)	8.4	-1.5	-17.9%								
Grand Rapids HCD	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	11.8	0.3	2.5%								
		S	R is married and living with spouse at 5-yr FU (%)	20.5	-0.2	-1.0%								
Riverside HCD	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	10.9	1.6	14.7%								
		S	R is married and living with spouse at 5-yr FU (%)	18.1	3.7	20.4%								
Columbus Integrated	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	9.0	1.1	12.2%								
Columbus Traditional	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	9.0	0.9	10.0%								
Detroit	Recipients and applicants	S	R is married and living with spouse at 2-yr FU (%)	7.6	-3.4	-44.7%								
Oklahoma City	Applicants	S	R is married and living with spouse at 2-yr FU (%)	19.1	-3.4	-17.8%								

Table 7.1—Continued

			Marital Status				Change in Marital Status				Household Size			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities														
MFIP	Urban single-parent recipients	S	R is married at the 36-mo FU (%)	5.8	2.8	48.3%								
		S	R is married or living with partner at the 36-mo FU (%)	20.8	3.2	15.4%								
	Urban single-parent applicants	S	R is married at the 36-mo FU (%)	15.1	1.7	11.3%								
		S	R is married or living with partner at the 36-mo FU (%)	29.6	4.1	13.9%								
E. Programs that focus on other individual reforms														
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)														
EMPOWER (b)	Recipients	S	R is married at 3-yr FU(%)	28.9	-0.9	-3.1%	R changed marital status since RA as of 3-yr FU (%)	7.7	0.0	0.0%				
ABC	Single parent recipients and applicants	S	R is married and living with spouse at 4-19-mo FU (%)	7.6	1.4 *	18.4%								
FTP	Recipients and applicants	S	R is married and living with spouse at 4-yr FU (%)	19.1	-1.9	-9.9%					Total number of HH members (including R) at 18-mo FU	3.9	0.0	0.0%
Jobs First	Recipients and applicants	S	R is married and living with spouse at 18-mo FU (%)	7.0	-1.2	-17.1%	R changed marital status since RA as of 18-mo FU (%)	19.9	-1.7	-8.5%	Total number of HH members at 18-mo FU	3.3	0.2 ***	7.6%
		S	R is married and living with spouse at 3-yr FU (%)	10.8	-1.6	-14.8%					Total number of HH members at 3-yr FU	3.4	0.1	2.9%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; R=respondent; RA=random assignment.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) New Brunswick and British Columbia combined.

(b) Phoenix site only, cash assistance.

husband works, this will usually result in a lower benefit and thus would be expected to discourage marriage. In contrast, SSP disregards income contributed by a husband or common-law spouse when calculating the earnings supplement, thereby removing the disincentive and encouraging marriage. However, the higher household income under SSP (discussed further in Chapter 8) might have been expected to induce some women to choose to live on their own, thus decreasing marriage.

SSP includes a measure of marital status as well as a broader measure that includes both formal marriage and Canadian common-law relationships.[58] Using this combined marriage and common-law relationship concept, SSP has insignificant impacts. Marriage is slightly less common; common-law relationships are slightly more common; neither effect is statistically different from zero.

The interpretation of the SSP results is, however, complicated by considering the two provinces—British Columbia and New Brunswick—separately. For almost all outcomes considered in the SSP evaluation, impacts do not differ significantly across the two provinces. Marriage is the exception. Using the broad SSP definition, marriage significantly decreases in British Columbia (by 3.1 percentage points, or 18 percent of the value for the control group); while marriage significantly increases in New Brunswick (by 4.1 percentage points; or 20 percent). Using a narrow definition of marriage that excludes common law relationships, the effect in British Columbia is still negative, but at $p < 0.10$ (but not at $p < 0.05$). The effect in New Brunswick is still positive, but not statistically different from zero. The difference is significant at $p < 0.10$, but not at $p < 0.05$.

Michalopoulos et al. (2000) discuss the possible reasons for the difference in results across provinces. They note that the results within each province are robust across subgroups, so that the small differences in baseline characteristics between the two provinces do not explain the differences in impact. They also note that the impacts on income and full-time employment were similar across the two provinces and the policy changes removing the marriage penalty were identical.

They suggest two other plausible reasons for the divergence across provinces. A first reason relates to the marriage market. During the period of the experiment, the unemployment rate for men was considerably higher in New Brunswick than in British Columbia. They speculate that these poor job prospects for men made the additional employment, earnings, and income provided by SSP more attractive. It should be noted that this argument—that poor economic prospects for men encourage them to marry—is the opposite of the standard argument that marriage among American black women is low because there are few marriageable men (Wilson and Neckerman, 1987). A second reason concerns cultural differences. New Brunswick is more rural, and the majority is Catholic; British Columbia is more urban, and there are fewer Catholics. With only two sites and nominally identical programs, more definite conclusions are not possible. They conclude: "The opposite direction of impacts by province underscores the importance of geographic and cultural context in translating employment and earnings impacts into effects on family structure."

[58] In Canada, couples who live together for at least one year and are not legally married are considered common-law partners, with rights that are akin to marriage (Michalopoulos et al., 2000).

Programs That Focus on Mandatory Work-Related Activities

With one exception, the programs that evaluate mandatory work-related activities show no significant impacts on the fraction married and living with their spouse as of the two-year follow-up survey. As seen in Panel C of Table 7.1, the 12 insignificant impacts are evenly divided in sign and most involve a small percentage point change. Only Riverside LFA has a marginally ($p < 0.10$) statistically significant negative impact on the likelihood of being married. For seven of the NEWSS sites, there are also five-year follow-up results. In none of them (including Riverside LFA) can we reject the hypothesis of no effect. Again, the sites are divided in sign and the point estimates are small.

Programs That Focus on Financial Work Incentives and Mandatory Work-Related Activities

Among programs that combine financial work incentives with mandatory work-related activities, only MFIP assesses the impact on marriage defined as marriage alone and a broader measure that includes cohabitation (see Panel D of Table 7.1). For both urban single parent recipients and applicants, the MFIP impacts on the narrow (marriage) and broad (cohabitation) measures are positive, but none are statistically significant.[59]

In addition, the MFIP evaluation considered the impact of the full program on marriage for the sample of two-parent families (results not shown). For that study sample, the MFIP intervention increased the fraction remaining married by nearly 40 percent, from 48.3 percent for the control group to 67.4 percent for the treatment group, and the result is statistically significant. Analyses of other outcomes suggest that the effect is concentrated among those married (rather than cohabiting) at random assignment and works partially through a drop in the divorce rate (about 6.5 percentage points). The balance of the effect appears to be higher rates of married couples living together. Furthermore, these results are confirmed and strengthened by an analysis of official divorce records. Five years post-randomization, the control group had a 20 percent divorce rate, while the experimental families had an 11 percent divorce rate.

Finally, we note that these results are consistent with the MFIP-IO results that also find an effect on marriage. Since the studies of mandatory work-related activities alone find no effect on marriage, it seems reasonable to interpret the main MFIP results (including mandatory work-related activities) as a financial incentive effect, lending more support to the inference that financial work incentives increase marriage.

[59]Although we rely throughout our analysis on MFIP results for the urban sample only, results for the pooled urban and rural MFIP single-parent recipient sample show a somewhat larger positive marriage impact (10.6 percent currently married and living with the spouse in the treatment group versus 7.0 percent in the control group), a difference of 3.6 percentage points that is statistically significant at the 5 percent level (Miller et al., 2000). There is no statistically significant impact on the same measure for the pooled sample of urban and rural single-parent applicants (and the impact estimate is actually negative). In contrast to the pooled result, the lack of significance for the subset of urban single-parent recipients reported in Table 7.1 may be due to the small sample sizes (less than 400 in each of the treatment and control groups).

Programs That Focus on TANF-Like Bundles of Reforms

Finally, four of the programs that involve TANF-like bundles of reforms assess marriage and, in two cases each, changes in marital status (EMPOWER and Jobs First) and household size (FTP and Jobs First). The follow-up periods range from as little as four months (ABC, for those entering latest) to four years (FTP). EMPOWER, FTP, and Jobs First have negative but insignificant impacts on the likelihood of being married (or married and living with their spouse). Changes in marital status in EMPOWER and Jobs First and are also insignificant. The impact on household size is zero for FTP but small, positive, and significant for Jobs First, where household size increases by 0.2 persons. Disaggregation by adults and children (not shown) shows the increase is evenly split between the two types of household members.

ABC is the only study to show a statistically significant ($p < 0.10$) increase in marriage, and this occurs even though the follow-up period averages 12 months, with a range from 4 to 19 months. Analyses for subgroups show a significant positive impact on marriage for women under 25, those who are capable of having additional children, those never married, and those with less than 12 years of schooling. The differences between age and education groups are also statistically significant. There are no significant impacts for subgroups defined by length of prior welfare receipt. A broader measure that includes living with a spouse or the respondent expects to marry shows no significant impact overall. For this broader measure of marriage and marriage expectations, the only significant difference for subgroups is by education, again with the least educated having the largest impact.

7.2.2. Fertility

Like marriage, with a few exceptions, the results for births since random assignment summarized in Table 7.2 are small and insignificant. For this outcome, results are available only for programs that focus on mandatory work-related activities, on family caps, and on TANF-like bundles of reforms.

Programs That Focus on Mandatory Work-Related Activities

Of the 12 studies that focus on mandatory work-related activities, only Columbus Traditional has a borderline statistically significant negative impact on births in the two years following random assignment. Against the prediction of the theory, the signs of the impacts in the other sites are more often positive than negative. For seven of the sites (but not Columbus), there are also five-year follow-up results. For none of these sites can we reject the hypothesis of no effect. Again, the signs are mixed, with more positive point estimates than negative point estimates.

Programs That Focus on Family Caps

Two experiments, FDP and AWWDP, evaluated a family cap. Both studies rely on administrative data from the welfare system to identify births after random assignment. They therefore analyze only the effect of the experiment on births while on welfare. This is a different concept from that analyzed by the other studies of fertility effects.

Table 7.2—Estimated Impact of Welfare Reform on Fertility: Random Assignment Studies

Name	Cases served	Data	Fertility: Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives						
B. Programs that focus on financial work incentives tied to hours of work						
C. Programs that focus on mandatory work-related activities						
LA Jobs-1st GAIN	Single-parent recipients and applicants	S	R had child since RA as of 2-yr FU (%)	9.3	-0.2	-2.2%
Atlanta LFA	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	6.4	0.5	7.8%
		S	R had new baby present in HH as of 5-yr FU (%)	12.4	-0.8	-6.5%
Grand Rapids LFA	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	11.1	1.9	17.1%
		S	R had new baby present in HH as of 5-yr FU (%)	21.7	0.9	4.1%
Riverside LFA	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	12.7	-0.2	-1.6%
		S	R had new baby present in HH as of 5-yr FU (%)	22.1	3.4	15.4%
Portland	Recipients and applicants; no cases with substantial barriers	S	R had child since RA as of 2-yr FU (%)	10.7	-1.2	-11.2%
		S	R had new baby present in HH as of 5-yr FU (%)	22.7	-5.3	-23.3%
Atlanta HCD	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	6.4	1.4	21.9%
		S	R had new baby present in HH as of 5-yr FU (%)	12.4	0.1	0.8%
Grand Rapids HCD	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	11.1	2.4	21.6%
		S	R had new baby present in HH as of 5-yr FU (%)	21.7	0.5	2.3%
Riverside HCD	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	13.6	0.7	5.1%
		S	R had new baby present in HH as of 5-yr FU (%)	23.1	1.0	4.3%
Columbus Integrated	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	7.9	1.7	21.5%
Columbus Traditional	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	7.9	-3.2 *	-40.5%
Detroit	Recipients and applicants	S	R had child since RA as of 2-yr FU (%)	12.3	-2.6	-21.1%
Oklahoma City	Applicants	S	R had child since RA as of 2-yr FU (%)	14.9	0.7	4.7%

Table 7.2—Continued

Name	Cases served	Data	Fertility: Measure	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities						
E. Programs that focus on other individual reforms						
AWWDP	Recipients and applicants	A	Avg. number of births since RA as of 5-yr FU	0.16	0.0	-12.5%
FDP	Recipients	A	Regression-projected likelihood of R having a child since RA as of 17-Q FU (%)	34.9	-3.2 **	-9.2%
	Applicants	A	Regression-projected likelihood of R having a child since RA as of 17-Q FU (%)	30.3	-3.7 **	-12.2%
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)						
EMPOWER (a)	Recipients	S	Case head had child since RA as of 3-yr FU (%)	18.0	-1.0	-5.6%
		S	Case head conceived a child since RA as of 3-yr FU (%)	11.3	0.1	0.9%
		S	Unwed minor had child since RA as of 3-yr FU (%)	4.0	-2.4 **	-60.0%
		S	Unwed minor conceived a child since RA as of 3-yr FU (%)	2.9	-1.8 *	-62.1%
ABC	Single parent recipients and applicants	S	R conceived a child since RA as of 4-19-mo FU (%)	13.8	-0.3	-2.2%
FTP	Recipients and applicants	S	R had child since RA as of 4-yr FU (%)	22.7	1.2	5.3%
Jobs First	Recipients and applicants	S	R had child since RA as of 18-mo FU (%)	24.3	-0.2	-0.8%
		S	R had child since RA as of 3-yr FU (%)	20.7	0.1	0.5%

NOTES:

For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; R=respondent; RA=random assignment.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) Phoenix site only, cash assistance.

Like the results for other outcomes (e.g., welfare use and earnings), the evaluation of the AWWDP in Arkansas finds no effect on fertility. In addition, there was no statistically significant effect on participation in family planning or use of birth control. However, several methodological issues suggest caution in interpreting these findings. First, the sample size used for the analysis of fertility is very small: the researchers use a 5 percent random subsample of the population available for study. Thus, for their analysis of births, the samples sizes in the treatment and control groups are 86 and 88, respectively. Such small samples make it difficult to detect even moderate sized effects.

Second, the AWWDP evaluators report that “a substantial portion of workers explained the cap on benefits to clients in both the experimental and control groups” (Turturro, Benda, and Turney, 1997, p. 2). It is therefore not surprising that the family cap appears to have been only poorly understood. In a small survey of study participants (N = 102), about half did not know how their benefits would change with an additional child (45.7 in the experimental group versus 51.8 percent in the control group). Inasmuch as members of the control group believed that they were subject to the family cap, the experiment will underestimate its true effect.

Results for New Jersey are quite different. In New Jersey, the family cap was instituted as part of FDP, a wide-ranging waiver package including enhanced welfare-to-work services, financial work incentives, transitional Medicaid, and elimination of some marriage penalties. Comparisons of the experimental and control groups imply that for recipients the entire package of reforms led to a statistically significant decline in fertility of 9 percent, but there was no effect on abortion (not shown). For applicants, FDP resulted in a statistically significant 12 percent decline in fertility. In addition, abortions increased 14 percent, but this effect appears to be concentrated in the early months of the experiment, with convergence by the end of the analysis period (1996, four years later).

The experimental analysis of FDP also found effects on family planning. Survey questions indicate that, compared to those in the control group, those in the treatment group were 4 percentage points more likely to use family planning in the last year (30.9 percent versus 26.6 percent). Regression analyses of sterilization and family planning visits from Medicaid files are also consistent with a moderate to large effect on fertility practices, and the timing of these effects is also plausible.

Like AWWDP, however, methodological issues suggest concern in interpreting the findings. First, randomization does not appear to have been performed properly. More than one-quarter of case workers admitted to evaluators that they used discretion when making assignments to the treatment and control groups (Camasso et al., 1996). Second, like the Arkansas demonstration, the FDP client survey suggests that understanding of the program was very poor.[60] Combining the groups that reported either that their cash benefits would not increase or that none of their benefits (including food stamps and Medicaid) would increase, survey results suggest that only 3.5 percent more of the experimental group believed that the cash benefit would not increase with the birth of a new child.

[60]In addition, a small number of cases (21, well under 1 percent of the control cases) were informed that they were subject to the family cap when they were not.

If understanding of the program was truly this weak, then the large fertility and abortion effects that were found are surprising. Poor recipient understanding of the family cap would be expected to bias the effects of the program downward relative to more complete understanding. These results would then imply even larger effects when the program was understood. Another interpretation is possible. FDP was broader than the family cap. It also involved an enhanced earnings disregard, enhanced case management, and relaxation of the marriage penalty. Thus, even if recipients did not understand the family cap, fertility effects might have resulted from these other program components.

Nevertheless, less than perfect understanding by the treatment and control groups of the policies that applied to them would still lead to a downward bias in the estimated program impact. Partially to address this concern, the New Jersey evaluation also conducted a before-and-after econometric analysis. In particular, again using the administrative data, Camasso et al. (1999) estimated a standard regression model for fertility with controls for demographic characteristics (e.g., age, marital status, education, and number of children), earnings, history of AFDC use, the unemployment rate, the FDP participation rate, county dummies, and a linear time trend. The effect of FDP was estimated as the deviation from the time trend implied by this regression model. Again, large negative effects of FDP on fertility were detected, as were moderate positive effects on abortions. Note, however, that by our standards for judging observational studies, this is a weak design. If fertility began to decline (or the decline accelerated) nationally for welfare recipients (as Figure 7.1 suggests), this approach would have attributed that decline to FDP. A stronger design would have included some form of control for trends in other states (which did not implement a family cap); however, as a New Jersey-specific evaluation, the evaluators did not have easy access to such data.

Programs That Focus on TANF-Like Bundles of Reforms

The four programs that focus on TANF-like bundles of reforms all find small and insignificant impacts on births or conceptions for the recipient for a follow-up interval ranging from four months (ABC) to four years (FTP). Three of the six impact estimates are negative. ABC also included an analysis of fertility desires (results not shown) by asking whether the respondent wants to have more children. Overall the impact estimate is negative but insignificant. Subgroup analyses for ABC showed a significant reduction in conceptions only for those on welfare between one and two years in the past five years. There was also a significant negative impact on fertility desires for this subgroup. In addition, the impact on fertility desires was significantly negative for women age 25 and above and for those ever married.

EMPOWER also measures births and conceptions for unwed minors and finds statistically significant negative impacts for both measures. As seen in Table 3.5, EMPOWER's reforms included a family cap, as well as a minor residency requirement and a provision removing the exemption from JOBS participation for teens under age 16 (those age 13 and above must now participate). Because these three reforms are bundled with the program's other reforms; it is not possible to ascribe the reduction in unwed teen fertility to these specific policies. It is also worth noting that the control group in EMPOWER became subject to the treatment group provisions two years into the three-year follow-up period. Thus, some of the measured impact of the EMPOWER reforms on adult and teen fertility may have been diluted by the control-group crossover.

7.3. ECONOMETRIC ANALYSES OF THE EFFECTS OF WELFARE REFORM ON FAMILY STRUCTURE

The effects of waivers and TANF on family structure have also been explored using econometric methods. As noted earlier, since welfare reform's effect on family structure may be expected to operate primarily through entry effects that are not captured by random assignment studies, econometric approaches are likely to be more appropriate.

Table 7.3 summarizes the results of the two econometric studies that consider marriage and living arrangements using CPS data, and all but one of the studies considering fertility and abortion. Table 7.4 provides additional results from another study of a fertility outcome—the nonmarital fertility ratio. We begin by discussing results for marriage and living arrangements, followed by those for fertility.

7.3.1. Marriage and Living Arrangements

Schoeni and Blank (2000) consider the propensity to be married and the propensity to be a female head of household using the March CPS. (See the discussion of their analyses of other outcomes in earlier chapters.) As seen in Section A of Table 7.3, their DoD specification suggests that for high school dropouts, any implemented waiver increases marriage (by about 2 percentage points) and depresses female headship (also by about 2 percentage points). For those with exactly 12 years of schooling, waivers have a significant negative effect on marriage (not what would be expected) and a positive (but not statistically significant) effect on female headship. For those with more than 12 years of schooling, waivers again have a significant positive effect on marriage, but not female headship.

They also estimate the effect of TANF using interstate variation in the date of implementation of each state's TANF program. These models show almost no significant effect of TANF on marriage or female headship. The only exception is a 2 percentage-point increase in female headship for those with more than a high school diploma (not shown in the tables). There is also a small (less than 1 percent), but statistically insignificant, increase in marriage, although in alternative specifications there is a small, but statistically significant (at the 10 percent level), negative effect on marriage (the opposite of the expected sign).

Bitler, Gelbach, and Hoynes (2001) use the March CPS to explore the effect of welfare reform on living arrangements (Panel B of Table 7.3). In addition to the possibility that welfare reform might increase marriage, they hypothesize that welfare reform might also increase "doubling up" (i.e., moving in with other relatives, such as an aunt or grandmother of the children). Consistent with their hypothesis, they find evidence that welfare waivers increase several measures of doubling up: The number of persons in the household (at the 5 percent level), the number of children (only at 10 percent level), the number of families, the number of females, the number of males (only at the 10 percent level), and the number of "families" with kids ($p < 0.05$).

Related research (not shown in Table 7.3) also contributes to our understanding of the potential impact of welfare reform on marriage. Rosenbaum (2000) takes a more structural approach to the effect of government policies—including waivers, but primarily the EITC—on marriage (similar to that of Meyer and Rosenbaum, 2000, on employment). Using both the CPS and the

SIPP, he finds strong effects of financial work incentives on marriage. A $1,000 marriage penalty from a combination of welfare and taxes (including the EITC) decreases the fraction of women married by about 5 percentage points, and the effects appear to be concentrated in entries into marriage, not exits from marriage. He also considers three explicit reform measures—(1) any waiver application, (2) a broadly defined time limit combining what we refer to as “time limits” (leading to a decrease in or termination of the welfare benefit) and what we refer to as “work triggers” (leading to a requirement for work or participation in a welfare-to-work activity), and (3) a binding time limit (i.e., benefits have been cut because at least some recipients have reached what we refer to as a “time limit”). Only one of the waiver proxies significantly affects marriage rates, and only in the CPS specification (not the SIPP specification). Moreover, the estimated effect is contrary to expectation: A time limit is estimated to lower marriage rates. Similarly, in the SIPP models of entry to and exit from welfare, a time limit counterintuitively lowers the probability of entry into marriage. Rosenbaum notes that these results appear to be quite sensitive to the details of the specification, suggesting caution in using these results for policy.

Ellwood (2000) uses CPS data from 1986 to 1998 (mostly the waiver period) to explore the effect of public policy on marriage. He parameterizes states by the “aggressive[ness] of welfare reform policies” but finds no effect of either the EITC or welfare policy on marital status.

Finally, two studies have explored the effect of child support enforcement on marital status. The theoretically expected effect is ambiguous. Better child support enforcement makes divorce more attractive for mothers, but less attractive for fathers. The interactions with the welfare system are complicated. Under the baseline AFDC rules, mothers kept only the first $50 of child support paid and the possibility of under-the-table payments further complicates the analysis. In net, Nixon (1997) argues that the deterrent effect on divorce is likely to be larger.

Nixon (1997) estimates the effect using cross-sectional variation from two March–April CPS matches. She finds a negative and robust, but small, effect of child support enforcement. For the largest of her proxies, a 1 percent increase in child support enforcement only decreases the probability of divorce by 0.16 percent (or, assuming linearity, a 10 percent increase would decrease the probability of divorce by 1.6 percent). Even given the baseline divorce rate of 12 percent, this is not a large effect. Note, however, that with only two years of CPS data, Nixon does not have sufficient variation to estimate a full DoD specification. As discussed in Chapter 3, her analysis is thus potentially biased by unmeasured state-specific characteristics that are correlated with the policies implemented.

Heim (2001) explores the effect of child support enforcement on the annual state-specific divorce rate using Vital Statistics data for 1989 to 1995. These data allow him to include a full DoD specification (i.e., fixed effects for state and year). He specifies five proxies for child support enforcement, child support collections, paternity establishment, efforts to find fathers, and average child support orders. Like Nixon, when no state fixed effects are included, child support enforcement is found to decrease divorce. However, once state fixed effects are included, there is no statistically significant effect of child support enforcement on divorce.

Table 7.3—Estimated Impact of Welfare Reform on Marital Status, Headship, Living Arrangements, Fertility, and Abortion: Econometric Studies

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr. and Geogr.	Fixed Effects	Policy
A. Marriage and Headship													
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	Percent married	Level	Any waiver	0.0229 (0.0073)	5.4	U, U-1, EG, each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	-0.0144 (0.0060)	-2.2				
		women 16-54, educ>12					Any waiver	0.0075 (0.0049)	1.3				
		women 16-54, educ<12					TANF	-0.0004 (0.0171)	-0.1				
		women 16-54, educ=12					TANF	-0.0161 (0.0150)	-2.5				
		women 16-54, educ>12					TANF	0.0034 (0.0114)	0.6				
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	Percent head of household	Level	Any waiver	-0.0171 (0.0070)	-8.2	U, U-1, EG, each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	0.0052 (0.0058)	2.3				
		women 16-54, educ>12					Any waiver	-0.0014 (0.0047)	-0.5				
		women 16-54, educ<12					TANF	-0.0133 (0.0165)	-6.4				
		women 16-54, educ=12					TANF	-0.0025 (0.0144)	-1.1				
		women 16-54, educ>12					TANF	0.0239 (0.0110)	8.5				

Table 7.3—Continued

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr. and Geogr.	Fixed Effects	Policy
B. Living Arrangements													
Bitler, Gelbach and Hoynes (2001)	CPS micro data	women 16-54	84	98	Number of persons in household	Level	Any waiver	0.055 (0.020)	1.2	U, U-1, EG	R, MSA, CC	S, Y	B
							TANF and ever had waiver	0.100 (0.037)	2.2				
							TANF and never had waiver	0.042 (0.038)	0.9				
Bitler, Gelbach and Hoynes (2001)	CPS micro data	women 16-54	84	98	Number of children in household	Level	Any waiver	0.030 (0.017)	1.3	U, U-1, EG	R, MSA, CC	S, Y	B
							TANF and ever had waiver	0.065 (0.028)	2.8				
							TANF and never had waiver	0.025 (0.027)	1.1				
Bitler, Gelbach and Hoynes (2001)	CPS micro data	women 16-54	84	98	Number of families in household	Level	Any waiver	0.018 (0.007)	1.6	U, U-1, EG	R, MSA, CC	S, Y	B
							TANF and ever had waiver	0.023 (0.011)	2.0				
							TANF and never had waiver	0.026 (0.011)	2.2				

Table 7.3—Continued

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff.	(s.e.)	% effect	Economy	Demogr. and Geogr.	Fixed Effects	Policy
C. Fertility														
Levine (2001)	State-level vital statistics	women 15-44	85	96	Birth rate	Log	Any waiver	-0.030	(0.007)	-3.0	U	R, A, E, MS	S, Y, ST	B, PI, MD
							Family cap	0.050	(0.010)	5.0				
Kearny (2001)	State-level vital statistics	women 15-34	89	98	Number of births	Log	Any waiver	0.003	(0.003)	0.3	U	A	S, Y, ST	B, WE
							TANF	0.007	(0.005)	0.7				
							Family cap	0.001	(0.004)	0.1				
							Time limit	-0.002	(0.003)	-0.2				
Kaushal and Kaestner (2001)	CPS micro data	unmarried women 18-44 with high school or less	95	99	Had a birth in last year	quasi-DoDoD w/married women high school or less	Low Intensity Reforms (waiver or TANF)	-0.010	(0.005)	-24.4	U, U-1, U-2	A, R, N<6, N>=6, UI	S, Y	
							Medium Intensity Reforms (waiver or TANF)	-0.001	(0.005)	-2.4				
							High Intensity Reforms (waiver or TANF)	0.020	(0.009)	48.8				
							Family cap	0.009	(0.006)	22.0				
							Time limit	0.003	(0.006)	7.3				

Table 7.3—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr. and Geogr.	Other controls: Fixed Effects	Other controls: Policy
Kaushal and Kaestner (2001)	CPS micro data	unmarried mother with high school or less	95	99	Had a birth in last year	quasi-DoDoD w/married mothers high school or less	Low Intensity Reforms (waiver or TANF)	0.001 (0.007)	1.8	U, U-1, U-2	A, R, N<6, N>=6, UI	S, Y	
							Medium Intensity Reforms (waiver or TANF)	0.009 (0.007)	15.8				
							High Intensity Reforms (waiver or TANF)	0.027 (0.013)	47.4				
							Family cap	0.011 (0.008)	19.3				
							Time limit	0.013 (0.008)	22.8				
D. Abortion													
Levine (2001)	AGI survey	women 15-44	85	96	Abortion rate	Log	Any waiver	-0.022 (0.031)	-2.2	U	R, A, E, MS	S, Y, ST	B, PI, MD
							Family cap	0.086 (0.063)	8.6				

NOTES: Abbreviations: s.e. = standard error; U=unemployment rate; U-1=lagged unemployment rate; U-2=twice lagged unemployment rate; EG=employment growth; A=age, E=education, R=race, MS=marital status; N<6=Number of children less than 6; N>=6=Number of children 6 or older; *UI* -Unearned Income; MSA=Metropolitan Statistical Area (urban), CC=Central city; B=maximum welfare benefit; PI=Parental involvement in abortion to minors; MD=Mandatory delay in abortion; WE=Work exemption; S=state; Y=year; ST=state trends.

Table 7.4—Additional Estimates of Impact of Welfare Reform on Fertility: Econometric Studies

			Welfare Waiver Policy Variable (Implemented)								
			Any waiver	Family Cap	Time Limit	Work Requirement	Expanded Income Disregard and Asset Limit	AFDC-UP Expansion	Strengthen Child Support	Minor Parent Provision	School Attendance and Performance Requirement
Study	Data (Years)	Outcome	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect	Marginal effect (t-statistic) % effect
Horvath and Peters (1999)	State-level vital statistics (1984-1996)	Non-marital birth ratio for women 15-19	-0.011 (1.99) -1.6%	-0.052 (3.55) -7.6%	-0.092 (3.70) -13.5%	0.008 (0.59) 1.2%	-0.001 (0.03) -0.1%	-0.075 (4.83) -11.0%	0.035 (2.16) 5.1%	0.143 (6.06) 21.0%	0.018 (1.88) 2.6%
		Non-marital birth ratio for white women 15-19	-0.014 (2.11) -2.4%	-0.058 (3.28) -10.0%	-0.102 (3.40) -17.6%	0.001 (0.59) 0.2%	0.003 (0.18) 0.5%	-0.078 (4.15) -13.4%	0.041 (2.08) 7.1%	0.167 (5.87) 28.8%	0.009 (0.74) 1.6%
		Non-marital birth ratio for black women 15-19	-0.012 (2.24) -1.4%	-0.024 (1.63) -2.8%	-0.104 (4.08) -12.0%	0.003 (0.23) 0.3%	-0.002 (0.12) -0.2%	-0.047 (2.98) -5.4%	0.008 (0.51) 0.9%	0.115 (4.88) 13.2%	0.008 (0.83) 0.9%
		Non-marital birth ratio for women 20-49	-0.008 (3.43) -3.8%	-0.032 (5.55) -15.2%	0.0001 (0.01) 0.0%	0.004 (0.63) 1.9%	0.011 (2.03) 5.2%	-0.019 (2.73) -9.0%	0.022 (4.19) 10.5%		
		Non-marital birth ratio for white women 20-49	-0.006 (3.10) -4.3%	-0.024 (4.60) -17.1%	-0.010 (1.06) -7.1%	0.001 (1.64) 0.7%	0.004 (0.86) 2.9%	-0.006 (0.97) -4.3%	0.022 (4.68) 15.7%		
		Non-marital birth ratio for black women 20-49	-0.016 (3.72) -3.1%	-0.038 (3.38) -7.5%	-0.045 (2.25) -8.8%	-0.003 (0.31) -0.6%	0.016 (1.40) 3.1%	-0.023 (1.70) -4.5%	0.040 (3.78) 7.8%		

NOTE: Dependent variable is log of odds ratio transformations of race and age group specific ratios of non-marital births to total births for each state. Coefficients transformed to reported marginal effect. Absolute value of t-statistic in parentheses. Regressions are weighted by state population due to heteroskedasticity, and lagged nine months to account for natural lag associated with childbearing.

All models include state and year fixed effects and the following controls: state poverty rate; race-specific female unemployment rate; race-specific teenage unemployment rate; race and gender-specific wages; number of AIDS cases weighted by state population; ratio of whites to blacks in state population; number of abortion providers per 1000 women of childbearing age; fundamentalist adherents as proportion of state population; high school completion rate among 18-24 yr. olds not currently in high school; proportion of population living in urban area; race and age-specific marriage market opportunities; percent of children in single parent homes lagged 24 years (post-teen regressions only); race-specific ratio of teen births lagged 17 years (teen regressions only); maximum welfare plus food stamp benefit for a family of three; sex education required in public schools; and parental consent required for teen abortion (teen regressions only).

7.3.2. Fertility

Four econometric studies explore the effect of welfare reform on fertility (see Panel C of Table 7.3 and Table 7.4). Kearney (2001) estimates the effect of the family cap using DoD methods (with state-specific time trends) and birth certificate data. She finds no systematic effect of the family cap (see Panel C of Table 7.3). The point estimate is very small as is the standard error, so that the basic analysis can reject an effect of even 0.5 percent. This conclusion is robust to the inclusion of state-specific time trends, alternative coding of the family cap, alternative timing of the effect on fertility, using the birth rate rather than the number of births, and the inclusion of lead effects (which as expected are zero). Analyses by parity, race-ethnicity, and age (not shown) also show no consistent evidence for an effect of the family cap. In specifications that disaggregate by race, education, marital status, and parity, the point estimates for additional births to high school dropouts age 20 to 34 are positive and significant for blacks ($p < 0.10$) and whites ($p < 0.01$) for both marital and nonmarital births but not significant for other groups (high school graduates and first births). Estimates for teenagers are positive and significant ($p<0.01$) for additional births to unmarried blacks, but insignificant for the other groups (first births to unmarried blacks, married blacks, and whites).

While not the primary focus of her analysis, Kearney includes dummy variables for any waiver and time limits in some of her models. Like the results for the family cap, her results provide no evidence of an effect of either any waiver or a time limit on fertility. The point estimates are small and not significantly different from zero.

Levine (2001) estimates DoD models of births and birth rates using Vital Statistics data. His basic models including state-specific time trends suggest that welfare reform as a bundle decreases the birth rate by about 3 percent (and is highly significant). However, if this result were causal, we would expect it to be larger for the less educated, who are more likely to receive welfare. Levine, however, finds the effect is constant or grows with education. Thus, this pattern across subgroups suggests caution in interpreting the estimated negative effect as causal. Levine also considers the effect of the family cap. In his models with state-specific time trends, he finds that, contrary to the theory, the family cap raises fertility (and the effect is clearly statistically significant), and the effect is consistent across the age, education, and parity subgroups. Finally, Levine uses the same methods and Alan Guttmacher Institute data on abortions (see Panel D of Table 7.3). He finds no evidence of an effect of any waiver or a time limit on the abortion rate. Furthermore, these results are consistent with disaggregation by age. The data do not allow disaggregation by education or parity.

Kaushal and Kaestner (2001) use CPS data to estimate the effect of time limits, the family cap, and welfare reform as a bundle (characterized as “low intensity,” “medium intensity,” and “high intensity”). Their estimates can be interpreted as a restricted difference-of-difference-of-differences (DoDoD) specification; they include year and state fixed effects, and they interact the policy with a dummy variable for the population assumed to be most affected by welfare reform.[61] Their first affected group is appropriate for considering effects of fertility potentially leading to welfare entry. It consists of unmarried women with 12 or fewer years of education,

---

[61] With three levels, the full DoDoD specification would include not merely state and year fixed effects, but a fixed effect for every state-year combination. They do estimate that model.

with two alternative corresponding unaffected groups—married women with 12 or fewer years of education and unmarried women with an associate degree. Their second affected group is appropriate for considering the effects of a family cap on subsequent fertility. It considers unmarried women with at least one child and 12 or fewer years of schooling, with two alternative corresponding unaffected groups—married women with children and 12 or fewer years of schooling and unmarried women with children and an associate degree. Their findings are not consistent with an effect of welfare reform on fertility. Across each of the individual policies they consider—time limits and family caps—and across each of the four comparison groups they consider, they find no statistically significant effect.

Kaushal and Kaestner do find effects of reform bundles, but the sign patterns are difficult to interpret. The theory suggests that reform should lower fertility. However, the only statistically significant negative effect is for low-intensity reforms and then only for the first comparison group. If there were truly an effect of reform, we would expect to find larger (in absolute value) effects with more-intensive reforms. Kaushal and Kaestner, however, find no statistically significant effect of medium-intensity reforms. Furthermore, against the theory, for their first and second comparison groups high-intensity reforms increase fertility ($p < 0.05$ and $p < 0.10$, respectively).

Finally, Horvath and Peters (1999) explore the effect of waivers on a different outcome (see Table 7.4). While the previously discussed studies analyze fertility (number of births or the birth rate), Horvath and Peters analyze the nonmarital fertility ratio, defined as the fraction of births that are to unmarried women. They compute the marital fertility ratio from Vital Statistics data (i.e., from birth certificates), and their models include state and year fixed effects. As seen in Table 7.4, they find that any implemented waiver decreases the nonmarital fertility ratio in all subgroups, teenagers and nonteenagers, whites and blacks. In almost all specifications, the effect is statistically significant at the 5 percent level or better.

They also present results for the impact of specific waiver policies on the nonmarital birth ratio. Family caps are estimated to lower the nonmarital fertility rate among teenagers by 6 percentage points for whites and 2 percentage points for blacks; for adults age 20 and above, the corresponding effects for whites and blacks are 2 and 4 percentage points. With the exception of the black teenager effect (significant at the 10 percent level), these effects are statistically significant at the 5 percent level.

These results for the effects of the family cap diverge from those found in the other observational studies. One possible explanation is that Horvath and Peters analyze the nonmarital fertility ratio, while the other studies analyze births or the birth rate. Kearney (personal communication 3/12/02) reports that when she applies her basic models to the nonmarital fertility rate, she does not find an effect of the family cap. Thus, it seems unlikely that the different outcome explains the divergence. Kearney also notes that Horvath and Peters do not appear to have adjusted for changes in the coding of marital status during this period, and that their coding of waivers differs substantially from those used in other studies.

Horvath and Peters also find effects for other specific waivers. The magnitudes are often even larger than these effects for the family cap, but they sometimes have the opposite of the expected sign. Consistent with the expectation, time limit waivers decrease the nonmarital fertility ratio by 10 percentage points for teenagers, and 5 percentage points for black adults. AFDC-UP expansion waivers also decrease the nonmarital fertility ratio, especially for teenagers and especially for white teens (effects as high as 8 percent). Waivers to strengthen child support raise the nonmarital fertility ratio (by 2 to 4 percentage points for adults and teens, respectively). This would be the expected sign if women were now more confident of support from the father; but not if fathers were now more cautious about conceiving a child. Counter to intuition, however, minor parent provisions are estimated to raise the nonmarital fertility ratio by over 10 percentage points, with effects that are statistically significant at the 5 percent level or better. Finally, work requirement waivers, benefit structure waivers, and school attendance and performance requirement waivers appear to have no effect (almost all the estimated effects are statistically insignificant and small in magnitude).

7.4. EVALUATING THE EFFECTS OF WELFARE REFORM ON FAMILY STRUCTURE

In this section we synthesize the findings from experimental and econometric studies that aim to measure the impact of welfare reform on marriage and fertility, the two outcomes that are the focus of the bulk of the studies that consider family structure impacts. We first discuss impacts for specific reform policies, and then for welfare reform as a bundle.

7.4.1. Effects of Specific Reforms

The demonstration studies provide the strongest basis for assessing the impact of specific reforms on the family structure decisions of current recipients (but not on potential entrants); yet, the reform policies that can be evaluated are somewhat limited. Strong financial work incentives alone and, when tied to hours worked, or in combination with mandatory work-related activities, have only been evaluated in terms of their impact on marriage. None of the demonstrations that combine weaker incentives with work requirements consider either marriage or fertility. While programs with TANF-like bundles of reforms have evaluated both marriage and fertility, they do not allow us to draw solid inferences about the marginal contribution of the time-limit feature. As might be expected, family caps have only been evaluated in terms of their impact on fertility, but as the earlier discussion reveals, the studies with this focus have a number of potential flaws. It is striking that we have the most evidence regarding the impact of work requirements on marriage and fertility since these policies might be expected to have the weakest impact on these outcomes, and the evidence indeed bears this out.

Marriage

Based on the results presented in Section 7.2, there does not appear to be any effect of mandatory work-related activities on marriage. The fact that nearly all the impact estimates are statistically insignificant and almost evenly divided in their sign suggests that work activity requirements alone have no effect on marriage rates.

The findings from the studies that evaluate the impact of financial work incentives present a number of puzzles. The findings for MFIP provide some evidence that financial work incentives can raise marriage, but these findings are not consistent for recipients and applicants, and there are differences when the program is limited to the financial work incentives component. In interpreting the MFIP findings, it is important to note that MFIP did more than simply enhance the financial work incentives of the welfare program. It also broadened eligibility for AFDC-UP and changed the treatment of stepparent earnings.

The MFIP-IO results suggest that strong financial work incentives alone may raise marriage rates, at least for recipients. When combined with mandatory work requirements, MFIP still produces positive impacts on marriage, for both recipients and applicants, but they are no longer statistically significant. In addition, the two-parent sample in MFIP demonstrated a large and significant impact on the likelihood of staying married. For the MFIP two-parent sample, almost all these cases had a married spouse present at the time of randomization. The control group means suggest that three years later, less than half are still married. Thus, there is considerable potential for improvement simply by maintaining the current marital status. In contrast, while there is considerable scope for improving marriage rates among the one-parent cases, such an improvement would require a change in marital status (rather than simply maintaining the previous status).

Another puzzle is associated with the findings for SSP. In that case, there is no significant impact on marriage for the pooled sample, but significant and opposing effects for the two study areas, British Columbia and New Brunswick. (Unlike MFIP, there are no two-parent results for SSP.)

As a whole, these results suggest the possibility that financial work incentives alone or in combination with mandatory work-related activities may both promote marriage and discourage divorce. There is, however, evidence in the opposite direction from British Columbia. Given the prominent role of marriage in PRWORA, additional random assignment evaluation of the effect of financial work incentives on marriage seems warranted.

Fertility

As with marriage, it appears that there is no impact of mandatory work-related activities on fertility behavior. The individual point estimates are not statistically different from zero, and they are of both signs.

In contrast to financial work incentives, work requirements, and time limits, family caps were instituted with the express goal of reducing subsequent childbearing for those already on aid. Here also, the evidence is mixed. One experimental study (AWWDP) finds no effect; another (FDP) finds a large negative effect. One observational study (Horvath and Peters, 1999) finds a decrease in the fraction of births that are nonmarital; three other studies find no effect (Kearney, 2001; Kaushal and Kaestner, 2001; Levine, 2001).

The quality of the studies that evaluate family caps is not uniform. While random assignment usually yields robust estimates of policy effects, the methodological issues surrounding the two random assignment analyses of family caps are so severe as to require that those results be strongly discounted. Among the observational studies, Horvath and Peters (1999) appear to be the outlier, and Kearney does not find an effect even when she uses the nonmarital fertility ratio. Thus, the available evidence appears to be inconsistent with an effect of the family cap on fertility.

7.4.2. Effects of Reform as a Bundle

The econometric studies and random assignment studies that evaluate TANF-like bundles of reform present a mixed picture of the overall effect of welfare reform on marriage. The econometric studies summarized in Section 7.3 generally suggest that welfare reform as a bundle increases marriage. The econometric studies provide evidence of both negative and positive effects, again typically small in magnitude. Of these studies, only ABC finds a positive impact on marriage, and it is marginally statistically significant. This is also the only study to look at subgroup differences and provides some indication that the positive impacts for marriage may be strongest for younger, less educated women who have yet to marry as of random assignment.

It is not clear why ABC's findings differ from either FTP or Jobs First with which it is most directly comparable. FTP and Jobs First share similar and sometimes stricter features with ABC (for example, the shorter time limit in Jobs First), although the sanctions in ABC may be viewed as stronger and the financial incentives weaker. In terms of other areas of program impact, ABC and Jobs First had similar impacts on earnings (see Table 5.1). As we will see next in Chapter 8, ABC had no impact on a broader measure of income. In contrast, FTP and Jobs First each have sizeable positive impacts, at least prior to time limits becoming binding (see Table 8.1). The absence of any income gains in ABC may provide part of the explanation. In fact, Fein (1999) associates the positive marriage impact for ABC with work requirements and strong sanctions that placed pressure on women to find alternative sources of income support. If FTP and Jobs First allowed women to increase income on their own, at least prior to time limits setting in, they may depress marriage rates relative to a program like ABC, which has no impact on recipient income.

The evidence on fertility is also mixed. The evidence from econometric studies with respect to reform as a bundle is not consistent, with studies suggesting no effect or a negative impact. As we have discussed in other chapters, the lack of statistically significant impacts for TANF as a bundle on family structure may be the result of limited variation in the timing of the implementation of TANF across states. When models include state and year fixed effects, there is too little variation left to precisely estimate TANF effects.

The random assignment studies that evaluate TANF-like bundles of reform also find no consistent impact on fertility, with a mixture of positive and negative insignificant impacts for the case head. The one exception is the statistically significant negative impact on births and conceptions to unwed minors found in EMPOWER. Arizona had other reforms that may explain the fertility impact for minors, namely a family cap, a minor residency requirement, and a requirement for teen JOBS participation. However, the separate impact of these provisions is not known.

7.5. CONCLUSIONS

TANF and the welfare reforms under waivers in the pre-TANF period aimed specifically to change family formation—to increase marriage, to decrease separation or divorce, and to decrease nonmarital fertility. Unfortunately, with the exception of mandatory work-related activities, the research base is comparatively weaker than for outcomes considered in earlier chapters. We, therefore, are more limited in our ability to draw firm conclusions about the impact of specific reform policies or welfare reform as a bundle on family structure. Furthermore, marriage and fertility may be two outcomes where the impact of welfare reform policies will be more pronounced over a longer horizon than what is available with most of our current research base.

In terms of marriage, the evidence from both random assignment and econometric studies is insufficient to draw any conclusions about the effect of welfare reform as a bundle. In terms of specific policy reforms, the experimental studies are quite clear that there is no effect of programs that focus on work-related activities. There is some suggestive evidence from MFIP that programs that provide generous financial work incentives, either alone or with work requirements, may increase marriage or keep existing marriages intact. However, the mixed results for the Canadian SSP suggest caution in interpreting the MFIP results. The contrast with the earlier Negative Income Tax Experiments suggests that those interested in affecting marital status through welfare policy give careful consideration to the relative attractiveness of welfare programs for one-parent and two-parent cases. Relevant program features may include who keeps the benefits if the marriage breaks up and how a new partner's earnings (perhaps not the father of the child) would affect the benefit received.

Likewise, for fertility, the evidence on whether there is an effect of welfare as a bundle on fertility is inconclusive, and the demonstration studies are quite clear that there is no effect of work-related activities programs. There is no basis for evaluating the effect of financial work incentives alone or in combination with mandated work-related activities on fertility. The available evidence on the family cap is limited and contradictory, but the best of the studies finds no effect.

Chapter Eight

INCOME AND POVERTY

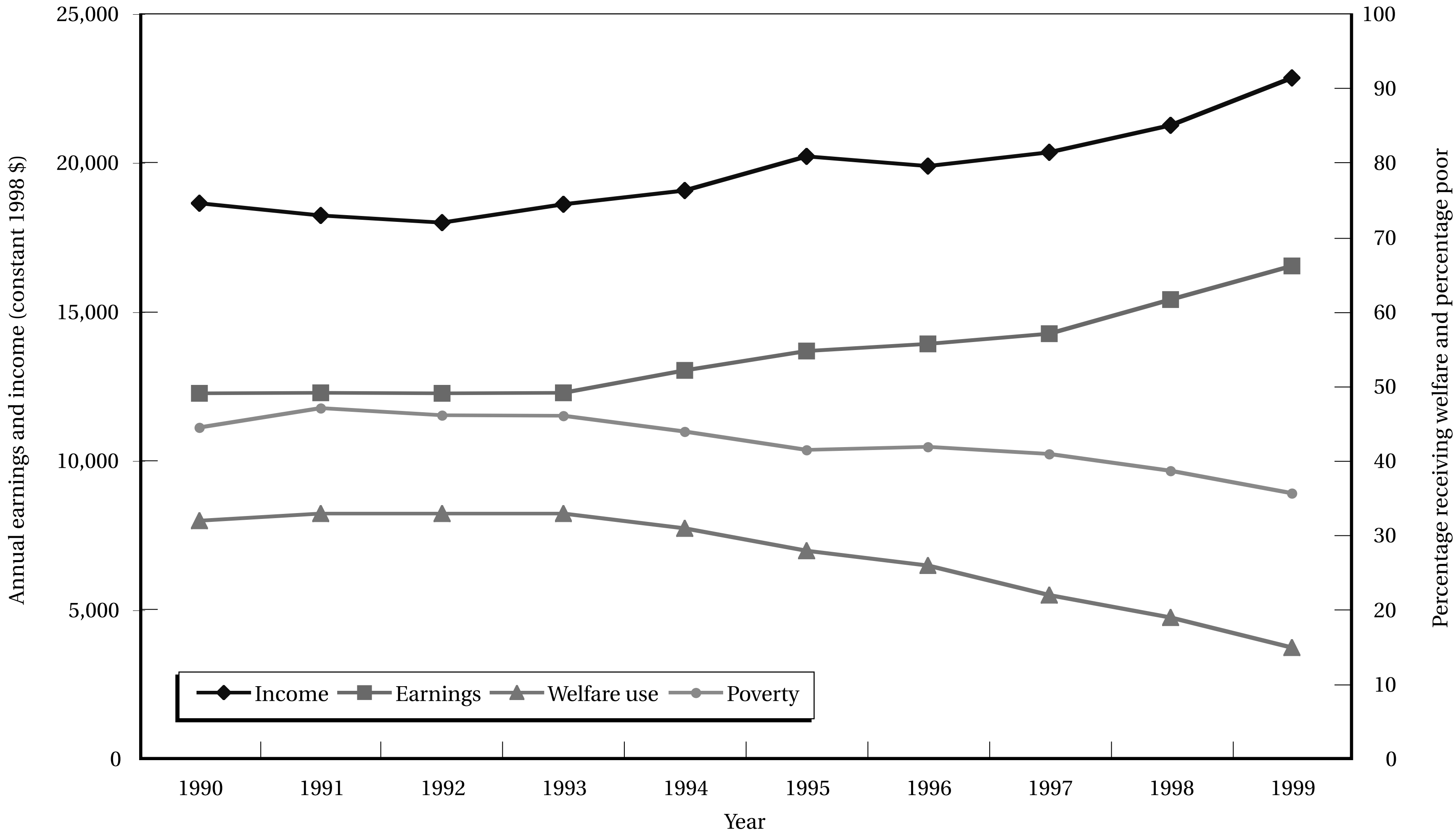
8.1. BACKGROUND

As we move from the topics covered by the prior four chapters—welfare participation, participation in other social welfare programs, labor market outcomes, and fertility and marriage—to the focus of this chapter—income and poverty—we begin to consider the broader impacts of welfare reform on families and children. (Other outcomes that also capture broader concepts of well-being, such as material well-being and child well-being, are covered in subsequent chapters.) Income is one gauge of a family's command over resources, and poverty is one widely used metric to identify the fraction of families with resources below a specified needs standard. Thus, for some, these outcomes are among the most important to consider when evaluating the effects of welfare reform.

The impact of welfare reform on income and poverty will be determined, in part, by the effects on the outcomes covered in Chapters 4, 5, and 6. The two most significant sources of income for low-income families with children are earnings and cash transfers from means-tested social welfare programs. If welfare reform raises both income sources, we would expect total family income to increase and poverty rates to fall. Conversely, income would rise and poverty would fall if welfare reform increased earnings and transfer payments. The relationship will not be exact because other sources of income (e.g., earnings from other family members; other non-means-tested public transfers, such as unemployment insurance or disability payments; or private transfers, such as child support payments) may increase or decrease as well, thereby reinforcing or undoing the contribution of changes in earnings and means-tested programs to the changes in income and poverty. In addition, if some policies work to lower dependence on welfare while others raise earnings from work, incomes may rise or fall depending on which change is greatest.

Whether poverty rates change depends on where in the income distribution any income changes take place. If income changes are small or occur only among those already above or below the poverty line, then the poverty rate would remain unchanged. Alternatively, income changes may be small on average but still lead to changes in the proportion of families classified as poor. For example, small increases in income may be associated with reductions in poverty if the income gains occur among those with incomes near the poverty threshold.

Considering only the trends in income levels and poverty rates for the population at risk of welfare utilization (i.e., female-headed households), we find that the trends since welfare reform are quite favorable. Figure 8.1 plots the levels of annual earnings and annual income



SOURCE: Welfare, income, earnings: Grogger (2001); poverty: U.S. Census Bureau (2000).

Figure 8.1—Income, Earnings, Welfare Use, and Poverty for Female-Headed Families: 1990–1999

(measured off the left y-axis) and the rates of poverty and welfare utilization (measured off the right y-axis) for female-headed families between 1990 and 1999. As discussed in earlier chapters, the welfare participation rate has been steadily declining since the mid-1990s, and earnings have been on an upward trajectory over the same period. Total family income shows the same pattern as earnings, indicating that the earnings increases have been large enough to offset the decline in welfare benefits (or that other income sources have increased).[62] The income gains have resulted in a decline in the poverty rate for female-headed families, which in 1999 stood at its lowest level—35.7 percent—since 1959. The child poverty rate has also fallen over this period (Haskins, 2001).

A more focused look at the welfare population, and those who leave welfare, suggests that this positive assessment on average may not tell the whole story. Data from the eight USDHHS-funded welfare leaver studies with survey data on household income indicate that post-exit incomes remain relatively low both soon after leaving welfare (6–8 months) and also after more time has passed (26–34 months)—in the range of $1,000 to $1,500 per month (USDHHS, 2001a).[63] Over one-half of the sample (57 to 58 percent) of leavers in two states and one county (Missouri and Washington, and Cuyahoga County, Ohio) were classified as poor after leaving welfare, in both the short time horizon (6–8 months) and longer time horizon (26–34 months), although this was lower than the rate of poverty for the stayers as calculated with data for Washington State.[64] The post-exit poverty rate was somewhat lower in Iowa: 47 percent based on cash income, and 41 percent including the value of food stamp benefits.

There is also some evidence from these leaver studies that those whose income is above the poverty line after leaving welfare would still be considered "near poor," although some may be better off than they appear once the full range of support available to the working poor are taken into account (e.g., Medicaid, food stamps, child care, and EITC) (Haskins, 2001). At the same time, there is a small- to moderate-sized group that has very low income after leaving welfare, even accounting for these forms of support which they may or may not take advantage of. As would be expected, most income (70–80 percent) derives from the earnings of the former recipient or other family members.

A comparison of an early cohort of leavers in Wisconsin (1995) with a cohort that left two years later, shows that the later cohort—which had more barriers to work—had lower earnings and higher poverty rates after leaving welfare than the earlier cohort of leavers (Cancian et al., 2000). By comparing outcomes prior to welfare exit with those after welfare exit, Cancian et al. (2000) find that nearly two-thirds of leavers in both the early and late cohorts had higher earnings in the year after welfare exit, but only one-third had higher combined income from their own work and public assistance benefits in the post-exit period. Thus, the earnings gains were not sufficient to offset the loss of welfare benefits for most leavers. (The impact on earnings of other household members is not known in this study.) For both cohorts, poverty rates based on combined recipient income were above 65 percent in the year after exit. The early cohort,

---

[62]Haskins (2001) also documents similar trends for single mothers with children.

[63]Of the USDHHS-funded leaver studies, only those conducted in five states (Arizona, Illinois, Iowa, Missouri, and Washington), the District of Columbia, and two counties (Cuyahoga, Ohio, and San Mateo, California) collected information on income. Poverty rates were constructed only for Iowa, Missouri, Washington State, and Cuyahoga County.

[64]Similar evidence is obtained from other recent reviews of leaver studies (see, e.g., Loprest, 1999; GAO, 1999c; and Cancian et al., 1999a).

which could be tracked over three years post-exit, did show a modest improvement in own income and poverty rates over time.

While the leaver studies provide a useful perspective on the experiences of those who left welfare at a particular point in time, they are not designed to assess the contribution that welfare reform made to the observed changes. To what extent are the changes observed in the national data or those captured in the leaver studies the direct result of the welfare reforms that began with waivers and continued as part of PRWORA? Like the other outcomes we have considered thus far, confounding factors such as the economy and other policy changes (e.g., the minimum wage and EITC) may have contributed to the observed changes. Thus, our goal in this chapter is to assess what we know about the effects of welfare reform on income and poverty after accounting for these and other confounding factors.

The causal impact of welfare reform on income and poverty has been evaluated using both experimental and econometric evidence. These two outcomes typically cannot be measured accurately by relying solely on administrative data, so both experimental and observational studies also rely on survey data to construct measures of income and to calculate poverty rates. Conceptually, the ideal income measure would capture all sources available to the recipient in her own name, as well as sources available through income pooled at the family or household level. Those income sources would include earnings, cash and noncash government means-tested transfers (e.g., welfare, food stamps, Supplemental Security Income, general assistance, and Medicaid), other government transfers (e.g., unemployment insurance, disability insurance, Social Security), private transfers (e.g., child support or alimony), and income from assets (e.g., interest and dividends). Income would also be measured net of taxes paid and tax credits received (e.g., EITC).[65] Family income is typically defined as all income sources for the unit of individuals who are related by blood, marriage, or adoption. A cohabiting partner of the recipient might be considered part of the family group as well, especially if income is pooled. To the extent that income pooling occurs within the household, between related or unrelated individuals—for example, in a three-generation household—the income for the entire household would be measured.

The complexities of income measurement, both conceptually and in practice, mean that such an ideal measure is rarely available. In the studies we review in this chapter, the measures of income and poverty vary between and within the experimental and econometric studies. All the econometric studies we review rely on the CPS; thus, they use a fairly comprehensive measure of family income, although they typically do not account for taxes and tax credits and noncash transfers (including food stamps).

To draw on administrative data, experimental studies often measure income just for the recipient and count only earnings, cash assistance (e.g., welfare payments and any financial work incentives such as wage subsidies from the program), and food stamps. We refer to this concept as "combined income." In some cases, taxes and tax credits such as the EITC are imputed based on the administrative data, and some evaluations have access to administrative earnings records for other family members. Some experimental studies also collect information

---

[65]Some would also argue that income should be measured net of work-related expenses such as out-of-pocket child care and transportation costs, but this is rarely done in practice. (See for example, Citro and Michael, 1995, for a proposed modification to the official poverty measure that takes this approach.) If these costs are deducted from income, a single mother moving from welfare to work may experience a decline in net income if the increase in her earnings is not large enough to offset the loss of welfare benefits and the increased out-of-pocket work-related expenses.

on a broader array of income sources directly from participants. These survey data alone, or in combination with administrative data, are used to measure a broader concept of income that might include other public and private transfers sources for the recipient, and earnings and unearned income of other family members, possibly just the spouse or partner and perhaps for all other adults in the household.

Thus, in comparing studies, it is important to keep in mind that the measurement of income (and hence poverty) may be incomplete (i.e., not all sources are measured or imputed) or suffer from underreporting of certain income sources for those that are measured.[66] Measures of family income in some studies may include cash sources only, while income measures in other studies may include the value of in-kind benefits such as food stamps. Income may be measured before taxes, so that taxes paid or tax credits received (e.g., the EITC) are not accounted for in the measure of well-being. Income data collected through survey methods may be measured using a recall period as short as one month, or as long as a year. Finally, the unit of observation may vary, from the narrowest perspective of the recipient's income, to the broadest perspective of income for the family or household.

In those studies that calculate a poverty rate, the standard approach is to compare the measure of income to a needs standard that is usually specific to the family type (size and composition in terms of the number of adults and children). Almost all studies use the poverty cutoffs defined by the Census Bureau to calculate the official poverty rate, even though the income concept they use may differ from the official measure used by the Census Bureau (family pretax cash income). In some cases, the measure of income is more comprehensive than the official measure (e.g., the EITC and value of food stamps are included). In other cases, it is less comprehensive (e.g., income from assets, private transfers, and non-welfare-related public transfers are excluded). The income measure may also be based on a different unit of analysis, for example, individual income only, whereas the Census Bureau poverty cutoffs are intended to apply to a measure of family income. Thus, the concept of poverty as applied in these studies does not necessarily capture the same concept as the official measure, and the concepts are not necessarily comparable across studies.[67]

With this background, the next section focuses first on the results from random assignment studies with respect to income, income sources, and poverty. In that section, we also briefly summarize the results in Appendix A with respect to differential impacts for subgroups. A discussion of the econometric studies that consider these outcomes follows in the third section. The results from both types of studies are synthesized in the fourth section. A final section concludes the chapter.

8.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECTS OF WELFARE REFORM ON INCOME, INCOME SOURCES, AND POVERTY

With the exception of the four programs that focus on other reforms (AWWDP, FDP, PPI and PIP), all the experimental studies reviewed in previous chapters also include at least one

[66]Some would argue that income, especially the way it is usually measured in surveys, may not be the best indicator of material well-being (Edin and Lein, 1997; Meyer and Sullivan, 2001; Haskins, 2001). More limited information on other measures of well-being, such as material hardship, food insecurity, and housing problems, will be covered in Chapter 9.

[67]The lack of correspondence with the official U.S. poverty measure is not necessarily problematic, given the concerns with the validity of that measure (see Citro and Michael, 1995). The lack of comparability across studies is more of an issue.

measure of income in their impact analyses, while some also include a measure of poverty. Table 8.1 summarizes the findings from this literature. For the relevant experiment and population served, the table indicates whether administrative or survey data are used to measure a particular income concept and then whether the income concept is used to calculate a poverty rate. In some cases, the same income measure is reported for different points in time to determine if impacts fade out or grow stronger with time; some studies report multiple income measures for the same or different follow-up periods. Since incomes are measured over varying time intervals, we also normalize all estimated impacts to a monthly concept.

In addition, we also consider the sources of income to the extent they are reported in the experimental studies. Since earnings are reported in Chapter 5, we do not repeat those results here. We do tabulate in Table 8.2 the impacts for the amount of welfare payments and food stamp payments, ideally for the same follow-up interval as the income measures. (Chapters 4 and 6 previously covered participation rates in welfare and food stamps, respectively.) Welfare payments are reported in all the studies that report income, and the same is true for food stamp payments with two exceptions: MFIP (where the Food Stamp Program was cashed out) and the Canadian SSP (where food stamps do not apply). For both income sources, the measures are typically based on administrative data for the recipient. In a few cases, survey data are used to collect information about benefits received for other family or household members. In a few studies, data on other sources of income (e.g., earnings from other household members, child support payments, and other public transfers) are collected as part of a follow-up survey. While we do not report those results in a table, we note in the text when significant impacts on these other income sources are measured.

Finally, to assess individual or family self-sufficiency, we also report in Table 8.2 program impacts for earnings' share of total income. In some cases, this outcome is measured directly. In other cases, the fraction of the sample with 50 percent or more of income from earnings is reported. The earnings share is not available for all the studies that report income.

We follow the structure used in previous chapters and review studies by the reform policy or policies they evaluate, considering income, poverty, welfare and food stamp payments, and earnings share for each group of studies.

8.2.1. Programs That Focus on Financial Work Incentives

Three studies assess the impact of financial work incentives: WRP-IO, CWPDP, and MFIP-IO. CWPDP does not report a combined income measure, so the estimated impact in year three reported in Table 8.1 is based on the separate impacts for the recipient's earnings (Table 5.1) and welfare and food stamp payments (Table 8.2). (Consequently, the significance level is not known.) The impact estimate for recipient combined income is close to 0, a negative $3 per month. Food stamp payments are the only combined income component that has a significant positive impact; earnings and welfare payments have negative but insignificant impacts. The negative welfare payment impact is despite the fact that there is a positive (but insignificant) impact on welfare use (Table 4.1), which reflects the fact that this program also reduced benefit levels.

Vermont's WRP-IO program shows no gain for the treatment group over the control group in recipient combined income, measured using administrative data over the full four-year follow-up or in the last quarter only. However, a large and significant treatment-control difference in household income is measured at the 42-month follow-up survey, equal to $139–$145 per month depending on the income measure. Panel A of Table 8.2 reveals that WRP-IO did not result in a significant difference in welfare or food stamp payments, and the program had no significant impact on earnings at the 42-month follow-up (Table 5.1). The only income source that shows a sizeable increase is reported earnings from other family members, but the treatment-control difference for this income component was not significant (not shown). Likewise, while the fraction of treatment group members with more than one-half of household income from earnings was higher than the control group, the difference was not significant (Table 8.2).

MFIP-IO shows some favorable impacts for both income and poverty. The MFIP-IO results are strongest for long-term recipients, with an estimated statistically significant increase in recipient combined income of $80–$100 a month (depending on whether measures are pre- or post-tax, including the EITC) and a reduction in the poverty rate of 7–8 percentage points. However, there is no gain in family income for the recipient group measured at the 36-month follow-up survey, a result that may be due to differential reporting bias between the treatment and control groups.[68] Smaller effects were found for the income gains and poverty rate reductions in the MFIP applicant group, with the exception of family income at the 36-month follow-up where the impact was larger (i.e., a positive impact compared with the negative impact estimate for recipients). To the extent that MFIP-IO generated income gains, they were the result of higher welfare and (cashed-out) food stamp payments generated by the financial work incentives: MFIP-IO raised welfare and food stamp payments for both applicants and recipients by about the same dollar amount ($91–$97 per month, as shown in Table 8.2). There were no significant impacts on earnings for either MFIP-IO recipients or applicants (Table 5.1). Consequently, the earnings share measure declined, indicating a lower level of self-sufficiency as a result of MFIP-IO, an effect that was largest and significant for the MFIP-IO applicants (Table 8.2).

8.2.2. Programs That Focus on Financial Work Incentives Tied to Hours of Work

The two programs that evaluate financial work incentives tied to hours of work, with follow-up periods that range from 18 months to three years, both show positive and significant income impacts and corresponding significant negative poverty impacts. (See Panel B of Table 8.2.) In the case of New Hope, the effects are evident only for families not employed full-time at randomization. In the case of the Canadian SSP evaluations, the program with the most

[68]For the same MFIP-IO sample with administrative data and survey data on combined income (earnings plus welfare, including the food stamp cash-out), the mean survey report of combined income for the AFDC (control) group exceeds the administrative data mean for combined income, while the reverse is true for the MFIP-IO (treatment) group (Miller et al., 2000, Table 4.6). Thus, the impact estimate (treatment-control difference) for monthly combined income is $109 ($p < 0.05$) based on the administrative data versus $12 based on the survey data. There is almost no difference between the treatment and control groups in the mean value of the other income sources (e.g., earnings of other household members, child support, and other income) collected in the survey data.

Table 8.1—Estimated Impact of Welfare Reform on Income and Poverty: Random Assignment Studies

Name	Cases served	Data	Measure	Income				Poverty		
				Control mean	Impact	%	Normalize to monthly	Control mean	Impact	%
A. Programs that focus on financial work incentives										
CWPDP	Single-parent recipients	A	Avg. annual recipient E+W+FS income in year 3 of 42-mo FU	$8,421	-$30 n.a.	-0.4%	-$3			
WRP-IO	Single-parent recipients and applicants	A	Avg. quarterly recipient E-W-FS income over 42-mo FU	$2,256	$5	0.2%	$2			
		A	Avg. quarterly recipient E-W-FS income in last 3 mos. of 42-mo FU	$2,307	$7	0.3%	$2			
		S	Avg. mo. HH income in mo. prior to 42-mo FU	$1,410	$145 **	10.3%	$145			
		S	Avg. mo. HH income + EITC in mo. prior to 42-mo FU	$1,501	$139 **	9.3%	$139			
MFIP-IO	Urban single parents recipients	A	Avg. quarterly recipient E-W-FS income in last 3 quarters of 10-quarter FU	$2,525	$243 ***	9.6%	$81	77.7	-8.3 ***	-10.7%
		A	Avg. quarterly recipient E-W-FS income + EITC/taxes in last 3 quarters of 10-quarter FU	$2,613	$299 ***	11.4%	$100	70.5	-6.6 ***	-9.4%
		S	Avg. mo. family income in month prior to 36-mo FU	$1,459	-$11	-0.8%	-$11			
	Urban single parents applicants	A	Avg. quarterly recipient E-W-FS income in last 3 quarters of 10-quarter FU	$2,578	$138	5.4%	$46	66.0	-4.0 ***	-6.1%
		A	Avg. quarterly recipient E-W-FS income + EITC/taxes in last 3 quarters of 10-quarter FU	$2,556	$177 **	6.9%	$59	59.7	-3.6 ***	-6.0%
		S	Avg. mo. family income in month prior to 36-mo FU	$1,838	$86	4.7%	$86			

Table 8.1—Continued

Name	Cases served	Data	Measure	Income: Control mean	Income: Impact	Income: %	Income: Normalize to monthly	Poverty: Control mean	Poverty: Impact	Poverty: %
B. Programs that focus on financial work incentives tied to hours of work										
New Hope (a)	Poor families employed FT at RA	A	Avg. annual recipient E+W+FS +EITC for year 1 of 2-yr FU	$14,561	$187	1.3%	$16	58.5	-5.2	-8.9%
		A	Avg. annual recipient E+W+FS +EITC for year 2 of 2-yr FU	$15,294	-$1,148	-7.5%	-$96	56.2	-6.9	-12.3%
	Poor families not employed FT at RA	A	Avg. annual recipient E+W+FS +EITC for year 1 of 2-yr FU	$9,843	$1,347 ***	13.7%	$112	89.3	-5.6 **	-6.3%
		A	Avg. annual recipient E+W+FS +EITC for year 2 of 2-yr FU	$9,915	$1,298 ***	13.1%	$108	81.4	-8.2 ***	-10.1%
SSP (b)	Single-parent recipients	A	Avg. mo. recipient E-W income in Q 5 and 6	$932	$179 ***	19.2%	$179			
		A, S	Avg. mo. family income at 18-mo FU	$1,286	$199 ***`	15.5%	$199	89.8	-12.2 ***	-13.6%
		A, S	Avg. mo. family income in 6 mos. prior to 36-mo. FU	$1,432	$153 ***	10.7%	$153	86.2	-9.4 **	-10.9%
SSP Plus (b)	Single-parent recipients	A, S	Avg. mo. family income in 6 mos. prior to 18-mo FU	$1,171	$156 ***	13.3%	$156			
SSP Applicants (b)	Single-parent applicants	A, S	Avg. mo. family income in 6 mos. prior to 30-mo FU	$1,686	$286 ***	17.0%	$286	68.5	-11.3 ***	-16.5%

Table 8.1—Continued

Name	Cases served	Data	Measure	Income				Poverty		
				Control mean	Impact	%	Normalize to monthly	Control mean	Impact	%
C. Programs that focus on mandatory work-related activities										
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$9,920	$136	1.4%	$11			
		A	Avg. annual recipient E+W+FS +EITC/payroll taxes in year 2 of 2-yr FU	$10,262	$206	2.0%	$17	75.6	-4.5	-6.0%
		A, S	Avg. mo. HH income + EITC/payroll taxes in last month of 2-yr FU	$1,001	$86 *	8.6%	$86	67.6	-9.7 ***	-14.3%
Atlanta LFA	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,549	$191	2.5%	$16	87.1	-1.6	-1.8%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$699	$24	3.4%	$24	80.1	-1.2	-1.5%
Grand Rapids LFA	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,746	-$303 **	-3.9%	-$25	86.5	-1.2	-1.4%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$833	-$42	-5.0%	-$42	68.6	-0.4	-0.6%
Riverside LFA	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,874	-$358 ***	-4.5%	-$30	83.5	-1.0	-1.2%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$867	$19	2.2%	$19	72.6	-6.3 ***	-8.7%
Portland	Recipients and applicants; no cases with substantial barriers	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$8,110	$238	2.9%	$20	83.4	-4.0 ***	-4.8%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$843	$59	7.0%	$59	70.8	-6.4	-9.0%
Atlanta HCD	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,549	$235	3.1%	$20	87.1	-2.0 *	-2.3%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$699	$26	3.7%	$26	80.1	-0.6	-0.7%
Grand Rapids HCD	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,746	-$91	-1.2%	-$8	86.5	-0.3	-0.3%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$833	-$28	-3.4%	-$28	68.6	0.1	0.1%
Riverside HCD	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$7,768	-$619 ***	-8.0%	-$52	86.4	-0.2	-0.2%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$859	-$10	-1.2%	-$10	76.8	-3.5	-4.6%

Table 8.1—Continued

				Income				Poverty		
Name	Cases served	Data	Measure	Control mean	Impact	%	Normalize to monthly	Control mean	Impact	%
Columbus Integrated	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$8,332	-$41	-0.5%	-$3	79.3	0.0	0.0%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$806	-$9	-1.1%	-$9	76.1	-3.3	-4.3%
Columbus Traditional	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$8,332	$29	0.3%	$2	79.3	-0.3	-0.4%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$806	$17	2.1%	$17	76.1	-4.5	-5.9%
Detroit	Recipients and applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$8,892	$101	1.1%	$8	84.1	-1.2	-1.4%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$766	$10	1.3%	$10	79.1	-3.7	-4.7%
Oklahoma City	Applicants	A	Avg. annual recipient E+W+FS in year 2 of 2-yr FU	$5,238	-$137	-2.6%	-$11	92.8	-0.5	-0.5%
		S	Avg. mo. Recipient income +EITC + CC in last month of 2-yr FU	$737	-$40	-5.4%	-$40	74	0.6	0.8%
IMPACT Basic Track	Recipients and applicants, less job ready	A	Avg. annual recipient E-W-FS income over year 2 of 2-yr FU	$5,645	$336	6.0%	$28	91.2	-3.0	-3.3%

Table 8.1—Continued

Name	Cases served	Data	Measure	Income: Control mean	Income: Impact		Income: %	Income: Normalize to monthly	Poverty: Control mean	Poverty: Impact		Poverty: %
D. Programs that focus on financial work incentives and mandatory work-related activities												
WRP	Single-parent recipients and applicants	A	Avg. quarterly recipient E-W-FS income over 42-mo FU	$2,256	$41		1.8%	$14				
		A	Avg. quarterly recipient E-W-FS income in last 3 mos. of 42-mo FU	$2,307	$25		1.1%	$8				
		S	Avg. mo. HH income in mo. prior to 42-mo FU	$1,410	$10		0.7%	$10				
		S	Avg. mo. HH income + EITC in mo. prior to 42-mo FU	$1,501	$27		1.8%	$27				
MFIP	Urban single-parent recipients	A	Avg. quarterly recipient E-W-FS income in last 3 quarters of 10-quarter FU	$2,525	$296	***	11.7%	$99	77.7	-12.4	***	-16.0%
		A	Avg. quarterly recipient E-W-FS income + EITC/taxes in last 3 quarters of 10-quarter FU	$2,613	$382	***	14.6%	$127	70.5	-12.1	***	-17.2%
		S	Avg. mo. family income in month prior to 36-mo FU	$1,459	-$24		-1.6%	-$24				
	Urban single-parent applicants	A	Avg. quarterly recipient E-W-FS income in last 3 quarters of 10-quarter FU	$2,578	$162	**	6.3%	$54	66	-6.9	***	-10.5%
		A	Avg. quarterly recipient E-W-FS income + EITC/taxes in last 3 quarters of 10-quarter FU	$2,556	$187	***	7.3%	$62	59.7	-6.1	***	-10.2%
		S	Avg. mo. family income in month prior to 36-mo FU	$1,838	$75		4.1%	$75				
TSMF	Single parent recipients	A	Avg. annual family E+W+FS over 4-yr FU	$8,849	$118	***	1.3%	10				
	Single parent applicants	A	Avg. annual family E+W+FS over 1-yr FU	$8,558	$10		0.1%	1				
		A	Avg. annual family E+W+FS over 2-yr FU	$8,414	-$163		-1.9%	-14				
FIP	Recipients	A	Avg. quarterly recipient E-W income in Q4 of 2-yr FU	$1,721	$98	***	5.7%	$33				
		A	Avg. quarterly recipient E-W income in Q8 of 2-yr FU	$1,907	$37		1.9%	$12				
	Applicants	A	Avg. quarterly recipient E-W income in Q4 of 2-yr FU	$2,004	$218	***	10.9%	$73				

Table 8.1—Continued

Name	Cases served	Data	Measure	Income: Control mean	Income: Impact	Income: %	Income: Normalize to monthly	Poverty: Control mean	Poverty: Impact	Poverty: %
E. Programs that focus on other individual reforms										
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
EMPOWER (c)	Recipients	A,S	Avg. mo. HH income in survey month approx. at 30-mo FU	$1,339	$80	6.0%	$80			
IMPACT Placement Track	Recipients and applicants, more job ready	A	Avg. annual recipient E-W-FS income over year 2 of 2-yr FU	$7,502	$77	1.0%	$6	80.2	-1.0	-1.2%
VIP/VIEW	Recipients	A	Avg. annual recipient E-W-FS in year 2 of 2-yr FU	$6,482	$128	2.0%	$11			
ABC	Single parent recipients and applicants	S	Avg. mo. HH income in month before survey 12-18-mo FU	$778	$0	0.0%	$0			
FTP	Recipients and applicants	A	Avg. total recipient E-W-FS income in year 2	$6,358	$351 *	5.5%	$29			
		A	Avg. total recipient E-W-FS income in year 3	$6,137	$496 **	8.1%	$41			
		A	Avg. total recipient E-W-FS income in year 4	$6,310	$253	4.0%	$21			
		A	Avg. total recipient E-W-FS income in 2nd Q of year 5	$1,674	-$52	-3.1%	-$17			
		S	Avg. mo. HH income in month before 4-yr FU	$1,379	$89	6.5%	$89			
Jobs First	Recipients and applicants	A	Avg. annual recipient E-W-FS income in year 2	$10,037	$1,121 ***	11.2%	$93			
		A	Avg. annual recipient E-W-FS income in year 3	$10,647	$172	1.6%	$14			
		A	Avg. annual recipient E-W-FS income in year 4	$11,249	-$132	-1.2%	-$11			
		A	Avg. annual recipient E-W-FS + EITC/taxes income in yr. 3 and 4	$10,828	$150	1.4%	$13			
		S	Avg. total recipient income in month before 3-yr FU	$1,022	$74 ***	7.2%	$74			
		S	Avg. total income of other HH members in month before 3-yr FU	$442	$12	2.7%	$12			

NOTES: For full program names and citations, see Table 3.4. In calculating the poverty rates, the official poverty lines are applied to the given income measure. This will not necessarily correspond to the official poverty definition. Abbreviations: A=administrative data; S=survey data; E=earnings; W=cash welfare payments; FS=Food Stamp payments; EITC=Earned Income Tax Credit; CC=out-of-pocket child care expenses; FU=follow-up; HH=household; Q=quarter.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) Poverty line based on earnings-related income only (earnings, EITC, earnings supplement).

(b) Results in Canadian dollars.

(c) Phoenix site only, cash assistance.

Table 8.2—Estimated Impact of Welfare Reform on Income Sources: Random Assignment Studies

			Welfare payments					Food Stamp payments					Earnings share			
Name	Cases served	Data	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Contro l mean	Impact	%
A. Programs that focus on financial work incentives																
CWPDP	Single-parent recipients	A	Avg. annual R AFDC payment in year 3 of 42-mo FU	$4,643	-$26	-0.6%	-$2	Avg. annual R FS payment in year 3 of 42-mo FU	$1,406	$156 **	11.1%	$7				
WRP-IO	Single-parent recipients and applicants	A	Avg. quarterly R ANFC payment for last 3 mos. of 42-mo FU	$539	-$17	-3.2%	-$6	Avg. quarterly R FS payment for last 3 mos. of 42-mo FU	$266	$10	3.8%	$3				
		S	Avg. mo. HH ANFC payment in mo. prior to 42-mo FU	$193	-$27	-14.0%	-$27	Avg. mo. HH FS payment in mo. prior to 42-mo FU	$104	-$3	-2.9%	-$3	50% or more of HH income from HH earnings at 42-mo FU (%)	55.1	2.9	5.3%
MFIP-IO	Urban single parents recipients	A	Avg. quarterly R welfare (AFDC, FS, GA) payment in last 3 quarters of 10-quarter FU	$1,227	$291 ***	23.7%	$97	(results combined with those for welfare payments)					50% or more of R income from R earnings in last Q of 10-Q FU (%)	34.1	-1.0	-2.9%
	Urban single parents applicants	A	Avg. quarterly R welfare (AFDC, FS, GA) payment in last 3 quarters of 10-quarter FU	$561	$274 ***	48.8%	$91	(results combined with those for welfare payments)					50% or more of R income from R earnings in last Q of 10-Q FU (%)	49.5	-5.0 ***	-10.1%
B. Programs that focus on financial work incentives tied to hours of work																
New Hope	Poor families employed FT at RA	A	Avg. annual R AFDC payments in year 1 of 2-yr FU	$1,396	-$56	-4.0%	-$5	Avg. annual R FS payments in year 1 of 2-yr FU	$1,305	-$67	-5.1%	-$6				
		A	Avg. annual R earnings supplement in year 1 of 2-yr FU	$0	$630		$53									
		A	Avg. annual R AFDC payments in year 2 of 2-yr FU	$1,181	-$445 **	-37.7%	-$37	Avg. annual R FS payments in year 2 of 2-yr FU	$1,167	-$274 **	-23.5%	-$23				
		A	Avg. annual R earnings supplement in year 2 of 2-yr FU	$0	$496		$41									
	Poor families not employed FT at RA	A	Avg. annual R AFDC payments in year 1 of 2-yr FU	$2,962	-$11	-0.4%	-$1	Avg. annual R FS payments in year 1 of 2-yr FU	$1,837	-$10	-0.5%	-$1				
		A	Avg. annual R earnings supplement in year 1 of 2-yr FU	$0	$418		$35									
		A	Avg. annual R AFDC payments in year 2 of 2-yr FU	$1,690	$26	1.5%	$2	Avg. annual R FS payments in year 2 of 2-yr FU	$1,242	$176 **	14.2%	$15				
		A	Avg. annual R earnings supplement in year 2 of 2-yr FU	$0	$396		$33									
SSP (a)	Single-parent recipients	A	Avg. mo. R IA payments in Q 5 and 6	$723	-$103 ***	-14.2%	-$103									
		A	Avg. mo. R SSP supplement payments in Q 5 and 6	$0	$196 ***	n.a.	$196									
		A	Avg. mo. R IA payments in 6 mos. prior to 36-mo FU	$573	-$67 ***	-11.7%	-$67									
		A	Avg. mo. R SSP supplement payments in 6 mos. prior to 36-mo FU	$0	$156 ***	n.a.	$156									
SSP Plus (a)	Single-parent recipients	A	Avg. mo. R IA payments in 6 mos. prior to 18-mo FU	$607	-$144 ***	-23.7%	-$144									
		A	Avg. mo. R SSP supplement payments in 6 mos. prior to 18-mo FU	$1	$219 ***	n.a.	$219									
SSP Applicants (a)	Single-parent applicants	A	Avg. mo. R IA payments in 6 mos. prior to 30-mo FU	$449	-$97 ***	-21.6%	-$97									
		A	Avg. mo. R SSP supplement payments in 6 mos. prior to 30-mo FU	$0	$154 ***	n.a.	$154									

Table 8.2—Continued

			Welfare payments					Food Stamp payments					Earnings share			
Name	Cases served	Data	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%
C. Programs that focus on mandatory work-related activities																
LA Jobs-1st GAIN	Single-parent recipients and applicants	A	Avg. annual R AFDC/TANF payment in year 2 of 2-yr FU	$4,269	-$540 ***	-12.6%	-$45	Avg. annual R FS payment in year 2 of 2-yr FU	$1,713	-$192 ***	-11.2%	-$16				
		A, S	Avg. mo. HH AFDC/TANF payment in last month of 2-yr FU	$336	-$58 ***	-17.3%	-$58	Avg. mo. HH FS payment in last month of 2-yr FU	$138	-$16 *	-11.6%	-$16				
Atlanta LFA	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$4,922	-$369 ***	-7.5%	-$15	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,934	-$88	-1.8%	-$4	R earnings as % of total R income in last month of 2-yr FU	26.5	3.3 *	12.5%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$9,946	-$881 ***	-8.9%	-$15	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$11,089	-$428 **	-3.9%	-$7				
Grand Rapids LFA	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$7,347	-$1,404 ***	-19.1%	-$59	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$3,695	-$279 ***	-7.6%	-$12	R earnings as % of total R income in last month of 2-yr FU	36.3	5.7 **	15.7%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$12,966	-$2,552 ***	-19.7%	-$43	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$6,966	-$615 ***	-8.8%	-$10				
Riverside LFA	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$9,600	-$1,308 ***	-13.6%	-$55	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$2,725	-$353 ***	-13.0%	-$15	R earnings as % of total R income in last month of 2-yr FU	23.5	6.4 ***	27.2%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$18,294	-$2,710 ***	-14.8%	-$45	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$5,870	-$888 ***	-15.1%	-$15				
Portland	Recipients and applicants; no cases with substantial barriers	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$7,014	-$1,196 ***	-17.1%	-$50	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,359	-$405 ***	-9.3%	-$17	R earnings as % of total R income in last month of 2-yr FU	26.2	12.2 ***	46.6%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$11,686	-$2,746 ***	-23.5%	-$46	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$7,753	-$827 ***	-10.7%	-$14				
Atlanta HCD	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$4,922	-$288 ***	-5.9%	-$12	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,934	-$3	-0.1%	$0	R earnings as % of total R income in last month of 2-yr FU	26.5	1.4	5.3%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$9,946	-$710 ***	-7.1%	-$12	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$11,089	-$159	-1.4%	-$3				
Grand Rapids HCD	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$7,347	-$835 ***	-11.4%	-$35	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$3,695	-$103 *	-2.8%	-$4	R earnings as % of total R income in last month of 2-yr FU	36.3	2.3	6.3%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$12,966	-$1,767 ***	-13.6%	-$29	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$6,966	-$387 ***	-5.6%	-$6				
Riverside HCD	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$10,302	-$1,049 ***	-10.2%	-$44	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$2,929	-$286 ***	-9.8%	-$12	R earnings as % of total R income in last month of 2-yr FU	18.7	5.3 **	28.3%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$20,126	-$2,949 ***	-14.7%	-$49	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$6,504	-$1,013 ***	-15.6%	-$17				
Columbus Integrated	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$5,469	-$694 ***	-12.7%	-$29	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,710	-$432 ***	-9.2%	-$18	R earnings as % of total R income in last month of 2-yr FU	31.9	9.4 ***	29.5%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$9,005	-$1,523 ***	-16.9%	-$25	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$8,185	-$1,025 ***	-12.5%	-$17				
Columbus Traditional	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$5,469	-$530 ***	-9.7%	-$22	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,710	-$312 ***	-6.6%	-$13	R earnings as % of total R income in last month of 2-yr FU	31.9	2.7	8.5%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$9,005	-$1,105 ***	-12.3%	-$18	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$8,185	-$648 ***	-7.9%	-$11				
Detroit	Recipients and applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$8,615	-$158	-1.8%	-$7	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$4,829	-$92 *	-1.9%	-$4	R earnings as % of total R income in last month of 2-yr FU	21.4	9.5 ***	44.4%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU	$16,247	-$561 **	-3.5%	-$9	Avg. total R FS payment for years 1 to 5 of 5-yr FU	$9,519	-$334 **	-3.5%	-$6				
Oklahoma City	Applicants	A	Avg. total R AFDC payment for years 1 to 2 of 5-yr FU	$3,624	-$233 ***	-6.4%	-$10	Avg. total R FS payment for years 1 to 2 of 2-yr FU	$3,554	-$69	-1.9%	-$3	R earnings as % of total R income in last month of 2-yr FU	38.1	-0.5	-1.3%
		A	Avg. total R AFDC payment for years 1 to 5 of 5-yr FU			-11.9%		Avg. total R FS payment for years 1 to 5 of 2-yr FU	$1,881	$63	3.3%	$5				20.3%
IMPACT Basic Track	Recipients and applicants	A	Avg. total R AFDC/TANF payments for year 2 of 2-yr FU	$1,084	-$4	-0.4%	$0	Avg. total R FS payments for year 2 of 2-yr FU	$1,443	$115	8.0%	$10	R earnings as % of total R income in year 2 of 2-yr FU	41.4	-0.4	-1.0%

Table 8.2—Continued

			Welfare payments					Food Stamp payments					Earnings share			
Name	Cases served	Data	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities																
WRP	Single-parent recipients and applicants	A	Avg. quarterly R ANFC payment for last 3 mos. of 42-mo FU	$539	-$106 ***	-19.7%	-$35	Avg. quarterly R FS payment for last 3 mos. of 42-mo FU	$266	-$5	-1.9%	-$2				
		S	Avg. mo. HH ANFC payment in mo. prior to 42-mo FU	$193	-$71 ***	-36.8%	-$71	Avg. mo. HH FS payment in mo. prior to 42-mo FU	$104	-$9	-8.7%	-$9	50% or more of HH income from HH earnings at 42-mo FU (%)	55.1	9.1 ***	16.5%
MFIP	Urban single-parent recipients	A	Avg. quarterly R welfare (AFDC, FS, GA) payment in last 3 quarters of 10-qtr FU	$1,227	$154 ***	12.6%	$51	(results combined with those for welfare payments)					50% or more of R income from R earnings in last Q of 10-Q FU (%)	34.1	1.7	5.0%
	Urban single parents applicants	A	Avg. quarterly R welfare (AFDC, FS, GA) payment in last 3 quarters of 10-qtr FU	$561	$147 ***	26.2%	$49	(results combined with those for welfare payments)					50% or more of R income from R earnings in last Q of 10-Q FU (%)	49.5	-1.5	-3.0%
TSMF	Single-parent recipients	A	Avg. annual R AFDC/SFA payments over 4-yr FU	$3,442	-$101 ***	-2.9%	-$8	Avg. annual R FS payments over 4-yr FU	$1,920	-$31 **	-1.6%	-$3				
	Single-parent applicants	A	Avg. annual R AFDC/SFA payments over 2-yr FU	$2,857	-$104 **	-3.6%	-$9	Avg. annual R FS payments over 2-yr FU	$1,540	-$42 *	-2.7%	-$4				
		A	Avg. annual R AFDC/SFA payments over 1-yr FU	$3,218	-$152	-4.7%	-$13	Avg. annual R FS payments over 1-yr FU	$1,737	-$56	-3.2%	-$5				
FIP	Recipients	A	Avg. quarterly R FIP payments in Q8 of 2-yr FU	$570	-$38 ***	-6.7%	-$13	Avg. quarterly R FS payments in Q8 of 2-yr FU	$400	-$34 ***	-8.5%	-$11				
	Applicants	A	Avg. quarterly R FIP payments in Q4 of 2-yr FU	$319	$5	1.6%	$2	Avg. quarterly R FS payments in Q4 of 2-yr FU	$246	-$12	-4.9%	-$4				

Table 8.2—Continued

			Welfare payments					Food Stamp payments					Earnings share			
Name	Cases served	Data	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%	Norm. to mo.	Measure	Control mean	Impact	%
E. Programs that focus on other individual reforms																
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)																
EMPOWER (b)	Recipients	A	Avg. mo. R cash assistance payments in year 2 of 3-yr FU	$144	-$13 **	-9.0%	-$13	Avg. mo. R FS payments in year 2 of 3-yr FU	$129	-$3	-2.3%	-$3				
		A	Avg. mo. R cash assistance payments in year 3 of 3-yr FU	$82	-$16 ***	-19.5%	-$16	Avg. mo. R FS payments in year 3 of 3-yr FU	$81	-$6		-$6				
		A, S	Avg. mo. HH cash assistance and FS in survey month approx. at 30-mo FU	$196	-$19	-9.7%	-$19									
IMPACT Placement Track	Recipients and applicants	A	Avg. total R AFDC/TANF payments for year 2 of 2-yr FU	$1,043	-$264 ***	-25.3%	-$22	Avg. total R FS payments for year 2 of 2-yr FU	$1,516	-$219 ***	-14.4%	-$18	R earnings as % of total R income in year 2 of 2-yr FU	48.1	5.9 ***	12.3%
VIP/VIEW	Recipients	A	Avg. annual R TANF payments in year 2 of 2-yr FU	$1,682	-$17	-1.0%	-$1	Avg. annual R FS payments in year 2 of 2-yr FU	$2,077	-$92 ***	-4.4%	-$8				
ABC	Single-parent recipients and applicants	A	Avg. quarterly R cash welfare payment in Q5 of 5-quarter FU	$516	-$91 ***	-17.6%	-$30									
FTP	Recipients and applicants	A	Avg. total R AFDC/TANF payments in year 2	$1,288	-$136 ***	-10.6%	-$11	Avg. total R FS payments in year 2	$1,792	-$174 ***	-9.7%	-$15				
		A	Avg. total R AFDC/TANF payments in year 3	$870	-$289 ***	-33.2%	-$24	Avg. total R FS payments in year 3	$1,416	-$125 **	-8.8%	-$10				
		A	Avg. total R AFDC/TANF payments in year 4	$549	-$277 ***	-50.5%	-$23	Avg. total R FS payments in year 4	$1,122	-$37	-3.3%	-$3	50% or more of R E+W+FS from R earnings over 4-yr FU (%)	44.7	5.4 ***	12.1%
		A	Avg. total R AFDC/TANF payments 2nd Q of year 5	$94	-$45 ***	-47.9%	-$15	Avg. total R FS payments in 2nd Q of year 5	$251	-$23	-9.1%	-$8	50% or more of R E+W+FS from R earnings in 2nd Q of year 5 (%)	45.0	-1.0	-2.1%
		S	Avg. mo. HH AFDC/TANF payments in month before 4-yr FU	$54	-$28 ***	-52.2%	-$28	Avg. mo. HH FS payments in month before 4-yr FU	$122	-$5	-4.2%	-$5	R earnings as % of total HH income in month before 4-yr FU	63.3	3.7 *	-4.2%
Jobs First	Recipients and applicants	A	Avg. annual R AFDC/TFA payments in year 2	$3,019	$363 ***	12.0%	$30	Avg. annual R FS payments in year 2	$1,553	$118 ***	7.6%	$10				
		A	Avg. annual R AFDC/TFA payments in year 3	$2,259	-$422 ***	-18.7%	-$35	Avg. annual R FS payments in year 3	$1,333	-$10	-0.8%	-$1				
		A	Avg. annual R AFDC/TFA payments in year 4	$1,645	-$479 ***	-29.1%	-$40	Avg. annual R FS payments in year 4	$1,113	-$17	-1.5%	-$1				

NOTES: For full program names and citations, see Table 3.4. Abbreviations: A=administrative data; S=survey data; FU=follow-up; HH=household; R=recipient; Q=quarter; SFA=State Family Assistance program (Michigan); IA=Income Assistance (Canada); RA=random assignment; FT=full-time.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

(a) Results in Canadian dollars.

(b) Phoenix site only, cash assistance.

generous financial work incentives to reward full-time work, longer-term recipients (evaluated in two provinces), and new applicants (evaluated in one province) show large percentage gains in recipient combined income or family income—from 11 to 19 percent—for up to 36 months post-randomization.[69] The SSP Plus results fall in between the results for the main SSP. The poverty rate, when reported, also falls by a substantial magnitude, between 9 and 12 percentage points.

New Hope and SSP, by operating outside the traditional U.S. and Canadian welfare systems, were designed to replace welfare payments with an earnings supplement. Both programs show significant decreases in welfare payments and increases in the earnings supplement. Food stamp payments, relevant only for New Hope, showed a significant decrease for families employed full time at random assignment, while the reverse is true for those not employed full time. Recall that the earnings impact for New Hope participants employed full time at random assignment were insignificant and negative. By year two, the positive impact on the earnings supplement just exceeded the negative impact on AFDC payments. These effects for earnings and welfare/earnings supplement combined to produce a negative impact overall on recipient combined income for this group. In contrast, those in New Hope not employed full-time at random assignment experienced significant positive earnings gains, and the average size of the earnings supplement more than offset the impact on AFDC benefits. Food stamp benefit payments also increased by the second year of follow-up. Consequently, this group experienced large and significant income gains overall for both years one and two. Finally, SSP produced increases in both cash transfer income (earnings supplements net of lost income assistance payments) and earnings. This combination of impacts on the sources of income produced the large increase in recipient combined income and family income.

8.2.3. Programs That Focus on Mandatory Work-Related Activities

All the programs that focus on mandatory work-related activities report measures of income and poverty at the end of two years, for both administrative and survey data in all but one case, where only administrative data is available (results summarized in Panel C of Table 8.1). Measures of combined recipient income based on administrative data show a significant negative impact for three of the NEWWS programs, ranging from –$25 to –$52 per month. The other evaluations all show smaller insignificant impacts, either negative (three NEWWS programs) or positive (five NEWWS programs, L.A. Jobs-First GAIN, and IMPACT). With the exception of Riverside LFA, the survey-based measures of recipient combined income (in some cases, accounting for the EITC, payroll taxes, and child care expenses) for the last month of the follow-up period all have the same sign as the measures based on administrative data and are similar in magnitude. The only statistically significant survey-based impact estimate is for L.A. Jobs-First GAIN, a 9 percent increase in income.

With a few exceptions, most of the poverty impacts are insignificant and small in magnitude.[70] They are almost all negative, including the three significant impacts, suggesting these programs are somewhat more effective at raising incomes near the poverty threshold than at the bottom

---

[69] The impacts are shown in Canadian dollars. Converting to U.S. dollars using the rate $1Canada = $0.75U.S., the absolute changes in income are among the largest for the programs in Panel C.

[70] The results for the NEWWS programs are similar when household income is considered instead of recipient income (Freedman et al., 2000a).

of the income scale. At the same time, several of the NEWWS programs resulted in a slight increase in the fraction with incomes below 50 percent of the poverty line as of the two-year follow-up, suggesting that those near the bottom of the income scale may be worse off (Freedman et al., 2000a).

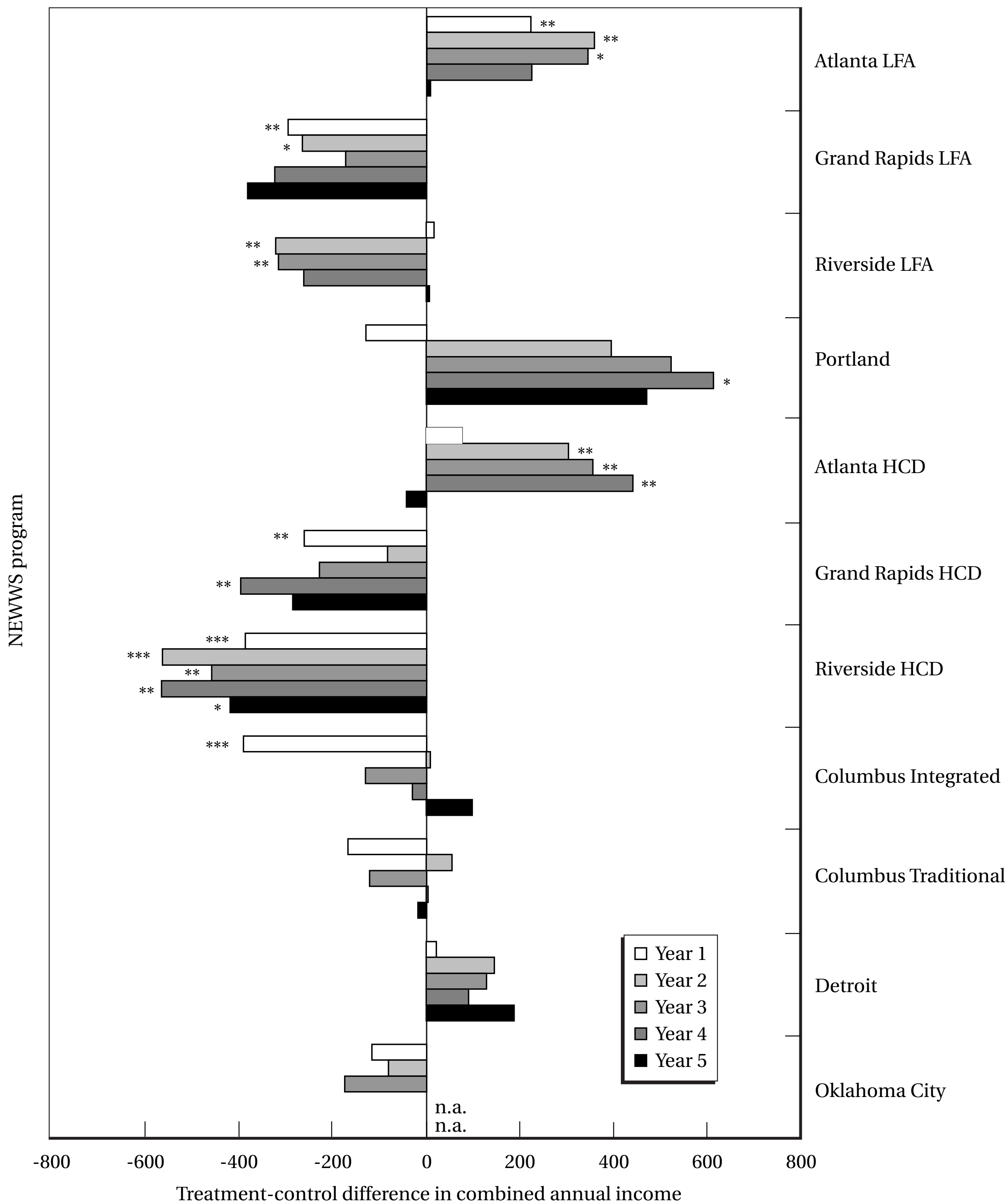
Newly available data for the 11 NEWWS programs provides information on combined income (income from earnings, welfare, and food stamps plus the EITC, less payroll taxes based on administrative data) five years after the program began. The impact estimates are plotted in Figure 8.2 for follow-up years one to five. By year five, there is only one significant income impact in any of the sites (a negative impact in Riverside HCD; data are not available for Oklahoma). The year five impact estimates are almost equally divided between negative impacts, impacts close to zero, and positive impacts.

Based on Figure 8.2, there is no clear pattern in NEWWS site impacts associated with the program orientation or service delivery approach. Portland's employment-focused, mixed–first activity model has the largest five-year income and two-year poverty impacts, but it excluded some recipients with substantial barriers to work. Thus, it is not clear if similar impacts would be found for a more disadvantaged population. The pattern over time for the various NEWWS programs is also not consistent, although the majority of the sites exhibit a fade-out effect over time. For example, Portland has increasing positive impacts in years two and three followed by the largest positive impact of all the sites in year 4 (just over $600 in annual combined income; $p < 0.10$). But the impact shrinks in year five. At the other extreme, Riverside HCD has negative and significant impacts exceeding $400 in all five years, although the year five impact is somewhat less negative than in year four. Some of the fade-out may be due to the control-group crossover that took place in years four and five, but this crossover did not apply to the Portland or Riverside sites and was probably not a significant factor in the other sites either (Hamilton et al., 2001).

Whether measured at the two-year or five-year follow-up, the modest, if any, income and poverty effects are consistent with the combination of the positive earnings gains produced by these programs (Table 5.1 and Figure 5.2) and high benefit reduction rates under the old AFDC rules that led to a significant reduction in welfare payments as earnings rose (Panel C of Table 8.2). Likewise, food stamp payment declines were also almost always significant. Since income, by and large, did not change, but the composition shifted from welfare benefits to earnings, it is not surprising that the majority of the programs also raised self-sufficiency as measured by the earnings share, available only as of the two-year follow-up (Table 8.2). Even so, the average share of income from earnings for treatment group members never exceeds 50 percent.

8.2.4. Programs That Focus on Financial Work Incentives and Mandatory Work-Related Activities

Four programs combine financial work incentives and work requirements with results for income and poverty rates for follow-up periods that range from two to four years. (See Panel D of Table 8.1.) With the exception of WRP, the combined programs produce statistically significant income gains and, when measured, poverty reductions for at least part of the population served over both short and longer horizons. Recipients benefit more than applicants



SOURCE: Hamilton et al. (2001), Table E.1.

NOTE: Combined income is income from earnings, welfare and Food Stamps, plus the EITC, less payroll taxes. Treatment-control difference is statistically significant at the *=10%, **=5%, ***=1% level.

Figure 8.2—Impact Estimates for Combined Annual Income in 11 NEWWS Programs, Years 1 to 5

in the MFIP and TSMF studies, while the reverse is true for Iowa's FIP. The significant impact for FIP recipients as of the fourth quarter fades by the eighth quarter, from $33 per month (significant) to $12 per month (insignificant). [71] In the case of MFIP, a more comprehensive survey-based measure of family income shows a negative impact for recipients, compared with the sizeable gain in combined income measured with administrative data, and a somewhat larger, but insignificant, impact for applicants. As discussed above, the difference in estimated impacts for recipients using administrative data compared with survey data is likely a result of differential reporting biases for treatment versus control group members (Miller et al., 2000). The MFIP impacts based on administrative data are similar in magnitude and are not statistically different from those for MFIP-IO.

The changes in income sources that produce the overall income impacts differ to some extent across these programs. In WRP, which had among the least generous financial work incentives, there were significant declines in welfare payments, and food stamp payments showed no change (Panel D of Table 8.2). The significant earnings increase essentially offset the transfer payment decline, leading to a significant increase in the share of income from earnings, but no change in income. TSMF and FIP also lowered welfare and food stamp payments, but somewhat larger earnings increases resulted in modest income gains overall. (The impact on the earnings share was not measured for these two programs.) In contrast, MFIP raised earnings but also increased combined welfare and food stamp payments, so that the gains in combined income were even larger than programs that decreased welfare income. Consequently, there is no change in MFIP in self-sufficiency as measured by the earnings share.

8.2.5. Programs That Focus on TANF-Like Bundles of Reforms

As shown in Panel F of Table 8.1, four of the six programs that include TANF-like bundles of reforms—EMPOWER, IMPACT, VIP/VIEW, and ABC—find no significant impacts on income or poverty (which is only measured in one study), with follow-up periods that range from 12 to 30 months. Of these four programs, IMPACT had the largest positive (and significant) impact on earnings, but it also resulted in a significant reduction in welfare and food stamp payments (Panel F of Table 8.2). It also resulted in increased self-sufficiency as measured by the earnings share. As noted in Table 3.5 (see notes b, c, and f), a sizeable fraction of the control group in three of these four studies believed that the time limits applied to them, which may bias the estimated program impacts toward zero.

FTP and Jobs First—programs for which the time-limit feature was better understood—demonstrate positive income results earlier in the follow-up period that tend to disappear as time limits are reached. For FTP, the positive and significant impact on recipient combined income peaks in year three—the year recipients first begin to reach the time limit. It becomes insignificant but still positive in year four, as more recipients reach the time limit, and turns negative and insignificant in the first part of year five. Jobs First showed an even larger positive income impact for recipient combined income as late as year two, but this positive impact

[71]The fade-out of the FIP income effects for recipients by year two are likely to be repeated in the follow-up at three and one-half years post-randomization. Results for quarter 14 reported in Fraker and Jacobson (2000) show that average quarterly earnings are $8 more for the treatment group, while average quarterly FIP welfare benefits are $24 less. A combined recipient income result is not reported for the follow-up through quarter 14.

dissipated by year three, and turned negative by year four.[72] More comprehensive measures of income based on administrative or survey data show the same pattern for both programs. The pre– and post–time limit impacts for FTP and Jobs First suggest that reaching the time limit is associated with an income decline, as welfare benefits are exhausted and recipients must rely more on their own earnings.

Both programs had a significant positive impact on earnings (Table 5.1) and a significant reduction in welfare payments, at least by the final follow-up period, three to five years past randomization (Table 8.2). FTP also measured a significant increase in child support payments (not shown). Only FTP analyzed the earnings share, and the results show a significant increase in self-sufficiency as measured by recipient earnings as a fraction of recipient combined income (earnings, welfare, and food stamps) and as a fraction of household income by year four. This impact is no longer significant, however, by the second quarter of year five. The structure of the benefit levels and income disregards in the Connecticut program made it more generous than FTP in Florida, which helps explain the larger initial increase in income and welfare and food stamp use in Jobs First (as of the two-year follow-up).

8.2.6. Subgroup Differences

As summarized in Appendix A, the experimental studies demonstrate both similarities and differences in the impacts of reform policies on income and income sources between subgroups, defined by various measures of disadvantage. Because the studies often analyze different subgroups, it is difficult to draw broad inferences of the impacts of individual reform policies or welfare reform more generally for subgroups with specific characteristics.

Only one of the evaluations that focus exclusively on financial work incentives considers any subgroups, and it is hard to generalize from the specific patterns from that study. The two programs that focus on financial work incentives tied to hours of work are not consistent, with some impacts larger for the more disadvantaged while the reverse is true for other impacts. A larger number of subgroup analyses for programs that focus on mandatory work requirements, including the pooled analysis of the NEWWS programs (combined with nine others), suggests that these programs can have differential impacts on subgroups. The pooled NEWWS analyses indicate that income gains are largest and that welfare payment declines are smallest for the least disadvantaged. The L.A. Jobs-First GAIN results differ from this pooled finding in that the program impacts did not vary by subgroups. The subgroup patterns for programs that combine financial work incentives and work requirements or that focus on TANF-like bundles of reform are mixed, with examples of larger impacts for both the least and the most disadvantaged.

8.3. ECONOMETRIC STUDIES OF THE EFFECTS OF WELFARE REFORM ON INCOME AND POVERTY

Compared with the large number of econometric studies that have examined welfare caseloads (reviewed in Chapter 4), only a handful of econometric studies have examined family income,

---

[72] When income data for Jobs First are examined by quarter, the quarterly combined income impact falls from \$266 ($p < 0.01$) to \$150 ($p < 0.01$) between quarters 7 and 8 as the first recipients begin to reach the time limit, a difference that is equal to about 5 percent of the control group mean in quarter 7. The impact declines further from \$152 ($p < 0.05$) in quarter 9 to \$16 in quarter 10. The treatment-control difference remains insignificant through quarter 16 and is sometimes negative.

and even fewer have examined poverty. (See Table 8.3.) All the studies in this literature rely on the CPS, the primary nationally representative data source with information on annual individual and family income. The Annual Demographic Supplement to the March CPS is the source used by the Census Bureau to calculate poverty rates on an annual basis; hence, these studies can implement a measure of income (annual family pretax cash income) and poverty that follows the concepts employed by the Census Bureau in its calculations.

Like the caseload studies that rely on survey data, these analyses use aggregated microdata or individual-level microdata to model the level or log of income and the poverty rate as a function of the existence (approval or implementation) of a waiver or TANF as a bundle. In one case, the effect of a specific TANF policy—time limits—is also considered. The welfare policy variables generally follow the "modified dummy variable" approach used in the caseload literature. The study population either includes all women or female-headed families in a given age range (typically 16–54). In one case, the sample is children under 16. The models typically include controls for the business cycle (e.g., current and lagged unemployment rate), demographic characteristics of the recipient (e.g., age, education, and race/ethnicity), other policy variables (e.g., maximum welfare benefit level, minimum wage, and EITC), and state and year fixed effects. Some models also include state-specific time trends.

In the remainder of this subsection, we first discuss the findings from these studies for a specific welfare policy, time limits; we then discuss the findings for waivers or TANF as a bundle. Results for income and poverty rates are discussed in turn.

8.3.1. Effects of Specific Reforms: Time Limits

Grogger's (forthcoming) study is the only econometric analysis of income to consider the effect of a specific TANF reform, in this case, time limits. In addition to estimating the impact of reform as a bundle (discussed below), Grogger's model of income is estimated with a dummy for the implementation of a time limit and an age interaction.[73] The point estimates suggest that time limits lower incomes by 3–6 percent (linear and log model, respectively) when the age of the youngest child is 13 or above (Panel A of Table 8.3). For each year below age 13, time limits further reduce income or leave it unchanged. However, none of these estimates is statistically significant. Given that Grogger finds significant negative effects of time limits on welfare use and only modest positive effects on earnings, family income might be expected to decline in states with time limits in effect. The insignificant effects may mean that other sources of income increase enough to offset the welfare declines. The estimated impact is also limited to the period before most recipients began reaching the time limit. Thus, the impact may change as more recipients exhaust their benefits.

[73]See the discussion in Chapter 4 of the interpretation of the age interaction with the time-limit dummy in Grogger's (forthcoming) analysis.

Table 8.3—Estimated Impact of Welfare Reform on Income and Poverty: Econometric Studies

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr. and Geogr.	Fixed Effects	Policy
A. Income													
Moffitt (1999)	CPS aggregated	women 16-54	77	95	Annual pre-tax family cash income	Level	Any waiver	393 (474)	1.3	U, U-1		S, Y, State time trends	B
	CPS aggregated	women 16-54, educ<12	77	95	Annual pre-tax family cash income	Level	Any waiver	-240 (909)	-0.8 (a)	U, U-1	A, E	S, Y, State time trends	B
		women 16-54, educ=12					Any waiver	569 (909)	1.9 (a)				
		women 16-54, educ=13-15					Any waiver	870 (909)	2.9 (a)				
		women 16-54, educ>16					Any waiver	-898 (909)	-3.0 (a)				
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	Annual pre-tax family cash income	Log	Any waiver	0.061 (0.013)	6.1	U, U-1, EG, each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	-0.006 (0.011)	-0.6				
		women 16-54, educ>12					Any waiver	-0.011 (0.009)	-1.1				
		women 16-54, educ<12					TANF	0.031 (0.031)	3.1				
		women 16-54, educ=12					TANF	0.022 (0.027)	2.2				
		women 16-54, educ>12					TANF	-0.011 (0.021)	-1.1				

Table 8.3—Continued

										Other controls			
Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Economy	Demogr. and Geogr.	Fixed Effects	Policy
Bitler, Gelbach and Hoynes (2001)	CPS micro data	women 16-54	84	98	Annual pre-tax family cash income, CPS families	Level	Any waiver	-78 (480)	-0.2	U, U-1, EG	R, MSA, CC	S, Y	B
							TANF and ever had waiver	-781 (1,412)	-1.5				
							TANF and never had waiver	-602 (1,304)	-1.2				
		women 19-54, educ < 12			Annual pre-tax family cash income, CPS families		Any waiver	1,466 (536)	5.5				
							TANF and ever had waiver	-70 (1,090)	-0.3				
							TANF and never had waiver	-1,431 (974)	-5.4				
		children under age 16			Annual pre-tax family cash income. CPS families		Any waiver	-331 (647)	-0.7				
							TANF and ever had waiver	-313 (1,854)	-0.7				
							TANF and never had waiver	-833 (1,746)	-1.8				
Grogger (forthcoming)	CPS micro data	female headed families 16-54	78	99	Annual pre-tax family cash income	Level	Any reform (waiver or TANF)	702 (570)	3.6	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-55 (81)	-0.3				
							Modified dummy>0 if time limit in place	-480 (986)	-2.5				
							Modified dummy>0 if time limit in place * (age of youngest child - 13)	31 (85)	0.2				
			78	99	Annual pre-tax family cash income	Log	Any reform (waiver or TANF)	0.098 (0.028)	9.8	U	A, E, R Young child A, # kids	S, Y, State time trends	B, MW, EITC
							Any reform * Age of youngest child	-0.007 (0.003)	-0.7				
							Modified dummy>0 if time limit in place	-0.056 (0.040)	-5.6				
							Modified dummy>0 if time limit in place * (age of youngest child - 13)	0.000 (0.003)	0.0				

Table 8.3—Continued

Study	Data	Sample population	Begin	End	Outcome	Dep. var.	Policy var.	Coeff. (s.e.)	% effect	Other controls: Economy	Other controls: Demogr. and Geogr.	Other controls: Fixed Effects	Other controls: Policy
B. Poverty													
Schoeni and Blank (2000)	CPS aggregated	women 16-54, educ<12	76	98	Poverty rate (annual pre-tax family cash inc.)	Level	Any waiver	-0.024 (0.006)	-8.2	U, U-1, EG, each *E	A, E, A*E, R	S, Y, state time trends, Y*E	B, B*E
		women 16-54, educ=12					Any waiver	0.001 (0.005)	1.0				
		women 16-54, educ>12					Any waiver	0.001 (0.004)	1.6				
		women 16-54, educ<12					TANF	-0.022 (0.013)	-7.8				
		women 16-54, educ=12					TANF	-0.011 (0.012)	-8.4				
		women 16-54, educ>12					TANF	0.003 (0.009)	4.6				
Bitler, Gelbach and Hoynes (2001)	CPS micro data	women 16-54	84	98	Annual pre-tax family cash inc. is below poverty line (CPS families)	Level	Any waiver	-0.006 (0.003)	-4.2	U, U-1, EG	R, MSA, CC	S, Y	B
							TANF and ever had waiver	-0.006 (0.007)	-4.0				
							TANF and never had waiver	-0.004 (0.007)	-2.6				
		women 19-54, educ < 12			Annual pre-tax family cash inc. is below poverty line (CPS families)		Any waiver	-0.031 (0.010)	-8.3				
							TANF and ever had waiver	-0.046 (0.022)	-12.3				
							TANF and never had waiver	0.004 (0.021)	1.1				
		children under age 16			Annual pre-tax family cash inc. is below poverty line (CPS families)		Any waiver	-0.008 (0.005)	-3.5				
							TANF and ever had waiver	-0.018 (0.009)	-8.3				
							TANF and never had waiver	-0.013 (0.008)	-5.9				

NOTES: Abbreviations: U=unemployment rate; U-1=lagged unemployment rate; EG=employment growth; A=age, E=education, R=race, MSA=Metropolitan Statistical Area (urban), CC=Central city; B=maximum welfare benefit, MW=minimum wage, EITC=Earned Income Tax Credit; S=state; Y=year.

(a) Percentage effects calculated using mean for entire sample.

8.3.2. Effects of Reform as a Bundle

Income

All four studies summarized in Panel A of Table 8.3 estimate models of the determinants of annual pretax family cash income, either in a linear or log model. Moffitt (1999) covers the shortest interval of time and finds no statistically significant effects of waivers (as approved) on income for women age 16–54. Schoeni and Blank (2000), with four additional years in their time series, find that the existence of a waiver (as implemented) raises family income by 6 percent for women with fewer than 12 years of schooling (a statistically significant effect), while the effect for TANF (as implemented) is 3 percent for the same group but statistically insignificant. There is no effect of waivers or TANF on the incomes of women with higher levels of education, which would be expected given that this group is less likely to be affected by welfare reform.

Like Moffitt (1999), Bitler, Gelbach, and Hoynes (2001) (hereafter BGH) find no effect of a major state waiver (or TANF) as implemented on income for all women. There is also no significant impact on income for the sample of children in the CPS. Like Schoeni and Blank, BGH do find a significant positive effect of any waiver on income for women with less than a high school education, but no effect of TANF (where the estimated impact of TANF is differentiated by whether the state previously had implemented a waiver). The estimated waiver impact on income for less educated women based on their model estimated in levels—a nearly 6 percent increase—is comparable to Schoeni and Blank's estimate based on a log-linear model. BGH also estimate alternative models using different living arrangement concepts to measure income. In addition to the results based on the traditional CPS concept of family reported in Table 8.3, they also consider income for the household, personal income (not relevant for the child sample), and family income, where income from related subfamilies is either pooled or not pooled with the primary family's income. The results show some sensitivity to the assumption about income pooling, with more pooling in the sample of less educated women leading to bigger impacts of waivers or TANF in waiver states. This result follows from BGH's estimated impacts of welfare reform on the number of adults in the household and families "doubling up," discussed in Chapter 7.

Grogger (forthcoming) combines waivers and TANF into one measure of any reform and finds that family income among female family heads rises by 4–10 percent for women whose youngest child is under age one, depending on whether the model is estimated in levels or logs (although the coefficient is only statistically significant at conventional levels in the log model).[74] The interaction between the reform dummy and age of the youngest child suggests that the impact on family income declines with the youngest child's age. (Again, this effect is significant only in the log model.) The larger effect Grogger finds may result from the fact that he limits his sample to female-headed families. This population is likely to be most affected by welfare reform, hence, the larger estimated impact.

---

[74]In Grogger's specification, the main effect of any reform (waivers or TANF) applies only to women whose youngest child is less than one. For these women, the interaction term between any reform and age of youngest child is zero.

All these models suffer from the collinearity problem discussed in earlier chapters. There simply is too little variation left to precisely estimate a TANF effect once state and year fixed effects are included in these models. Schoeni and Blank (2000) address this problem using changes in estimated year effects (in regression models with other controls) between 1995 and 1998 for less educated versus more educated women. This strategy does not lead to any more precisely estimated effects of TANF on the income of low-skill women.

Schoeni and Blank (2000) also consider the distributional effects of waivers and TANF by estimating equations of the log of the 20th and 50th percentiles of family income for women 16 to 54, and the 20th/50th ratio (not shown). Waivers are estimated to raise family incomes at the 20th and 50th percentiles by about 8–10 percentage points (a statistically significant effect) for women who drop out of high school. TANF, however, has a statistically significant and positive effect only on the 50th percentile of family income for women in the lowest education category. The lack of an effect at the 20th percentile leads to an estimated widening of the 20th–50th gap because of the implementation of TANF.

Poverty

Only Schoeni and Blank (2000) and BGH (2001) model the poverty rate, with results shown in Panel B of Table 8.3. Consistent with their findings for income, Schoeni and Blank report a statistically significant decline in the poverty rate for women with fewer than 12 years of schooling, equal to about 8 percent associated with the implementation of waivers and 8 percent associated with the implementation of TANF. They find a similar size effect for TANF using their alternative residual change methodology. There is no evidence to suggest that waivers or TANF affected the poverty rate for women with 12 or more years of schooling

BGH likewise find that the implementation of waivers significantly reduced poverty rates for less educated women by 8 percent, but there was no significant reduction for all women. The implementation of TANF on top of waivers further reduces poverty for the sample of less educated women, but the effect is zero for TANF implemented in states with no prior waivers and for TANF as a whole for all women. For the sample of children, only TANF implemented on top of waivers results in a statistically significant reduction in the poverty rate. Again, there is some sensitivity of the magnitude and significance of the impacts, depending on whether income is measured at the household or family level and depending on the concept of the family.

8.4. EVALUATING THE EFFECTS OF WELFARE REFORM ON INCOME, INCOME SOURCES, AND POVERTY

The econometric studies and demonstration studies reviewed thus far provide a varied picture of the impact on both income and poverty of specific welfare reform policies and groups of policies implemented simultaneously. What lessons can we learn from these varied studies, and can the different results be reconciled? To what extent can we expect welfare reform as a whole and the specific policy and program components to affect the incomes of recipients and the incomes of their families, as well as the poverty rate?

8.4.1. Effects of Specific Reforms

Of the studies reviewed in this chapter, the experimental evaluations are the most informative about the effects of specific reforms, but then only for a limited group of policies, principally financial work incentives, work requirements, and time limits. One econometric study also considered the impact of time limits on income only. However, these studies are not informative about the impact of other policy and program components of welfare reform, including sanctions.

Financial Work Incentives

Based on the results of eight high-quality experimental studies, we conclude that generous financial work incentives alone or in combination with work requirements, or those that are tied to hours of work, have the effect of raising incomes, with a corresponding reduction in poverty, an effect that is sustained up to three or four years (the length of the longest follow-up period). This effect is particularly strong in a program like MFIP (about \$100–\$125 per month in recipient combined income) with or without mandated work-related activities or a program like SSP (up to \$200–\$300 Canadian per month in family income or roughly \$150–\$200 U.S.) where income supports were linked to a requirement of full-time work (30 hours per week). Although these programs also produced significant reductions in poverty, the poverty rate for the treatment group that benefits from the higher income still exceeds 50 percent or more.

Part of the income gains result from increased benefit payments, given more generous disregards and benefit reduction rates (or the earnings supplement), and part of the income gains may stem from earnings gains, especially in programs that promote greater work effort through the incentive structure and work mandates. At the same time, the increase in welfare payments may mean that self-sufficiency as measured by the share of income from earnings will not necessarily improve and may even move in the direction of greater dependency. When incentives are weak, however, as in the case of the Vermont WRP program, there may be no income gains or poverty declines.

Mandatory Work-Related Activities

For work requirements alone—at least as implemented in the NEWWS evaluation, L.A. Jobs-First GAIN, and Indiana IMPACT—the results for income and poverty are mixed. Most programs show no statistically significant or economically meaningful effects on income or poverty, particularly compared to the large gains observed for programs that include generous financial work incentives. In the case of the five-year NEWWS follow-up, the impact estimates for recipient combined income are evenly divided in sign, and only one negative impact estimate is statistically significant. The two-year impact estimate for a broader measure of household income in L.A. Jobs-First GAIN is the only statistically significant positive impact. Impact estimates for poverty, which are only available for the 13 studies for the two-year follow-up, are more consistently negative, but at most two estimates are statistically significant depending on the income measure used. Thus, there is some evidence that these programs may modestly reduce poverty by raising incomes for those just below the poverty line. Evidence from NEWWS that deep poverty may increase in some programs is a counter to this more favorable assessment. These results on income, poverty, and self-sufficiency should not be too surprising, given that the earnings gains from these mandatory work programs are accompanied by reductions in welfare payments. As a result, these programs tend to raise self-sufficiency as measured by the share of income from earnings.

The programs that combine work requirements and financial work incentives produce gains in both earnings and cash assistance, thereby contributing to more favorable effects for income and poverty in contrast to the programs with work requirements only. This suggests that it is the financial work incentives that can be credited with this result. Two of the programs—MFIP and WRP—with their two treatment contrasts allow a more direct test of the role of the incentives component versus the work requirement component. For these two programs, when the impacts on income and poverty based on incentive-only programs shown in Panel A of Table 8.1 are compared with the alternative programs that combine incentives and work requirements shown in Panel D, the major share of the strong income gains and poverty reductions for MFIP and WRP can be attributed to the incentive component of the program. For example, the gain in recipient combined income for long-term recipients is $81 per month in the MFIP incentives-only program and $99 when incentives are combined with work requirements. In the case of WRP, the broader measure of household income is statistically significant and positive only for the incentives-only program, unlike MFIP where the combination of financial work incentives and work requirements produces favorable effects as well.

Time Limits

Programs that only implement time limits have yet to be experimentally evaluated. The econometric evidence provided in Grogger (forthcoming) suggests that time limits serve to reduce incomes. However, given the time period covered by the study, these estimates pertain to the impact of time limits before they are binding for most recipients. Moreover, none of Grogger's point estimates are statistically significant.

FTP and Jobs First, with follow-up periods that include the period prior to time limits being reached and after time limits are reached, provide some insights into the mechanical effects of time limits. They both suggest that the favorable income gains observed in the pre–time limit period fade and are then reversed in the post–time limit period. We are not able to assign a significance level to these estimated negative impacts. These inferences, while not derived from the experimental design of these two studies, are nevertheless suggestive that time limits serve to reduce income as recipients begin to exhaust their benefits. The impact of FTP and Jobs First on poverty is not reported and therefore cannot be assessed with these studies.

8.4.2. Effects of Reform as a Bundle

The econometric studies reviewed in Section 8.3 suggest that welfare reform as a whole, and specifically TANF, has resulted in an increase in family pretax cash income for low-skilled women and single women with children, where the estimated impact is largest for the latter group. Focusing on the point estimates, Schoeni and Blank's (2000) results suggest that family incomes for women with less than a high-school education increased by about 3 percent in the post-TANF period, and by a smaller amount (2 percent) for women who completed high school. Bitler, Gelbach, and Hoynes (2001) find a nearly 6 percent increase in income for women with less than a high school degree from implementing a waiver (and no significant effect for TANF implementation). Grogger's (forthcoming) estimate translates into an 8 percent increase in income for a female-headed family whose youngest child is age 3 as a result of implementing a waiver or TANF. The effect falls as the age of the youngest child increases, reaching an impact of 5 percent for single mothers whose youngest child is age 7 and zero when the youngest child is age 14.

Grogger's larger estimate is consistent with the fact that more of the sample of single-headed families are at risk of welfare compared with the sample of all women. To the extent that women without children and married women are mostly unaffected by welfare reform, Schoeni and Blank's estimates will be biased downward from the true impact on women at risk of welfare receipt. Then again, by selecting on marital status and the presence of children, Grogger's approach may be biased to the extent that welfare reform changes the composition of the at-risk population. Grogger argues that this bias will be negative, so that his estimates may be, if anything, too low. It must be stressed, however, that all of the econometric studies essentially measure the impact of reform as a bundle prior to the period when time limits become binding. Thus, they must be viewed as "pre–time limit" impacts.

Assuming these studies are correct in placing a bound on the impact of welfare reform as a bundle on family income from 3–9 percent in the pre–time limit period, how does this inference compare with the findings from the demonstration studies? Several methodological differences between the econometric analyses and the demonstration studies should be reiterated. First, the income concepts may differ, with less comprehensive income concepts often used in the demonstration studies compared with the CPS-based studies. However, in most demonstration studies, the most significant drivers of change in family income are changes in the recipient's earnings and cash transfers; there are few instances where earnings from other household members or income from other sources, as reported in surveys, differ between the treatment and control groups during the follow-up period. Thus, the absence of data on other sources of income for the recipient or other family members will not necessarily have a large negative bias on the income impacts. Second, the demonstration studies, as noted in Chapter 3, cannot be used to estimate the full effects of TANF, because they do not capture program entry effects, i.e., they do not capture the impacts for potential welfare recipients. They also do not represent the weighted combination of reforms implemented under TANF. Even if they did, a demonstration program may not reproduce the effects of a large-scale program.

With these caveats in mind, we cautiously compare findings across the two types of studies. Since the last year of data analyzed in any of the econometric studies is 1999, only a small share of sample members in the CPS are likely to have had their benefits cut off because of time limits. Thus, the econometric results might be usefully compared with the impact estimates from the pre–time limit impacts of the demonstration studies that evaluated time limits combined with other policies (e.g., work requirements and financial work incentives). Panel F of Table 8.1 provides examples of such programs and their early impacts, which range from small (0–2 percent for VIP/VIEW and ABC up to two years post-randomization) to moderate (6–11 percent for recipient combined income in FTP and Jobs First). Since few states have implemented financial work incentives as generous as Connecticut's, the impacts on the high end are not likely to be observed for TANF as implemented. Thus, the impacts estimated using the CPS data seem plausible given the range of the small and moderate impacts in the experimental studies with more modest financial work incentives.

One caution, however, is that the positive impact of TANF as a bundle on incomes of either recipients or the at-risk population may disappear once the TANF time limits have had an opportunity to produce both behavioral and mechanical effects. The demonstration studies of programs that include time limits suggest that the initial income gains are not sustained once time limits become effective and income from welfare payments goes to zero. At the same time, it is possible that the earnings gains may increase with time off welfare as former recipients obtain additional labor market experience. Such longer-term earnings gains may be sufficient to offset the welfare benefit losses. Additional years of data for the CPS can be used to explore which effect may dominate for the larger population at risk of welfare.

This same analysis applies to the estimated impact of welfare reform on the poverty rate. Schoeni and Blank's econometric estimates suggest a 2 percentage-point reduction in the poverty rate for women with less than a high school education, which translates into an 8 percent reduction in the poverty rate. Bitler, Gelbach, and Hoynes (2001) find a similar impact estimate for the poverty rate for women who drop out of high school based on implementing waivers and an additional reduction in the poverty rate when TANF is implemented in states that had a major waiver. Unfortunately, none of the demonstration studies in Panel F of Table 8.1 reported effects for poverty. Judging from the income effects, we would certainly expect the demonstration studies with time limits to generate reductions in the poverty rate in the short run corresponding to the income gains. Likewise, if incomes fall after time limits become binding, the poverty rate may be expected to rise. Thus, while the estimated effect from the econometric studies is plausible in light of the experimental evidence, the antipoverty effect of TANF as a whole may be short-lived.

8.5. CONCLUSIONS

The studies reviewed in this chapter suggest that some welfare reform components can raise incomes and reduce poverty, although this result is not associated with all policy components, and there is reason to believe that some of the initially favorable effects will not persist over a longer horizon. The econometric and experimental evidence suggests that welfare reform as a bundle has raised incomes for disadvantaged women and lowered their poverty rate, at least in the short run before time limits have become binding. As time limits begin to affect a larger share of the recipient population, this outcome may not be sustained with the decline in welfare income that accompanies reaching the time limit. In addition, these favorable effects in the short run may mask important distributional changes. There is some limited evidence from econometric and experimental studies that reductions in poverty may be accompanied by an increase in the rate of deep poverty.

Generous financial work incentives—high earned income disregards and low benefit reduction rates inside the welfare system or earnings supplements outside the welfare system—generate the strongest income gains and antipoverty effects, especially when the incentive structure encourages full-time work. However, these programs can lower self-sufficiency because the increased welfare payments or earnings supplements may exceed the increased earnings. Work requirements alone have relatively weak effects on family income and poverty, but they do raise self-sufficiency by increasing the fraction of income from earnings. Finally, time limits, once they become binding, may erase income gains made possible by generous financial work incentives associated with working while on welfare.

While the antipoverty effectiveness of policy components such as financial work incentives appears to be quite robust, the income levels are relatively low and consequently the rates of poverty are relatively high among those who benefit from all the welfare reforms considered in this chapter. For example, in the experimental studies that evaluate financial work incentives (Panels A, B, and D of Table 8.1), the poverty rate for the treatment group falls below 50 percent just one time for the study populations and time periods measured in the table. The fraction with very low incomes may not move much at all, and many of those raised above the poverty line still remain "near poor." Changes in poverty status over time as a result of welfare reform remain relatively unexplored.

It is important to keep in mind that much of this chapter has focused on income measures that are rarely as comprehensive as would be desired to evaluate changes in well-being. Most econometric studies consider only family income before taxes and exclusive of in-kind benefits. Many of the experimental evaluations likewise use a concept of income limited to earnings and social welfare benefits (e.g., welfare and food stamp benefits in the U.S. experiments) for the recipient. Even if a comprehensive income measure were available, it may not fully reflect the individual's or family's command over resources. For that reason, in the next chapter, we focus on results for broader measures of well-being, which may provide a better gauge of living standards than what can be gleaned from examining income alone.

Chapter Nine

OTHER MEASURES OF WELL-BEING

9.1. BACKGROUND

In seeking to understand the full impact of welfare reform, policymakers are interested in broader measures of well-being beyond the outcomes typically considered, such as welfare use, employment and earnings, or income and poverty. For example, income may not fully capture a family's command over resources if it has savings available to draw on during periods of low income, if it is able to borrow money from family or friends, or if it can incur debt to pay for unexpected costs. In addition, in the transition from welfare to work, there may be an increase in work-related expenses (e.g., child care, transportation, and clothing) that will affect the ability to consume other goods and services but that will not be reflected in income. For this reason, consumption or expenditures are often considered a better gauge of a family's well-being, because this measure reflects the value of what a family actually consumes in total or for specific categories of purchases such as food. The importance of food consumption for overall nutrition and health status has led to the development of specific measures of food insecurity to capture problems with having enough money to buy food or experiencing periods when there is not enough food or meals are skipped. Likewise, access to health care is considered by some to be a key measure of well-being. Health care coverage for adult and child family members, whether from public or private sources, is one indicator of whether a family can afford visits to medical professionals for preventative care or to treat acute or chronic conditions. Housing conditions and the quality of the neighborhood are other indicators of the circumstances under which families live.

These various indicators of well-being may be affected by welfare reform through either direct or indirect mechanisms. Some features of state waivers or state TANF plans, for example, affect health care coverage directly by providing transitional medical assistance. In other cases, leaving welfare may mean a loss in public health insurance coverage (either due to changes in eligibility or administrative problems in obtaining coverage when eligible) that may not be replaced by employment-based health insurance. The influences may also be more indirect, with changes in welfare use and employment leading to changes in family income that, in turn, affect decisions about consumption and savings or about residential location. For example, if incomes rise or income flows become more stable as a result of welfare reform, we would expect food insecurity to be less of a problem. Moreover, having more income may allow a family to move to a better quality home or safer neighborhood, to put money in a savings account, or to purchase or maintain an automobile. Then again, if incomes fall or remain the same on average but with more fluctuation, there may be an increase in food or housing insecurity, and assets may decline or a family may incur new debt.

Recent descriptive analyses provide some perspective on how former welfare recipients and the low-income population more generally are faring, as captured by various measures of well-being. A recent summary of 15 state "leaver" studies funded by USDHHS reveals that former welfare recipients are at risk of various forms of material hardship (USDHHS, 2001a). For example, 7–45 percent of adult leavers have no health insurance. The comparable range for children in households of leavers is 8–33 percent. Various measures of food insecurity indicate that one-fifth to one-half of former recipients experience problems with having enough money to buy food, running out of food, skipping meals, and other forms of food insecurity and hunger.[75] The prevalence of forms of housing insecurity and medical hardships was somewhat lower. In five of the state studies, 13–30 percent of single-parent leavers reported they were worse off financially overall after leaving welfare, while 46–68 percent reported they were better off.

While the collection of leaver studies is informative, these studies are not designed to capture the causal impact of welfare reform as a whole, or of specific policy components in particular, on other measures of well-being. Like the preceding chapters, the remainder of this chapter is devoted to a summary of the causal studies from the experimental and econometric literatures. In particular, we focus on the following domains of well-being: material hardship and food insecurity, health insurance coverage, housing hardships and neighborhood quality, and asset ownership. For these four outcome domains, all the evidence derives from the experimental studies summarized in Tables 3.5 and 3.6; we are not aware of any econometric studies that use the DoD methodology to assess the causal impact of waivers or TANF as a bundle or specific policy reforms on these measures of well-being.[76] [77] [78]

---

[75]The problem of food insecurity among low-income female-headed households is confirmed in other recent survey and ethnographic data (e.g., Polit, London, and Martinez, 2000).

[76]Meyer and Sullivan (2001) examine consumption data (in aggregate and for subcomponents) for the Consumer Expenditure Survey and Panel Study of Income Dynamics. However, they do not explicitly estimate the impact of welfare waivers or TANF as a bundle, or of specific welfare reform policies. Rather, they test for differences across four time intervals: 1984–1990, 1991–1993, 1994–1995, and 1996–1998. These time intervals are designed to capture the period prior to waivers; the initial welfare waiver period when the EITC also began to expand; the more active waiver period leading up to TANF when the EITC also continued to expand; and the post-TANF period. Their DoD methodology compares consumption trends for single mothers with single women and married mothers. Their findings indicate that consumption for single mothers has not declined in the 1990s, either in absolute terms or relative to the comparison groups of women. In fact, consumption may have improved somewhat, even for women with the least education. They interpret their results to suggest that recent changes in tax and welfare policy have not had detrimental effects on the material conditions of single mothers and their children.

[77]Borjas (2001a) uses data from the CPS for 1994 to 1998 to examine the differential impact on food insecurity for immigrants in states that did and did not extend state-funded assistance to immigrants after PRWORA eliminated eligibility for TANF and food stamps for some groups of immigrants (specifically, nonrefugee, noncitizen households that arrived in the United States after 1996). DoDoD models are estimated for affected and unaffected immigrant groups, in more and less generous states, for pre- versus post-PRWORA years, but Borjas does not consider the impact of welfare waivers, of TANF as a bundle, or of specific reform policies. His results show that food insecurity increased the most between 1994 and 1998 among those immigrants affected by the PRWORA changes in eligibility and living in states that did not extend state funding to cover immigrant groups ineligible under PRWORA.

[78]Hurst and Ziliak (2001) use data from the 1994 and 1999 PSID asset modules to examine the relationship between changes in household savings over the five-year period and an array of welfare reform variables, including benefit levels, time limits, asset limits for liquid assets and vehicles, and availability of individual development accounts (IDA). Since they essentially have a single cross section, they cannot employ the DoD methodology. Their results suggest a modest positive effect on liquid savings for those at risk of welfare receipt from policies that increased asset limits on liquid savings and instituted IDAs. In addition to ordinary-least-squares estimation, they also use instrumental variables to control for possible policy endogeneity, but the coefficients on the policy variables are not precisely estimated.

All the experimental studies we review in this chapter base their impact analysis on survey-based measures of well-being; administrative data simply cannot fully capture these broader measures of a family's circumstances.[79] For example, material hardship is typically captured by recipients' perceptions of financial strain, or experiences with specific problems affecting their housing conditions (e.g., a leaky roof or ceiling), neighborhood (e.g., crime, assault, or burglaries), material needs (e.g., could not pay rent or mortgage), and food security (e.g., did not have enough food to eat or used a food bank).

The reliance on survey data raises a few methodological concerns. First, in many cases, the sample sizes available for analysis are smaller than those shown in Table 3.5. Survey samples in demonstration studies are often smaller than the overall study population by design, and survey nonresponse further reduces the sample of respondents. Smaller samples will reduce the statistical power of the study for detecting small- to moderate-sized effects, as well as differences for subgroups.

Second, compared with the outcomes reviewed in prior chapters, a larger number of the experimental studies do not include any broader measures of well-being in their impact analyses (or their analyses to date). Of the studies we consider, CWPDP, SSP Plus, SSP Applicants, IMPACT, TSMF, FIP, VIP/VIEW, AWWDP, FDP, PPI, PIP, and ABC do not assess measures in the domains we list above. For those studies that do focus on other measures of well-being, many include only a few of the various measures that could be collected. For example, the inclusion of a measure of health insurance coverage for adults and children is quite common, while only a handful of the studies collect data on one or more asset measures, such as a savings account or car ownership. Because all measures are not available for all studies, it is more difficult to draw solid conclusions about the impact of the welfare reform policy or policies being evaluated in each demonstration study.

Third, many of these measures of well-being can be conceptualized and measured in a number of different ways, and there is no assurance of uniformity across studies in the measures actually used. Health insurance coverage is at one end of the spectrum, with most demonstration studies measuring whether the respondents or their children have any form of public or private health care coverage at the time of follow-up. The amount of savings or whether the respondent owns a vehicle is also measured in a similar way across the few studies that include such indicators. At the other extreme, food insecurity is measured in a different way in almost every study that includes one or more measures.[80] In some cases, a specific question is asked (Did you use a food bank in the last three months? Has your family had enough to eat in the last month?); other studies report a multi-item scale of food insecurity, such as the one developed by the U.S. Department of Agriculture (USDA).[81] As the questions above suggest, for some measures, a positive impact would indicate a favorable outcome, whereas for other measures, a negative impact would indicate a favorable outcome. All these factors make comparisons across studies within and between the classifications we have

---

[79]In some cases, administrative data are combined with survey data. For example, the former may provide information on Medicaid coverage while the latter are used to determine whether other sources of public or private coverage are available. The data are then combined to form a measure of whether the individual has any health care coverage.

[80]Food insecurity has been defined to exist when "the availability of nutritionally adequate and safe foods or the ability to acquire acceptable foods in socially acceptable ways is limited or uncertain" (as quoted in Polit, London, and Martinez, 2001, p. 48).

[81]The short version of the USDA scale (now administered annually in the CPS) includes six items to assess food-related hardship.

defined more problematic. We will revisit these issues when we synthesize the findings across the various studies at the end of this chapter.

We turn to the experimental literature in the next section, summarizing the findings from the studies that include other measures of well-being. The results from the available studies are synthesized in the third section. The final section concludes the chapter. The limited amount of information regarding subgroup differences in the outcomes covered in this chapter is summarized in Appendix A.

9.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECT OF WELFARE REFORM ON OTHER MEASURES OF WELL-BEING

Tables 9.1, 9.2, 9.3, and 9.4 summarize the results for the random assignment studies in terms of the following measures of well-being: material hardship and food insecurity (9.1); health insurance coverage (9.2); residential moves, housing hardships, and neighborhood quality (9.3); and assets (savings and vehicle ownership) (9.4). In each case, the tables report the specific measure or measures available for the population served, the control group mean, and the impact estimate and its statistical significance.

In this section, we organize our discussion by the four outcome domains. In the section that follows, we synthesize these results by the reform policy or policies being evaluated, using the classification scheme outlined in Chapter 3 (see Table 3.5), considering all the well-being measures within each class of studies. For our six-way classification scheme, the studies in the fifth group (shown in Panel E) do not include any of the measures of well-being covered in this chapter.

9.2.1. Material Hardship and Food Insecurity

As seen in Table 9.1, measures of material hardship and food insecurity are available for at least one program in Panels A through D and F, with follow-up periods that range from 18 months to four years. Of the two programs that focus on financial work incentives (Panel A), only MFIP-IO provides measures of material hardship. Neither of these two measures are statistically significant with the exception of a significant favorable impact on the number of material hardships in the last year as of the three-year follow-up for MFIP applicants. With one exception, MFIP-IO also had no statistically significant impact on two measures of food insecurity. Again, for applicants in the study, MFIP-IO had a statistically significant favorable impact on whether the family had enough to eat in the last month, a measure of food insecurity.[82] WRP-IO had no statistically significant effects on the two measures of food insecurity included in the 42-month follow-up survey. Although most of the effects for WRP-IO and MFIP-IO are statistically insignificant, with one exception (a measure of meals skipped by children for MFIP-IO long-term recipients), they are all in the favorable direction.

[82]Note that the food insecurity measure is specific to the group of families with children age 5–12 at the 36-month follow-up, an even smaller study population.

Both New Hope and SSP (Panel B), in their evaluations of financial work incentives tied to hours of work, consider food insecurity, while New Hope also includes a measure of material hardship. For New Hope, the only significant impact is a favorable effect on the number of material hardships at the two-year follow-up interview for poor families not employed full-time at random assignment. For the food insecurity measures, SSP shows significant favorable effects for two of the three measures assessed at the 18- and 26-month follow-up interviews.

Among the evaluations of programs that focus on mandatory work-related activities (Panel C), L.A. Jobs-First GAIN is the only one to include measures in this domain. That study shows an unfavorable effect as of the two-year follow-up on the multi-item measures of food insecurity and food insecurity with hunger, with a statistically significant impact on the latter outcome. In this case, 13.3 percent of the single-parent recipient control group report five or more problems associated with food insecurity and hunger, compared with 18.8 percent for the treatment group, a 41 percent difference.

For the studies that focus on financial work incentives and mandatory work-related activities (Panel D), only WRP and MFIP assess measures in these domains. In terms of material hardship, MFIP produces significant favorable effects on at least one measure of financial strain and/or the number of material hardships for both recipients and applicants. None of the food insecurity measures for WRP or MFIP are statistically significant, and the signs are mixed.

Of the studies that focus on TANF-like bundles of reforms (Panel F), both FTP and Jobs First consider material hardship and food insecurity using similar measures, while EMPOWER includes only a measure of food insecurity. In the case of FTP, the three hardship measures show favorable effects as of the four-year follow-up, two of them statistically significant. In particular, the fraction with three or more “severe” hardships was 8.8 percent in the treatment group compared with 14.1 percent for the control group, mostly because of the reduction in housing and neighborhood problems (discussed below).[83] Sixty-nine percent of the treatment group reported at the four-year follow-up that they usually had enough money at the end of the month compared with 63 percent of the control group. The two hardship measures reported for Jobs First as of the three-year follow-up are not statistically significant and mixed in sign. There is no impact of FTP or Jobs First on food insecurity, which is measured three ways.[84] The impact on the use of food banks and soup kitchens is also insignificant in the EMPOWER evaluation.

9.2.2. Health Insurance Coverage

Measures of health insurance coverage for adults and children, reported in Table 9.2, are included in a larger number of studies, with a follow-up period as long as five years. Of the programs reporting impacts in this domain, only WRP and Jobs First provide transitional health benefits that exceed what is available to control group members. Lower rates of coverage for

---

[83]There are five possible “severe” hardships, specifically four or more neighborhood problems, two or more housing problems, four or more material hardships, two or more social services used, and food insecurity with hunger.

[84]The three measures are based on a subset of the USDA Household Food Security Scale administered in the CPS. Respondents are classified as food secure (no or only one hardship); food insecure without hunger (two to four hardships); or food insecure with hunger (five or six hardships).

Table 9.1—Estimated Impact of Welfare Reform on Material Hardship and Food Insecurity: Random Assignment Studies

			Material hardship				Food insecurity			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives										
WRP-IO	Single-parent recipients and applicants	S					Sometimes or often not enough food in past 12 mos at 42-mo FU (%)	22.9	-2.6	-11.4%
		S					Often true in last 12 mos that food bought didn't last and didn't have money to get more at 42-mo FU (%)	18.0	-1.9	-10.6%
MFIP-IO	Urban single parents recipients	S	Perception of financial strain at 36-mo FU (range of 1=least strain to 4=most strain)	3.0	-0.1	-3.3%	Family had enough to eat in last month at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	80.1	4.8	6.0%
		S	Number of material hardships (7 items) during past 12 mos at 36-mo FU	1.5	-0.1	-6.7%	In last month, any children skip a meal because not enough money for food at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	3.9	1.1	28.2%
	Urban single parents applicants	S	Perception of financial strain at 36-mo FU (range of 1=least strain to 4=most strain)	2.8	0.0	-1.4%	Family had enough to eat in last month at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	85.6	7.5 **	8.8%
		S	Number of material hardships (7 items) during past 12 mos at 36-mo FU	1.5	-0.3 *	-16.6%	In last month, any children skip a meal because not enough money for food at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	4.1	-2.2	-53.7%
B. Programs that focus on financial work incentives tied to hours of work										
New Hope	Poor families employed FT at RA	S	Number of material hardships at 2-year FU	1.9	0.0	-1.6%	Insufficient food in last month at 2-year FU (%)	7.7	1.2	15.6%
	Poor families not employed FT at RA	S	Number of material hardships at 2-year FU	2.4	-0.3 ***	-12.5%	Insufficient food in last month at 2-year FU (%)	13.8	-1.6	-11.6%
SSP	Single-parent recipients	S					Used food bank in last 3 months at 18-mo FU (%)	21.1	-2.0 *	-9.5%
		S					Used food bank in last 3 months at 36-mo FU (%)	18.8	-1.0	-5.3%
		S					Could not get groceries at 36-mo FU (%)	34.4	-4.2 ***	-12.2%
C. Programs that focus on mandatory work-related activities										
LA Jobs-1st GAIN	Single-parent recipients and applicants	S					Experienced food insecurity (2 or more problems) at yr 2 (%)	48.6	4.5	9.3%
		S					Experienced food insecurity with hunger (5 or more problems) at yr 2 (%)	13.3	5.5 **	41.4%

Table 9.1—Continued

Name	Cases served	Data	Material hardship: Measure	Control mean	Impact	%	Food insecurity: Measure	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities										
WRP	Single-parent recipients and applicants	S					Sometimes or often not enough food in past 12 mos at 42-mo FU (%)	22.9	-1.2	-5.2%
		S					Often true in last 12 mos that food bought didn't last and didn't have money to get more at 42-mo FU (%)	18.0	-1.5	-8.3%
MFIP	Urban single-parent recipients	S	Perception of financial strain at 36-mo FU (range of 1=least strain to 4=most strain)	3.0	-0.1 *	-3.3%	Family had enough to eat in last month at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	80.1	-0.3	-0.4%
		S	Number of material hardships (7 items) during past 12 mos at 36-mo FU	1.5	0.1	6.7%	In last month, any children skip a meal because not enough money for food at 42-mo FU (for those with child aged 5 to 12 at FU) (%)	3.9	2.0	51.3%
	Urban single-parent applicants	S	Perception of financial strain at 36-mo FU (range of 1=least strain to 4=most strain)	2.8	-0.1 *	-3.2%	Family had enough to eat in last month at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	85.6	4.5	5.3%
		S	Number of material hardships (7 items) during past 12 mos at 36-mo FU	1.5	-0.2 *	-10.6%	In last month, any children skip a meal because not enough money for food at 42-mo FU (for those with child aged 5 to 12 at FU) (%)	4.1	0.2	4.9%
E. Programs that focus on other individual reforms										
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
EMPOWER (a)	Recipients	S					Used food bank or soup kitchen since RA as of 30-mo FU (%)	30.9	-1.2	-3.9%
FTP	Recipients and applicants	S	Four or more material hardships at 4-year FU (%)	19.9	-1.7	-8.5%	Food secure (USDA 6-item scale) at 4-year FU (%)	64.2	1.8	2.7%
		S	Three or more "severe" hardships at 4-year FU (%)	14.1	-5.3 ***	-37.6%	Food insecure (USDA 6-item scale) at 4-year FU (%)	18.8	-0.5	-2.7%
		S	Usually has enough money at end of mo. at 4-year FU (%)	63.0	6.0 ***	9.5%	Food insecure with hunger (USDA 6-item scale) at 4-year FU (%)	17.0	-1.3	-7.4%
Jobs First	Recipients and applicants	S	Four or more material hardships at 3-year FU (%)	16.9	-0.8	-5.0%	Food secure (USDA 6-item scale) at 3-year FU (%)	59.8	1.5	2.5%
		S	Three or more "severe" hardships at 3-year FU (%)	11.8	1.1	9.2%	Food insecure (USDA 6-item scale) at 3-year FU (%)	18.3	-1.2	-6.7%
		S					Food insecure with hunger (USDA 6-item scale) at 3-year FU (%)	21.8	-0.3	-1.1%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: S=survey data; FU=follow-up.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.
(a) Phoenix site only, cash assistance.

Table 9.2—Estimated Impact of Welfare Reform on Health Insurance Coverage: Random Assignment Studies

			Recipient health insurance coverage				Children health insurance coverage			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives										
WRP-IO	Single-parent recipients and applicants	S	Respondent covered by any HI at 42-mo FU (%)	81.6	0.1	0.1%	Child covered by any HI at 42-mo FU (%)	84.2	-1.5	-1.8%
MFIP-IO	Urban single parents recipients	S	Respondent has HCC at 36-mo FU (%)	83.9	1.6	1.9%				
		S	Respondent had HCC continuously for 36 mos. after RA (%)	61.3	13.6 ***	22.2%	Children had HCC continuously for 36 mos after RA (for those with child aged 5 to 12 at FU) (%)	67.0	11.7 ***	17.5%
	Urban single parents applicants	S	Respondent has HCC at 36-mo FU (%)	73.9	5.0	6.8%				
		S	Respondent had HCC continuously for 36 mos. after RA (%)	50.0	17.9 ***	35.8%	Children had HCC continuously for 36 mos after RA (for those with child aged 5 to 12 at FU) (%)	62.7	13.3 **	21.2%
B. Programs that focus on financial work incentives tied to hours of work										
New Hope	Poor families employed FT at RA	S	Any periods without HI at 2-year FU (%)	55.2	-8.5	-15.4%				
	Poor families not employed FT at RA	S	Any periods without HI at 2-year FU (%)	60.5	-11.3 ***	-18.7%				
C. Programs that focus on mandatory work-related activities										
LA Jobs-1st GAIN	Single-parent recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	93.4	-1.3	-1.4%	Children have HCC at end of 2-year FU (%)	92.9	-0.3	-0.3%
Atlanta LFA	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	86.0	-2.4	-2.8%	All dependent children have HCC at end of 2-year FU (%)	85.6	0.5	0.6%
		S	Respondent has HCC at end of 5-year FU (%)	72.4	-1.4	-1.9%	All dependent children have HCC at end of 5-year FU (%)	84.5	0.6	0.7%
Grand Rapids LFA	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	86.0	-3.3	-3.8%	All dependent children have HCC at end of 2-year FU (%)	85.7	-1.4	-1.6%
		S	Respondent has HCC at end of 5-year FU (%)	77.7	-2.6	-3.3%	All dependent children have HCC at end of 5-year FU (%)	81.8	-3.0	-3.7%
Riverside LFA	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	87.3	-1.8	-2.1%	All dependent children have HCC at end of 2-year FU (%)	88.4	-3.3 **	-3.7%
		S	Respondent has HCC at end of 5-year FU (%)	80.3	-2.0	-2.5%	All dependent children have HCC at end of 5-year FU (%)	83.2	-1.3	-1.6%
Portland	Recipients and applicants; no cases with substantial barriers	S	Respondent has HCC at end of 2-year FU (%)	90.4	-3.3	-3.7%	All dependent children have HCC at end of 2-year FU (%)	88.6	-4.8	-5.4%
		S	Respondent has HCC at end of 5-year FU (%)	80.6	-6.0	-7.4%	All dependent children have HCC at end of 5-year FU (%)	80.2	-4.7	-5.9%

Table 9.2—Continued

Name	Cases served	Data	Recipient health insurance coverage: Measure	Control mean	Impact	%	Children health insurance coverage: Measure	Control mean	Impact	%
Atlanta HCD	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	86.0	-2.2	-2.6%	All dependent children have HCC at end of 2-year FU (%)	85.6	-0.8	-0.9%
		S	Respondent has HCC at end of 5-year FU (%)	72.4	1.6	2.2%	All dependent children have HCC at end of 5-year FU (%)	84.5	-1.1	-1.3%
Grand Rapids HCD	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	86.0	-1.7	-2.0%	All dependent children have HCC at end of 2-year FU (%)	85.7	0.5	0.6%
		S	Respondent has HCC at end of 5-year FU (%)	77.7	0.1	0.1%	All dependent children have HCC at end of 5-year FU (%)	81.8	-2.5	-3.1%
Riverside HCD	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	87.5	-0.8	-0.9%	All dependent children have HCC at end of 2-year FU (%)	88.8	-0.7	-0.8%
		S	Respondent has HCC at end of 5-year FU (%)	80.0	0.3	0.4%	All dependent children have HCC at end of 5-year FU (%)	82.1	3.2	3.9%
Columbus Integrated	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	85.0	-5.2 *	-6.1%	All dependent children have HCC at end of 2-year FU (%)	86.3	-6.3 **	-7.3%
Columbus Traditional	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	85.0	0.8	0.9%	All dependent children have HCC at end of 2-year FU (%)	86.3	0.2	0.2%
Detroit	Recipients and applicants	S	Respondent has HCC at end of 2-year FU (%)	92.0	-0.9	-1.0%	All dependent children have HCC at end of 2-year FU (%)	90.9	-0.6	-0.7%
Oklahoma City	Applicants	S	Respondent has HCC at end of 2-year FU (%)	70.9	-3.3	-4.7%	All dependent children have HCC at end of 2-year FU (%)	72.5	-9.0 **	-12.4%
D. Programs that focus on financial work incentives and mandatory work-related activities										
WRP	Single-parent recipients and applicants	S	Respondent covered by any HI at 42-mo FU (%)	81.6	-2.3	-2.8%	Child covered by any HI at 42-mo FU (%)	84.2	-4.2 *	-5.0%
MFIP	Urban single-parent recipients	S	Respondent has HCC at 36-mo FU (%)	83.9	1.6	1.9%				
		S	Respondent had HCC continuously for 36 mos. after RA (%)	61.3	7.9 **	12.9%	Children had HCC continuously for 36 mos after RA (for those with child aged 5 to 12 at FU) (%)	67.0	8.5 **	12.7%
	Urban single-parent applicants	S	Respondent has HCC at 36-mo FU (%)	73.9	4.4	6.0%				
		S	Respondent had HCC continuously for 36 mos. after RA (%)	50.0	12.9 ***	25.8%	Children had HCC continuously for 36 mos after RA (for those with child aged 5 to 12 at FU) (%)	62.7	7.2 *	11.5%
E. Programs that focus on other individual reforms										
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
FTP	Recipients and applicants	S	Respondent has no HI at 4-year FU (%)	38.4	0.9	2.3%	Children have no HI at 4-year FU (%)	15.7	1.2	7.6%
Jobs First	Recipients and applicants	S	Respondent has no HI at 3-year FU (%)	18.4	-4.4 ***	-24.2%	Children have no HI at 3-year FU (%)	4.6	-0.7	-14.5%

NOTES: For full program names and citations, see Table 3.4. Abbreviations: S=survey data; FU=follow-up; RA=random assignment.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

treatment group members may indicate a loss of eligibility for public programs (e.g., Medicaid) that is not replaced by private programs (e.g., employer-provided coverage), or it may result from administrative problems obtaining or maintaining coverage for public programs for which the individual is eligible

In terms of programs that focus on financial work incentives (Panel A), WRP-IO had a very small and statistically insignificant impact on health insurance coverage as of the 42-month follow-up for the respondent and the child. This is despite the fact that WRP-IO provided three years of transitional Medicaid benefits for recipients leaving welfare for work in contrast to only one year for AFDC recipients in the control group. However, since Vermont has an array of health insurance programs available for low-income families, recipients in the control group were able to obtain coverage from other sources, so there is little treatment-control difference in coverage rates for adults or children.

MFIP-IO also shows no significant impact on coverage at the time of the 36-month follow-up for the respondent. (A comparable measure is not available for children.) In contrast, the fraction of adults and children with continuous health care coverage during the entire 36-month period following random assignment is higher by 12–18 percentage points, differences that are statistically significant. This effect is consistent with the higher rates of welfare usage, which automatically qualified the family for Medicaid coverage among MFIP-IO participants at the time MFIP was in place, as discussed in Chapter 4.

New Hope is the only study in Panel B (programs that focus on financial work incentives tied to hours of work) that reports impacts on health insurance coverage. Poor families not employed full-time at the time of random assignment into New Hope were less likely by 11 percentage points to experience any periods without health insurance in the two-year interval since random assignment compared with the control group. The impact estimate for the group employed full-time at random assignment is nearly as large but not statistically significant. New Hope provided subsidized health insurance, which is attributed with reducing gaps in coverage. Even so, a sizeable fraction (47–49 percent) of the treatment group experienced one or more gaps in health insurance coverage over two years.

Among the programs that focus on work mandates (Panel C), measures of health care coverage for adults and children are available for L.A. Jobs-First GAIN two years after randomization and all 11 NEWWS programs up to five years after randomization. None of these programs provided transitional Medicaid coverage for the treatment group that differed from the comparison group. Almost all the point estimates are negative, indicating that programs that require mandatory work activities tend to reduce the probability of health insurance coverage after two years for both adults and children. However, all the effects are very small and only 4 of the 38 estimates are statistically significant. For the 7 NEWWS programs with five-year follow-up impacts, there is little change between years two and five. The reductions in health insurance coverage are consistent with the move off welfare to employment associated with mandatory work requirements. Medicaid coverage received while on welfare is not entirely made up by transitional Medicaid coverage, coverage under the poverty-related Medicaid expansions, or transitions to employment-based coverage.

Among the programs that combine financial work incentives and mandatory work-related activities (Panel D), MFIP has a favorable effect on health insurance coverage, while the reverse is true for WRP, the program with less generous financial work incentives. For example, urban single-parent applicants in MFIP were 13 percentage points more likely to have had continuous health care coverage in the three years since random assignment compared with the control group. As noted earlier, the favorable MFIP impact is probably the result of the increase in welfare use, which automatically qualified the family for Medicaid coverage at the time MFIP was in place. Even though WRP is the only program of this group to offer transitional Medicaid benefits, the fact that it did not result in a more favorable impact is attributable to the availability of public insurance coverage through other programs in the state.

For the programs that evaluate TANF-like bundles of reforms, health insurance coverage is reported only for FTP and Jobs First (Panel F). The former program shows statistically insignificant effects as of the four-year follow-up. In contrast, Jobs First raises health insurance coverage rates as of the three-year follow-up for children and adults respectively, with an effect that is statistically significant only for the adult recipient. The difference may be due to the fact that Jobs First provided two years of transitional Medicaid coverage compared with only one year for the control group. FTP did not offer any additional transitional Medicaid benefits for program participants compared with the controls.

9.2.3. Residential Moves, Housing Hardships, and Neighborhood Quality

Table 9.3 reports impact estimates for residential moves, housing hardships, and neighborhood quality using various measures. Among the studies that focus on financial work incentives, only MFIP-IO reports results for measures of residential moves and neighborhood quality, with each reported only for the sample with a child age 5–12 at the 36-month follow-up. For both recipients and applicants, MFIP-IO reduces the number of moves that take place following random assignment for families with primary-school-age children, with an effect for the applicant group that is just under one-half a move and statistically significant. Fewer moves may be indicative of reduced housing instability. Alternatively, it may indicate a diminished ability to upgrade housing or neighborhood quality. On this point, reported neighborhood safety does not appear to be affected much for either the recipients or applicants, although the point estimates are opposite in sign for the two groups. When financial work incentives are combined with work mandates (Panel D), MFIP shows no statistically significant impacts for either residential moves or neighborhood quality.

Of the studies classified under Panel B (those that focus on financial work incentives tied to hours of work), impacts on residential moves, housing hardships, and neighborhood quality are reported for New Hope and SSP. For SSP, the number of moves is higher for the treatment group, but the impact is statistically significant only for SSP families with a child age 6–11 at follow-up. Again, whether the larger number of moves is favorable or unfavorable depends upon the motivation for the change in residence. Measures of housing hardships are reported in SSP, with favorable effects for two different measures (one of which is statistically significant). On the other hand, neighborhood quality as measured by SSP is no different for families with children in the first two age categories (3–5 and 6–11 at follow-up), and it is negatively affected for families with the oldest children (age 12 –18 at follow-up). Thus, it would appear that the higher number of moves associated with SSP are not leading to improved neighborhood conditions (and they may actually be worse for families with teenagers), but

Table 9.3—Estimated Impact of Welfare Reform on Residential Moves, Housing Hardships, and Neighborhood Quality: Random Assignment Studies

			Residential moves				Housing hardships				Neighborhood quality			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives														
MFIP-IO	Urban single parents recipients	S	Number of moves since RA at 36-mo FU (for those with child aged 5 to 12 at FU)	1.7	-0.1	-5.9%					Live in a safe neighborhood at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	74.0	2.5	3.4%
	Urban single parents applicants	S	Number of moves since RA at 36-mo FU (for those with child aged 5 to 12 at FU)	1.6	-0.4 **	-25.0%					Live in a safe neighborhood at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	83.1	-2.0	-2.4%
B. Programs that focus on financial work incentives tied to hours of work														
New Hope	Poor families employed FT at RA	S					Living in an overcrowded dwelling at 2-year FU (%)	16.7	-4.3	-25.7%				
		S					Having had utilities cut off at 2-year FU (%)	34.3	1.3	3.8%				
		S					Experiencing one or more housing defects at 2-year FU (%)	37.1	6.4	17.3%				
	Poor families not employed FT at RA	S					Living in an overcrowded dwelling at 2-year FU (%)	15.2	-1.4	-9.2%				
		S					Having had utilities cut off at 2-year FU (%)	43.0	-1.1	-2.6%				
		S					Experiencing one or more housing defects at 2-year FU (%)	49.7	-3.7	-7.4%				
SSP	Single-parent recipients	S	Any residential moves since RA at 36-mo FU (for those with child aged 3 to 5 at FU) (%)	75.0	4.4	5.8%	Structural problems in house at 36-mo FU (%)	12.3	-2.0 **	-16.3%	Good neighborhood quality at 36-mo FU (for those with child aged 3 to 5 at FU) (%)	76.4	0.3	0.4%
		S	Any residential moves since RA at 36-mo FU (for those with child aged 6 to 11 at FU) (%)	63.4	4.5 **	7.2%	Things not working properly in house at 36-mo FU (%)	12.7	-1.2	-9.4%	Good neighborhood quality at 36-mo FU (for those with child aged 6 to 11 at FU) (%)	75.3	0.3	0.4%
		S	Any residential moves since RA at 36-mo FU (for those with child aged 12 to 18 at FU) (%)	51.1	2.9	5.8%					Good neighborhood quality at 36-mo FU (for those with child aged 12 to 18 at FU) (%)	78.6	-5.9 **	-7.6%
C. Programs that focus on mandatory work-related activities														
LA Jobs-1st GAIN	Single-parent recipients and applicants	S									Neighborhood is bad place to raise children at yr 2 (%)	32.1	-4.4	-13.7%
		S									Neighborhood is unsafe for children to play outside at yr 2 (%)	27.0	-3.1	-11.5%
Atlanta LFA		S	Any residential moves since RA at 5-yr FU (%)	66.2	0.8	1.2%								
Grand Rapids LFA		S	Any residential moves since RA at 5-yr FU (%)	78.0	7.6 ***	9.7%								
Riverside LFA		S	Any residential moves since RA at 5-yr FU (%)	84.0	2.4	2.9%								
Portland		S	Any residential moves since RA at 5-yr FU (%)	85.7	-0.9	-1.1%								
Atlanta HCD		S	Any residential moves since RA at 5-yr FU (%)	66.2	1.5	2.3%								
Grand Rapids HCD		S	Any residential moves since RA at 5-yr FU (%)	78.0	5.4 **	6.9%								
Riverside HCD		S	Any residential moves since RA at 5-yr FU (%)	81.8	0.0	0.0%								

Table 9.3—Continued

			Residential moves				Housing hardships				Neighborhood quality			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
D. Programs that focus on financial work incentives and mandatory work-related activities														
MFIP	Urban single-parent recipients	S	Number of moves since RA at 36-mo FU (for those with child aged 5 to 12 at FU)	1.7	0.2	11.8%					Live in a safe neighborhood at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	74.0	-0.6	-0.8%
	Urban single-parent applicants	S	Number of moves since RA at 36-mo FU (for those with child aged 5 to 12 at FU)	1.6	0.1	6.3%					Live in a safe neighborhood at 36-mo FU (for those with child aged 5 to 12 at FU) (%)	83.1	0.1	0.1%
E. Programs that focus on other individual reforms														
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)														
EMPOWER (a)	Recipients	S					Used emergency shelter since RA as of 30-mo FU (%)	1.0	-0.1	-10.0%				
FTP	Recipients and applicants	S	Any residential moves since RA at 4-year FU (%)	69.6	2.9	4.2%	2 or more housing problems at 4-year FU (%)	18.4	-4.3 **	-23.4%	4 or more neighbor-hood problems at 4-year FU (%)	21.0	-3.8 *	-18.1%
		S					Crowding (more than 1 person per room) at 4-year FU (%)	13.8	0.7	5.3%				
Jobs First	Recipients and applicants	S	Any residential moves since RA at 3-year FU (%)	65.4	-0.1	-0.2%	2 or more housing problems at 3-year FU (%)	18.1	-0.4	-2.1%	1 ore more neighbor-hood problems at 3-year FU (%)	70.6	-6.1 ***	-8.6%
		S	Number of moves since RA at 3-year FU	1.4	0.0	-3.4%	Ever homeless and living on street in last year at 3-year FU (%)	1.5	1.1 *	76.2%				
		S					Lived in homeless, emergency or DV shelter in last year at 3-year FU (%)	3.2	-0.4	-13.5%				

NOTES: For full program names and citations, see Table 3.4. Abbreviations: S=survey data; FU=follow-up; DV=domestic violence.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.
(a) Phoenix site only, cash assistance.

housing quality may be somewhat better. Finally, New Hope considers only housing hardships and finds impacts that are statistically insignificant and mixed in sign.

Measures in this domain are available for a few of the studies that focus on mandatory work-related activities (Panel C). In the case of L.A. Jobs-First GAIN, there is no statistically significant difference in neighborhood quality after two years. Impact estimates on the number of residential moves as of the five-year follow-up are reported for 7 of the NEWWS programs. With one exception, the impact estimates are all positive—indicating a higher fraction of the treatment group made any move since random assignment—and two of the impacts are statistically significant. Again, it is unclear whether this increased mobility is desirable or not.

A few measures of residential moves, housing hardships, and neighborhood quality are reported in EMPOWER, FTP, and Jobs First (Panel F), three studies that focus on TANF-like bundles of reforms. Four years after random assignment, FTP reduces the incidence of two or more housing problems and four or more neighborhood problems, with effects that are significant at the 5 and 10 percent level, respectively. There is no statistically significant effect of FTP on crowding (defined as more than one person per room) or on the prevalence of making one or more residential moves. Jobs First also lowers the number of neighborhood problems after three years, but it results in a statistically significant increase in the incidence of homelessness in the last year. Use of an emergency shelter is rare, even for the EMPOWER control group (about 1 percent). The treatment group is lower by one-tenth of one percent, but the difference is not significant.

9.2.4. Assets

Table 9.4 records various measures of the level and distribution of financial assets, and ownership of physical property such as a house or automobile. These measures are reported for an even smaller number of studies. WRP-IO is the only program that focuses on financial work incentives with outcomes in this domain. One feature of WRP-IO was that the asset limit (specifically the vehicle value) that determines welfare eligibility was increased, although it is difficult to attribute changes in outcomes to this particular program features in itself. While the effect on the average level of savings for more than three years after random assignment is small and insignificant, the fraction with savings over $500 increases by more than half, from 9.2 to 14.4 percent, an effect that is significant at the 5 percent level. The impact on the fraction owning a vehicle is not statistically significant. When financial work incentives are combined with work mandates as for WRP (Panel D), the impact estimates for these measures have the same sign but are smaller, and hence none are statistically significant.[85]

Among the programs that focus on financial work incentives tied to hours worked, only SSP includes asset measures. (As part of the SSP reforms, there were no relevant changes in asset rules.) As of the 36-month follow-up, average savings levels are estimated to decrease by $27 (or 5 percent) for SSP. Neither effect is a statistically insignificant impact.[86] SSP does result in a reduction in the fraction with no savings and an increase in the fraction with savings above $500, with the latter effect being statistically significant. At an earlier follow-up (18 months) for

[85]There is no statistically significant impact on the fraction with debts exceeding $500 in WRP-IO or WRP (not shown).

[86]There is also no statistically significant difference in debt levels or in the distribution of debt for SSP (not shown).

Table 9.4—Estimated Impact of Welfare Reform on Assets: Random Assignment Studies

			Financial assets				Other assets			
Name	Cases served	Data	Measure	Control mean	Impact	%	Measure	Control mean	Impact	%
A. Programs that focus on financial work incentives										
WRP-IO	Single-parent recipients and applicants	S	Average savings at 42-mo FU ($)	281	6	2.1%	Respondent owns car, van or truck at 42-mo FU (%)	70.5	3.7	5.2%
		S	No savings at 42-mo FU (%)	67.4	-2.0	-3.0%				
		S	Savings of $500 or more at 42-mo FU (%)	9.2	5.2 **	56.5%				
B. Programs that focus on financial work incentives tied to hours of work										
SSP	Single-parent recipients	S	Respondent has a savings account at 18-mo FU (%)	46.5	4.4 ***	9.5%	Respondent owns a car at 18-mo FU (%)	24.6	1.6	6.5%
		S	Respondent has a checking account at 18-mo FU (%)	62.1	0.3	0.5%				
		S	Respondent has registered retirement savings plan at 18-mo FU (%)	1.2	1.2 ***	100.0%				
		S	Average savings at 36-mo FU ($)	511	-27.0	-5.3%				
		S	No savings at 36-mo FU (%)	34.3	-1.8	-5.2%				
		S	Savings of $500 or more at 36-mo FU (%)	9.9	2.9 ***	29.3%				
C. Programs that focus on mandatory work-related activities										
D. Programs that focus on financial work incentives and mandatory work-related activities										
WRP	Single-parent recipients and applicants	S	Average savings at 42-mo FU ($)	281	41	14.6%	Respondent owns car, van or truck at 42-mo FU (%)	70.5	2.4	3.4%
		S	No savings at 42-mo FU (%)	67.4	-0.1	-0.1%				
		S	Savings of $500 or more at 42-mo FU (%)	9.2	3.2	34.8%				
E. Programs that focus on other individual reforms										
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)										
FTP	Recipients and applicants	S	Average savings at 4-year FU ($)	198.0	86.0	43.4%	Respondent owns car, van or truck at 4-year FU (%)	60.2	-1.1	-1.8%
		S	No savings at 4-year FU (%)	73.5	-2.1	-2.9%				
		S	Savings of $1,00 or more at 4-year FU (%)	4.1	0.9	22.0%				
Jobs First	Recipients and applicants	S	Average savings at 3-year FU ($)	182	-31	-16.9%	Respondent owns car, van or truck at 3-year FU (%)	36.7	4.2 **	11.5%
		S	No savings at 3-year FU (%)	77.7	1.8	2.3%	Respondent owns home at 3-year FU (%)	0.5	0.2	4.7%
		S	Savings of $501 or more at 3-year FU (%)	7.6	-0.5	-6.8%				

NOTES: For full program names and citations, see Table 3.4. Abbreviations: S=survey data; FU=follow-up.

* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.

SSP, there was a significant increase in the fraction of the treatment group with a savings account and with a registered retirement savings plan. In the case of the retirement plan, the rate doubled from 1.2 percent for the control group to 2.4 percent for the treatment group. This is indicative of how modest the asset levels are for the welfare population. In addition, SSP had no significant impact on the fraction with a car, although the point estimate is positive.

Finally, of the programs that examine TANF-like bundles of reform, both FTP and Jobs First examine changes in financial and other assets. In both cases, these programs increased the asset limit (including vehicle value) that is used to determine eligibility for welfare receipt, from $1,000 to $5,000 in each program for liquid assets and from $1,500 to $8,150 (FTP) and $9,500 (Jobs First) for the value of a vehicle. Again, the experimental designs of these studies do not make it possible to identify the separate impact of changing asset rules.

Vehicle ownership is examined for both programs, with FTP using a broader measure (car, van, or truck versus car only). While there is no difference for FTP at the four-year follow-up, Jobs First shows a 4 percentage point higher rate of car ownership for the treatment group as of the three-year follow-up, a statistically significant difference from the control group. There is no impact in Jobs First on home ownership rates, from a base of less than 1 percent in the control group.

Both programs also measure impacts on the level and distribution of savings. No significant impacts for FTP are found as of the four-year follow-up, nor for Jobs First as of the three-year follow-up. In the case of Jobs First, the point estimates suggest a negative impact on the level of savings and the fraction with savings of $500 or more, and an increase in the fraction with no savings. This stands in contrast to the impact estimates for these same measures as of the 18-month follow-up, where there was a significant reduction in the fraction with no savings and a significant increase in the fraction with $500 or more in savings (results not shown, see Bloom et al., 2000b). In the latter case, the rate more than doubled, from 3.9 percent to 8.1 percent. The early impacts on asset accumulation for Jobs First may be attributable to the more generous financial work incentives, which also resulted in a significant initial impact on family income. However, after year two, the effect of Jobs First on income fades (as the time limits go into effect). It would appear that the favorable impact on asset accumulation fades at the same time.

9.3. EVALUATING THE EFFECTS OF WELFARE REFORM ON OTHER MEASURES OF WELL-BEING

In synthesizing the findings about the impact of welfare reform on other measures of well-being, it is important to recognize that only experimental studies are available for drawing broader inferences. Unlike the outcomes we analyze in other chapters, there are no econometric studies with which to cross-validate the experimental findings. Thus, the limitations of the experimental approach for estimating causal impacts of welfare reform must be kept in mind, including the inability to capture effects on nonparticipants, the limits on the generalizability of findings, and the potential for deviations from ideal experimental conditions that may bias the impact estimates. With these concerns in mind, we aim in this section to draw broader lessons about the impacts of specific reform policies and reform as a bundle on the domains of well-being considered in this chapter.

9.3.1. Effects of Specific Reforms

The demonstration studies with results for other measures of well-being evaluate three broad policies alone or in combination: financial work incentives, mandatory work-related activities, and time limits. The studies of family caps or parental responsibility requirements do not include these broader outcome measures.

Financial Work Incentives

Taken together, the two experimental studies that focus primarily on financial work incentives cover the four domains of well-being reviewed here. Generally, they show favorable impacts, with several statistically significant impacts, most notably for continuous health care coverage under MFIP-IO. There are no statistically significant unfavorable impacts. Unfortunately, MFIP-IO and WRP-IO have almost no comparable measures in the domains covered by Tables 9.1 to 9.4, so it is difficult to determine whether more generous financial work incentives have a more favorable impact on other measures of well-being.

Favorable impacts are also the norm for the two programs that focus on financial work incentives tied to hours of work, and for the two programs that combine financial work incentives with work requirements, especially when the financial incentives are more generous. A comparison of New Hope and SSP with WRP and MFIP shows that both classes of programs, with a few exceptions, produce favorable impacts on multiple measures of well-being. The fact that these programs also tie their incentives to work does not appear to dampen the modestly beneficial impacts. When the incentives are weaker, however, the impacts may be less favorable. For example, there is a significant negative impact on children's health insurance coverage in WRP.

Mandatory Work-Related Activities

More limited information is available on the impact of mandatory work-related activities on a broad array of measures of well-being. Health care coverage is the most widely reported measure, with most of the studies that focus on mandatory work activities showing small unfavorable, but insignificant, effects. In addition, L.A. Jobs-First GAIN provides mixed results on the measures of food insecurity and neighborhood quality included in the study. In sum, the results suggest a very small unfavorable impact, if any, on other measures of well-being for the programs that focus primarily on work-related mandates. This is consistent with the lack of sizeable effects on income and poverty, as discussed in Chapter 8.

Time Limits

None of the demonstration studies focus exclusively on the impact of time limits. Rather, time limits are combined with other policies, such as work requirements and financial work incentives. For this reason, it is impossible to tease out the marginal contribution of time limits on the broader measures of well-being reviewed in this chapter. A comparison of the results for demonstrations in Panel D with those in Panel F does not reveal any striking differences that might be attributed to time limits, a feature of the second group of studies only. Such a comparison is hampered by the inconsistencies in the measures available across studies and the differences in the length of the follow-up periods. The differences in outcomes for Jobs First between the 18-month and 36-month follow-ups (e.g., on asset measures) are suggestive at least that this outcome is negatively affected as time limits become binding.

9.3.2. Effects of Reform as a Bundle

In the absence of any econometric studies that assess the causal impact of welfare reform as a bundle on other measures of well-being, only the experimental studies are available to try to make inferences about how welfare reform as a whole may affect broader measures of well-being. As noted above, these studies are limited in that they do not necessarily represent the combination of policies as implemented in any given state under TANF, nor the weighted combination of policies implemented across states. The demonstration studies also do not capture the impact on potential welfare recipients, in other words, those who are diverted from or who choose not to participate in TANF.

Among the demonstration studies we consider, those shown in Panel F of Tables 9.1 to 9.4 come the closest to representing the combination of policies implemented by most states under TANF, with a combination of time limits, financial work incentives, and work requirements, along with other features such as family caps and parental responsibility requirements. Of the three studies that include other measures of well-being, we place the least weight on EMPOWER. As noted in Chapter 3, the control and treatment groups in EMPOWER were almost equally likely to think they were affected by time limits. Furthermore, by the 30-month follow-up, the EMPOWER control group had become subject to the treatment conditions, so the impact estimate is potentially biased by the control group crossover. Jobs First in Connecticut and FTP, by comparison, maintained the experimental conditions through the three- and four-year follow-up periods, respectively, and there was better differentiation between the treatment and control groups in their understanding of the policies that applied to them.

Focusing then on these two studies, the three-year follow-up period for Jobs First suggests that, with a few exceptions, the combination of policies implemented in Connecticut had little effect in the domains of well-being considered in this chapter. The exceptions are statistically significant favorable impacts for adult health insurance coverage, reports of neighborhood problems, and car ownership. The fact that the positive income effects for Jobs First evident in the first two years disappear by the third and fourth years (see Table 8.1) suggests that favorable impacts for these broader measures of well-being may also diminish over time.

The four-year follow-up available in FTP provides a similar picture. While FTP shows unfavorable impacts for two of the four domains (health care coverage and vehicle ownership), these effects are very small and statistically insignificant. In contrast, statistically significant favorable effects are found for measures of material hardship, housing problems, and neighborhood quality. However, none of these impacts are especially large. For the measures that are available, it appears that FTP at worst has no large unfavorable impacts and, at best, may lead to a small improvement in some domains. However, even as of the four-year follow-up, only 17 percent of families in the FTP group had reached their time limits, accumulating either 24 or 36 months of welfare receipt.[87] Whether the patterns observed as of the four-year follow-up persist as more families reach their time limit or when there is a longer time interval since the time limit was reached is an unanswered question.

While this assessment is fairly positive, we must be cautious in making a strong association between the findings from FTP and Jobs First and the likely overall impact of welfare reform. The results from these two experimental studies are suggestive that welfare as a bundle does not necessarily lead to large negative impacts on other measures of well-being up to four years after reform. However, this inference is based on only a limited set of measures available for two studies that are inherently limited in their focus on current and former recipients in a particular location and point in time. Moreover, the full impact of time limits may not have been observed with the length of follow-up available for analysis. Ideally, these findings would be validated by high quality econometric studies of nationally representative samples for a broad range of measures of well-being.

9.4. CONCLUSIONS

Limitations in the ability of family income to fully capture well-being led us to consider in this chapter other measures of well-being, specifically four domains: material hardship and food insecurity; health insurance coverage; residential moves, housing hardships, and neighborhood quality; and asset ownership. However, the types of measures reviewed in this chapter are not available in all the experimental studies, and when they are, the specific measures often differ from study to study. Smaller sample sizes in demonstration study surveys mean that the statistical precision of impact estimates will be reduced. Moreover, no econometric studies to date have used the DoD approach to examine the outcomes in the domains we consider in this chapter. The inability to cross-validate findings between econometric and experimental approaches is a significant limitation. Consequently, we are more constrained in our ability to make broader inferences about these important dimensions of family well-being.

Nevertheless, the random assignment studies provide some indication of the likely impact of specific reform elements and possibly even welfare reform as a bundle. There is evidence that financial work incentives, especially those that are more generous and increase family income, have mostly favorable, but modest, impacts on other measures of well-being. Programs with mandatory work-related activities tend to have smaller and often unfavorable impacts. Often, these programs lead to lower welfare use and little change or even a decline in family income. Health insurance coverage typically declines because the loss of Medicaid coverage received by welfare participants is not fully replaced in the transition from welfare to work by transitional Medicaid, poverty-related Medicaid coverage, or employer-based coverage. Less is known about the specific impact of time limits on other measures of well-being, because the demonstration studies do not isolate this program feature and because a comparison across studies with and without time limits (among other program features) is limited by a lack of comparability in the available measures of well-being. Unfortunately, there is almost no basis for determining whether any of these conclusions would be different for different subpopulations.

---

[87] Some who reached the time limit were granted extensions for medical reasons.

Finally, we cautiously suggest that some of the experimental studies may be indicative of the types of impacts associated with welfare reform as a bundle. In particular, two of the higher-quality demonstrations that evaluate time limits, financial work incentives, and work requirements suggest that the short-term impacts on broader measures of well-being are not likely to be substantially unfavorable, but neither are they associated with sizeable improvements. This rather neutral conclusion may not hold up when outcomes are assessed over a longer horizon, especially if time limits are associated with eventual declines in family incomes, as families leave welfare for work and no longer benefit from the welfare system's financial work incentives.

Chapter Ten

CHILD OUTCOMES

10.1. BACKGROUND

As part of the debate that preceded the passage of PRWORA, there was considerable discussion about the potential for both negative and positive impacts of the new TANF program on child well-being. Some were concerned that increased work effort by welfare-reliant mothers would be harmful to their children. A related concern was that the loss of welfare income might further increase child poverty, again with negative consequences for children. Others suggested that the transition from dependency to self-sufficiency would increase income and provide a positive role model for disadvantaged children and youth. Promoting marriage and family stability was also viewed as beneficial for children.

There are a number of reasons to expect welfare reform to affect child well-being. First, some welfare reform policies are directly aimed at changing parental behavior or investments in their children through features such as parental responsibility requirements regarding school attendance or immunizations and well-child care, and through requirements for parenting classes. Second, welfare reform policies may change other behaviors that have indirect effects on child well-being. For example, as we have seen in earlier chapters, work effort by mothers with children may change as a result of welfare reform policies. Research suggests that the relationship between maternal employment and child health and development depends on the nature of the mother's job, the change in family resources, the quality of child care and activities for older children, and the mother's psychological well-being (Morris et al., 2001). Family income may also be affected by welfare reform through changes in welfare payments received, earnings, and other transfers. Again, there is a body of research that indicates that family income can affect child development, showing a link between childhood poverty and detrimental outcomes for children. (See, for example, the studies in Duncan and Brooks-Gunn, 1997, and the reviews provided by Haveman and Wolfe, 1995; Mayer, 1997.) Other outcomes relevant for children's development that might be affected by welfare reform include maternal schooling, child care utilization and quality, access to health insurance, and living arrangements. Indeed, we have seen evidence in prior chapters that welfare reform as a whole, and specific policies and programs embedded in the reforms in particular, has almost certainly affected work effort, welfare receipt, and family income.

As investigations of the linkages between child outcomes and welfare reform have multiplied, researchers have developed a model of the pathways through which changes in welfare policy might affect child well-being (see Duncan and Chase-Lansdale, 2001). That model generally posits that welfare reform will have immediate or direct effects on parental work effort, welfare receipt, family income, child care, family structure, and educational attainment. These outcomes, in turn, affect the amount and composition of the resources—financial and otherwise—available for raising children. In addition to resources, other intervening or intermediate behaviors and outcomes can change as a result of the direct impacts. Examples include parent psychological well-being (e.g., self-esteem, sense of self-efficacy, stress, depression, substance abuse), parent-child interaction (i.e., the quantity and quality of time available for positive interaction and supervision), child socialization (e.g., for older children, messages about work, responsibility, and self-sufficiency), and access to services (e.g., health care).

These direct and intermediate impacts are then expected to affect child health and development in a number of domains, including cognitive development, behavioral and emotional adjustment, school achievement and attainment, antisocial and delinquent behavior, child safety, and physical and mental health. Within this framework, welfare reform might be expected to have both negative and positive effects on children's outcomes, and some outcomes might remain unchanged because of opposing forces. This framework also suggests that child impacts might vary with the age or gender of the child, as well as with other family background characteristics. Finally, some aspects of child development may be more responsive within a short period to the effects of welfare reform, while other indicators of child development may take time for the impacts to cumulate. For example, child behavior problems at both younger and older ages may manifest themselves within a short time frame, while it may take longer for effects on child health to become apparent.

Many of the child outcomes of interest are not measured routinely as part of large nationally representative surveys (e.g., the CPS). In addition, more specialized smaller-scale surveys may not have sufficient sample sizes to implement the DoD methodology reviewed in Chapter 3. Likewise, there are few administrative databases that track relevant outcomes in a consistent manner over both geographic space and time. Consequently, in contrast to outcomes like welfare caseloads, employment and earnings, and income, there are considerably fewer econometric studies that employ the DoD methodology to examine child well-being; in fact, we are aware of just one study that analyzes direct measures of child well-being.[88]

Instead, much of the research on child outcomes and welfare reform is conducted in the context of experimental evaluations. A child outcome component has been included in a number of the welfare experiments, with most studies relying primarily on data collected from parents (and sometimes teachers or the children themselves). A few experimental studies also use administrative data for outcomes such as child maltreatment and foster care. However, across the experimental studies, data on child outcomes are not universally available. Of the studies listed in Table 3.5, the programs in Arizona (EMPOWER), Arkansas (AWWDP), Indiana (IMPACT), Iowa (FIP), New Jersey (FDP), and Virginia (VIP/VIEW) do not report results for child outcomes in the domains we list above.[89] The other evaluations report results for at least one child outcome for at least one treatment-control contrast.[90]

---

[88]Haider, Jacknowitz, and Schoeni (2002) use aggregate state-level data on rates of breast-feeding from 1990 to 2000 to estimate the impact of various features of state work requirements under waivers and TANF on breast-feeding rates. They find that the most stringent work requirements that apply to women with infants have significantly reduced the prevalence of breast-feeding six months after birth for all mothers and for women on WIC. Since the incidence of breast-feeding is a more indirect measure of child well-being, we do not include their study in our synthesis.

[89]In some cases, like Indiana and Iowa, future reports are planned with impact estimates for child outcomes.

[90]In the case of the multitreatment SSP study, child outcomes are reported for only the primary study.

Since our synthesis of the impact of welfare reform on child well-being will, of necessity, draw almost exclusively on random assignment studies—and then only on a subset of the studies covered in this report—the caveats discussed in Chapter 9 for other measures of well-being are equally relevant here. As with other measures of well-being, the sample sizes available for analysis are often smaller than those available for the full evaluation. Studies focus on various child outcome domains, and the measures used in a given domain are not always comparable across the studies that cover that domain.[91] Again, this affects our ability to draw more general inferences from the collection of studies. The limitations of experimental studies—the inability to capture program entry effects, questions about generalizing from a local or state demonstrations to national reform, and problems with maintaining ideal experimental conditions—will also affect the broader conclusions we draw in this chapter.

With these concerns in mind, the remainder of this chapter will first discuss the findings from the experimental studies that focus on child well-being. Our understanding of the differences by subgroups gleaned from these studies are discussed in Appendix A and are also summarized in this second section. We then turn to a discussion of the one relevant econometric study. The results from the experimental and econometric studies are synthesized in the fourth section. A concluding section ends the chapter.

10.2. RANDOM ASSIGNMENT STUDIES OF THE EFFECTS OF WELFARE REFORM ON CHILD WELL-BEING

For the studies that do include child outcome measures, there is a wealth of information, which presents a number of challenges. Unlike some of the other chapters in this report, where there is considerable uniformity in the outcome measure across studies, child well-being can be conceptualized in many ways, with a myriad of indicators for any given domain, whether it be child behavioral problems, academic success, physical and mental health, or some other area of functioning. Very few studies measure the exact same sets of indicators, or even the same indicators within the same domains.

In some cases, the measures are straightforward indicators of child outcomes. Good examples include whether a grade has been repeated since random assignment or whether a child has made an emergency room visit since random assignment. In other cases, the indicators are standard scales or test batteries with well-understood psychometric properties (e.g., reliability and validity).[92] For example, the Behavioral Problems Index (BPI) is a frequently used measure of child problem behavior in both small- and large-scale studies, and it appears in the child outcome impact analyses for several of the demonstration studies. However, the BPI is not the only measure of problem behavior used in these studies, so differences across studies in child behavior measures may arise from the scales themselves and from the dimensions of behavior they measure rather than from true differences in behavior. As another example, the Peabody Picture Vocabulary Test-Revised (PPVT-R), a standard measure of receptive vocabulary, is used in just one of the demonstration studies we review. Other studies that measure language or reading ability do so with different measures. The reliance on the same well-validated child outcome measures in many demonstrations is an advantage for making cross-study

[91]Issues associated with measuring child outcomes are discussed more fully in the next section.

[92]A test is reliable if an individual has similar scores on repeated applications of the test. A valid test is one that measures what it purports to measure.

comparisons. However, cross-study comparisons are made more difficult when studies do not focus on the same domains, or when they do not use a similar set of indicators within a domain or the same metric or scale for a given indicator.

Another feature of this literature is that some outcome measures represent favorable outcomes, so a positive numerical impact is desirable; in other cases, the metric represents an unfavorable outcome, so a negative impact is the goal. Our analysis is further complicated by the expectation, as noted in the introductory section, that the impact of welfare reform may vary with the age of the child. Hence, we are interested in differential impacts for children by age, with age groups typically defined as preschool, primary school age, and adolescents.

In light of these challenges, we have organized the tabular presentation of the results from the experimental literature in a format different from the one we used in earlier chapters. First, for any given study, we have grouped the measures of child outcomes into four broad categories: behavior, school performance, health, and other.[93] The first captures both positive and negative aspects of behavior, ranging from such measures as the BPI or an index of positive social behaviors for younger children, to being suspended or expelled from school, to being involved in criminal or delinquent activity (for older children). The second outcome category includes various measures of school performance and achievement, including subject-specific test scores (e.g., reading and math), parental reports of school performance, the extent of grade repetition, and use of special education. General measures of health status are included in the third category, along with other indicators of health and safety, such as reports of child abuse and neglect. A fourth residual category captures other outcomes such as foster care placements and participation in clubs or organizations.

Many of these measures relate to the child's current status at the time of data collection (e.g., health status), while others capture outcomes over a child's lifetime (e.g., ever repeated a grade in school). Others ask about behavior or outcomes since random assignment. We have tried to clearly indicate whether a measure is cumulative ("ever") or measured since random assignment ("since RA"). Those not explicitly designated are assumed to relate to current status.

Given the multiple measures that are often not comparable across studies, we do not record the numeric level of the outcome for the control group or the impact estimate (treatment-control difference), as we have done in previous tables. Instead, we record impact estimates in a favorable direction (whether numerically positive or negative) as "F," while those in an unfavorable direction are recorded as "U." Impact estimates that are zero are recorded as "0." The statistical significance of each impact is indicated along with the effect size for the impact

[93]These categories overlap to some extent given that some outcomes could be easily classified in more than one of the areas we have defined. For example, school suspensions could be a behavior problem or a school outcome. We have classified it as the former, but we recognize that others might place it in the latter category. For our purposes, the goal was to be consistent rather than rigid in defining these broad outcome categories.

estimate when it is available.[94] The directional indicator and the effect size are in bold type for those impact estimates that are statistically significant.

Finally, where possible, we record outcomes first for all children and then stratified by age group. Since studies often use different age cutoffs in their age strata, we separately record results for the youngest age group (typically preschool), middle age group (typically primary grades), and oldest age group (typically preteens and teens in secondary grades and above) and indicate for each study what age cutoffs are used. The ages recorded are the ages at follow-up, so they can be compared across studies with different follow-up intervals. Given the outcome domains considered and the metrics available, infants and toddlers are often excluded from the analyses.

We now turn to a summary of the results for experimental evaluations of programs grouped by their policies. The results are recorded in Table 10.1 using the approach we have just outlined.

10.2.1. Programs That Focus on Financial Work Incentives

Of the programs that primarily evaluated financial work incentives (Panel A of Table10.1), both MFIP-IO and WRP-IO evaluated child outcomes. The MFIP study focuses on a cohort in a more narrow age range (5 to 12), while WRP reports results for all children, and for children age 10 and above. The follow-up periods were between three and three and one-half years.

Vermont's WRP evaluation included a follow-up telephone and in-person survey with both recipients and applicants in WRP and WRP-IO (Bloom, Hendra, and Michalopoulos, 2000). Compared with MFIP, WRP examined a somewhat more limited set of child indicators. In general, the results show no consistent effect of the WRP-IO program for the sample of nearly 1,200 children studied. Children in the treatment group had a significantly higher likelihood of missing a day or more of school in the last month, but a significantly higher rate of participation in clubs and organizations. For those age 10 and above, the treatment group reported a statistically significant higher rate of ever being in trouble with the police (26.8 percent versus 17.2 percent for controls), but there were no significant differences in school dropout behavior or behavior problems (although the impact estimates were in the favorable direction).

The MFIP child analysis (both the Incentives Only and full program) focuses on a random subset of families in the evaluation sample who entered the study in the first six months (April to October 1994) and who had at least one child age 5 to 12 at the time of the survey, three years after random assignment (Gennetian and Miller, 2000). Much of the data collected refers to a "focal" child rather than to all children in the family.[95] With this narrower age group, MFIP-IO does not provide results for infants and adolescents, unlike some of the other evaluations. As with other MFIP analyses, results are available for long-term recipients and applicants, with about 600 and 400 children in the combined treatment and control groups, respectively.

---

[94]The effect size is a standardized measure of impact and is defined as the program impact (treatment minus control group difference) divided by the standard deviation of the outcome for treatment and control groups combined. In Table 10.1, we report the absolute value of the effect size for those studies that report it. Since standard deviations are typically not provided, it is not possible to calculate effect sizes when they are not reported by the study authors.

[95]Some measures are available for all children. Impacts for the full MFIP for adolescents age 13 and above, in addition to those for the focal children age 5 to 12, are discussed in Section 10.2.4 below.

Table 10.1—Estimated Impact of Welfare Reform on Child Behavior, Schooling, Health, and Other Outcomes: Random Assignment Studies

Name	Population Followup length	Outcome Domain	All children: Measure	Impact	Signif.	Effect size	Youngest age group (ages at FU): Measure	Impact	Signif.	Effect size	Middle age group (ages at FU): Measure	Impact	Signif.	Effect size	Oldest age group (ages at FU): Measure	Impact	Signif.	Effect size
A. Programs that focus on financial work incentives																		
															Age 10 and above			
WRP-IO	Single parents R&A 3.5 years	Behavior	• Expelled/suspended since RA	U		–									• Behavior probs. since RA	F		–
															• Ever in trouble w/police	U	**	–
		School	• Absences last mo.	U	*	–									• Ever dropped out	F		–
			• Repeated grade since RA	F		–												
		Other	• CU participates clubs/organ.	F	*	–												
											Age 5 – 12							
MFIP-IO	Urban single parent recipients 3 years	Behavior									• Current total BPI	F	*	0.15				
											• CU externalizing BPI	F	*	–				
											• CU internalizing BPI	F	*	–				
											• CU high level behav./emot. probs.	F		–				
											• CU pos. social behaviors (PBS)	F	**	0.18				
											• CU behav. probs. at school	F		–				
											• Suspended/expelled since RA	U		–				
		School									• CU avg. performance in school	F	*	0.14				
											• CU perf. in school below avg.	F		–				
											• CU engagement in school	F	**	0.20				
											• Repeated grade since RA	U		0.02				
											• In spec. educ. since RA	U		–				
		Health									• ERV for accident/injury since RA	U	**	–				
											• Current overall health	F		0.06				
	Urban single parent applicants 3 years	Behavior									• Current total BPI	U		0.12				
											• CU pos. social behaviors (PBS)	U		0.11				
											• CU behav. probs. at school	F		–				
											• Suspended/expelled since RA	U	*	–				
		School									• CU avg. performance in school	U	*	0.20				
											• CU perf. in school below avg.	U		–				
											• CU engagement in school	U	**	0.28				
											• Repeated grade since RA	U		0.06				
											• In spec. educ. since RA	F		–				
		Health									• Current overall health	F		0.06				
B. Programs that focus on financial work incentives tied to hours of work																		
							Age 3 – 5				Age 6 – 12				Age 12 – 18			
New Hope	Poor families with 1 or more children age 1-11 2 years	Behavior	• CU total positive behavior (P)	F		0.03	• CU total behavior problems (P)	F		0.12	• CU total behavior problems (P)	F		0.03	• Contacts for behav. prob. since RA	U	***	0.3
			• CU total positive behavior (T)	F	**	0.25	• CU externalizing problems (P)	F		0.11	• CU externalizing problems (P)	F		0.02	• Suspended/expelled since RA	U		0.0
							• CU internalizing problems (P)	F		0.09	• CU internalizing problems (P)	F		0.02				
											• CU total social competency (C)	U		0.06				
											• CU total aggression score (C)	F		0.07				
		School	• CU normal school progress (P)	F		0.09					• CU cognitive competence (6-8) (C)	U		0.12	• CU avg. performance in school	U	*	0.2
			• CU school achievement (P)	F		0.09					• CU physical competence (6-8) (C)	U		0.06	• CU perf. In school below avg.	U	*	0.2
			• CU not making normal progress (T)	F		0.10					• CU schol. competence (9-12) (C)	F		0.03	• Current use of special education	F	**	0.2
			• CU SSRC academic subscale (T)	F	**	0.25					• CU athletic competence (9-12) (C)	F		0.07	• Current in gifted/talented program	F	*	0.2
											• CU global self-worth (9-12) (C)	F		0.17	• Repeated a grade since RA	U	***	0.3
							Age 3 – 5				Age 6 – 11				Age 12 – 18			
SSP	Single parent recipients 3 years	Behavior					• Current behavior problems	U		0.01	• Current behavior problems	F		0.03	• CU school behavior problems	U	*	0.09
							• CU positive social behaviors	U		0.06	• CU positive social behaviors	U		0.02	• CU delinquent activity (12-14) (C)	F		0.06
											• CU school behavior problems	F		0.01	• CU delinquent activity (15-18) (C)	U	**	0.21
															• CU any smoking (C)	U	*	0.11
															• CU drinks once/week or more (C)	U	***	0.20
															• CU any drug use (C)	U	*	0.12
		School					• Current reading score (PPVT-R)	F		0.05	• CU reading score (6-7) (PPVT-R)	F		0.13	• CU math score	U		0.03
											• CU math score (7-11)	F	**	0.14	• CU avg. sch. subject achvmt	U	*	0.11
											• CU avg. sch. subject achvmt (P)	F	**	0.11	• Ever repeated grade	U		0.03
											• Ever repeated grade (P)	0		0.05	• Ever dropped out (15-18)	U		0.09
		Health					• Current average health	U		0.05	• Current average health	F	**	0.11	• Current average health	F		0.05
							• CU any long-term hlth problems	F		0.04	• CU any long-term hlth problems	F	**	0.09	• CU any long-term hlth problems	U		0.02
							• Injuries in last 6 mos.	F		0.04	• Injuries in last 6 mos.	U		0.01	• Injuries in last 6 mos.	F		0.03

Table 10.1—Continued

Name	Population Followup length	Outcome Domain	All children				Youngest age group (ages at FU)				Middle age group (ages at FU)				Oldest age group (ages at FU)			
			Measure	Impact	Signif.	Effect size	Measure	Impact	Signif.	Effect size	Measure	Impact	Signif.	Effect size	Measure	Impact	Signif.	Effect size
C. Programs that focus on mandatory work-related activities																		
							Age 5 – 7				Age 8 – 11				Age 12 – 20			
LA Jobs-1st GAIN	Single parent R&A 2 years	Behavior	• Ever expelled/suspended	**F**	*	–					• Ever expelled/suspended	F		–	• Ever expelled/suspended	F		–
			• Ever class for behav. prob.	U		–	• Ever class for behav. prob.	U		–	• Ever class for behav. prob.	**U**	*	–	• Ever class for behav. prob.	U		–
			• Ever behav. prob. affect P work	U		–	• Ever behav. prob. affect P work	**U**	*	–	• Ever behav. prob. affect P work	F		–	• Ever behav. prob. affect P work	U		–
		School	• Current school performance	F		–					• Current school performance	F		–	• Current school performance	U		–
			• Ever honor roll/award	U		–					• Ever honor roll/award	U		–	• Ever honor roll/award	U		–
			• Ever repeated grade	F		–	• Ever repeated grade	**U**	***	–	• Ever repeated grade	F		–	• Ever repeated grade	F		–
			• Ever dropped out	F		–									• Ever dropped out	F		–
		Health	• Ever ERV for accident/injury	U		–	• Ever ERV for accident/injury	U		–	• Ever ERV for accident/injury	F		–	• Ever ERV for accident/injury	0		–
			Any child in family				Age (6 – 7) 8 – 10				Age 11 – 14				Age 15 – 23			
Atlanta LFA	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	F		0.07	• Expelled/suspended in yrs 3-5	U		0.01	• Expelled/suspended in yrs 3-5	U		0.00	• Expelled/suspended in yrs 3-5	**F**	**	**0.16**
		School	• Repeated grade in yrs 3-5	**F**	*	**0.10**	• Repeated grade in yrs 3-5	0			• Repeated grade in yrs 3-5	F		0.06	• Repeated grade in yrs 3-5	F		0.05
			• Current use of special ed.	U		0.02	• Current use of special ed.	U		0.00	• Current use of special ed.	F		0.04	• Current use of special ed.	U		0.04
			• Ever dropped out	F		0.03				0.04	• Ever dropped out	n/a		–	• Ever dropped out	F		0.03
		Health	• ERV for accident/injury in yrs 3-5	F		0.07	• ERV for accident/injury in yrs 3-5				• ERV for accident/injury in yrs 3-5	0		0.00	• ERV for accident/injury in yrs 3-5	F		0.04
			• CU condition requires med. care	U		0.04	• CU condition requires med. care	F **U**		**0.20**	• CU condition requires med. care	U		0.02	• CU condition requires med. care	U		0.03
			• CU condition affects M wk/school	U		0.05	• CU condition affects M wk/school	**U**	**	0.02 **0.26**	• CU condition affects M wk/school	U		0.06	• CU condition affects M wk/school	F		0.04
			• Had baby as teen (< 18) since RA	**F**	**	**0.12**			***						• Had baby as teen (< 18) since RA	F		0.11
Grand Rapids LFA	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	F		0.04	• Expelled/suspended in yrs 3-5	(**F**) F	(***)	(**0.27**) 0.06	• Expelled/suspended in yrs 3-5	**U**	*	**0.18**	• Expelled/suspended in yrs 3-5	F		0.03
		School	• Repeated grade in yrs 3-5	U		0.10	• Repeated grade in yrs 3-5	(F) U		(0.05) 0.10	• Repeated grade in yrs 3-5	U		0.14	• Repeated grade in yrs 3-5	**U**	**	**0.16**
			• Current use of special ed.	U		0.01	• Current use of special ed.	(F) U		(0.10) 0.07	• Current use of special ed.	F		0.06	• Current use of special ed.	U		0.10
			• Ever dropped out	U		0.04					• Ever dropped out	F		0.08	• Ever dropped out	U		0.08
		Health	• ERV for accident/injury in yrs 3-5	F		0.02	• ERV for accident/injury in yrs 3-5	(F) U		(0.03) 0.02	• ERV for accident/injury in yrs 3-5	F		0.09	• ERV for accident/injury in yrs 3-5	F		0.09
			• CU condition requires med. care	U		0.03	• CU condition requires med. care	(**F**) U	(**)	(**0.20**) 0.08	• CU condition requires med. care	F		0.04	• CU condition requires med. care	U		0.11
			• CU condition affects M wk/school	F		0.05	• CU condition affects M wk/school	(F) U		(0.05) 0.01	• CU condition affects M wk/school	F		0.06	• CU condition affects M wk/school	U		0.09
			• Had baby as teen (< 18) since RA	F		0.07									• Had baby as teen (< 18) since RA	F		0.00
Riverside LFA	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	F		0.06	• Expelled/suspended in yrs 3-5	F		0.08	• Expelled/suspended in yrs 3-5	**F**	**	**0.14**	• Expelled/suspended in yrs 3-5	F		0.01
		School	• Repeated grade in yrs 3-5	U		0.01	• Repeated grade in yrs 3-5				• Repeated grade in yrs 3-5	U		0.06	• Repeated grade in yrs 3-5	**U**	**	**0.14**
			• Current use of special ed.	U		0.06	• Current use of special ed.	**F**	***		• Current use of special ed.	F		0.01	• Current use of special ed.	U		0.10
			• Ever dropped out	U		0.05		U		**0.18** 0.03	• Ever dropped out	F		0.06	• Ever dropped out	U		0.01
		Health	• ERV for accident/injury in yrs 3-5	F		0.03	• ERV for accident/injury in yrs 3-5				• ERV for accident/injury in yrs 3-5	U		0.03	• ERV for accident/injury in yrs 3-5	F		0.05
			• CU condition requires med. care	F		0.09	• CU condition requires med. care	U			• CU condition requires med. care	U		0.03	• CU condition requires med. care	**F**	**	**0.12**
			• CU condition affects M wk/school	F		0.06	• CU condition affects M wk/school	F **F**		0.01 **0.13**	• CU condition affects M wk/school	F		0.01	• CU condition affects M wk/school	U		0.01
			• Had baby as teen (< 18) since RA	U		0.07			**	0.10					• Had baby as teen (< 18) since RA	U		0.07
Portland	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	F		0.01	• Expelled/suspended in yrs 3-5	(U) F		(0.03) 0.16	• Expelled/suspended in yrs 3-5	U		0.20	• Expelled/suspended in yrs 3-5	U		0.04
		School	• Repeated grade in yrs 3-5	U		0.14	• Repeated grade in yrs 3-5	(U) F		(0.19) 0.03	• Repeated grade in yrs 3-5	U		0.03	• Repeated grade in yrs 3-5	F		0.06
			• Current use of special ed.	U		0.10	• Current use of special ed.	(U) U		(0.08) 0.02	• Current use of special ed.	U		0.27	• Current use of special ed.	U		0.06
			• Ever dropped out	U		0.04					• Ever dropped out	U		0.00	• Ever dropped out	U		0.11
		Health	• ERV for accident/injury in yrs 3-5	U		0.01	• ERV for accident/injury in yrs 3-5	(F) U		(0.15) 0.20	• ERV for accident/injury in yrs 3-5	U		0.00	• ERV for accident/injury in yrs 3-5	U		0.09
			• CU condition requires med. care	U		0.07	• CU condition requires med. care	(U) U		(0.09) 0.16	• CU condition requires med. care	U		0.04	• CU condition requires med. care	F		0.02
			• CU condition affects M wk/school	U		0.04	• CU condition affects M wk/school	(U) F		(0.01) 0.06	• CU condition affects M wk/school	U		0.03	• CU condition affects M wk/school	U		0.10
			• Had baby as teen (< 18) since RA	F		0.12									• Had baby as teen (< 18) since RA	F		0.10
Atlanta HCD	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	U		0.01	• Expelled/suspended in yrs 3-5	U		0.00	• Expelled/suspended in yrs 3-5	U		0.02	• Expelled/suspended in yrs 3-5	U		0.05
		School	• Repeated grade in yrs 3-5	**F**	*	**0.11**	• Repeated grade in yrs 3-5				• Repeated grade in yrs 3-5	F		0.10	• Repeated grade in yrs 3-5	F		0.04
			• Current use of special ed.	F		0.05	• Current use of special ed.	U			• Current use of special ed.	F		0.09	• Current use of special ed.	F		0.02
			• Ever dropped out	U		0.03		U		0.05 0.04	• Ever dropped out	n/a		–	• Ever dropped out	U		0.05
		Health	• ERV for accident/injury in yrs 3-5	F		0.09	• ERV for accident/injury in yrs 3-5				• ERV for accident/injury in yrs 3-5	F		0.08	• ERV for accident/injury in yrs 3-5	F		0.01
			• CU condition requires med. care	F		0.06	• CU condition requires med. care	F **U**		**0.18**	• CU condition requires med. care	F		0.09	• CU condition requires med. care	F		0.04
			• CU condition affects M wk/school	U		0.05	• CU condition affects M wk/school		**	0.05	• CU condition affects M wk/school	U		0.06	• CU condition affects M wk/school	F		0.00
			• Had baby as teen (< 18) since RA	F		0.03		U		0.10					• Had baby as teen (< 18) since RA	U		0.02
Grand Rapids HCD	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	**F**	*	**0.10**	• Expelled/suspended in yrs 3-5	(F) F	(**)	(0.19) 0.02	• Expelled/suspended in yrs 3-5	F		0.05	• Expelled/suspended in yrs 3-5	U		0.05
		School	• Repeated grade in yrs 3-5	U		0.13	• Repeated grade in yrs 3-5	(F) **U**	*	(0.07) **0.17**	• Repeated grade in yrs 3-5	U		0.09	• Repeated grade in yrs 3-5	**U**	**	**0.16**
			• Current use of special ed.	U		0.00	• Current use of special ed.	(F) U		(0.01) 0.09	• Current use of special ed.	F		0.14	• Current use of special ed.	U		0.09
			• Ever dropped out	U		0.02					• Ever dropped out	F		0.14	• Ever dropped out	U		0.07
		Health	• ERV for accident/injury in yrs 3-5	**F**	*	**0.11**	• ERV for accident/injury in yrs 3-5	(F) F		(0.11) 0.08	• ERV for accident/injury in yrs 3-5	F		0.10	• ERV for accident/injury in yrs 3-5	F		0.00
			• CU condition requires med. care	F		0.10	• CU condition requires med. care	(F) U		(0.12) 0.06	• CU condition requires med. care	**F**	*	**0.16**	• CU condition requires med. care	U		0.02
			• CU condition affects M wk/school	**F**	**	**0.12**	• CU condition affects M wk/school	(F) U		(0.10) 0.00	• CU condition affects M wk/school	**F**	**	**0.19**	• CU condition affects M wk/school	U		0.02
			• Had baby as teen (< 18) since RA	F		0.00									• Had baby as teen (< 18) since RA	U		0.02
Riverside HCD	R&A w/child < 18 at RA 5 years	Behavior	• Expelled/suspended in yrs 3-5	F		0.05	• Expelled/suspended in yrs 3-5	U		0.04	• Expelled/suspended in yrs 3-5	**F**	***	**0.21**	• Expelled/suspended in yrs 3-5	F		0.02
		School	• Repeated grade in yrs 3-5	U		0.06	• Repeated grade in yrs 3-5				• Repeated grade in yrs 3-5	U		0.11	• Repeated grade in yrs 3-5	**U**	**	**0.18**
			• Current use of special ed.	U		0.04	• Current use of special ed.	F			• Current use of special ed.	F		0.01	• Current use of special ed.	**U**	*	**0.15**
			• Ever dropped out	U		0.11		F		0.08 0.07	• Ever dropped out	F		0.05	• Ever dropped out	**U**	*	**0.14**
		Health	• ERV for accident/injury in yrs 3-5	F		0.03	• ERV for accident/injury in yrs 3-5				• ERV for accident/injury in yrs 3-5	**F**	*	**0.15**	• ERV for accident/injury in yrs 3-5	F		0.08
			• CU condition requires med. care	F		0.00	• CU condition requires med. care	U			• CU condition requires med. care	U		0.06	• CU condition requires med. care	F		0.07
			• CU condition affects M wk/school	**U**	*	**0.13**	• CU condition affects M wk/school	F		0.02	• CU condition affects M wk/school	F		0.03	• CU condition affects M wk/school	**U**	*	**0.17**
			• Had baby as teen (< 18) since RA	U		0.09		U		0.05 0.01					• Had baby as teen (< 18) since RA	U		0.11

Table 10.1—Continued

Name	Population Followup length	Outcome Domain	All children: Measure	Impact	Signif.	Effect size	Youngest age group (ages at FU): Measure	Impact	Signif.	Effect size	Middle age group (ages at FU): Measure	Impact	Signif.	Effect size	Oldest age group (ages at FU): Measure	Impact	Signif.	Effect size
Columbus Integrated	R&A with all children 6+ 2 years	Behavior	• CU help for behav/emot probs	F		–												
			• CU class for behav/emot probs	**F**	*	–												
			• Expelled/suspended since RA	F		–												
		School	• Repeated grade since RA	F		–												
			• Current use of special ed.	**F**	**	–												
		Health	• ERV for accident/injury since RA	U		–												
Columbus Traditional	R&A with all children 6+ 2 years	Behavior	• CU help for behav/emot probs	U		–												
			• CU class for behav/emot probs	F		–												
			• Expelled/suspended since RA	U		–												
		School	• Repeated grade since RA	F		–												
			• Current use of special ed.	U		–												
		Health	• ERV for accident/injury since RA	U		–												
Detroit	R&A with all children 6+ 2 years	Behavior	• CU help for behav/emot probs	U		–												
			• CU class for behav/emot probs	U		–												
			• Expelled/suspended since RA	F		–												
		School	• Repeated grade since RA	F		–												
			• Current use of special ed.	U		–												
		Health	• ERV for accident/injury since RA	U		–												
Oklahoma City	R&A with all children 6+ 2 years	Behavior	• CU help for behav/emot probs	**U**	**	–												
			• CU class for behav/emot probs	U		–												
			• Expelled/suspended since RA	U		–												
		School	• Repeated grade since RA	U		–												
			• Current use of special ed.	F		–												
		Health	• ERV for accident/injury since RA	F		–												
D. Programs that focus on financial work incentives and mandatory work-related activities																		
WRP	Single parents R&A 3.5 years														Age 10 and above			
		Behavior	• Expelled/suspended since RA	F		–									• Behavior probs. since RA	F		–
															• Ever in trouble w/police	U		–
		School	• Absences last mo.	U		–									• Ever dropped out	F		–
			• Repeated grade since RA	U		–												
		Other	• CU participates clubs/organ.	**F**	***	–												
MFIP	Urban single parent recipients 3 years										Age 5 – 12				Age 13 and above			
		Behavior									• Current total BPI	**F**	*	**0.14**	• CU behav. probs. at school	U		
											• CU externalizing BPI	**F**	**	–				
											• CU internalizing BPI	F		–				
											• CU high level behav./emot. probs.	**F**	***	–				
											• CU pos. social behaviors (PBS)	F		0.01				
											• CU behav. probs. at school	F		–				
											• Suspended/expelled since RA	F		–				
		School									• CU avg. performance in school	**F**	*	**0.15**	• CU avg. performance in school	0		
											• CU perf. in school below avg.	**F**	**	–	• CU perf. in school below avg.	F		
											• CU engagement in school	**F**	**	**0.17**	• Repeated grade since RA	F		
											• Repeated grade since RA	F		0.10				
											• In spec. educ. since RA	F		–				
		Health									• ERV for accident/injury since RA	**U**	*	–				
											• Current overall health	U		0.07				
	Urban single parent applicants 3 years	Behavior									• Current total BPI	U		0.13	• CU behav. probs. at school	**U**	***	–
											• CU pos. social behaviors (PBS)	U		0.10				
											• CU behav. probs. at school	F		–				
											• Suspended/expelled since RA	**U**	*	–				
		School									• CU avg. performance in school	U		0.11	• CU avg. performance in school	**U**	**	–
											• CU perf. in school below avg.	U		–	• CU perf. in school below avg.	**U**	*	–
											• CU engagement in school	U		0.13	• Repeated grade since RA	U		
											• Repeated grade since RA	F		0.13				
											• In spec. educ. since RA	F		–				
		Health									• Current overall health	U		0.04				
TSMF	Single parents R&A 1 to 4 years														Age 12 and above			
		Health	• Substantiated reports A/N	U		–												
		Other	• Foster care placements	U		–									• Employment	F		–
															• Earnings	F		–

Table 10.1—Continued

Name	Population Followup length	Outcome Domain	All children: Measure	Impact	Signif.	Effect size	Youngest age group (ages at FU): Measure	Impact	Signif.	Effect size	Middle age group (ages at FU): Measure	Impact	Signif.	Effect size	Oldest age group (ages at FU): Measure	Impact	Signif.	Effect size
E. Programs that focus on other individual reforms																		
							Age 2 – 4											
PPI	R&A 2 years	Health					• Number of well-child visits	U		–								
							• At least 1 well-child visit/year	F		–								
							• Up-to-date in vaccinations	U		–								
											Age 4 – 10							
PIP	Recipients 4 years	Health									• Up-to-date in vaccinations	**F**	***	–				
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)																		
ABC	Single parent R&A 1 to 3 years	Health	• Sub. rpts. of neglect-yr 1, 3	**U**	**	–												
			• Sub. rpts. of neglect-yr 2	F		–												
			• Sub. rpts. of other abuse-yr 1,2,3	F		–												
		Other	• Foster care placements - yr 3	F		–												
											Age 5 – 12				Age 13 – 17			
FTP	R&A 4 years	Behavior									• Suspended since RA	F		0.02	• Suspended since RA	**U**	**	**0.17**
											• Expelled since RA	U		0.13	• Expelled since RA	U		0.02
											• Current total BPI	U		0.01	• Arrested since RA	U		0.01
											• CU pos. social behaviors (PBS)	**U**	*	**0.11**	• Found guilty since RA	U		
		School									• Current achievement	F		0.09				
											• Current engagement in school	0		0.00	• Current achievement	**U**	*	**0.14**
											• In special education since RA	U		0.07	• In special education since RA	U		0.09
											• Repeated a grade since RA	U		0.02				
		Health									• Current general health	**F**	*	**0.09**	• Had a baby since RA	F		0.03
											• ERV for accident/injury since RA	U		0.01				
											Age 5 – 12				Age 13 – 17			
Jobs First	R&A 3 years	Behavior									• Suspended since RA	F		0.04				
											• Current total BPI	**F**	**	**0.10**	• Suspended since RA	F		0.00
											• CU externalizing BPI	**F**	*	**0.09**	• Expelled since RA	U		0.11
											• CU internalizing BPI	**F**	**	**0.10**	• Arrested since RA	F		0.09
											• CU pos. social behaviors (PBS)	**F**	*	**0.09**	• Found guilty since RA	**F**	**	**0.14**
		School									• Current achievement	0		0.01	• Current achievement	**U**	***	**0.24**
											• Current engagement in school	F		0.07				
											• In special education since RA	U		0.04	• In special education since RA	U		0.12
											• Repeated a grade since RA	U		0.03	• Repeated a grade since RA	U		0.07
		Health									• Current general health	F		0.05	• Had a baby since RA	U		0.05

NOTES: For full program names and citations, see Table 3.4.
0=impact estimate is zero; F = impact estimate is in favorable direction; U = impact estimate is in unfavorable direction.
* = statistically significant at the 10 percent level; ** = statistically significant at the 5 percent level; *** = statistically significant at the 1 percent level.
Effect size = absolute value of difference between program and control group outcome expressed as a proportion of the standard deviation of the outcome for the control group or for both groups combined.
Abbreviations:

A/N=abuse or neglect
BCS/SRC = Braken Basic Concept Scale/School Readiness Composite
BPI=Behavior Problems Index
C=child report
CU=Current
ERV = Emergency room visit
FU=followup
M=maternal report
P=parent/parent report
PBS/SCS=Positive Behavior Scale/Social Competence Subscale
PPVT-R=Peabody Picture Vocabulary Test-Revised
RA=random assignment
R&A=recipients and applicants
SSRC=Social Skills Rating System
T=teacher report
– = not available

In general, MFIP-IO produced largely favorable effects for children of long-term recipients. Statistically significant beneficial impacts were concentrated in the behavioral and schooling domains. The impact for one health indicator, emergency room visits for any child in the family, was unfavorable and statistically significant for recipients. For applicants, most effects were not significant, but a few in the school performance domain were unfavorable. Other results (not shown) indicate that MFIP-IO improved the physical home environment and reduced harsh parenting for recipient children while the reverse was true for applicants (Gennetian and Miller, 2000). Recall that in Chapter 8, MFIP-IO had stronger effects on income for long-term recipients than for recent applicants. Also, applicant children in the control group performed better than recipient children in the control group, indicating there was less room for the program to improve child outcomes among the more advantaged applicant children.

10.2.2. Programs That Focus on Financial Work Incentives Tied to Hours of Work

Child impacts are available for two of the programs—New Hope and SSP— that are classified in Panel B of Table 10.1 as evaluating financial work incentives tied to hours of work. An analysis of child outcomes is expected in 2002 for the SSP Applicant study.

The New Hope Child and Family Study administered questionnaires to parents and teachers for the sample of families at random assignment with at least one child age 1–10 (Bos et al., 1999). At the two-year assessment, a focal child, then age 3–12, was the subject of parental reports on child behavior and school progress. Teachers also provided ratings on indicators in these domains for children in kindergarten and above. Information was also collected directly from children starting at age 6. In addition to presenting results for all children, results are often stratified into three age groups: 3–5, 6–8, and 9–12. The sample sizes for many of the impact results shown in Table 10.1 are among the smallest of the studies we consider, ranging from under 250 combined treatment and control children aged 3–5 and 6–8, to over 600 children of all ages with parental reports of school behavior. In addition to these results, Bos and Varga (2001), in a separate analysis, report results for adolescents age 12 to 18 based on data collected for all children in the New Hope sample.

Even though the New Hope parental outcomes, such as employment, earnings, and income, shown in previous chapters differed by employment status at the time of random assignment, there were few differences in child outcomes across the two groups. Pooled results are summarized in Table 10.1. With the exception of the adolescent results, only two of the outcomes recorded in the table show a significant favorable effect, while none shows a significant unfavorable effect. The teacher's report of school performance and positive social behavior both favor the treatment group, with an effect size that equals about 0.25 of a standard deviation in each case. Parents' reports of total positive behavior are also more favorable for the treatment group. None of the outcomes measures shown separately for the two younger age groups are statistically significant, but small sample sizes mean that small differences are unlikely to be detected. The impacts for New Hope adolescents are mixed, with both favorable and unfavorable statistically significant impacts across the behavior and school achievement domains. On the positive side, New Hope adolescents are less likely to be in special education and more likely to be in a gifted or talented program. On the negative side, they are more likely to have repeated a grade, have a higher number of contacts by the school for behavior or academic problems, and are reported by their parents to be performing more poorly in school.

The SSP evaluation assessed child outcomes three years post-randomization through a survey of participants and their children (Morris and Michalopoulos, 2000).[96] Outcome measures covered the child's social and antisocial behavior, school progress and achievement, and health and safety. In addition, it is among the few studies to administer achievement tests to children—the PPVT-R for children age 4 to 7 and a math skills test for children age 7 to 15—to directly assess academic performance. SSP reports results stratified into three age groups (3 to 5, 6 to 11, and 12 to 18), with about 1,000 children each in the youngest and oldest age strata, and about 450 in the middle strata.

The impacts for SSP are striking in how consistent the results are within the age strata. For the youngest children, none of the outcomes were significantly affected by the program. Of the significant impacts for children age 6–11, all were favorable and centered on measures of school achievement (math score and maternal report of achievement in specific subjects) and health (general health and presence of long-term health problems).[97] The effect sizes were generally small, however. In contrast, for the oldest age group, all the statistically significant effects were in the unfavorable direction, with detrimental effects concentrated in the behavior domain. For instance, children age 12–18 at follow-up had higher rates of school behavior problems, minor delinquent activity (15–18-year-olds only), and use of tobacco, alcohol, and drugs, with effect sizes that range from about 0.10 to 0.20 standard deviations. At the same time, there were no significant treatment-control differences in many of the other indicators measured for this age group, including math and reading test scores. It is worth noting that adult outcomes for SSP families with children age 12–18 at follow-up were as favorable as they were for families with children in the youngest and middle cohorts.

10.2.3. Programs That Focus on Mandatory Work-Related Activities

Panel C of Table 10.1 records the child outcomes measured as part of L.A. Jobs-First GAIN (Freedman et al., 2000b), as well as the 11 NEWWS programs (Hamilton, Freedman, and McGroder, 2000; Hamilton et al., 2001).[98] These programs, which focus on mandatory work-related activities, collected information at the two-year or five-year follow-up on the children of recipients and applicants, either single parents (Los Angeles) or all parents (NEWWS). In 6 NEWWS programs (LFA and HCD programs in Atlanta, Grand Rapids, and Riverside), the Child Outcomes Study (COS) collected additional measures for a focal child age 3 to 5 at randomization (ages 8 to 10 as of the five-year follow-up). We discuss those findings in the context of the measures reported in Table 10.1 for children in this age range, but we do not report the additional COS measures in the table.

[96]For some outcomes, reports were made by both parents and children, while others were collected from parents only or children only. When results are available from both parent and child reports, we record the parent result in Table 10.1. Unless otherwise noted, the parent and child impact results were very similar.

[97]There was no statistically significant difference in the child report of average school subject achievement and in general health status. Both measures were reported by those age 10–11 only. For both of these indicators, the favorable impact measured in the parental reports is strongest for children age 6–8, suggesting that the difference in parent and child reports results from a lower impact among the older children in this age range.

[98]Child outcomes were not assessed as part of the Indiana IMPACT evaluation.

L.A. Jobs-First GAIN collected child-level information through a client survey two years past baseline for close to 1,600 children (Freedman et al., 2000b). Information collected covered such areas as academic achievement and school performance, behavioral and emotional adjustment, and safety. Results are recorded for all children, as well as for children classified by age at the time of follow-up: 5 to 7, 8 to 11, and 12 to 20, with over 400 children in each group. For the pooled sample of children, there was only one statistically significant effect (and that only at the 10 percent level): a more favorable outcome on school expulsions/suspensions (9.3 percent treatment versus 12.9 percent for controls).

Children in the youngest age group (5 to 7 at the time of the follow-up) were significantly more likely to repeat a grade (6.2 percent versus 0.4 percent for controls) and to have a special physical, emotional, or mental condition that made their parents' work difficult. This effect may be the result of the higher work effort among the treatment group parents. For children 8 to 11 at follow-up, the treatment group experienced a significantly higher rate of attending a special class for physical, emotional, or mental condition (15.5 percent versus 9.8 percent for the control group). The other impact estimates were mixed. For the oldest age group, up to 18 at random assignment and 20 at the follow-up, there were no statistically significant treatment-control differences on any of the indicators. In some cases, the impact estimates were favorable and in others not favorable.

Outcomes in the NEWWS evaluation—collected for 4 programs only for the 2-year follow-up and for the other 7 programs as of the five-year follow-up—focused on maternal reports of behavioral adjustments, school progress, and health and safety for all children age 18 or under at random assignment, with samples that range from 500 to 1,200 as of the final follow-up. The COS, also collected through a survey at the two-year and five-year follow-ups, focuses in more detail on academic functioning, social skills and behavior, and health and safety for young children age 3 to 5 at random assignment based on reports from mothers, teachers, and the children. Sample sizes range between 250 and 550 depending on the measure and the site. For the 4 programs with only two-year follow-up impact estimates, since many of the indicators are only relevant for school-age children, the results recorded in Table10.1 for all children are for the analyses conducted on the sample of families where all children were age 6 and above.

Overall, the NEWWS child outcome results show no clear pattern of beneficial or harmful effects for children up to age 14 at follow-up across domains within the same program or across the 11 programs. Both favorable and unfavorable effects are found across all the domains, sometimes for the same program. Typically, one or just two specific outcomes out of the seven reported outcomes in the five-year follow-up or five reported outcomes in the two-year follow-up have statistically significant impacts. Results for infants and toddlers (those age 1 to 2 at random assignment and 6 to 7 at follow-up), available for only two sites, show largely favorable effects for the two Grand Rapids programs and more unfavorable effects for the Portland program. The results of the COS (not shown) which focuses on pre-school-age children at random assignment also shows no clear pattern of favorable or unfavorable effects, and the impacts that were found were not related to the program approach (Hamilton et al., 2001). For the adolescents at random assignment (those age 15 to 23 as of the five-year follow-up), however, there is higher prevalence of statistically significant unfavorable effects, especially for schooling outcomes in the Riverside HCD program, the program with the largest income declines (see Figure 8.2).

10.2.4. Programs That Focus on Financial Work Incentives and Mandatory Work-Related Activities

Child impacts are available for three of the programs—WRP, MFIP, TSMF—that are classified in Panel D of Table 10.1 as combining financial work incentives and mandatory work-related activities.[99] The results for the full WRP evaluation show even fewer significant impacts on children in total and for those age 10 and above than the Incentives Only component of the program (Bloom, Hendra, and Michalopoulos, 2000). The only significant impact is a higher rate of participation for all children in clubs and organizations (34.2 versus 26.5 percent). There is no clear pattern with respect to the relative contribution of work requirements on top of the incentive program. The only statistically significant difference is in the measure of school absence, which has a more favorable outcome in the combined program.

The results for the combined MFIP closely mirror those seen earlier in Panel A for MFIP-IO for children age 5–12 at follow-up. In particular, children aged 5–12 of long-term recipients experienced several favorable impacts concentrated in the behavior and school domains. Many of the same indicators significant in Panel A are likewise significant in Panel D. When selected indicators were considered for recipient children younger than age 9 versus age 9 and above, the latter group (those who were school-age at random assignment) had stronger impacts (not shown). Two school performance measures and the BPI had favorable effects for the older subgroup. There were no significant effects for the youngest children, those who were pre-school-age at the start of the experiment. Compared with the older cohort, adult MFIP participants in this group experienced a larger increase in employment and income.

A comparison of the two MFIP interventions (full MFIP versus MFIP-IO) for long-term recipients indicates that the favorable effects on child outcomes in terms of behavior and school performance can be attributed to the financial work incentives component of the program. The addition of the work requirements had an unfavorable effect on a measure of positive behavior—the opposite of the effect of financial work incentives alone—so that the impact of the full MFIP was close to zero (not shown). At the same time, all other child indicators were unaffected by the addition of the work requirements. This suggests that adding mandatory work-related activities to a program with more generous financial work incentives may not be that harmful to children. This is despite the fact that adding the work requirements further increased full-time employment and lowered welfare benefits.

For recent applicants, the full MFIP generally had no effect on the child indicators measured for those aged 5 to 12. The one exception was a measure of whether the focal child had been suspended or expelled since random assignment. Small sample sizes make it more problematic to separate the effects of work requirements versus financial work incentives for the applicants in the study. Compared with the long-term recipients, children in the recent applicant group began with fewer disadvantages and were more heterogeneous, which may explain some of the differential impact.

MFIP also provides a more limited number of measures, mostly in the schooling domain, for children age 13 and above at the time of follow-up (age 10 and above at random assignment). As seen in Panel D of Table 10.1, the impacts for these adolescents are all insignificant for long-

---

[99] Iowa's report on child outcomes in FIP is expected later in 2002.

term recipients and mixed in sign, but they are all unfavorable and, with one exception, statistically significant for the adolescents of recent applicants. It is unclear whether these unfavorable impacts might be associated with the work requirements or financial work incentives components of the program, although recent applicants were not subject to the MFIP work requirements for much of the follow-up period.

The evaluation of the Michigan TSMF program is one of the few to rely exclusively on administrative data to assess child outcomes. The evaluation shows no significant differences between treatment and controls in any of the measures considered: substantiated reports of abuse and neglect, placement in foster care, and for older children (age 12 and above), employment and earnings. The labor market outcomes for youth were examined because Michigan's program allowed a 100 percent disregard of earnings from dependent children. Despite this program feature, it does not appear to have significantly affected work effort on the part of teens.

10.2.5. Programs That Focus on Other Reforms

Of the four studies classified in our residual category and listed in Panel E of Table 10.1, only PPI and PIP include analyses of child outcomes. Since these two demonstrations focused on parental responsibility requirements, specifically with regard to preventative health care (PPI) or immunizations (PIP), the main outcomes of interest pertain to the domain of child health for younger children.

The PPI program in Maryland required families with children age 3–24 months to verify that their children received preventative health care, including immunizations. Data collected through medical records abstraction for nearly 1,800 treatment and control children one and two years after randomization provided information on preventative visits per year and on whether children were up to date for three specific vaccinations (Minkovitz et al., 1999). The results show no significant treatment-control differences at either the first or second year follow-up in whether vaccinations were current, and no significant differences in the number of well-child visits to a primary care provider or whether at least one well-child visit was made a year.

Georgia's PIP, which required families to demonstrate proof of up-to-date immunization status semiannually (up to 1996) or annually (after 1996), served families with children age 6 or younger. Medical records were examined for about 2,800 treatment and control children in the demonstration four years after randomization to determine age-appropriate rates of immunization annually for five specific vaccinations (Kerpelman, Connell, and Gunn, 2000). In each of the four years post-randomization, children in the treatment group were significantly more likely to be current on at least four, and most often all five, of their vaccinations. For example, four years after randomization, 87.5 percent of treatment children were up to date on their polio vaccination compared with 80.1 percent of the control group. The impacts for this vaccination and two other vaccinations with relatively high rates of immunization even in the control group (specifically DTP, Diphtheria-Tetanus Toxoids-Pertussis, and MMR, Measles-Mumps-Rubella) are considered large given the potential for a "ceiling effect" (i.e., efforts to increase immunization rates often reach a "ceiling" beyond which further increases to 100 percent coverage are difficult to achieve).

10.2.6. Programs That Focus on TANF-Like Bundles of Reforms

Of the six programs classified in Panel F with a focus on TANF-like bundles of reforms, to date only ABC, FTP, and Jobs First report analyses of child outcomes.[100] Delaware's ABC evaluation examined child protective services administrative data for nearly 4,000 children in the treatment and control groups to assess the impact of the program on substantiated reports of child abuse and neglect (in aggregate and for subcomponents of maltreatment) and placements in foster care (Fein and Lee, 2000).[101] The results show a statistically significant increase in the incidence of child neglect in years one and three but not in year two.[102] For example, in year one, 2.6 percent of the control group had a substantiated report of child neglect compared with 4.1 percent for the treatment group. No significant differences were found for other types of maltreatment (e.g., physical and emotional abuse or sexual abuse) and foster care placements. The combination of benefit decreases and increased work effort in years one and three is suggested by Fein and Lee (2000) as an explanation for this result. In contrast, while benefits fell in the second year, earnings did not increase. This interpretation is consistent with Paxson and Waldfogel (1999) who estimate that increased maternal employment in single-parent families increases child maltreatment. Since income—which is sometimes negatively associated with the incidence of abuse and neglect in Paxson and Waldfogel's (1999) study—did not increase, the negative employment effect would be expected to dominate.

Florida's evaluation of FTP covered behavior problems, school outcomes, health status, and, for older children, delinquency and child bearing measured through a client survey four years post-randomization (Bloom et al., 2000). There are few statistically significant impact estimates on a range of child outcome measures, evaluated separately for about 1,100 children age 5–12, and nearly 750 children age 13–17. For the younger age group, FTP led to a statistically significant unfavorable outcome on the positive behavior scale, but the mother's report of the child's general health was significant and favorable. There were no statistically significant treatment-control differences for the youngest children in terms of current school achievement as reported by the mother, use of special education, or school suspensions, with impact estimates in both the favorable and unfavorable direction.

For the older children, FTP resulted in an increase in the rate of school suspensions (41 percent for the treatment group versus 33 percent for the control group), equal to 0.17 of a standard deviation. The maternal report of educational success was also unfavorable and marginally significant. No significant differences in other outcomes for the older children such as grade repetition, ever arrested, or having a baby were found in the four-year follow-up; all but one of these insignificant impact estimates was in the unfavorable direction, however. The contrast between the younger and older children in FTP is not as sharp as it was for the age differences observed in SSP.

---

[100]A child outcome study is included as part of the Indiana evaluation with results expected to be released in late 2002.

[101]The Delaware evaluation will include an analysis of child schooling outcomes in the future. ABC required parents to ensure children were attending school as part of a Contract of Mutual Responsibility (Fein, Lee, and Schofield, 1999). If attendance is unsatisfactory, the parent has agreed to cooperate with efforts to address the problem. Violations can lead to a reduction in the size of the cash grant, up to a permanent loss of benefits.

[102]Recall that in Delaware, the control group was enrolled in ABC 18 months after the initial randomization. Thus, by the second year of follow-up, some controls had received the treatment for up to 6 months. Also, by year three a substantial fraction of the treatment group had begun to hit their 24-month time limits (but none of those in the treatment group whose clocks started 18 months later).

The Jobs First three-year follow-up survey collected information on school achievement for all children under age 18, collected information on contact with the police and fertility for older children age 13 to 17 (a sample of about 1,000 adolescents), and collected more detailed information on behavior and functioning for about 1,500 "focal" children age 5 to 12 at follow-up (Bloom et al., 2002). (We do not include results based on a survey of the teachers of a subset of the focal children.) For the focal children, the impacts are largely favorable, and in the case of behavioral outcomes also statistically significant (although the effect sizes never exceed 0.1). For adolescents, however, the results are more mixed. Impacts in the schooling and health domains are all unfavorable, although only one impact (current school achievement) reaches statistical significance, perhaps because of smaller sample sizes. At the same time, adolescents of Jobs First participants were less likely to be convicted of a crime.

10.2.7. Subgroup Differences

Given the variation in the subgroups analyzed across random assignment studies (summarized in Appendix A), it is difficult to draw firm inferences about subgroup differences associated with different policies or programs. None of the programs that focus on financial work incentives alone included analyses by subgroups. Likewise, there are too few analyses by subgroups for the programs that focus on work requirements alone to draw solid conclusions about likely differential impacts.

Among the programs that combine financial work incentives and mandatory work-related activities, more differences by subgroup are analyzed, but they do not show a clear pattern. Impacts by child gender sometimes favor boys and at other times favor girls. MFIP appears to generate more favorable impacts for those at greater socioeconomic risk. At the same time, New Hope and SSP do not reveal any differences by characteristics that also capture risk of dependency on welfare.

10.3. ECONOMETRIC STUDIES OF THE EFFECTS OF WELFARE REFORM ON CHILD WELL-BEING

Given the data issues discussed in the introductory section, it should not be surprising that there are so few econometric studies of the impact of welfare reform on child outcomes. Most of the outcomes of interest—such as impacts on cognitive, emotional, and social development; behavior problems; school performance; and child health—are simply not collected for large nationally representative samples over time. Without such data, it is difficult to implement the DoD methodology required to control for unobserved confounding factors.

We are aware of just one relevant econometric study, one that uses administrative data on child maltreatment and the DoD methodology to investigate the impact of welfare reform in general and specific reform policies. Paxson and Waldfogel (2001) use state-level administrative data on child maltreatment and foster care to model the relationship between these outcomes and welfare policies, measured by the existence of a pre-TANF waiver, the existence of a family cap under AFDC or TANF, benefit levels under AFDC or TANF, and then specific policies post-TANF (work requirements, sanctions, and time limits). The outcomes they model by state and year from 1990 to 1998 include the log of reports of child abuse and neglect (in total and disaggregated), substantiated cases of abuse and neglect and the substantiation rate, and the number of children in out-of-home (primarily foster) care. Controls are included for the size, age, and race composition of the child population; the fraction of children in urban areas; the proportion of children whose mother has less than a high school degree; the unemployment rate; and state and year fixed effects. Controls are not included, however, for other state-level child welfare policy variables that were changing over this period.

Table 10.2 reports the estimated model parameters. Results for the six outcomes are provided in each row, each representing one regression model with controls for the six welfare policy variables recorded in each column. Across the models they estimate, there is some evidence to suggest that welfare policy may affect child maltreatment outcomes, but many of the estimated effects are not statistically significant. The statistically significant coefficients suggest that a family cap lowers substantiated cases but raises out-of-home care. Paxson and Waldfogel hypothesize that a family cap reduces abuse by limiting family size, but the other studies in Chapter 7 suggest that family caps do not have a major impact on childbearing. Among the other effects, immediate work requirements under TANF are associated with increased foster care, while a first full-family sanction under TANF is estimated to raise reports of physical abuse by 16 percent and substantiated cases by 22 percent. None of the coefficients on the TANF time limit measure are statistically significant. Likewise, the estimated parameters reported in Table 10.2 show no statistically significant effects of a state waiver for a work requirement, welfare time limit, or work incentives prior to implementing TANF.

Taken together, the welfare policy variables are only jointly significant in the model of substantiated cases and out-of-home placements. As with the econometric models reviewed in other chapters, there may be too little variation in the time period—with data extending only to 1998—to separately sort out the effects of the different welfare policy variables on child outcomes. In addition, the absence of controls for other potentially relevant policy variables may bias the estimated welfare policy impacts. For these reasons, we place less weight on the results from this one econometric study.

10.4. EVALUATING THE EFFECTS OF WELFARE REFORM ON CHILD WELL-BEING

The studies reviewed in the prior section paint a complex portrait of the multiple domains and varied indicators of child well-being that can be affected by welfare reform. Our objective in the remainder of this section is to bring these results together to provide a more coherent understanding of the impact on child outcomes of welfare reform as a whole, and the specific policies and programs embedded in TANF.[103]

10.4.1. Effects of Specific Reforms

Much of our information about the effects of specific reforms comes from the experimental studies, although the one econometric study reviewed in this chapter also looked at policy components. There are some clear patterns across some of the specific reforms, while other specific reforms suggest more uncertain or mixed relationships.

[103]Other recent syntheses of this literature include Michalopoulos and Berlin (2001), Bloom and Michalopoulos (2001), Duncan and Chase-Lansdale (2001), Hamilton, Freedman, and McGroder (2000), and Morris et al. (2001).

Table 10.2—Estimated Impact of Welfare Reform on Child Maltreatment: Econometric Studies

						Welfare Policy Variable					
						AFDC/TANF Family Cap	TANF Immediate Work Requirement	TANF First Full Family Sanction	TANF Time Limit	ln (AFDC/TANF Benefit Level)	Any waiver
Study	Data	Begin	End	Outcome	Dep var.	% effect (t-stat)	% effect (t-stat)	% effect (t-stat)	% effect (t-stat)	Elasticity (t-stat)	% effect (t-stat)
Paxson and Waldfogel (2001)	State Child Protective Systems aggregated data	90	98	Reports of child abuse and neglect	Log	0.0 (0.00)	4.9 (0.86)	9.7 (1.39)	6.2 (0.71)	-0.37 (0.72)	-6.8 (1.33)
				Reports of physical abuse	Log	-6.8 (1.12)	-3.9 (0.63)	15.7 (2.06)	4.3 (0.45)	-0.20 (0.36)	-4.9 (0.89)
				Reports of neglect	Log	-15.2 (1.19)	-2.9 (0.35)	13.9 (1.38)	4.4 (0.35)	-3.15 (4.29)	-4.3 (0.59)
				Substantiated cases	Log	-14.0 (2.39)	-9.4 (1.55)	22.1 (2.98)	4.2 (0.45)	-0.41 (0.75)	-5.2 (0.97)
				Substantiation rate	Log	-14.0 (2.04)	-14.3 (2.02)	12.4 (1.42)	-2.1 (0.19)	-0.04 (0.06)	1.6 (0.26)
				Out-of-home care	Log	15.6 (3.64)	8.7 (2.03)	2.1 (0.37)	-7.4 (1.16)	-0.80 (2.19)	0.6 (0.18)

NOTE: All models include state and year fixed effects and controls for the size, and age and race composition of the child population; the fraction of children in urban areas; the proportion of children whose mother has less than a high school degree; and the unemployment rate.

Financial Work Incentives

Strong financial work incentives that lead to greater work effort and higher earnings and income, either alone or when tied to hours worked or in combination with work requirements, are often beneficial for children who are pre-school and elementary-school age at the time they entered the program. However, there are a few examples of unfavorable impacts as well. Both MFIP and MFIP-IO show positive impacts in the behavior and schooling domains for children of the more disadvantaged group of recipients who were between the ages of 2 and 9 at random assignment. A few outcomes were unfavorable for the more advantaged MFIP-IO and MFIP applicant children in the same domains. For longer-term recipients in MFIP, the addition of work requirements on top of a program of generous financial work incentives appears to have little in the way of positive or negative consequences for children, at least those who are preteenagers at random assignment. Coupled with the financial work incentives, MFIP recipient children appear to benefit from the higher income that results from the welfare benefit structure, even with the work-related mandates. The absence of unfavorable outcomes for WRP and TSMF—two programs with weaker financial work incentives—further supports this view, although these two studies considered a more limited set of child indicators.

The patterns are similar when financial work incentives are tied to hours work as in SSP and New Hope. SSP, which includes a younger age cohort, shows no significant impacts on the youngest children, and only favorable significant impacts for the children in the middle age range. The New Hope findings of positive impacts for children overall in school performance and a measure of positive behavior, and no statistically significant unfavorable effects for the other measures (or across age groups), reinforce the view that financial work incentives can be beneficial, or at least not harmful, for pre-school-age and primary-school-age children.[104]

There is some evidence, but more limited, that programs with financial work incentives alone or those tied to hours worked or work requirements may have unfavorable impacts on adolescents even when they increase family income. WRP-IO and WRP produced one negative impact (significant only for the former program) in the behavioral domain (trouble with the police), although this program had small impacts on adult behavior and family income. Adolescents of the more advantaged MFIP applicants experienced unfavorable impacts in the behavioral and schooling domains that were not evident for the more disadvantaged adolescents of longer-term recipients. (Impacts for adolescents in MFIP-IO are not reported.) SSP and New Hope, both programs with financial work incentives tied to hours worked that produced significant increases in income, also find unfavorable impacts for teens in the behavioral and schooling domains, although New Hope also had a few favorable schooling outcomes.

Thus, the experimental evidence suggests that more work and more income resulting from financial work incentives alone, or in combination with work requirements, are generally neutral or favorable for younger children (preschool and elementary age) but may be detrimental for adolescents, at least for some areas of development. The favorable effects are concentrated in the behavior and school performance domains for the younger children, while

---

[104]New Hope also offered a generous child care subsidy for children up to age 13. The increased use of formal child care centers and after-school programs may explain some of the favorable impacts on child outcomes for this program.

the negative impacts for the older youth fall primarily in the behavior domain. These results are consistent with the broader literature that evaluates the relationship between family income and child outcomes across the life course. Duncan and Brooks-Gunn (1997), in their synthesis of recent studies on this topic, conclude that "[f]amily economic conditions in early and middle childhood appear to be far more important for shaping ability and achievement than they do during adolescence (p. 597)." The MFIP applicants, a more advantaged group, provide the one exception where even younger children experienced some unfavorable impacts for a program with incentives alone or combined with work requirements. Since the income gains for these families were smaller, it is not clear what other factors can explain these less favorable outcomes.

Mandatory Work-Related Activities

The L.A. Jobs-First GAIN study and NEWWS results demonstrate small but mixed effects on children of programs that require mandatory work-related activities in general and of programs with an employment versus education focus. Recall that L.A. Jobs-First GAIN led to a significant increase in earnings that was offset by falling cash welfare payments because of the high benefit reduction rate. Thus, pretax income changed little for the participants compared with the controls. Accounting for the EITC and other taxes resulted in a slight gain in income for the Los Angeles participants. Like the L.A. Jobs-First GAIN study, the NEWWS programs generally produced small, if any, income gains and little change in poverty. In general, these studies provide very mixed impacts for pre-school-age and primary-school-age children across all three domains, with both favorable and unfavorable impacts but many that were not statistically significant. Again, there is more consistent evidence of unfavorable impacts for adolescents, especially in the school achievement domain. In the NEWWS demonstrations, there was no clear relationship between program impacts on income and impacts on children, although there is some evidence to suggest worse child outcomes with higher employment and lower income. It would appear from these results that reductions in welfare dependency without significant gains in income result in ambiguous effects on child outcomes, with examples of both favorable and unfavorable impacts. The limited econometric evidence also suggests a weak or inconsistent impact of work requirements.

Time Limits

The evidence with respect to time limits is more limited. The one econometric study suggests no impact of time limits on child maltreatment, but we have placed less weight on this study because of methodological concerns. Of the experimental studies that include time limits, none is designed to estimate the specific impact of time limits separate from the other program features in the bundle of reforms. Moreover, unlike some of the other outcomes considered in earlier chapters, FTP and Jobs First do not provide impacts for child outcomes before and after time limits begin to become binding. Thus the DoD strategy employed in earlier chapters to infer the mechanical effects of time limits is not available to assess the impact on child outcomes. The pattern of impacts for FTP and Jobs First for both school-age children and adolescents is not markedly different from that for MFIP, which includes financial work incentives and work requirements but no time limit. Thus, it is not clear whether, on the margin, the addition of time limits in programs like FTP and Jobs First has favorable or unfavorable impacts on child well-being.

Parental Responsibility Requirements

Of the two studies that focus on parental responsibility requirements related to child health (i.e., preventative care or vaccinations), PPI had no effect on the required behaviors, while PIP had a sizeable and significant favorable impact. This difference may be attributed to the fact that PIP had larger sanctions compared with PPI. In the case of PIP, the sanction equaled a portion of the nonimmunized child's grant. PPI effectively levied a $10 per month sanction against a family that was out of compliance with the verification requirement.[105] Another explanation may be that recipients responded more to the PIP initiative because the intervention was focused only on changing immunization outcomes—with expectations that were easier to understand and comply with—compared with the broader set of requirements under PPI regarding health care (e.g., preventative care more generally and prenatal care) and school attendance (Kerpelman, Connell, and Gunn, 2000).

10.4.2. Effects of Reform as a Bundle

As noted in prior chapters, the econometric studies are potentially the best methodology for estimating the impact of waivers or TANF as a bundle. However, the limited number of studies on child outcomes using this approach makes it more difficult to ascertain the effect of welfare reform as a whole on the multidimensional concept of child well-being. The one econometric study reviewed in Section 10.3 suggests that waivers as a whole had very little impact on child maltreatment and placement in foster care. In the absence of similar econometric analyses of other child outcomes in domains such as behavior and cognition, school progress, other aspects of child health, and so on, it is not possible to conclude whether waivers as a whole had any effect on children.

In other chapters, we have also looked at the experimental evidence to gauge the possible effects of welfare reform as a bundle, especially for the demonstration studies that combine the three key features of most state TANF plans: time limits, financial work incentives, and mandated work-related activities. Panel F in Table 10.1 shows that three of the programs with these features have results for child outcomes. The ABC results suggest a possible unfavorable impact on child maltreatment (specifically neglect). For the school-age children at follow-up, FTP and Jobs First show both favorable and unfavorable impacts in the behavior domain, no effects on the school performance measures, and one favorable effect for health. Likewise, for adolescents at follow-up, there is mixed evidence in the behavior domain, more consistent evidence of unfavorable impacts on school performance, and no significant impact in the health domain (here a measure of teen fertility). These results, while far from conclusive, suggest that welfare reform as a package may affect several domains of child well-being, including antisocial and problem behavior, school achievement, and health, but the specific impacts and their differences by child age are less well understood.

---

[105]The stated sanction was $25 per month but compensatory policies in food stamps and housing vouchers effectively reduced the sanction to $10.

The subgroup results for the programs that evaluate TANF-like bundles of reform are also mixed, with less favorable effects in ABC for families at greater socioeconomic risk and better outcomes for families in FTP at the greatest risk of long-term dependency. There is some nonexperimental evidence that the parents in the lower risk group in Florida were less likely to closely supervise their children, so these children may have been more prone to problem behavior. This group had the largest earnings impacts, especially near the end of the follow-up period.

10.5. CONCLUSIONS

The studies reviewed in this chapter reveal that there is scope for both positive and negative effects on child well-being of various components of welfare reform policies and programs. Positive and negative effects were observed for indicators that capture socioemotional behavior, academic performance, and health. The most favorable effects are associated with financial work incentives, most likely because of the increase in family income that is accompanied by combining work and welfare. But even for these programs, there is some evidence of unfavorable impacts for some subgroups of participants, especially for adolescent children of participants and for younger children of participants who do not experience large income gains. Work requirements do not appear to have either strong favorable or unfavorable impacts on children, although again there is evidence of unfavorable impacts for adolescents, especially in the school performance domain. There is too little evidence regarding the specific impacts of time limits to draw firm conclusions.

There is also relatively little evidence on which to draw solid inferences about the impact of welfare reform as a bundle on child well-being, based either on econometric or experimental data. In the case of the econometric literature, there is just one study and it is limited to one outcome domain. It is also difficult to extrapolate from the three currently available random assignment studies that evaluated the impact on child well-being of TANF-like bundles of reform since the policy combinations evaluated are not representative of the full range and mix of policies implemented by the states under PRWORA. To the extent that there are favorable effects from these studies, they are concentrated in outcomes for children who are school aged at the time of follow-up. The unfavorable impacts, in contrast, are concentrated in outcomes for adolescents, particularly in the area of school performance.

Thus, the impacts of welfare reform appear to differ with the stage of the child's development, regardless of the policy component or bundle of reforms considered, and for a given age, impacts may be favorable or unfavorable depending on the outcome domain considered. Based on the experimental evaluations that assess child well-being, it appears that there are countervailing forces that both promote and diminish healthy child behavioral, social, cognitive, and physical development. The resulting impacts of welfare reform policies on child outcomes are likely to depend on the strength of the opposing forces and the child's stage of development and other circumstances. Moreover, it is possible that some consequences for children will not materialize until more time has passed under the new policy regime, with the potential for cumulative favorable and unfavorable impacts. Effects that are small now, whether positive or negative, may become more pronounced as more time passes.

Chapter Eleven

CONCLUSION

Beginning in the 1960s, concern about the unintended consequences of the AFDC program led to a sequence of reform efforts. The goals of these reforms were to promote work and reduce dependence while still alleviating need. These efforts culminated with PRWORA, which replaced the AFDC program with TANF. In addition to promoting work and reducing dependence, PRWORA also aimed to promote marriage and to reduce unwed childbearing. Welfare reform may also have implications for poverty and the well-being of low-income children.

The challenge for future reform efforts is not in achieving any one of these goals but in achieving them all simultaneously. As we will see below, most reform policies increase employment. Some raise both income and welfare use in the process, whereas others reduce welfare use but leave income unchanged. Likewise, some policies are more effective at improving children's outcomes or at least at not leaving children worse off. As lawmakers seek to refine the new welfare system, it is important that they understand the trade-offs that different policies entail.

Our task has been to synthesize the evidence on how recent welfare reform policies affect these goals, as measured by a series of outcomes. Each of the preceding chapters has focused on a particular set of outcomes, such as welfare use, employment and earnings, or income and poverty. In this concluding chapter, we evaluate the trade-offs among the different reform goals that arise from different policies. The next section synthesizes the literature across all outcomes and all policies. After that, we assess the strengths and limitations of the existing research base. We close by discussing directions for future research.

11.1. SYNTHESIZING THE LITERATURE ACROSS ALL OUTCOMES AND ALL POLICIES

To compare and contrast the impacts of our list of policies on our list of outcomes, we return to the idea of a matrix, like the one discussed in Chapter 1. In Table 11.1, we list policies we examined along the rows and outcomes we examined along the columns. The outcome columns appear in the order of the chapters in which the outcomes were discussed.[106] Many of the policy rows correspond to the entries in the tables in Chapters 4 to 10. In a few cases, we

[106]Table 11.1 lists only three of the outcomes considered in Chapter 9. The results for adult health insurance coverage are similar to what is shown for children's coverage. Many of the cells for the other outcomes considered in the chapter would be blank or have an asterisk.

Table 11.1—Impact of Welfare Reform as a Whole and Specific Reform Policies on Various Outcomes: A Synthesis of the Research

	Policy or Policy Bundle	Welfare Use	Employment	Earnings	Use of Other Government Programs		Marriage	Fertility	Income	Poverty	Other Measures of Well-being		
					Food Stamps	Medicaid					Food Security	Children's Health Coverage	Savings
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
(1)	Financial Work Incentives	INCREASE	INCREASE	*	*	*	INCREASE		INCREASE	DECREASE	INCREASE	INCREASE	INCREASE
(2)	Financial Work Incentives Tied to Hours Worked	INCREASE §	INCREASE	INCREASE	INCREASE		*		INCREASE	DECREASE	INCREASE		INCREASE
(3)	Mandatory Work-related Activities	DECREASE	INCREASE	INCREASE	DECREASE	DECREASE	NO CHANGE	NO CHANGE	MIXED	DECREASE	DECREASE	DECREASE	
(4)	Sanctions for non-compliance	DECREASE											
(5)	Mandatory Work-Related Activities and Strong Financial Work Incentives	INCREASE	INCREASE	INCREASE	DECREASE	MIXED	*		INCREASE	DECREASE	*	INCREASE	*
(6)	Mandatory Work-Related Activities and Weak Financial Work Incentives	DECREASE	INCREASE	INCREASE	DECREASE	DECREASE			INCREASE		*	DECREASE	
(7)	Time Limits (Before Recipients Reach Limit)	DECREASE	INCREASE	*				*	*				
(8)	Time Limits (After Recipients Reach Limit)	DECREASE	MIXED	*	*				*				*
(9)	Family Cap	MIXED				*		MIXED					
(10)	Parental Responsibility							*					
(11)	Reform as a Bundle (Before Recipients Reach Time Limits)	DECREASE	INCREASE	INCREASE	DECREASE	MIXED	MIXED	MIXED	INCREASE	DECREASE	*	*	*

	Policy or Policy Bundle	Child Well-being									
			Preschool Age at Follow-Up			Grade School Age at Follow-Up			Adolescents at Follow-Up		
		Child Abuse and Neglect (all ages)	Behavior Problems	School Achievement Problems	Health Problems	Behavior Problems	School Achievement Problems	Health Problems	Behavior Problems	School Achievement Problems	Health Problems
		(M)	(N)	(O)	(P)	(Q)	(R)	(S)	(T)	(U)	(V)
(1)	Financial Work Incentives					DECREASE	MIXED	INCREASE	INCREASE	*	
(2)	Financial Work Incentives Tied to Hours Worked		*	*	*	*	DECREASE	DECREASE	INCREASE	MIXED	*
(3)	Mandatory Work-related Activities	*	MIXED	MIXED	MIXED	MIXED	MIXED	DECREASE	MIXED	INCREASE	MIXED
(4)	Sanctions for non-compliance	*									
(5)	Mandatory Work-Related Activities and Strong Financial Work Incentives					DECREASE	DECREASE	INCREASE	INCREASE	INCREASE	
(6)	Mandatory Work-Related Activities and Weak Financial Work Incentives	*							*	*	
(7)	Time Limits (Before Recipients Reach Limit)	*									
(8)	Time Limits (After Recipients Reach Limit)										
(9)	Family Cap	*									
(10)	Parental Responsibility				*			DECREASE			
(11)	Reform as a Bundle (Before Recipients Reach Time Limits)	*				MIXED	*	DECREASE	MIXED	INCREASE	*

NOTES: * Cell has up to three moderate or high-quality studies with no significant impacts or a single moderate-quality study with a significant impact.

§ These programs increase the sum of welfare payments and the earnings supplement provided outside the welfare system, although welfare payments per se may decrease.

LEGEND

Knowledge base:	DIRECTION	DIRECTION	DIRECTION	
	Much evidence	Some evidence	Little evidence	No evidence

have disaggregated the policies further along dimensions where their effects may differ. For mandatory work-related activities combined with financial work incentives, we distinguish between programs with strong incentives (e.g., MFIP and FIP) and those with weaker incentives (e.g., WRP and TSMF). The latter programs involve implicit tax rates that may be higher than those under AFDC. For time limits, we distinguish pre–time limit effects from post–time limit effects. The effects of reform as a bundle, in row (11), pertain only to the pre–time limit period. The assignment of particular studies to the rows of Table 11.1 is detailed in Appendix B.

The entries in the cells of Table 11.1 are qualitative summaries of the effect of each policy on each outcome. The words in the cells indicate the direction of the effect, whereas the shading of the cells indicates the depth of the knowledge base, that is, how much is known about the effect of that policy reform on that outcome.[107] The entry “increase” indicates that a majority of the studies that analyze the policy-outcome pair in question show that the policy increases the outcome. The entry “decrease” indicates the opposite. The entry “mixed” indicates that there are roughly as many results showing a decrease as an increase. The entry “no change” indicates that the estimated impacts are mixed in sign and that nearly all are small and insignificant.

Cells with a deep knowledge base are indicated by the dark gray shading. These cells are populated by several high-quality studies, most of which yield similar and significant estimates. At the other end of the spectrum, cells with a shallow knowledge base are indicated by no shading. These cells are populated by a single high-quality study that yielded a significant estimate, two moderate-quality studies that yield similar and significant estimates, or similar constellations of evidence. Cells whose knowledge base falls between these two categories are indicated by intermediate gray shading. Cells populated by a single moderate-quality study, or one or more high-quality studies whose results were insignificant, are indicated by an asterisk, denoting a nearly empty knowledge base. Cells for which there are no studies are left blank.

The entries in column (B) indicate that most reforms or combinations of reforms considered in Table 11.1 increase employment, although we assign varying degrees of confidence to this qualitative assessment. Column (C) shows that many of these policies also increase earnings. Beyond employment and earnings, however, the impacts of specific policies vary to a greater extent, illustrating the trade-offs facing policymakers. Thus, we organize the remainder of our discussion of Table 11.1 by the table rows.

11.1.1. Financial Work Incentives

Although all the recent reform policies are capable of increasing employment (column (B)), they involve different trade-offs between reducing welfare use (column (A)) and increasing income (column (H)). Programs with generous financial work incentives generally increase welfare use, as seen in cell A1. This is also true for financial work incentives tied to hours of work through an earnings supplement outside the welfare system (cell A2), where transfer payments (welfare payments or the earnings supplement) increase. Welfare use also increases when generous financial incentives are combined with mandatory work-related activities (cell A5), but the opposite is true for programs with weaker incentives (cell A6).

[107]Appendix B provides the details on how we determined the indicators for the direction and depth of the knowledge base assigned to each cell in the table.

We are unable to assign a direction for financial work incentives alone on the use of food stamps or Medicaid (cells D1 and E1). A shallow evidence base suggests financial work incentives tied to hours of work may increase food stamp use (cell D2). When combined with work requirements, it appears that both strong and weak financial work incentives may decrease food stamp use (cells D5 and D6). Such programs decrease (weak incentives) or have an uncertain impact (strong incentives) on Medicaid use (cells E5 and E6). A very shallow knowledge base suggests that financial work incentives alone may increase marriage (cell F1), but we do not have enough evidence to say how marriage is affected when financial work incentives are tied to hours worked or combined with work requirements. There is some suggestive evidence from MFIP that programs that provide generous financial work incentives combined with work requirements may increase marriage or keep existing marriages intact. However, the mixed results for the Canadian SSP suggest caution in interpreting the MFIP results. There is no evidence base from which to assess the relationship between financial work incentives and fertility.

Since financial work incentives allow families to keep more of their welfare benefits as their earnings rise, they also increase income and decrease poverty, as shown in columns (H) and (I). However, these programs increase income by modest amounts. The most effective program was MFIP, which combined a strong financial incentive—stronger than most state TANF plans—with mandatory work-related activities for long-term recipients. MFIP increased the income of long-term recipients by 12–15 percent throughout the three-year follow-up period. However, even in MFIP, participants' incomes remained low, averaging $12,000 per year, even after accounting for the EITC. More than one-half of the participants had income from work and transfer programs that *still* fell below the official poverty line. At best, these programs abate poverty somewhat; none can be said to alleviate it altogether.

With one exception (cell K6), financial work incentives (alone, or tied to hours of work, or combined with work requirements) are also associated with improvements in other measures of well-being such as food security, children's health insurance coverage, and savings (the intersection of rows (1), (2), (5), and (6) with columns (J) to (K)). However, several of the relevant cells are empty, and those that indicate a favorable impact are derived from a shallow knowledge base.

The impact on child well-being is more uncertain, and, when we are able to assign a direction, it is almost always based on a shallow evidence base. For children who are school-aged at the time of follow-up, strong financial work incentives (alone, or tied to hours of work, or combined with work mandates) appear to decrease behavior problems and possibly also school achievement problems as well (the intersection of rows (1), (2), (5), and (6) with columns (Q) to (S)). The increase in health problems for this age group associated with strong financial work incentives comes from statistically significant unfavorable impacts of MFIP-IO and MFIP on emergency room visits, which may result from less parental supervision or better access to health care. Thus, it appears that for this age group, the increased income associated with reforms that incorporate strong financial work incentives may lead to some improvements in children's outcomes in certain domains.

In contrast, for adolescents at follow-up, the various policies that include financial work incentives consistently appear to increase behavior problems and school achievement problems (the intersection of rows (1), (2), (5), and (6) with columns (T) and (U)). (There is insufficient evidence regarding health problems for adolescents as seen in column (V).) The evidence base available to assess the impact of financial work incentives on outcomes for preschool-aged children at the time of follow-up is almost nonexistent although some of the impacts recorded in Table 11.1 for grade-school-aged children pertain to children who were preschoolers at the time of random assignment.

11.1.2. Mandated Work-Related Activities

Mandated work-related activities have been studied more than any other reform. Consequently, most of the cells with the darkest shading are in row (3). A substantial body of evidence shows that they generally reduce welfare use (cell A3). However, they have little effect on income, with 11 of 13 studies in the cell finding no significant impact and only one study each finding a positive or negative impact (cell H3). This is because, in the absence a financial work incentive, the increase in earnings generated by work mandates (cell C3) is offset nearly dollar-for-dollar by a decrease in benefit payments. The evidence base is also deep in indicating that mandated work-related activities reduce food stamp use (cell D3). These programs also appear to decrease Medicaid use, although the knowledge base for that conclusion is much shallower (cell E3).

However, viewed from a different perspective, row (5) of the table shows that it is possible to require work and raise income at the same time. The key is to combine the work requirement with a strong financial incentive, so that earnings rise more rapidly than benefits fall. The price for raising incomes is higher welfare use, which again illustrates the central trade-off facing efforts to reform welfare.

Turning to the poverty results in column (I), the more limited evidence available suggests that mandatory work-related activities decrease poverty somewhat (cell I3). Although these programs have no effect on mean income, they may be able to raise incomes for families just below the poverty line. This is consistent with the evidence presented in Appendix A that such programs have greater effects on income among relatively advantaged recipients than among disadvantaged recipients.

This policy has no effect on marriage or fertility (cells F3 and G3), a conclusion that is based on five years of follow-up data for seven programs and two years of follow-up data for five other programs (hence the dark shading). Regarding other measures of well-being, the available evidence suggests that mandated work-related activities reduce food security and children's health insurance coverage (cells J3 and K3). A somewhat more substantial evidence base provides a very mixed picture of the impact of these programs on child well-being for all three of the age groups shown in Table 11.1 (the intersection of row (3) with columns (N) to (V)). The only favorable assessment is for health problems of grade schoolers (cell S3), while the one clear unfavorable impact is for school achievement problems of adolescents (U3).

11.1.3. Sanctions

While financial work incentives and work mandates have been relatively well-studied, other policy reforms have been analyzed less thoroughly. As a result, their effects are less well understood. Sanctions are an important case in point (see row (4)). Many states have enacted sanctions that are substantially more stringent than those under JOBS. Moreover, many families have lost their aid, or at least part of their aid, because of sanctions. However, no experiments were conducted to isolate the effects of sanctions. Indeed, none of the experiments we consider involve any experimental variation in sanction policy, except in conjunction with other policy reforms. Some econometric studies of the caseload indicate that stricter sanctions have greater effects on welfare use, but evidence showing that substantial declines in welfare use preceded the imposition of such sanctions by several years clouds the interpretation of those findings (cell A4). With the exception of a single econometric study of child maltreatment (represented by the asterisk in cell M4), there are no studies of the effects of sanctions on any other outcome.

11.1.4. Time Limits

Time limits have been better studied than sanctions, but much less well-studied than mandatory work-related activities. Because the random assignment studies that involved time limits all involved other reforms as well, the experimental results from those studies do not isolate the effects of time limits. Several econometric studies have analyzed the behavioral effects of time limits, that is, how time limits affect behavior before recipients exhaust their benefits. These studies form the basis for the cell entries in row (7) (see Table B.1 in Appendix B). The cell entries in row (8) are based on nonexperimental estimates from two random assignment studies, along with one econometric study of employment, that provide some insights into how behavior changes once recipients begin reaching the time limit.

Most of the econometric studies suggest that time limits reduce welfare use during the pre–time limit period (cell A7). One set of studies reports results that are consistent with the notion that some families bank their months of eligibility for future use. Only two studies suggest that time limits also increase employment during the pre–time limit period (cell B7), so we place less confidence on this cell entry. There is insufficient evidence or no evidence available for assigning the direction of impact of time limits before recipients reach the limit for any of the other outcomes shown in Table 11.1, including child well-being.

The knowledge base regarding the post–time limit effects of time limits is even shallower. Two studies show that welfare use falls sharply once recipients begin to exhaust their benefits. Effects on employment are mixed, but none of the evidence suggests that it changes much, either up or down, once recipients start reaching the limit. Clearly, the post–time limit consequences of time limits could increase substantially once a higher proportion of the caseload reaches the limit.

11.1.5. Family Caps and Parental Responsibility

Table 11.1 also documents that we know relatively little about how family caps and parental responsibility requirements affect key outcomes. The limited available evidence points to a mixed impact of family caps on fertility (cell G9). An equally shallow evidence base also produces mixed evidence with respect to the impact of family caps on welfare use (cell A9). Parental responsibility requirements, specifically those related to well-baby and well-child services (e.g., vaccinations), have been assessed in terms of their direct impact on the behaviors they seek to change, with some limited evidence of favorable effects in terms of child health for young children (cell S10). How this policy affects other outcomes is unknown.

11.1.6. Welfare Reform as a Bundle

Looking beyond specific policy reforms, a number of econometric studies and the six random assignment studies that involved TANF-like bundles of reforms provide insights into the effects of reform as a bundle. For many outcomes, welfare reform as a bundle produces impacts similar to those seen for mandatory work-related activities with weak financial work incentives (compare rows (6) and (11)): a decline in welfare use and food stamp use, and an increase in employment, earnings, and income. This is plausible given that most states implemented weaker financial work incentives combined with mandatory work-related activities, and given that what is known about the behavioral impacts of time limits suggests that they operate in the same direction as weak financial work incentives and work mandates (compare cell A6 with A7, and B6 with B7).

For other outcomes, the knowledge base is very shallow. The impact on Medicaid use, marriage, and fertility is mixed (cells E11, F11, and G11), while poverty appears to decrease (cell I11). There is too little evidence to assess the impact of reform as a bundle on other measures of well-being in columns (J) to (L). In the case of the child well-being outcomes in columns (N) to (V), the limited available evidence appears to show a mixed impact on behavior problems of young children and adolescents, and an increase in school achievement problems for adolescents. There is some indication of reduced health problems for grade schoolers. However, the cells that are signed in these columns are based exclusively on results from two random assignment studies, FTP and Jobs First. The bundle of reforms implemented in these two states is not very representative of the reforms implemented in other states in terms of the length of the time limit (two years or less in both cases) or the generosity of the financial work incentives (notably in Connecticut). Thus, the impacts in row (11) for these columns should be interpreted cautiously.

It is also important to note that, regardless of the depth of the knowledge base, the entries in row (11) represent the effects of reform as a bundle during the pre–time limit period. Post–time limit evidence is very limited, and most studies summarized in this row cover time periods prior to when recipients could have exhausted their benefits. Once recipients reach the limit in substantial numbers, these effects could change.

11.1.7. Welfare Reform Effects on Subgroups

As detailed in Appendix A, for the most part, the effects of reform do not generally appear to be concentrated among any particular group of recipients. Many observers would view this as good news, since there was widespread concern when PRWORA was enacted that only relatively advantaged recipients would respond, leaving the most disadvantaged behind. The subgroup-specific analyses provide no consistent evidence about this effect. In some cases, subgroup-specific impacts are similar for persons of different levels of disadvantage. In other cases, different measures of disadvantage generate different patterns, some appearing to favor the relatively advantaged and some appearing to favor the relatively disadvantaged. In many cases, subgroup-specific estimates are insignificant, in part because subgroup-specific sample sizes are too small to generate precise results even when the program has a substantial effect.

11.2. STRENGTHS AND LIMITATIONS OF THE EXISTING KNOWLEDGE BASE

Another way to use Table 11.1 is to look across the whole table and assess the general state of the knowledge base (the types of shading or lack of shading in the various cells). Table 11.1 reveals that the knowledge base is strongest for understanding the impact of various welfare reform policies on welfare use, employment, earnings, and income. The base is weakest for assessing the impact of policies on broader measures of well-being, especially child outcomes, most notably those for pre-school-age children. Among the policies, a solid base of research exists to evaluate the impacts of mandatory work-related activities on most outcomes, and it is nearly as strong for financial work incentives, either alone or when tied to hours worked or in combination with mandatory work-related activities. As we have already discussed, several reform policy components have received less attention, most notably time limits, sanctions, family caps, and parental responsibility requirements. Overall, just under half the cells in our matrix (120 out of 242 cells) are empty, indicating no research base exists to assess the policy-outcome pair. Another 36 cells (those with an asterisk) are nearly empty.

Some of the gaps in the knowledge base are particularly relevant for policy. For example, there have been relatively few causal studies of how welfare reform has affected Medicaid participation, as shown in column (E), or the health care coverage of children more generally, as shown in column (K). This omission is particularly important in light of the initial decreases in Medicaid enrollment that occurred following the implementation of TANF—despite 15 years of policy initiatives designed to increase the coverage of poor children. As seen in columns (F) and (G), less is known about the impact of individual welfare reform policies and reform as a whole on marriage and fertility despite continued interest among many policymakers in policies to promote the formation and maintenance of two-parent families and to reduce out-of-wedlock childbearing. Columns (N)–(P) show that little is known about welfare reform and child development prior to school entry, which is troublesome given the increased emphasis on work for mothers of children as young as age one or younger. This is an issue that is particularly relevant for policies aimed at improving early care and education.

For the policy-outcome combinations where we have a more substantial knowledge base, a nearly universal limitation of our conclusions is that they apply mostly to the short run. Most of the studies present evidence from follow-up periods of roughly two years, although the 11 NEWWS programs and several others provide results based on four or five years of follow-up data. The limited available evidence suggests that some of the effects change over time for reasons that are not well understood.

The short-run nature of the evidence limits our understanding of whether reform has accomplished its goals of reducing unwed childbearing, encouraging marriage, and maintaining two-parent families. Marriage and fertility involve substantially more inertia than other aspects of behavior. As a result, we would expect the effects of welfare reform on such outcomes to become apparent only over a longer horizon. With mostly a short-run follow-up period to draw on, it should come as little surprise that most of the evidence from high-quality studies is mixed and statistically insignificant. One exception is MFIP-IO, which appears to have decreased marital disruption among married couples and perhaps increased marriage among previously unmarried recipients. However, the results from SSP suggest caution in interpreting the MFIP-IO results. In SSP, a similar program of financial work incentives provided through an earnings supplement, marriage increased in one site but decreased in the other. The different outcomes may be the result of the different treatment of recipients who marry, but more research is needed. The other exception is the NEWWS evaluation, which provides consistent evidence that mandatory work-related activities have no impact on marriage or fertility up to five years after random assignment.

The short-run nature of the data also poses a problem for assessing how welfare reform affects the well-being of children. Although some aspects of a child's well-being, such as behavior problems, may respond quickly in reaction to changes in his or her parent's behavior, other aspects, such as cognitive skills, are likely to take much longer to change. Furthermore, even effects apparent in the short-term may change as children are exposed to cumulatively lower levels of welfare use and higher levels of employment on the part of their parents. In the short run, there is some evidence of favorable impacts on grade-school-aged children in the behavior and school achievement domains associated with programs that include more generous financial work incentives, either alone or tied to hours worked or to mandatory work-related activities. But the available evidence shows both favorable and unfavorable impacts associated with work requirements and reform as a bundle in these same domains. In the case of adolescents, there is more consistent evidence of unfavorable behavioral and school achievement impacts associated with these same policies up to five years after reform. Whether these same patterns will continue in the longer run—or whether they will be attenuated or exacerbated—remains to be determined.

A more general omission is any understanding of how reform has affected families' decisions to go on welfare to begin with. Random assignment experiments are a powerful research design for revealing how policy reforms affect families' decisions to leave the welfare rolls, but they provide no information at all on how families decide to join the rolls. Econometric studies of welfare use reflect the effects of entry decisions, but they do not distinguish them specifically. To date, there have been few econometric studies that focus specifically on welfare entry.

This omission is significant because entry appears to be important. Theoretical considerations lead us to expect that most policy reforms affect both entry and exit. Recent empirical work indicates that as much as one-half of the recent decline in the caseload is attributable to declining rates of entry. To the extent that welfare entry is the point at which many families learn they are eligible for Medicaid, the effects of welfare reform on welfare entry may work against the policy goal of expanding health coverage among the poor. Finally, as we argued in Chapter 7, understanding the full effects of reform on marriage and fertility will require that we understand how reform affects the types of behavior that historically have triggered recipients' initial entry onto welfare.

This discussion highlights a number of gaps in our knowledge base and indicates that we do not know everything that ideally should be known as policymakers begin to debate the reauthorization of PRWORA. To some extent this is inevitable. Policy evaluations take time to conduct, and the policies being evaluated have been in place for at most the last decade.

Another conclusion that we draw from the same evidence is more constructive. It shows that, with sufficient time and sufficient resources, we can greatly expand our knowledge about the workings of new policies. NEWWS provides the leading case in point. Because of the NEWWS evaluation, the effects of mandatory work-related activities are among the best-understood of all social policies.

11.3. AN AGENDA FOR FURTHER RESEARCH

From this example we draw a broader lesson: Our knowledge base in 2002 is stronger because of research programs put in place in the late 1980s and early 1990s under the strong guidance of USDHHS, and that increase in knowledge occurred only as a result of major expenditures on program development and evaluation. Likewise, the inclusion of research funding in the PRWORA legislation supported a continuation of the research and evaluation studies that were initiated prior to federal reform. Consequently, the available knowledge base associated with the welfare reforms implemented in the last decade is superior in many respects to that available for many other areas of social policy.

To add to that knowledge base, it is desirable to learn about current policies that are poorly understood and about reforms that may be proposed in the future. To do so requires that we act now, putting in place a research agenda capable of bearing fruit in time for the next reauthorization. Since the research cycle is at least as long as the policy cycle, we need to continue to put research efforts in place now for what we will need to know when the nation next considers major welfare reform.

Several specific agenda items deserve priority. To begin, more long-run information on the effects of current policies is crucial. Current long-run studies should be continued and, where possible, extended. Long-term evaluations should include such outcomes as child well-being, where the impacts may take time to materialize or where they may vary with the stage of child development. Further research is also needed to understand the effectiveness of alternative strategies for promoting the transition from welfare to work for subgroups of the welfare population, such as for recipients with substance abuse problems and those who experience domestic violence.

Other policies that are less well understood need further evaluation. Time limits represent an important example. Although the time has probably passed for conducting experiments to understand their behavioral effects, their mechanical effects will soon become increasingly important: As of April 2001, roughly 120,000 families had hit their time limits, most of whom live in states with time limits that are shorter than the federal five-year maximum (Center on Budget and Policy Priorities, 2001). The number of families exhausting their benefits may grow sharply in the near future as recipients in other states reach the federal five-year time limit. Studies to assess how families respond are critical.

Sanctions are among the most poorly understood of all of the policy reforms. This is an area where both econometric and experimental work would be useful. Econometric analyses that incorporate information on the likelihood of sanctioning and the monetary value of sanctions would provide a more complete understanding of this policy than the studies that are currently available. Experimentation could also help reveal how different levels of sanctions affect a broad range of outcomes. In either case, future research should continue on the path of expanding the range of outcomes examined, in addition to welfare use, employment, and earnings, which have been the focus of most studies to date.

Entry effects need to be better understood, both to fully grasp how reform has affected welfare use and labor market behavior and to understand how it affects fertility and the utilization of important in-kind services. This is an area where experimentation has less to offer. What is needed are high-quality econometric studies that focus directly on entry decisions.

Evidence from future econometric studies would be more useful if researchers characterized the variation in specific policy reforms across the states. Sanctions can be characterized by their monetary penalties, as suggested above; financial work incentives can be characterized by the benefit payment available to working recipients, as some researchers have done. More fully characterizing the policy environment is essential if econometric studies are to move beyond estimating the effects of reform as a bundle. Although existing national databases pose limitations for such efforts, approaches that utilize richer representations of states' policies are more likely to yield success in estimating the effects of specific reforms than approaches that rely on policy-specific dummy variables (Moffitt and Ver Ploeg, 2001; Adams and Hotz, 2001).

Initiatives sponsored by USDHHS and other agencies in several of these areas will add to the current knowledge base. For example, follow-up studies continue for a number of the experimental evaluations we examined in Chapters 4 to 10, and evaluations are under way in a number of other states that implemented other bundles of reforms. Reports are expected soon with longer-term results for Indiana's IMPACT program, Iowa's FIP, and Vermont's WRP, including impacts on child well-being up to five years after randomization (like those already available for the NEWWS programs). Other studies are under way to understand issues regarding accessing Medicaid and the Food Stamp Program, to evaluate the effectiveness of programs serving particularly disadvantaged segments of the welfare population, and to evaluate alternative approaches to promoting job retention and advancement among TANF recipients.

As in the past, advancing such an ambitious research agenda will require substantial federal participation. Many of the experiments reviewed here were conducted to satisfy the requirement that waiver-era reforms be evaluated and because the federal government paid for a portion of the costs. TANF's devolution of discretion to the states removed the requirement for rigorous evaluation. If we are to increase our knowledge base between now and the next time the nation considers major welfare reform, federal funds need to be invested to continue the evaluation of state investments under TANF. Even given TANF's devolution of welfare policy to the states, a strong federal role in research and evaluation remains necessary. As this study demonstrates, knowledge gained in one state may be broadly applicable in others. Because of these knowledge spillovers, the states cannot be expected to finance and carry out the needed amount of evaluation research without federal assistance.

Although the current research base provides answers to many of the important questions about welfare reform, many others remain unanswered. With planning, resources, and action, many of the outstanding questions can be addressed in time for the next reauthorization debate. With further research, we can better understand the trade-offs that reform entails and determine whether it has met its goals.

Appendix A

EFFECTS OF WELFARE REFORM FOR SUBGROUPS

A key question for many policymakers is whether different subgroups respond to welfare reform differently. Since groups facing different barriers or advantages may respond differently to different reforms, it is possible that some groups could be left behind, even as others fare well under the new regime. Understanding such responses would enable legislators and administrators to refine their programs, targeting groups most in need of assistance and improving their programs' overall performance. In this appendix, we discuss what is known about the effects of the various welfare reforms and outcomes considered in Chapters 4 to 10 on different segments of the welfare population.

In the case of the random assignment studies, it is possible to consider impacts for subgroups of the populations served by a particular demonstration.[108] Table A.1 shows the subgroups analyzed across the random assignment studies, with codes to indicate which outcomes are analyzed for each subgroup. For example, Vermont's WRP examines impacts on welfare use (W), employment and earnings (E), use of other government programs (G), and income and poverty (I) for subgroups defined by a composite measure of disadvantage and a measure of prior welfare receipt. Depending on the study, the composite measure of disadvantage may be designed to reflect risk of long-term dependency, barriers to employment, or other combined measures of disadvantage. Other subgroups are defined by single dimensions that capture prior welfare use or labor market history, human capital accumulation, demographic characteristics, health and family structure, or child characteristics. All characteristics are measured as of the time of randomization.

It is evident from Table A.1 that some subgroups receive more analysis than others, and that most subgroup analyses are confined to the basic outcomes of welfare use, employment and earnings, and income and poverty. Subgroup differences for family structure and other measures of well-being are examined by only one study each. A number of studies do not analyze subgroups for any of the outcomes shown in Table A.1.

Given the differences in which subgroups are examined for a given outcome, it will often not be possible to cross-validate findings across multiple studies as to whether the impacts of a given policy or bundle of policies vary by subgroup. Even when two studies within our classification scheme consider differences for the same outcome and subgroup category, differences across studies in the way the subgroups are defined often make results less comparable. Some

[108]In principle, econometric studies could provide subgroup-specific estimates, but few do, probably because of the small samples that result from subsetting the data.

Table A.1—Subgroups Analyzed by Random Assignment Studies Included in Synthesis

Name	Composite measure of disadvantage	Prior welfare receipt	Employment history	Earnings history	Educational attainment	Age	Race	English proficiency	Physical and/or mental health	Age of youngest child	Number of children	Marital and fecundity status	Child characteristics: Age	Child characteristics: Sex	Child characteristics: Developmental risk
A. Programs that focus on financial work incentives															
CWPDP															
WRP-IO	W E G I	W E G I											K		
MFIP-IO		W E	E		E										
B. Programs that focus on financial work incentives tied to hours of work															
New Hope	W E G I		K				E						K	K	
SSP		E I	E I		E I K	K			K	E I	K		K	K	
SSP - Plus															
SSP - Applicants															
C. Programs that focus on mandatory work-related activities															
LA Jobs-1st GAIN	W E	W E	W E		W E		W E	W E					K	K	
Atlanta LFA	E I	E I		E I	E I						E I				K
Grand Rapids LFA	E I	E I		E I	E I						E I				K
Riverside LFA	E I	E I		E I	E I						E I				K
Portland	E I	E I		E I	E I						E I				
Atlanta HCD	E I	E I		E I	E I						E I				K
Grand Rapids HCD	E I	E I		E I	E I						E I				K
Riverside HCD	E I	E I		E I	E I						E I				K
Columbus Integrated	E I	E I		E I	E I						E I				
Columbus Traditional	E I	E I		E I	E I						E I				
Detroit	E I	E I		E I	E I						E I				
Oklahoma City	E I	E I		E I	E I						E I				
IMPACT Basic Track															
D. Programs that focus on financial work incentives and mandatory work-related activities															
WRP	W E G I	W E G I											K		
MFIP	K	W E K	E	K	E K		K							K	
TSMF															
FIP				W E I						W E I					
E. Programs that focus on other individual reforms															
AWWDP															
FDP															
PPI															
PIP															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
EMPOWER															
IMPACT Placement Track															
VIP/VIEW															
ABC		W E F K	W E		W E F K	W E F K	K			W E K		F			
FTP	W E G I M K									W E G			K		
Jobs First	W E G I	W E G I	W E G I	W E G I						W E G I					

NOTES: For full program names and citations, see Table 3.4. W=welfare caseload (Chapter 4); E=employment and earnings (Chapter 5); G=use of other government programs (Chapter 6); F=fertility and marriage (Chapter 7); I=income and poverty (Chapter 8); M=other measures of well-being (Chapter 9); K=child well-being (Chapter 10).

subgroups that might be of interest, for example immigrants are not covered at all by the studies in Table A.1.

Two other issues complicate our ability to examine differences by subgroups. The first involves sample sizes. In many random assignment studies, the overall sample size is chosen to ensure that the overall program impacts will be statistically significant if their magnitudes are economically meaningful. However, detecting differences in effects between subgroups requires samples sizes several times larger than those required to detect an overall effect. Thus, there may be important variation in effects across subgroups, but they will not be detected unless the sample is much larger than that needed to generate significant results in the full sample. Many studies are based on samples that are too small to reliably detect even important differences between groups.

The second issue involves hypothesis testing. All the studies report whether the subgroup-specific estimates are significantly different from zero. Of equal importance is whether they are significantly different from each other. However, tests for the heterogeneity of impacts across groups are reported in only a few studies, leaving the reader to draw conclusions about group-specific differences from less objective criteria.

The remainder of the appendix discusses the results for program impacts by subgroups for the outcomes examined in Chapters 4 to 10, following the same order as the chapters. In the case of the Chapter 7 analysis of fertility and marriage, only one study includes any subgroup analyses, so those results are reported as part of the main discussion in Chapter 7. For the outcomes covered in this appendix, we note that the subgroup analyses may be based on a shorter follow-up period than the main results presented in the body of the report. For example, throughout the appendix, our analysis of subgroup impacts for Jobs First is based on outcomes measured in the 18-month follow-up survey (Bloom et al., 2000) or administrative data through years two or three (Hendra, Michalopolous, and Bloom, 2001). Subgroup results based on the three-year follow-up survey and administrative data through year four provide a more limited set of analyses (specifically by level of disadvantage, welfare recipiency status, and race/ethnicity) than those reported here (Bloom et al., 2002). Likewise, the subgroup analyses for the NEWWS programs are based on a pooled analysis using data through the third follow-up year (see the discussion of Michalopoulos and Schwartz, 2000, below). Subgroup results are not available in the five-year follow-up study of the NEWWS programs (Hamilton et al., 2001).

A.1. WELFARE USE

Subgroup differences for welfare use are reported for a subset of the random assignment studies reviewed in Chapter 4. The available results for ten of the programs are reported in Table A.2. The table records impacts for up to four different subgroups, with subgroups arrayed, to the extent possible, from most to least disadvantaged. Some studies report only whether subgroup impacts are statistically significant (denoted using asterisks next to the impact estimate); others also report whether differences in impacts across subgroups are statistically significant (denoted in the first column by x's). We indicate when the statistical significance of subgroup differences is not available. Thus, when that cell in the table is empty, it means that the differences by groups are not statistically significant.

Table A.2—Estimated Impact of Welfare Reform on Welfare Use for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
A. Programs that focus on financial work incentives															
WRP-IO	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Received welfare in last 3 mos. of FU	45.0	0.7	1.6%	27.3	-0.1	-0.4%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Received welfare in last 3 mos. of FU	55.2	6.0	10.9%	40.6	-2.6	-6.4%	24.5	5.2	21.2%			
MFIP-IO	Urban single parent applicants		**By status at RA:**	Short-term recipient			New applicant								
		n.a.	Avg. quarterly welfare receipt, year 3	48.9	15.4 ***	31.5%	33.4	7.6 ***	22.8%						
B. Programs that focus on financial work incentives tied to hours of work															
New Hope	Poor families employed FT at RA		**By barriers to employment:**	Two (3)			One (3)			None (3)					
			Months on aid, year 1	6.9	0.3	3.9%	5.8	0.2	3.4%	5.2	-0.6	-11.5%			
			Months on aid, year 2	4.5	0.9 *	20.8%	3.1	0.4	14.5%	3.2	-0.5	-15.3			
C. Programs that focus on mandatory work-related activities															
LA Jobs-First GAIN	Single-parent recipients and applicants		**By race/ethnicity:**	African-American			Hispanic			Asian			White		
			Received welfare , Q 9	72.6	-4.1 ***	-5.6%	66.0	-6.4 ***	-9.7%	66.9	0.8	1.2%	55.1	-3.2	-5.8%
			By English proficiency:	Not proficient			Proficient								
			Received welfare , Q 9	71.1	-4.6 **	-6.5%	65.1	-4.9 ***	-7.5%						
			By education:	No diploma or GED			Diploma or GED								
			Received welfare , Q 9	70.2	-3.7 ***	-5.3%	61.4	-5.7 ***	-9.3%						
			By status at RA:	Long-term recipient			Short-term recipient			New applicant					
			Avg. quarterly welfare receipt, year 3	71.6	-4.7 ***	-6.6	52.0	-4.4 **	-8.5%	46.9	-1.3	-2.8%			
			By employment in year prior to RA	Not employed			Employed								
			Received welfare , Q 9	70.4	-5.1 ***	-7.2%	58.8	-4.0 ***	-6.8%						
			By disadvantage:	Most disadvantaged (2)											
			Received welfare , Q 9	75.5	-3.1 *	-4.1%									

Table A.2—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
D. Programs that focus on financial work incentives and mandatory work-related activities															
WRP	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Received welfare in last 3 mos. of FU	45.0	-3.4	-7.6%	27.3	0.2	0.7%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Received welfare in last 3 mos. of FU	55.2	-2.9	-5.3%	40.6	-2.4	-5.9%	24.5	-0.2	-0.8%			
MFIP	Urban single parent applicants		**By status at RA:**	Short-term recipient			New applicant								
		n.a.	Avg. quarterly earnings, year 3	48.9	8.7 **	17.8%	33.4	6.3 ***	18.9%						
FIP	Recipients		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Any welfare receipt, year 2	80.0	1.4	1.8%	69.2	3.3 *	4.8%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Any welfare receipt, year 2	77.7	0.5	0.6%	71.4	4.5 ***	6.3%						
	Applicants		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Any welfare receipt, year 1	70.9	3.8	5.4%	59.5	2.9	4.9%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Any welfare receipt, year 1	66.9	1.1	1.6%	59.4	5.8 *	9.8%						

Table A.2—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1			Group 2			Group 3			Group 4		
				Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%
E. Programs that focus on other individual reforms															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
ABC	Recipients and applicants		**By age of mother:**	Under 25			25 to 34			35 and over					
		n.a.	Percent receiving aid in March 1997	50.8	-1.9	-3.7	50.8	-6.1 **	-12.0	46.7	-1.6	-3.5			
			By age of youngest child:	Under 3			3 to 5			6 and older					
		n.a.	Percent receiving aid in March 1997	49.4	-3.4	-6.8	51.9	-0.4	-0.8	48.4	-6.7 **	-13.8			
			By when last employed:	24 or more months before RA			6 to 23 months before RA			Within 6 months of RA			Within 6 months of RA		
		n.a.	Percent receiving aid in March 1997	59.1	-6.9 **	-11.7	51.2	-7.5 **	-14.6	40.5	1.3	3.3	40.5	1.3	3.3
			By years of schooling:	Less than 12			12 or more								
		n.a.	Percent receiving aid in March 1997	55.0	-5.6	-10.1	45.4	-2.1	-4.5						
			By years on aid in past 5 years:	3 to 5			1 to 2			Less than 1					
		n.a.	Percent receiving aid in March 1997	58.9	-2.8	-4.8	46.7	-7.0 **	-14.9	40.2	-1.6	-4.0			
FTP	Recipients and applicants		**By length of time limit:**	36 months (4)			24 months (4)								
		n.a.	Ever received aid in years 1-2	87.8	0.8	0.9	79.6	1.0	1.2						
			By age of youngest child:	Under 3			3 or over								
		n.a.	Ever received aid in years 1-2	84.2	0.9	1.0	81.6	1.5	1.8						
			By level of risk:	Most risk (5)			Medium risk (5)			Least risk (5)					
			Quarterly aid receipt, year 2	69.0	-1.1	-1.6%	42.5	0.2	0.5%	24.2	-3.3	-13.6%			

Table A.2—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
			By status at RA:		Recipient			Applicant							
		n.a.	Ever received aid, years 1-2	74.1	3.5 ***	4.7	50.5	9.0 ***	17.7						
		n.a.	Ever received aid, year 3	50.2	-11.8 ***	-23.5	30.1	-3.7 **	-12.4						
			By disadvantage:		Most disadvantaged (6)						Least disadvantaged (6)				
		n.a.	Ever received aid, years 1-2	80.2	0.5	0.6				43.4	11.6 ***	26.7%			
		n.a.	Ever received aid, year 3	59.5	-11.2 ***	-18.8				23.7	0.1	0.4%			
			By age of youngest child:		Less than 6			6 to 11			12 to 18				
		n.a.	Avg. quarterly aid receipt, Q1-6	71.0	7.4 ***	10.4%	69.0	10.4 ***	15.1%	60.7	5.8 *	9.6%			
		n.a.	Any aid receipt, Q8	54.1	-5.9 ***	-10.9%	51.1	-4.9	-9.6%	38.3	-8.8 **	-23.0%			
JOBS First	Recipients and applicants		**By aid receipt in year before RA:**		Long-term recipient (7)			Short-term recipient (7)			New applicant (7)				
		n.a.	Avg. quarterly aid receipt, Q1-6	82.3	5.0 ***	6.1%	65.2	9.4 ***	14.4%	53.3	8.7 ***	16.3%			
		n.a.	Any aid receipt, Q8	66.1	-11.5 ***	-17.4%	46.6	-4.4	-9.4%	33.4	1.9	5.7%			
			By employment in year before RA:		Not employed			Employed							
		n.a.	Avg. quarterly aid receipt, Q1-6	75.4	3.2 ***	4.2%	63.2	10.9 ***	17.2%						
		n.a.	Any aid receipt, Q8	58.1	-7.2 ***	-12.4%	44.5	-4.2 **	-9.4%						
			By earnings in year before RA:		No earnings			$1 to 5000			Over $5000				
		n.a.	Avg. quarterly aid receipt, Q1-6	75.4	3.2 **	4.2%	72.6	7.1 ***	9.8%	48.6	16.7 ***	34.4%			
		n.a.	Any aid receipt, Q8	58.1	-7.2 ***	-12.4%	53.4	-9.7 ***	-18.2%	30.4	4.3	14.1%			

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%.

Abbreviations: Q=quarter; RA = random assignment; n.a. = not available.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantage d had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) "Most disadvantaged" is defined as long-term recipients who did not have a high school credential at RA and who did not work for pay in the year prior to RA.

(3) Potential barriers to employment are not having worked in the past six years; having been arrested since age 16; having either two or more children under age 6 or four children under age 12; having been fired from one's period of longest employment; and not having a high school credential.

(4) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(5) "Most at risk" have risk score in top quartile of dependency index; "Least at risk" have score in the bottom quartile; "Medium risk" are in between. Dependency index is based on prior quarter of employment, months employed prior to RA, AFDC recipiency status in quarter prior to RA, months of AFDC prior to RA, age of youngest child, and high school credential status.

(6) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. "Least disadvantaged" is defined as having none of these traits.

(7) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

A.1.1. Programs That Focus on Financial Work Incentives

As seen in Panel A of Table A.2, there are two programs in this category, WRP-IO and MFIP-IO, that provide subgroup results for welfare use. Both provide subgroup-specific impact estimates, but the MFIP impacts for subgroups are available only for the recent applicants.[109] None of the subgroup-specific impacts from WRP-IO are significant. In MFIP-IO, program impacts are significantly different from zero for both short-term recipients and new applicants. There is no way to determine whether the subgroup-specific estimates are significantly different from each other, however.

A.1.2 Programs That Focus on Financial Work Incentives Tied to Hours of Work

New Hope provides tests for subgroup differences; none of these tests rejects homogeneity across the three groups shown (Panel B). The evaluations of SSP Applicants and SSP Plus provide no subgroup analyses. The main SSP evaluation estimates subgroup impacts on employment (see below), but not on welfare use.

A.1.3. Programs That Focus on Mandatory Work-Related Activities

A different type of evidence on subgroup effects is available for policy reforms involving mandatory work-related activities (Panel C). For this type of reform policy, there are estimates from a large number of studies, all of which define subgroups in a consistent manner. Michalopoulos and Schwartz (2000) reanalyze the microlevel data from 20 studies that involve mandated work-related activities, including the eleven NEWWS programs discussed above as well as MFIP and FTP.[110] By systematically reanalyzing the microdata from these studies, they are able to construct more rigorous and powerful tests for subgroup differences than can be obtained by simply comparing group-specific program impacts from separate studies. Unfortunately, Michalopoulos and Schwartz (2000) do not analyze welfare use, but rather focus on welfare payments. As a result, we defer our discussion of their results to Section A.4 below, where we discuss components of income.

Subgroup results for L.A. Jobs-First GAIN, which was not included in the Michalopoulos and Schwartz analysis, are presented in Panel B of Table A.2. Few substantial between-group differences are readily apparent, with the possible exception of the results for Asians in the race-ethnicity breakdown. None of the differences are statistically significant.

---

[109]Miller et al. (2000) provide subgroup results for urban recipients for employment and earnings, but not for welfare use.

[110]The NEWWS programs focus primarily on work-related activity mandates, as discussed in Chapter 3. MFIP involved mandates for recipients once they had been on aid for 24 months and also included a financial work incentive. Although data from MFIP were included in the Michalopoulos-Schwartz analysis, Miller et al. (2000), who provide the estimates that are available for this synthesis, do not provide welfare impacts by subgroups for single-parent recipients. In FTP, the treatment and control groups were both subject to work requirements, but the treatment group had access to a wider range of services than the control group. The treatment group also faced a time limit and a financial work incentive. Although FTP and MFIP involve major reforms besides mandatory work-related activities, they account for only about 10 percent of the pooled sample, which implies that their other major reforms are unlikely to substantially affect the pooled results. The other studies included in Michalopoulos and Schwartz are from earlier time periods, including the GAIN evaluations of the late 1980s.

A.1.4. Programs That Combine Financial Work Incentives and Mandatory Work-Related Activities

Although a substantial number of studies focus on this particular combination of reforms, as a whole the studies provide little useful evidence about the combination's effects on different groups (Panel D). The limitations pertaining to WRP-IO and MFIP-IO discussed above pertain to the full WRP and MFIP programs as well. None of the subgroup estimates from WRP are significant. Both subgroup estimates from MFIP are significantly different from zero, but they are similar, and no test for their difference is available.

In FIP, none of the estimates for families without prior-year earnings or with children under three are significant, compared to three of the four estimates for families with prior-year earnings or with no children under three. There are no tests to determine whether the subgroup differences are statistically significant. The evaluation of TSMF provides no subgroup analyses for welfare use.

A.1.5. Programs That Focus on TANF-Like Bundles of Reforms

Among studies that combine time limits with other policy reforms, ABC, FTP, and Jobs First provide subgroup impacts on welfare use (Panel F). However, the subgroups are defined differently across the studies, as seen in Table A.2, although all the subgroup classifications are intended to reflect measures of disadvantage. As noted above, all are arrayed from most to least disadvantaged across the table. In the case of FTP, the length of the time limit provides a measure of disadvantage because participants with low levels of education, short employment histories, and lengthy welfare histories were given a 36-month time limit, whereas other participants were given a 24-month time limit. Using participants' applicant/recipient status at random assignment as a measure of disadvantage, as with Jobs First, is based on the observation that most new entrants will stay on welfare for a relatively short period, whereas most ongoing recipients are in the midst of what will become a lengthy spell on aid (Bane and Ellwood, 1994).

Across these groups, there is no clear tendency for program impacts to vary by the level of disadvantage. The effects of ABC were insignificant at both levels of schooling. They were largest for those with intermediate levels of prior aid use and were smallest for those with the most recent employment history. FTP's effects were similar for all groups. Jobs First increased aid use by a greater amount among disadvantaged groups during the pre–time limit period. By some measures, the program decreased aid use among more disadvantaged groups during the post–time limit period; by other measures, the pattern is less clear. The change in impacts between the pre– and post–time limit period is fairly similar across the groups. Based on this relatively small number of studies and differing definitions of disadvantage, we cannot conclude whether the effects of policy reforms involving time limits have greater or lesser effects on relatively disadvantaged groups.

A.2. EMPLOYMENT AND EARNINGS

Although only a subset of the random assignment studies provides subgroup analyses for employment and earnings, the subset is larger in this case than it was in the case of welfare use. Thus, we can consider subgroup differences for a broader range of policy reforms than we could above. Table A.3 presents subgroup-specific program impacts for employment and Table A.4

presents corresponding estimates for earnings.[111] As a whole, the available evidence on subgroups does not suggest that any of the reforms consistently work to the greater employment or earnings detriment—or benefit—of relatively disadvantaged groups.

A.2.1. Programs That Focus on Financial Work Incentives

The two programs that focus on financial work incentives provide no clear evidence that the effects of this policy vary according to recipients' level of disadvantage (Panel A). Although the group-specific employment impacts vary for WRP-IO, they are insignificant and display no consistent pattern. MFIP-IO has greater effects for participants with higher levels of schooling, but also for persons with shorter employment histories and longer welfare histories. The estimates in Table A.4 show that these programs had insignificant effects on earnings for all the subgroups, which is consistent with the finding that the programs had no significant effect on third-year earnings overall, as shown in Table 5.1.

A.2.2 Programs That Focus on Financial Work Incentives Tied to Hours of Work

As seen in Panel B, the first-year employment effects of New Hope are essentially uniform across groups, defined by differing numbers of barriers to employment, although there are significant differences in the second-year effects. In both years, employment impacts for whites are insignificant and smaller than the impacts for blacks and Hispanics. The effects of SSP are nearly uniform across the distribution of the age of the youngest child, the mother's education, employment history, and welfare history. Indeed, formal tests indicate that these effects are homogenous across levels of disadvantage for all the disadvantage measures considered.

Only New Hope provides subgroup estimates for earnings. The earnings effect of the program is strongest for the middle-disadvantage group. By race, the impacts of New Hope are positive for blacks and Hispanics, but negative (though always insignificant) for whites.

A.2.3. Programs That Focus on Mandatory Work-Related Activities

Employment and earnings impacts by subgroup from L.A. Jobs-First GAIN are presented in Panel C of Tables A.3 and A.4. Employment impacts differ significantly only among subgroups defined by English proficiency and by the recipient's overall level of disadvantage. None of the earnings impacts differ significantly across the different subgroups.

Panel C of Table A.4 also presents subgroup-specific earnings impacts from Michalopoulos and Schwartz (2000). As noted above, these estimates are based on pooled microdata from the NEWWS programs and several other experiments that involved mandatory work-related activities. Several of the subgroup differences are statistically significant. However, in some cases the impacts are larger for relatively disadvantaged groups, such as families with several

[111]For SSP, subgroup impacts are available only for employment.

Table A.3—Estimated Impact of Welfare Reform on Employment for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
A. Programs that focus on financial work incentives															
WRP-IO	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Employed in last 3 mos. of FU	46.0	2.5	5.4%	51.4	3.3	6.4%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Employed in last 3 mos. of FU	33.3	9.9	29.7%	44.9	0.3	0.7%	59.5	5.5	9.2%			
MFIP-IO	Urban single parents recipients		**By years of schooling:**	Less than 12			12 or more								
		n.a.	Quarterly employment rate, year 3	35.7	-0.9	-2.5%	49.6	4.8 *	9.7%						
			By employment in year before RA:	No work in year before RA			Worked in year before RA								
		n.a.	Quarterly employment rate, year 3	35.4	5.1 *	14.4%	57.8	1.0	1.7%						
			By years on aid prior to RA:	5 or more years			Less that 5 years								
		n.a.	Quarterly employment rate, year 3	42.9	5.1 *	11.9%	47.7	1.6	3.4%						
B. Programs that focus on financial work incentives tied to hours of work															
New Hope	Poor families employed FT at RA		**By barriers to employment:**	Two (3)			One (3)			None (3)					
			Quarters employed, year 1	2.1	0.6 ***	28.6%	2.3	0.5 ***	21.7%	2.5	0.4 **	16.0%			
		xx	Quarters employed, year 2	2.4	0.3	12.5%	2.3	0.5 ***	21.7%	2.8	-0.2	-7.1%			
	All poor families		**By race/ethnicity:**	African-American			Hispanic			White					
			Quarters employed, year 1	2.6	0.5 ***	19.2%	2.6	0.4 ***	15.4%	2.6	0.3	11.5%			
			Quarters employed, year 2	2.6	0.3 ***	11.5%	2.7	0.2	7.4%	2.8	0.0	0.0%			
SSP	Single-parent recipients		**By age of youngest child:**	0 to 2			3 to 11			12 to 15			16 and over		
			Months of FT employment	4.4	3.5 ***	79.5%	5.4	3.2 ***	59.3%	5.3	3.4 ***	64.2%	6.1	2.0 *	32.8%
			By education:	No high school credential			High school credential								
			Months of FT employment	3.3	3.3 ***	100.0%	7.2	3.1 ***	43.1%						
			By employment status at RA:	Out of labor force			Unemployed			Employed PT			Employed FT		
			Months of FT employment	2.6	3.0 ***	115.4%	5.1	4.1 ***	80.4%	7.7	4.6 ***	59.7%	20.0	2.7 **	13.5%
			By months on aid in prior 3 years:	36 months			24 to 35			10 to 23					
			Months of FT employment	3.9	3.3 ***	84.6%	5.5	3.6 ***	65.5%	6.4	3.0 ***	46.9%			

Table A.3—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
C. Programs that focus on mandatory work-related activities															
LA Jobs-First GAIN	Single-parent recipients and applicants		**By race/ethnicity:**	African-American			Hispanic			Asian			White		
			Ever employed in years 1-2	62.0	6.6 ***	10.6%	57.5	12.0 ***	20.9%	41.9	13.0 ***	31.0%	54.8	8.3 ***	15.1%
			By English proficiency:	Not proficient			Proficient								
		xx	Ever employed in years 1-2	46.7	12.4 ***	26.6%	60.3	9.0 ***	14.9%						
			By education:	No diploma or GED			Diploma or GED								
			Ever employed in years 1-2	52.3	10.1 ***	19.3%	64.0	9.0 ***	14.1%						
			By status at RA:	Long-term recipient			Short-term recipient			New applicant					
			Ever employed in years 1-2	55.5	10.2 ***	18.4%	62.5	8.5 ***	13.6%	67.5	4.7	7.0%			
			By employment in year prior to RA	Not employed			Employed								
			Ever employed in years 1-2	43.4	12.8 ***	29.5%	82.4	4.1 ***	5.0%						
			By disadvantage:	Most disadvantaged (2)											
		xxx	Ever employed in years 1-2	39.3	13.2 ***	33.6%									
D. Programs that focus on financial work incentives and mandatory work-related activities															
WRP	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Employed in last 3 mos. of FU	46.0	10.3 ***	22.4%	51.4	6.6 ***	12.8%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Employed in last 3 mos. of FU	33.3	10.9 *	32.7%	44.9	9.8 ***	21.8%	59.5	7.0 **	11.8%			
MFIP	Urban single parents recipients		**By years of schooling:**	Less than 12			12 or more								
		n.a.	Quarterly employment rate, year 3	35.7	9.5 ***	26.6%	49.6	12.2 ***	24.6%						
			By employment in year before RA:	No work in year before RA			Worked in year before RA								
		n.a.	Quarterly employment rate, year 3	35.4	16.5 ***	46.6%	57.8	4.3	7.4%						
			By years on aid prior to RA:	5 or more years			Less that 5 years								
		n.a.	Quarterly employment rate, year 3	42.9	11.5 ***	26.8%	47.7	10.8 ***	22.6%						

Table A.3—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
FIP	Recipients		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Any employment, year 2	59.2	2.2	3.7%	81.0	1.5	1.9%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Any employment, year 2	68.7	4.4 **	6.4%	71.9	-0.3	-0.4%						
	Applicants		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Any employment, year 1	48.9	6.3 *	12.9%	85.8	2.7	3.1%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Any employment, year 1	70.0	6.9 ***	9.9%	77.0	0.9	1.2%						
E. Programs that focus on other individual reforms															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
ABC	Recipients and applicants		**By age of mother:**	Under 25			25 to 34			35 and over					
		n.a.	Percent working at survey interview	44.6	1.3	2.8	42.2	9.9 ***	23.5	36.1	11.5 **	31.8			
			By age of youngest child:	Under 3			3 to 5			6 and older					
		n.a.	Percent working at survey interview	43.7	2.6	6.1	48.8	4.5	9.3	34.1	16.8 ***	49.3			
			By when last employed:	24 or more months before RA			6 to 23 months before RA			Within 6 months of RA					
		n.a.	Percent working at survey interview	34.8	6.3	18.1	39.1	3.7	9.3	49.3	9.1 **	18.5			
			By years of schooling:	Less than 12			12 or more								
		n.a.	Percent working at survey interview	34.3	7.4 **	21.6	47.9	6.5 **	13.5						
			By years on aid in past 5 years:	3 to 5			1 to 2			Less than 1					
		n.a.	Percent working at survey interview	37.8	6.4	17.0	42.1	11.8 ***	28.0	46.5	3.0	6.5			
FTP	Recipients and applicants		**By length of time limit:**	36 months (4)			24 months (4)								
		n.a.	Ever employed in years 1-2	70.4	5.8 **	8.2	72.0	4.8 **	6.6						
			By age of youngest child:	Under 3			3 or over								
		n.a.	Ever employed in years 1-2	71.2	6.6 ***	9.2	71.5	4.3 **	6.0						
			By level of risk:	Most risk (5)			Medium risk (5)			Least risk (5)					
		n.a.	Quarterly employment, year 2	37.8	6.4 **	16.9%	38.1	7.5 ***	19.7%	58.4	5.4 *	9.2%			

Table A.3—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
JOBS First	Recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Avg. employment rate, years 1-2	41.8	10.9 ***	26.0	49.9	2.8 *	5.6						
		n.a.	Avg. employment rate, year 3	50.7	8.0 ***	15.8	55.3	3.7 **	6.7						
			By disadvantage:	Most disadvantaged (6)						Least disadvantaged (6)					
		n.a.	Avg. employment rate, years 1-2	19.3	14.1 ***	73.0				65.9	3.7	5.6%			
		n.a.	Avg. employment rate, year 3	30.2	11.6 ***	38.6				71.3	3.2	4.5%			
			By age of youngest child:	Less than 6			6 to 11			12 to 18					
		n.a.	Avg. quarterly employment, Q1-6	44.1	7.2 ***	16.3%	46.6	10.0 ***	21.5%	38.0	7.6 **	20.0%			
		n.a.	Any employment, Q8	49.2	9.1 ***	18.5%	51.6	7.1 ***	13.8%	43.0	5.5	12.8%			
			By aid receipt in year before RA:	Long-term recipient (7)			Short-term recipient (7)			New applicant (7)					
		n.a.	Avg. quarterly employment, Q1-6	37.7	12.1 ***	32.1%	48.7	6.1 ***	12.5%	46.8	3.1 *	6.6%			
		n.a.	Any employment, Q8	44.4	11.3 ***	25.5%	53.7	6.3 **	11.7%	52.5	3.8	7.2%			
			By employment in year before RA:	Not employed			Employed								
		n.a.	Avg. quarterly employment, Q1-6	23.3	10.5 ***	45.1%	61.9	5.2 ***	8.4%						
		n.a.	Any employment, Q8	31.1	11.4 ***	36.7%	66.0	4.4 **	6.7%						
			By earnings in year before RA:	No earnings			$1 to 5000			Over $5000					
		n.a.	Avg. quarterly employment, Q1-6	23.3	10.5 ***	45.1%	55.0	6.9 ***	12.5%	72.8	3.2	4.4%			
		n.a.	Any employment, Q8	31.1	11.4 ***	36.7%	61.0	5.0 **	8.2%	73.8	3.6	4.9%			

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%.

Abbreviations: Q=quarter; RA = random assignment; n.a. = not available.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantaged had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) "Most disadvantaged" consists of long-term recipients who did not have a diploma or GED and did not work for pay in the year prior to random RA.

(3) Barriers include not having worked in the past six years; being arrested since age 16; having 2 or more children under six or four or more under 12; having been fired from one's longest-lasting jobs; and not having a high school credential.

(4) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(5) "Most at risk" sample members are those whose risk score is in the top quartile of the distribution of a "dependency index." "Least at risk" sample members are those whose risk score is in the bottom quartile of the distribution of the dependency index. "Medium risk" sample members are those whose risk score is in the interquartile range of the distribution of the dependency index.

(6) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. The "least disadvantaged" meet none of these criteria.

(7) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

Table A.4—Estimated Impact of Welfare Reform on Earnings for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
A. Programs that focus on financial work incentives															
WRP-IO	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Employed in last 3 mos. of FU	$1,440	-$58	-4.0%	$1,585	$89	5.6%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Earnings in last 3 mos. of FU	$801	$90	11.2%	$1,406	-$108	-7.7%	1897.0	214.0	11.3%			
MFIP-IO	Urban single parents recipients		**By years of schooling:**	Less than 12			12 or more								
		n.a.	Avg. quarterly earnings, year 3	$783	-$88	-11.2%	$1,576	-$73	-4.6%						
			By employment in year before RA:	No work in year before RA			Worked in year before RA								
		n.a.	Avg. quarterly earnings, year 3	$1,064	$12	1.1%	$1,655	-$197	-11.9%						
			By years on aid prior to RA:	5 or more years			Less that 5 years								
		n.a.	Avg. quarterly earnings, year 3	$1,208	$5	0.4%	$1,428	-$97	-6.8%						
B. Programs that focus on financial work incentives tied to hours of work															
New Hope	Poor families employed FT at RA		**By barriers to employment:**	Two (4)			One (4)			None (4)					
			Earnings, year 1	$3,850	$625	16.2%	$4,228	$1,254 **	29.7%	$5,319	$490	9.2%			
		xx	Earnings, year 2	$5,713	$173	3.0%	$5,555	$1,690 ***	30.4%	$7,449	-$1,022	-13.7%			
	All poor families		**By race/ethnicity:**	African-American			Hispanic			White					
			Earnings, year 1	$5,526	$1,110 ***	20.1%	$6,140	$1,305 **	21.3%	$6,499	-$744	-11.4%			
			Earnings, year 2	$6,816	$334	4.9%	$8,097	$754	9.3%	$8,024	-$814	-10.1%			

Table A.4—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
C. Programs that focus on mandatory work-related activities															
LA Jobs-First GAIN	Single-parent recipients and applicants		**By race/ethnicity:**		African-American			Hispanic			Asian			White	
			Total earnings in years 1-2	$6,909	$1,429 ***	20.7%	$6,214	$1,862 ***	30.0%	$4,421	$2,052 ***	46.4%	$6,306	$1,553 ***	24.6%
			By English proficiency:		Not proficient			Proficient							
			Total earnings in years 1-2	$4,264	$1,905 ***	44.7%	$6,936	$1,543 ***	22.2%						
			By education:		No diploma or GED			Diploma or GED							
			Total earnings in years 1-2	$4,647	$1,481 ***	31.9%	$8,444	$1,805	21.4%						
			By status at RA:		Long-term recipient			Short-term recipient			New applicant				
			Total earnings in years 1-2	$5,410	$1,802 ***	33.3%	$8,958	$1,244 **	13.9%	$9,241	$589	6.4%			
			By employment in year prior to RA		Not employed			Employed							
			Total earnings in years 1-2	$3,624	$1,750 ***	48.3%	$11,212	$1,405 ***	12.5%						
			By disadvantage:		Most disadvantaged (3)										
			Total earnings in years 1-2	$2,624	$1,526 ***	58.2%									
NEWWS Sites-pooled	Recipients and applicants		**By earnings in year prior to RA:**		No earnings			$5,000 or less			More than $5000				
			Avg. annual earnings, years1-3	$1,754	$571 ***	32.6%	$3,425	$399 ***	11.6%	$6,957	$548 ***	7.9%			
			By education:		No high school credential			High school credential							
		xx	Avg. annual earnings, years1-3	$1,867	$430 ***	23.0%	$3,751	$627 ***	16.7%						
			By number of children:		Three or more			Two			One				
		xxx	Avg. annual earnings, years1-3	$2,523	$682 ***	27.0%	$2,957	$663 ***	22.4%	$3,196	$328 ***	10.3%			
			By level of disadvantage:		Most disadvantaged (2)			Moderately disadvantaged (2)			Least disadvantaged (2)				
		xx	Avg. annual earnings, years1-3	$983	$404 ***	41.1%	$2,955	$599 ***	20.3%	$5,664	$421 ***	7.4%			
			By status at RA:		Long-term reciepient			Short-term recipient			Applicant				
		x	Avg. annual earnings, years1-3	$2,480	$544 ***	21.9%	$3,708	$534 ***	14.4%	$3,025	$1,106 ***	36.6%			

Table A.4—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
D. Programs that focus on financial work incentives and mandatory work-related activities															
WRP	Single-parent recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Employed in last 3 mos. of FU	$1,440	$93	6.5%	$1,585	$189 *	11.9%						
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Earnings in last 3 mos. of FU	$801	$263	32.8%	$1,406	$56	4.0%	$1,897	$304 **	16.0%			
MFIP	Urban single parents recipients		**By years of schooling:**	Less than 12			12 or more								
		n.a.	Avg. quarterly earnings, year 3	$783	$186	23.8%	$1,576	$107	6.8%						
			By employment in year before RA:	No work in year before RA			Worked in year before RA								
		n.a.	Avg. quarterly earnings, year 3	$1,064	$267 **	25.1%	$1,655	-$66	-4.0%						
			By years on aid prior to RA:	5 or more years			Less that 5 years								
		n.a.	Avg. quarterly earnings, year 3	$1,208	$104	8.6%	$1,428	$165	11.6%						
FIP	Recipients		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Earnings, year 2	$2,909	$537 ***	18.5%	$6,171	$195	3.2%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Earnings, year 2	$4,286	$554 **	12.9%	$4,834	$223	4.6%						
	Applicants		**By earnings in year before RA:**	No earnings			Some earnings								
		n.a.	Earnings, year 1	$2,484	$419	16.9%	$8,084	$768	9.5%						
			By children less than 3:	Any children less than 3			No children less than 3								
		n.a.	Earnings, year 1	$5,469	$981 ***	17.9%	$6,969	$199	2.9%						
E. Programs that focus on other individual reforms															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
FTP	Recipients and applicants		**By length of time limit:**	36 months (5)			24 months (5)								
		n.a.	Total earnings, years 1-2	$4,367	$733 *	16.8	$6,837	$830 **	12.1						
			By age of youngest child:	Under 3			3 or over								
		n.a.	Total earnings, years 1-2	$4,935	$1,165 ***	23.6	$6,469	$640 *	9.9						
			By level of risk:	Most risk (6)			Medium risk (6)			Least risk (6)					
		n.a.	Earnings, year 2	$2,337	$479 *	20.5%	$2,637	$550 **	20.9%	$5,442	$1,175 ***	21.6%			

Table A.4—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1			Group 2			Group 3			Group 4		
				Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%
JOBS First	Recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Avg. annual earnings, years 1-2	$3,890	$747 ***	19.2	$5,790	-$49	-0.8						
		n.a.	Avg. annual earnings, year 3	$6,393	$672 **	10.5	$8,048	$558	6.9						
			By disadvantage:	Most disadvantaged (7)						Least disadvantaged (7)					
		n.a.	Avg. annual earnings, years 1-2	$1,428	$985 ***	69.0				$8,627	-$408	-4.7%			
		n.a.	Avg. annual earnings, year 3	$2,973	$1,195 ***	40.2				$12,027	-$296	-2.5%			
			By age of youngest child:	Less than 6			6 to 11			12 to 18					
		n.a.	Avg. quarterly earnings, Q1-6	$1,030	$82 *	8.0%	$1,193	$83	7.0%	$992	$123	12.4%			
		n.a.	Earnings, Q8	$1,454	$233 ***	16.0%	$1,566	$168	10.7%	$1,388	$250	18.0%			
			By aid receipt in year before RA:	Long-term recipient (8)			Short-term recipient (8)			New applicant (8)					
		n.a.	Avg. quarterly earnings, Q1-6	$781	$190 ***	24.3%	$1,181	$48	4.1%	$1,315	-$9	-0.7%			
		n.a.	Earnings, Q8	$1,157	$243 ***	21.0%	$1,598	$199 *	12.5%	$1,831	$155	8.5%			
			By employment in year before RA:	Not employed			Employed								
		n.a.	Avg. quarterly earnings, Q1-6	$464	$184 ***	39.7%	$1,595	-$12	-0.8%						
		n.a.	Earnings, Q8	$796	$313 ***	39.3%	$2,099	$111	5.3%						
			By earnings in year before RA:	No earnings			$1 to 5000			Over $5000					
		n.a.	Avg. quarterly earnings, Q1-6	$464	$184 ***	39.7%	$1,033	$116 *	11.2%	$2,479	-$185	-7.5%			
		n.a.	Earnings, Q8	$796	$313 ***	39.3%	$1,592	$140	8.8%	$2,912	$64	2.2%			

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%.

Abbreviations: Q=quarter; RA = random assignment; n.a. = not available.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantaged had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) "Most disadvantaged" is classified as those with no earnings in the year prior to RA, without a high school credential, and received welfare two years or more prior to RA. Those classified as "Least disadvantaged" had none of these characteristics. All others were classified as "Moderately disadvantaged."

(3) "Most disadvantaged" consists of long-term recipients who did not have a diploma or GED and did not work for pay in the year prior to random RA.

(4) Barriers include not having worked in the past six years; being arrested since age 16; having 2 or more children under six or four or more under 12; having been fired from one's longest-lasting jobs; and not having a high school credential.

(5) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(6) "Most at risk" sample members are those whose risk score is in the top quartile of the distribution of a "dependency index." "Least at risk" sample members are those whose risk score is in the bottom quartile of the distribution of the dependency index. "Medium risk" sample members are those whose risk score is in the interquartile range of the distribution of the dependency index.

(7) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. The "least disadvantaged" meet none of these criteria.

(8) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

children, whereas in other cases, the impacts are larger for relatively advantaged groups, such as recent applicants.

A.2.4. Programs That Combine Financial Work Incentives with Mandatory Work-Related Activities

Three programs estimate group-specific employment impacts for programs that combine financial work incentives with mandatory work-related activities: WRP, MFIP, and FIP (Panel D). For the most part, these programs produce employment effects that are similar across the distribution of socioeconomic disadvantage. WRP has similar effects for all three levels of its composite measure of disadvantage, and MFIP has similar effects by education and prior welfare use. It has smaller effects for persons with recent employment histories than for persons without any recent employment. The only significant effects of FIP are for relatively disadvantaged groups, but like most of the other studies, FIP provides no tests for whether the subgroup differences are statistically significant.

WRP's impact on earnings is significant only for the least disadvantaged group, although the estimates for the most and least disadvantaged groups are similar. The only estimate from MFIP that is significant is the one for persons with no recent work history, which is substantially larger than the corresponding estimate for persons with recent work histories. Estimates are similar by education and welfare histories, albeit insignificant. The overall third-year earnings effect of MFIP, shown in Table 5.1, is of similar magnitude and marginally significant. Smaller samples by subgroup are probably the reason why the subgroup-specific estimates are insignificant. The earnings effects of FIP are generally stronger for the more disadvantaged groups.

A.2.5. Programs That Focus on TANF-Like Bundles of Reforms

ABC, FTP, and Jobs First provide subgroup-specific employment impacts (Panel F of Table A.3). There is little uniformity in their impacts by level of socioeconomic disadvantage. In ABC, impacts are larger for older mothers, mothers with older children, and mothers with more recent work experience, but fairly similar by level of education. They are larger for women with intermediate welfare histories than for women with more or less recent time on aid.

In FTP, the program impacts on employment are similar by the length of the recipient's time limit, by the age of her youngest child, and by her level of dependency risk. In Jobs First, they are generally greater among the more disadvantaged groups. Earnings impacts in Jobs First likewise are larger among more disadvantaged groups. However, in FTP, they are greatest for those at least risk, based on the composite risk scale, but also greater among those with shorter time limits and with children under age three.

A.3. USE OF OTHER GOVERNMENT PROGRAMS

As seen in Table A.5, the only random assignment studies to provide subgroup-specific impact estimates are WRP, New Hope, FTP, and Jobs First. Only New Hope provides statistical tests for subgroup differences. Moreover, these studies provide subgroup impacts only for food stamps.

Table A.5—Estimated Impact of Welfare Reform on Use of Food Stamp Program for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %
A. Programs that focus on financial work incentives												
			By status at RA:	Recipient			Applicant					
WRP-IO	Single-parent recipients and applicants	n.a.	Received FS in last 3 mos. of FU	57.8	-0.1	-0.2%	39.9	2.6	6.5%			
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)		
		n.a.	Received FS in last 3 mos. of FU	68.4	0.1	0.1%	52.8	-0.6	-1.1%	38.7	4.1	10.6%
B. Programs that focus on financial work incentives tied to hours of work												
			By barriers to employment:	Two (2)			One (2)			None (2)		
New Hope	Poor families employed FT at RA		Months receiving FS, year 1	8.4	0.4	4.8%	7.3	0.0	0.0%	6.7	-0.6	-9.0%
			Months receiving FS, year 2	6.2	1.1 **	17.7%	5.1	0.4	7.8%	4.2	0.1	2.4%
C. Programs that focus on mandatory work-related activities												
D. Programs that focus on financial work incentives and mandatory work-related activities												
			By status at RA:	Recipient			Applicant					
WRP	Single-parent recipients and applicants	n.a.	Received FS in last 3 mos. of FU	57.8	-1.9	-3.3%	39.9	-0.9	-2.3%			
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)		
		n.a.	Received FS in last 3 mos. of FU	68.4	0.5	0.7%	52.8	-1.3	-2.5%	38.7	-3.1	-8.0%
E. Programs that focus on other individual reforms												
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)												
			By length of time limit:	36 months (3)			24 months (3)					
FTP	Recipients and applicants	n.a.	Ever received FS in years 1-2	92.4	2.0	2.2	88.4	-0.8	-0.9			
			By age of youngest child:	Under 3			3 or over					
		n.a.	Ever received FS in years 1-2	91.0	1.1	1.2	89.2	0.1	0.1			

Table A.5—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %
			By status at RA:	Recipient			Applicant					
		n.a.	Avg. percent receiving FS, years 1-2	79.9	0.6	0.8	56.1	5.2 ***	9.3			
		n.a.	Avg. percent receiving FS, year 3	61.8	-3.7 **	-6.0	40.1	-0.4	-1.0			
			By disadvantage:	Most disadvantaged (4)						Least disadvantaged (4)		
		n.a.	Avg. percent receiving FS, years 1-2	84.5	0.3	0.4				50.4	3.9	7.7
		n.a.	Avg. percent receiving FS, year 3	70.0	-2.1	-3.0				36.1	-1.6	-4.4
			By age of youngest child:	Less than 6			6 to 11			12 to 18		
		n.a.	Avg. quarterly FS receipt, Q1-6	74.7	2.9 **	3.9%	74.1	7.0 ***	9.4%	75.4	-2.0	-2.7%
		n.a.	Any FS receipt, Q8	59.6	-0.4	-0.7%	59.3	3.4	5.7%	57.3	-7.6 *	-13.3%
JOBS First	Recipients and applicants		**By aid receipt in year before RA:**	Long-term recipient (5)			Short-term recipient (5)			New applicant (5)		
		n.a.	Avg. quarterly FS receipt, Q1-6	87.3	-0.3	-0.3%	70.6	7.0 ***	9.9%	57.9	3.6 *	6.2%
		n.a.	Any FS receipt, Q8	75.1	-4.6 **	-6.1%	52.7	3.6	6.8%	40.9	1.9	4.6%
			By employment in year before RA:	Not employed			Employed					
		n.a.	Avg. quarterly FS receipt, Q1-6	78.8	1.2	1.5%	69.6	4.0 ***	5.7%			
		n.a.	Any FS receipt, Q8	65.3	-2.3	-3.5%	52.9	1.0	1.9%			
			By earnings in year before RA:	No earnings			$1 to 5000			Over $5000		
		n.a.	Avg. quarterly FS receipt, Q1-6	78.8	1.2	1.5%	76.6	1.5	2.0%	58.8	7.3 ***	12.4%
		n.a.	Any FS receipt, Q8	65.3	-2.3	-3.5%	60.4	-2.4	-4.0%	41.2	5.7	13.8%

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%;xx= 5%; xxx=1%. Abbreviations: FS=Food Stamps; FU=follow-up; RA=random assignment; n.a.=not available.

(1) People in the "most disadvantaged" group: (1) were on aid at least 22 of the 24 months prior to RA; (2) had not worked in the past year; (3) had no high school credential. People in the "least disadvantaged" group had none of these characteristics. People in the "moderately disadvantaged" group had some, but not all three.

(2) Barriers include not having worked in the past six years; being arrested since age 16; having 2 or more children under six or four or more under 12; having been fired from one's longest-lasting jobs; and not having a high school credential.

(3) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(4) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. Results for less disadvantaged groups are not provided.

(5) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

There is no information in the literature on whether the various reform programs have affected Medicaid coverage or utilization of other nutrition programs in a manner that varies according to socioeconomic disadvantage.

Neither WRP-IO nor WRP had significantly different effects on food stamp use by subgroup (Panels A and D). This underscores the impression from Table 6.1 and earlier chapters that this program had little effect on behavior overall.

Only one of the subgroup-specific estimates from New Hope is statistically significant (Panel B). The subgroup impacts do not differ significantly from each other.

Subgroup results for FTP and Jobs First are presented in Panel F of Table A.5. The pre–time limit impacts of FTP on FSP participation are similar to FTP's impacts on welfare use. The impacts vary little by either the length of the recipient's time limit or the age of her youngest child. The subgroup-specific FSP impacts of Jobs First are smaller, and mostly less significant, than the corresponding effects on welfare use. During the pre–time limit period, Jobs First increased food stamp use significantly among applicants, families with younger children, shorter-term recipients, and recipients with more favorable work histories. In the post–time limit period, Jobs First reduced food stamp use significantly among recipients (as opposed to applicants) and long-term recipients.

In the case of food stamps, a particularly important subgroup to consider is immigrants, since PRWORA included a provision to remove most legal immigrants from the food stamp rolls. Although that provision of the law was later overturned, it has been suggested that the provision had a chilling effect, causing immigrants to leave the program, returning only in smaller numbers after the law was rescinded. Three sets of analysts consider the decline in food stamp use between 1994 and 1997, noting that the decline among immigrants was greater than the decline among natives. Borjas (2001b) observes that the decline in immigrant food stamp use relative to natives is almost entirely attributable to the substantial relative decline that took place in California. He conjectures that it may have as much to do with anti-immigrant sentiment there as with the terms of PRWORA. Lofstrom and Bean (2001), in contrast, factor in the role of the economy and conclude that the labor market conditions facing immigrants were responsible for their greater decline in aid use. Haider et al. (2001) reach similar conclusions.

A.4. INCOME, INCOME SOURCES, AND POVERTY

A number of the experimental studies that provide results for income and poverty also examine impacts for subgroups.[112] Tables A.6, A.7, and A.8 record results for the studies with subgroup impacts for income, welfare payments, and food stamp payments, respectively. We review the findings by the reform policy or policies evaluated in the experiment for the three outcomes covered by the tables.

---

[112]None of the econometric studies reviewed in Chapter 8 consider differences for subgroups other than the differences for groups defined by education discussed in Section 8.3.

Table A.6—Estimated Impact of Welfare Reform on Income for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
A. Programs that focus on financial work incentives															
WRP-IO	Single-parent recipients and applicants		**By level of disadvantage:**	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Avg. quarterly R E-W-FS income in last 3 mos. of 42-mo FU	$2,039	$172	8.4%	$2,284	-$170	-7.4%	$2,408	$299 *	12.4%			
			By status at RA:	Recipient			Applicant								
		n.a.	Avg. quarterly R E-W-FS income in last 3 mos. of 42-mo FU	$2,394	-$52	-2.2%	$2,190	$70	3.2%						
B. Programs that focus on financial work incentives tied to hours of work															
New Hope	Poor families not employed FT at RA		**By potential employment barriers (6)**	Two or more			One			None					
			Avg. annual R E+W+FS +EITC for year 1 of 2-yr FU	$10,449	$1,669 **	16.0%	$9,550	$1,653 ***	17.3%	$9,724	$439	4.5%			
		xx	Avg. annual R E+W+FS +EITC for year 2 of 2-yr FU	$10,385	$1,894 **	18.2%	$9,217	$2,325 ***	25.2%	$10,409	-$706	-6.8%			
SSP (a)	Single-parent recipients		**By age of youngest child:**	0 to 2			3 to 11			12 to 15					
			Cumulative R E+IA+SSP since RA at 36-mo FU	$32,087	$4,935 ***	15.4%	$32,712	$5,258 ***	16.1%	$31,645	$5,533 ***	17.5%			
			By education:	No high school credential			High school credential								
		xx	Cumulative R E+IA+SSP since RA at 36-mo FU	$31,021	$4,303 ***	13.9%	$33,206	$6,179 ***	18.6%						
			By employment status at RA:	Out of labor force			Unemployed			Employed PT			Employed FT		
		xxx	Cumulative R E+IA+SSP since RA at 36-mo FU	$30,361	$3,971 ***	13.1%	$31,231	$6,386 ***	20.4%	$36,803	$8,036 ***	21.8%	$41,943	$9,204 ***	21.9%
			By months on aid in prior 3 years:	36 months			24 to 35			10 to 23					
		xx	Cumulative R E+IA+SSP since RA at 36-mo FU	$31,541	$4,000 ***	12.7%	$32,166	$6,025 ***	18.7%	$32,459	$6,585 ***	20.3%			

Table A.6—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1			Group 2			Group 3			Group 4		
				Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%
C. Programs that focus on mandatory work-related activities															
			By earnings in year prior to RA	No earnings			$5,000 or less			More than $5,000					
			Avg. annual R E+W+FS income in years 1 to 3	$8,082	$41	0.5%	$8,707	-$58	-0.7%	$11,200	$143	1.3%			
		xx	**By education:**	No high school credential			High school credential								
			Avg. annual R E+W+FS income in years 1 to 3	$8,282	-$66	-0.8%	$8,989	$123 *	1.4%						
			By number of children:	Three or more			Two			One					
NEWWS Sites - pooled	Recipients and applicants		Avg. annual R E+W+FS income in years 1 to 3	$10,412	$93	0.9%	$8,769	$128 *	1.5%	$7,589	-$65	-0.9%			
			By level of disadvantage:	Most disadvantaged (8)			Moderately disadvantaged (8)			Least disadvantaged (8)					
			Avg. annual R E+W+FS income in years 1 to 3	$8,426	-$116	-1.4%	$8,591	$79	0.9%	$9,558	$41	0.4%			
		xx	**By status at RA:**	Long-term recipient			Short-term recipient			Applicant					
			Avg. annual R E+W+FS income in years 1 to 3	$9,027	$4	0.0%	$8,463	$94	1.1%	6819.0	773.0 ***	11.3%			

Table A.6—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
D. Programs that focus on financial work incentives and mandatory work-related activities															
WRP	Single-parent recipients and applicants		**By level of disadvantage:**	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
		n.a.	Avg. quarterly R E-W-FS income in last 3 mos. of 42-mo FU	$2,039	$133	6.5%	$2,284	-$81	-3.5%	$2,408	$258 *	10.7%			
			By status at RA:	Recipient			Applicant								
		n.a.	Avg. quarterly R E-W-FS income in last 3 mos. of 42-mo FU	$2,394	-$44	-1.8%	$2,190	$124	5.7%						
FIP	Recipients		**By earnings in year before RA:**	Without earnings			With earnings								
		n.a.	Avg. annual R E-W income in year 2 of 2-yr FU	$5,818	$475 *	8.2%	$8,458	$150	1.8%						
			By age of youngest child:	Under 3			3 and above								
		n.a.	Avg. annual R E-W income in year 2 of 2-yr FU	$7,158	$401 *	5.6%	$7,178	$255	3.6%						
	Applicants		**By earnings in year before RA:**	Without earnings			With earnings								
		n.a.	Avg. annual R E-W income in year 1 of 2-yr FU	$4,316	$695 ***	16.1%	$9,540	$739 **	7.7%						
			By age of youngest child:	Under 3			3 and above								
		n.a.	Avg. annual R E-W income in year 1 of 2-yr FU	$7,237	$1,037 ***	14.3%	$8,356	$274	3.3%						
E. Programs that focus on other individual reforms															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
FTP	Recipients and applicants		**By risk of welfare dependency**	Most at risk (4)			Medium risk (4)			Least at risk (4)					
			Avg. total recipient E-W-FS income over 4-yr FU	$28,832	-$82	-0.3%	$22,353	$807	3.6%	$28,831	$3,200 **	11.1%			
			Avg. total recipient E-W-FS income in year 4 of 4-yr FU	$6,828	-$180	-2.6%	$5,548	$47	0.8%	$7,334	$1,050 *	14.3%			
			Avg. mo. HH income in month before 4-yr FU	$1,241	$32	2.6%	$1,352	$43	3.2%	$1,601	$231	14.4%			
			By disadvantage:	Highly disadvantaged (5)											
			Avg. total recipient E-W-FS income over 4-yr FU	$29,170	-$2,040	-7.0%									
			Avg. total recipient E-W-FS income in year 4 of 4-yr FU	$6,776	-$737	-10.9%									

Table A.6—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
Jobs First	Recipients and applicants		**By status at RA:**	Recipient			Applicant								
		n.a.	Avg. annual recipient E-W-FS income in years 1 and 2	$9,948	$1,379 ***	13.9%	$9,536	$870 ***	9.1%						
		n.a.	Avg. annual recipient E-W-FS income in year 3	$10,697	$100	0.9%	$10,516	$399	3.8%						
			By disadvantage:	Most disadvantaged (3)						Least disadvantaged (3)					
		n.a.	Avg. annual recipient E-W-FS income in years 1 and 2	$8,455	$1,120 ***	13.2%				$11,614	$700	6.0%			
		n.a.	Avg. annual recipient E-W-FS income in year 3	$8,303	$480	5.8%				$13,870	-$154	-1.1%			
			By age of youngest child:	Less than 6			6 to 11			12 to 18					
		n.a.	Avg. quarterly recipient E-W-FS income in Q1-6	$2,490	$324 ***	13.0%	$2,550	$411 ***	16.1%	$2,122	$246 **	11.6%			
		n.a.	Avg. quarterly recipient E-W-FS income in Q8	$2,608	$185 **	7.1%	$2,605	$169	6.5%	$2,145	$99	4.6%			
			By aid receipt in year before RA:	Long-term recipient (7)			Short-term recipient (7)			New applicant (7)					
		n.a.	Avg. quarterly recipient E-W-FS income in Q1-6	$2,501	$377 ***	15.1%	$2,417	$331 ***	13.7%	$2,282	$228 ***	10.0%			
		n.a.	Avg. quarterly recipient E-W-FS income in Q8	$2,573	$100	3.9%	$2,515	$180	7.2%	$2,482	$229 *	9.2%			
			By employment in year before RA:	Not employed			Employed								
		n.a.	Avg. quarterly recipient E-W-FS income in Q1-6	$2,034	$315 ***	15.5%	$2,769	$310 ***	11.2%						
		n.a.	Avg. quarterly recipient E-W-FS income in Q8	$2,025	$228 ***	11.3%	$2,982	$106	3.6%						
			By earnings in year before RA:	No earnings			$1 to 5000			Over $5000					
		n.a.	Avg. quarterly recipient E-W-FS income in Q1-6	$2,034	$315 ***	15.5%	$2,381	$368 ***	15.5%	$3,381	$238 **	7.0%			
		n.a.	Avg. quarterly recipient E-W-FS income in Q8	$2,025	$228 ***	11.3%	$2,647	$44	1.7%	$3,527	$194	5.5%			

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%. n.a. = not available. Abbreviations: E=earnings; W=cash welfare payments; FS=Food Stamp payments; IA=Income Assistance (Canada); EITC=Earned Income Tax Credit; CC=out-of-pocket child care expenses; FU=follow-up; HH=household; Q=quarter; R=recipient.

(a) Results in Canadian dollars.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantaged had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.
(3) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. "Least disadvantaged" is defined as having none of these traits.

(4) "Most at risk" have risk score in top quartile of dependency index; "Least at risk" have score in the bottom quartile; "Medium risk" are in between. Dependency index is based on prior quarter of employment, months employed prior to RA, AFDC recipiency status in quarter prior to RA, months of AFDC prior to RA, age of youngest child, and high school credential status.

(5) "Highly disadvantaged" consists of those in "most at risk" subgroup with no HS diploma or GED, no UI-reported earnings in year prior to RA, and 2 or more years of reported AFDC/TANF receipt prior to RA.

(6) Potential barriers to employment are not having worked in the past six years; having been arrested since age 16; having either two or more children under age 6 or four children under age 12; having been fired from one's period of longest employment; and not having a high school credential.

(7) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

(8) "Most disadvantaged" is classified as those with no earnings in the year prior to RA, without a high school credential, and received welfare two years or more prior to RA. Those classified as "Least disadvantaged" had none of these characteristics. All others were classified as "Moderately disadvantaged."

Table A.7—Estimated Impact of Welfare Reform on Welfare Payments for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1			Group 2			Group 3			Group 4		
				Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%	Control mean	Impact	%
A. Programs that focus on financial work incentives															
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
WRP-IO	Single-parent recipients and applicants	n.a.	Quarterly ANFC payment for last 3 mos. of 42-mo FU	$847	$56	6.6%	$593	-$55	-9.3%	$328	$44	13.4%			
			By status at RA:	Recipient			Applicant								
		n.a.	Quarterly ANFC payment for last 3 mos. of 42-mo FU	$638	$4	0.6%	$406	-$37	-9.1%						
B. Programs that focus on financial work incentives tied to hours of work															
			By potential employment barriers (6)	Two or more			One			None					
New Hope	Poor families not employed FT at RA		Avg. annual R AFDC payments in year 1 of 2-yr FU	$3,705	$288	7.8%	$2,818	$29	1.0%	$2,336	-$309	-13.2%			
			Avg. annual R AFDC payments in year 2 of 2-yr FU	$2,286	$332	14.5%	$1,490	$1	0.1%	$1,280	-$136	-10.6%			

Table A.7—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
C. Programs that focus on mandatory work-related activities															
LA Jobs-1st GAIN	Single-parent recipients and applicants		**By race and ethnicity:**	White			African-American			Hispanic			Asian		
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$8,903	-$822 ***	-9.2%	$10,394	-$874 ***	-8.4%	$10,211	-$1,180 ***	-11.6%	$10,211	-$626 **	-6.1%
			By education:	No high school credential			High school credential								
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$10,739	-$964 ***	-9.0%	$9,272	-$992 ***	-10.7%						
			By English proficiency:	Not proficient			Proficient								
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$11,006	-$1,034 ***	-9.4%	$9,842	-$977 ***	-9.9%						
			By emp. status in year prior to RA:	Not employed			Employed								
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$10,754	-$1,029 ***	-9.6%	$8,865	-$881 ***	-9.9%						
			By level of disadvantage:	Most disadvantaged (7)											
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$11,702	-$975 ***	-8.3%									
			By status at RA:	Long-term recipient			Short-term recipient			Applicant					
			Avg. annual R AFDC/TANF payments in years 1 and 2 of 2-yr FU	$10,766	-$1,034 ***	-9.6%	$8,229	-$856 ***	-10.4%	$7,665	-$507	-6.6%			
NEWWS Sites - pooled	Recipients and applicants		**By earnings in year prior to RA**	No earnings			$5,000 or less			More than $5,000					
			Avg. annual R AFDC payments in years 1 to 3	$4,675	-$416 ***	-8.9%	$3,696	-$359 ***	-9.7%	$2,967	-$305 ***	-10.3%			
			By education:	No high school credential			High school credential								
			Avg. annual R AFDC payments in years 1 to 3	$4,708	-$395 ***	-8.4%	$3,749	-$389 ***	-10.4%						
			By number of children:	Three or more			Two			One					
		x	Avg. annual R AFDC payments in years 1 to 3	$5,604	-$458 ***	-8.2%	$4,185	-$408 ***	-9.7%	$3,268	-$326 ***	-10.0%			
			By level of disadvantage:	Most disadvantaged (8)			Moderately disadvantaged (8)			Least disadvantaged (8)					
			Avg. annual R AFDC payments in years 1 to 3	$5,570	-$411 ***	-7.4%	$4,066	-$414 ***	-10.2%	$2,677	-$282 ***	-10.5%			
			By status at RA:	Long-term recipient			Short-term recipient			Applicant					
		x	Avg. annual R AFDC payments in years 1 to 3	$4,791	-$433 ***	-9.0%	$3,400	-$337 ***	-9.9%	$2,611	-$218 **	-8.3%			

Table A.7—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
D. Programs that focus on financial work incentives and mandatory work-related activities															
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)					
WRP	Single-parent recipients and applicants	n.a.	Quarterly ANFC payment for last 3 mos. of 42-mo FU	$847	-$134	-15.8%	$593	-$128 ***	-21.6%	$328	-$51	-15.5%			
			By status at RA:	Recipient			Applicant								
		n.a.	Quarterly ANFC payment for last 3 mos. of 42-mo FU	$638	-$123 ***	-19.3%	$406	-$75 **	-18.5%						
FIP	Recipients		**By earnings in year before RA:**	Without earnings			With earnings								
		n.a.	Avg. annual R FIP payments in year 2 of 2-yr FU	$2,909	-$62	-2.1%	$2,288	-$45	-2.0%						
			By age of youngest child:	Under 3			3 and above								
		n.a.	Avg. annual R FIP payments in year 2 of 2-yr FU	$2,873	-$153 *	-5.3%	$2,344	$32	1.4%						
	Applicants		**By earnings in year before RA:**	Without earnings			With earnings								
		n.a.	Avg. annual R FIP payments in year 1 of 2-yr FU	$1,832	$276 **	15.1%	$1,456	-$28	-1.9%						
			By age of youngest child:	Under 3			3 and above								
		n.a.	Avg. annual R FIP payments in year 1 of 2-yr FU	$1,767	$55	3.1%	$1,387	$75	5.4%						
E. Programs that focus on other individual reforms															
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)															
FTP	Recipients and applicants		**By length of time limit:**	36 months (2)			24 months (2)								
			Avg. total R AFDC/TANF payments in year 1 and 2 of 4-yr FU	$3,989	-$52	-1.3%	$2,783	-$207 *	-7.4%						
			By age of youngest child:	Under 3			3 or over								
			Avg. total R AFDC/TANF payments in year 1 and 2 of 4-yr FU	$3,667	-$303 **	-8.3%	$3,001	-$29	-1.0%						
			By risk of welfare dependency	Most at risk (4)			Medium risk (4)			Least at risk (4)					
			Avg. total R AFDC/TANF payments over 4-yr FU	$7,982	-$1,087 ***	-13.6%	$4,311	-$664 ***	-15.4%	$2,216	-$490 **	-22.1%			
		xxx	Avg. total R AFDC/TANF payments in year 4 of 4-yr FU	$969	-$518 ***	-53.5%	$503	-$249 ***	-49.5%	$217	-$87 **	-40.1%			
		xx	Avg. mo. HH AFDC/TANF payments in month before 4-yr FU	$78	-$44 ***	-56.4%	$56	-$27 ***	-48.2%	$22	-$12 *	-54.5%			
			By disadvantage:	Highly disadvantaged (5)											
			Avg. total R AFDC/TANF payments over 4-yr FU	$9,474	-$1,893 ***	-20.0%									
			Avg. total R AFDC/TANF payments in year 4 of 4-yr FU	$1,269	-$806 ***	-63.5%									

Table A.7—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %	Group 4 Control mean	Group 4 Impact	Group 4 %
			By status at RA:		Recipient			Applicant							
		n.a.	Avg. annual R AFDC/TFA payments in years 1 and 2	$4,059	$492 ***	12.1%	$2,538	$702 ***	27.7%						
		n.a.	Avg. annual R AFDC/TFA payments in year 3	$2,708	-$546 ***	-20.2%	$1,550	-$191 *	-12.3%						
			By disadvantage:		Most disadvantaged (3)						Least disadvantaged (3)				
		n.a.	Avg. annual R AFDC/TFA payments in years 1 and 2	$4,761	$77	1.6%				$2,015	$867 ***	43.0%			
		n.a.	Avg. annual R AFDC/TFA payments in year 3	$3,439	-$679 ***	-19.7%				$1,114	$99	8.9%			
			By age of youngest child:		Less than 6			6 to 11			12 to 18				
		n.a.	Avg. quarterly R AFDC/TFA payments in Q1-6	$991	$200 ***	20.2%	$900	$244 ***	27.1%	$735	$98 **	13.3%			
		n.a.	Avg. quarterly R AFDC/TFA payments in Q8	$761	-$51 *	-6.7%	$667	-$29	-4.3%	$455	-$109 **	-24.0%			
Jobs First	Recipients and applicants		**By aid receipt in year before RA:**		Long-term recipient (9)			Short-term recipient (9)			New applicant (9)				
		n.a.	Avg. quarterly R AFDC/TFA payments in Q1-6	$1,157	$156 ***	13.5%	$835	$212 ***	25.4%	$667	$183 ***	27.4%			
		n.a.	Avg. quarterly R AFDC/TFA payments in Q8	$920	-$127 ***	-13.8%	$602	-$34	-5.6%	$424	$43	10.1%			
			By employment in year before RA:		Not employed			Employed							
		n.a.	Avg. quarterly R AFDC/TFA payments in Q1-6	$1,072	$99 ***	9.2%	$787	$256 ***	32.5%						
		n.a.	Avg. quarterly R AFDC/TFA payments in Q8	$813	-$87 ***	-10.7%	$567	-$18	-3.2%						
			By earnings in year before RA:		No earnings			$1 to 5000			Over $5000				
		n.a.	Avg. quarterly R AFDC/TFA payments in Q1-6	$1,072	$99 ***	9.2%	$920	$204 ***	22.2%	$578	$332 ***	57.4%			
		n.a.	Avg. quarterly R AFDC/TFA payments in Q8	$813	-$87 ***	-10.7%	$688	-$83 **	-12.1%	$377	$80 *	21.2%			

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%. n.a. = not available. Abbreviations: FU=follow-up; HH=household; Q=quarter; R=recipient.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantaged had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(3) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. "Least disadvantaged" is defined as having none of these traits.

(4) "Most at risk" have risk score in top quartile of dependency index; "Least at risk" have score in the bottom quartile; "Medium risk" are in between. Dependency index is based on prior quarter of employment, months employed prior to RA, AFDC recipiency status in quarter prior to RA, months of AFDC prior to RA, age of youngest child, and high school credential status.

(5) "Highly disadvantaged" consists of those in "most at risk" subgroup with no HS diploma or GED, no UI-reported earnings in year prior to RA, and 2 or more years of reported AFDC/TANF receipt prior to RA.

(6) Potential barriers to employment are not having worked in the past six years; having been arrested since age 16; having either two or more children under age 6 or four children under age 12; having been fired from one's period of longest employment; and not having a high school credential.

(7) "Most disadvantaged" is defined as long-term recipients who did not have a high school credential at RA and who did not work for pay in the year prior to RA.

(8) "Most disadvantaged" is classified as those with no earnings in the year prior to RA, without a high school credential, and received welfare two years or more prior to RA. Those classified as "Least disadvantaged" had none of these characteristics. All others were classified as "Moderately disadvantaged."

(9) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

Table A.8—Estimated Impact of Welfare Reform on Food Stamp Payments for Subgroups: Random Assignment Studies

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %
A. Programs that focus on financial work incentives												
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)		
WRP-IO	Single-parent recipients and applicants	n.a.	Quarterly FS payment for last 3 mos. of 42-mo FU	$391	$26	6.6%	$285	-$7	-2.5%	$183	$41 *	22.4%
			By status at RA:	Recipient			Applicant					
		n.a.	Quarterly FS payment for last 3 mos. of 42-mo FU	$317	$2	0.6%	$198	$19	9.6%			
B. Programs that focus on financial work incentives tied to hours of work												
			By potential employment barriers (6)	Two or more			One			None		
New Hope	Poor families not employed FT at RA		Avg. annual R FS payments in year 1 of 2-yr FU	$2,274	$249	10.9%	$1,806	-$84	-4.7%	$1,363	-$79	-5.8%
		xx	Avg. annual R FS payments in year 2 of 2-yr FU	$1,568	$589 ***	37.6%	$1,259	-$13	-1.0%	$817	$123	15.1%
C. Programs that focus on mandatory work-related activities												
D. Programs that focus on financial work incentives and mandatory work-related activities												
			By level of disadvantage:	Most disadvantaged (1)			Moderately disadvantaged (1)			Least disadvantaged (1)		
WRP	Single-parent recipients and applicants	n.a.	Quarterly FS payment for last 3 mos. of 42-mo FU	$391	$4	1.0%	$285	-$9	-3.2%	$183	$5	2.7%
			By status at RA:	Recipient			Applicant					
		n.a.	Quarterly FS payment for last 3 mos. of 42-mo FU	$317	-$13	-4.1%	$198	$10	5.1%			

Table A.8—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %
E. Programs that focus on other individual reforms												
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)												
			By length of time limit:	36 months (2)			24 months (2)					
			Avg. total R FS payments in year 1 and 2 of 4-yr FU	$4,744	-$188	-4.0%	$3,685	-$464 ***	-12.6%			
			By age of youngest child:	Under 3			3 or over					
			Avg. total R FS payments in year 1 and 2 of 4-yr FU	$4,541	-$392 ***	-8.6%	$3,800	-$339 ***	-8.9%			
			By risk of welfare dependency	Most at risk (4)			Medium risk (4)			Least at risk (4)		
FTP	Recipients and applicants		Avg. total R FS payments over 4-yr FU	$10,280	-$473	-4.6%	$6,175	-$549 **	-8.9%	$3,901	-$531 **	-13.6%
			Avg. total R FS payments in year 4 of 4-yr FU	$1,928	$50	2.6%	$1,032	-$106	-10.3%	$504	-$11	-2.2%
			Avg. mo. HH FS payments in month before 4-yr FU	$185	$7	3.8%	$115	-$8	-7.0%	$66	-$12	-18.2%
			By disadvantage:	Highly disadvantaged (5)								
			Avg. total R FS payments over 4-yr FU	$12,249	-$1,721 **	-14.1%						
			Avg. total R FS payments in year 4 of 4-yr FU	$2,440	-$283	-11.6%						

Table A.8—Continued

Name	Cases served	Signif. of group diff.	Measure	Group 1 Control mean	Group 1 Impact	Group 1 %	Group 2 Control mean	Group 2 Impact	Group 2 %	Group 3 Control mean	Group 3 Impact	Group 3 %
			By status at RA:	Recipient			Applicant					
		n.a.	Avg. annual R FS payments in years 1 and 2	$1,999	$140 ***	7.0%	$1,209	$217 ***	17.9%			
		n.a.	Avg. annual R FS payments in year 3	$1,596	-$26	-1.6%	$917	$32	3.5%			
			By disadvantage:	Most disadvantaged (3)						Least disadvantaged (3)		
		n.a.	Avg. annual R FS payments in years 1 and 2	$2,267	$57	2.5%				$972	$241 ***	24.8%
		n.a.	Avg. annual R FS payments in year 3	$1,892	-$36	-1.9%				$729	$43	5.9%
			By age of youngest child:	Less than 6			6 to 11			12 to 18		
		n.a.	Avg. quarterly R FS payments in Q1-6	$468	$42 ***	9.0%	$457	$84 ***	18.4%	$395	$24	6.1%
		n.a.	Avg. quarterly R FS payments in Q8	$395	$3	0.8%	$369	$32	8.7%	$304	-$41	-13.5%
Jobs First	Recipients and applicants		**By aid receipt in year before RA:**	Long-term recipient (7)			Short-term recipient (7)			New applicant (7)		
		n.a.	Avg. quarterly R FS payments in Q1-6	$563	$31 ***	5.5%	$402	$71 ***	17.7%	$300	$53 ***	17.7%
		n.a.	Avg. quarterly R FS payments in Q8	$498	-$21	-4.2%	$315	$13	4.1%	$226	$31 *	13.7%
			By employment in year before RA:	Not employed			Employed					
		n.a.	Avg. quarterly R FS payments in Q1-6	$497	$32 ***	6.4%	$387	$65 ***	16.8%			
		n.a.	Avg. quarterly R FS payments in Q8	$422	-$3	-0.7%	$314	$12	3.8%			
			By earnings in year before RA:	No earnings			$1 to 5000			Over $5000		
		n.a.	Avg. quarterly R FS payments in Q1-6	$497	$32 ***	6.4%	$428	$48 ***	11.2%	$323	$91 ***	28.2%
		n.a.	Avg. quarterly R FS payments in Q8	$422	-$3	-0.7%	$366	-$15	-4.1%	$235	$51 **	21.7%

NOTES: For full program names and citations, see Table 3.4. Significance tests for treatment-control differences is indicated by: *=10%, **=5%; ***=1%. Significance of test for subgroups differences is indicated by: x = 10%; xx= 5%; xxx=1%. n.a. = not available. Abbreviations: FS=Food Stamp; FU=follow-up; HH=household; Q=quarter; R=recipient.

(1) Persons classified as "most disadvantaged" (1) had been on aid for at least 22 of the 24 months prior to random assignment; (2) had not worked in the prior year; and (3) did not have a high school credential. Persons classified as "least disadvantaged had none of these traits; persons classified as "moderately disadvantaged" had one or two.

(2) Participants with low levels of education, short employment histories, and long welfare histories were assigned a 36-month time limit; other participants were assigned a 24-month time limit.

(3) "Most disadvantaged" is defined as having no high school credential, not having worked in the year prior to RA, and having been on aid at least 21 of the 24 months prior to RA. "Least disadvantaged" is defined as having none of these traits.

(4) "Most at risk" have risk score in top quartile of dependency index; "Least at risk" have score in the bottom quartile; "Medium risk" are in between. Dependency index is based on prior quarter of employment, months employed prior to RA, AFDC recipiency status in quarter prior to RA, months of AFDC prior to RA, age of youngest child, and high school credential status.

(5) "Highly disadvantaged" consists of those in "most at risk" subgroup with no HS diploma or GED, no UI-reported earnings in year prior to RA, and 2 or more years of reported AFDC/TANF receipt prior to RA.

(6) Potential barriers to employment are not having worked in the past six years; having been arrested since age 16; having either two or more children under age 6 or four children under age 12; having been fired from one's period of longest employment; and not having a high school credential.

(7) "Long term" recipients are those who received aid for at least 22 of the 24 months prior to RA. "Short term" recipients are those who received aid for 1 to 22 months during the 24 months prior to RA. "New applicants" are those who received no aid during the 24 months prior to RA.

A.4.1. Programs That Focus on Financial Work Incentives

Only WRP-IO provides results for income, welfare payments, and food stamp payments by subgroups (Panel A of Tables A.6, A.7, and A.8), in this case defined by a composite measure of disadvantage and for recipients versus applicants. For the entire sample, impacts on income and transfer payments were small and insignificant. For the subgroups considered, impacts are only significant for the least disadvantaged group, which experienced an increase in income and an increase in food stamp payments. The direction and magnitude of the effects are similar (but smaller) for the most disadvantaged group; the moderately disadvantaged group is the outlier. There are no strong differences between recipients and applicants.

A.4.2. Programs That Focus on Financial Work Incentives Tied to Hours Worked

Both New Hope and SSP report subgroup differences in income impacts (Panel B of Table A.6). New Hope shows the largest income gains for the more disadvantaged, defined by barriers to employment (one or more barriers compared with those with none). In contrast, SSP consistently finds larger income gains for the least disadvantaged, with significant differences among groups defined by education, employment status at random assignment, and welfare use history. For example, the income gains over three years are more than two times as large for SSP participants employed full-time at random assignment compared with those out of the labor force.

Only New Hope reports results for welfare payments and food stamp payments for different subgroups (Panel B of Tables A.7 and A.8, respectively). New Hope finds no differences in welfare payments for groups defined by employment barriers. Differences in food stamp payments are sharper, with a significant increase in benefits to those with two or more employment barriers in the second year of follow-up, a difference that is significant from that measured for families with fewer employment barriers.

A.4.3. Programs That Focus on Mandatory Work-Related Activities

None of the programs in this group consider differences in food stamp payments by subgroup (Panel C of Table A.8). L.A. Jobs-First GAIN considers only welfare payment impacts by various subgroups (Panel C of Table A.7). Overall, this program resulted in a significant reduction in welfare payments, and this result also holds for all the subgroups reported in Panel C of Table A.7; none of the differences between groups are statistically significant. This finding suggests that the program impacts in terms of reduced welfare benefit payments apply to all the subgroups analyzed, from the most to the least disadvantaged.

As above, a pooled analysis of the 11 NEWWS experiments along with 9 others (including FTP and MFIP) considered impacts for various subgroups. Panel C of Tables A.6 and A.7 reports the results for this pooled analysis in terms of impacts on income and welfare payments, respectively (Michalopoulos and Schwartz, 2000). The results suggest that income gains are strongest for the least disadvantaged, for example those with a high school credential, with two children, and applicants. The differences in income impacts among groups defined by education and recipiency status are statistically significant, but the size of the annual income difference is meaningful only for recipients versus applicants (over $700 per year higher for applicants).

In the case of welfare payments, all subgroups experience a statistically significant decline, from $200 to $500 per year. Even when the impacts differ by group (and the differences are significant when defined by the number of children and recipiency status), the magnitude of the differences tends to be small. For example, the programs reduced welfare use by an average of $218 a year among new applicants, compared to $433 a year among long-term recipients.

A.4.4. Programs That Focus on Financial Work Incentives and Mandatory Work-Related Activities

Two programs—WRP and FIP—consider subgroup differences in income impacts (Panel D of Table A.6) and welfare payments (Panel D of TableA.7), but only WRP reports subgroup differences in food stamp payments (Panel D of Table A.8).

WRP, which had no effect on income overall, shows little differences in the program impacts on income for the two subgroups considered. Only the least disadvantaged group, based on a composite measure, shows a significant positive impact on income. Subgroup differences are more pronounced for FIP. In three of the four contrasts, FIP finds larger income gains for the most disadvantaged groups, defined by earnings history and the age of the youngest child.

Differences in welfare and food stamp payment impacts are not as pronounced and the pattern of impacts by the level of disadvantage is less clear. FIP and WRP do not report the significance of between-group differences, and both positive and negative impacts on welfare payments are recorded. WRP shows no differences in food stamp payment impacts for groups, defined by a composite measure of disadvantage or recipiency status.

A.4.5. Programs That Focus on TANF-Like Bundles of Reforms

Only FTP and Jobs First report differences for subgroups in the impacts on income, welfare payments, and food stamp payments (Panel F of Tables A.6, A.7, and A.8). In FTP, income differences are larger and statistically significant for those least at risk of welfare dependency (a composite measure). However, the differences among the groups defined by dependency risk are not statistically significant. Jobs First in Connecticut, in contrast, appears to raise income most for the most disadvantaged in the first two years of the program or the first six quarters by most of the measures of disadvantage recorded in Table A.6. The statistical significance of the differences among groups is not reported for Jobs First.

In terms of welfare payments and food stamp payments, FTP generally shows the largest negative transfer payment impacts for those most at risk, defined by age of the youngest child, risk of welfare dependency, and a composite measure of disadvantage. The pattern is different only when subgroups are defined by the risk of welfare dependency (food stamp payments only) and by the length of the time limit (for both welfare and food stamp payments), which is longest for those with the most barriers to work. In the case of groups defined by the length of the time limit, the larger negative welfare payment and food stamp payment impacts for those with the 24-month time limit may be a behavioral effect of having a shorter time limit.

For the various measures of disadvantage, the Jobs First results tend to show that those with less disadvantage experience the largest increase in welfare or food stamp payments in the first two years or first six quarters, and the smallest decline in payments in the third year or the eighth quarter. One exception to this pattern is when groups are defined by the age of the youngest child. For many of the measures, the differences in the impact estimates among the groups are quite small, and since the statistical significance of the between-groups differences are not reported, it is difficult to firmly establish a pattern.

A.5. OTHER MEASURES OF WELL-BEING

While many of the demonstration studies consider impact estimates for separate population subgroups, the other measures of well-being are generally not included in these analyses. There are three exceptions to this generalization, but even these studies offer little basis for drawing broader conclusions about subgroup differences in other measures of well-being as a result of specific reform policies or policies as a bundle.

For example, the two-year NEWWS follow-up discusses some subgroup differences for health insurance coverage. Overall, there are some subgroup differences by disadvantage status, as measured by a composite measure and by the level of education; the number of work barriers; and recent work experience. However, no clear patterns by disadvantage status emerge overall, or between the different program models (e.g., LFA versus HCD).

MFIP includes an analysis of a subset of the other well-being measures available for MFIP urban single-parent families for a sample of urban and rural two-parent recipient families. (Comparable results are not provided for two-parent recent applicants.) Again, however, there are no clear patterns from these analyses. For example, the impact estimate for perceptions of financial strain is negative and is the same magnitude as the effect for single parents, but it is not statistically significant. The effect on the index of material hardship is also small, negative, and statistically insignificant. For two-parent families, MFIP is estimated to significantly increase the likelihood of currently having health insurance coverage (impact estimate of 12.4 percentage points) but the effect on continuous health insurance coverage over the three-year follow-up is less than one-half as large and statistically insignificant. This is the opposite of the pattern found for single-parent families, where the larger and statistically significant effect was found for the measure of continuous coverage.

Finally, FTP analyzes differences in a subset of other well-being measures by the risk of long-term welfare dependency. Although FTP's favorable income impacts were concentrated among the group least at risk of long-term dependency, there were no systematic differences in measures of material hardship for the same subgroups. There was also no clear pattern of differences for groups defined by employment barriers.

A.6. CHILD WELL-BEING

As part of the review of the experimental studies conducted in Chapter 10, we have already seen that the impacts of various component policies of welfare reform can vary by the age of the child. MFIP also demonstrates that there can be differences in program impacts for long-term welfare recipients compared with recent applicants. In this section, we review the evidence that impacts vary with other characteristics of the child or family. Table A.9 summarizes the

Table A.9—Estimated Impact of Welfare Reform on Child Behavior, Schooling, Health, and Other Outcomes for Subgroups: Random Assignment Studies

Name	Child Outcome Domains	Subgroups Analyzed Defined By	Result
A. Programs that focus on financial work incentives			
B. Programs that focus on financial work incentives tied to hours of work			
New Hope	Behavior problems School performance	• Child sex	• Favorable impacts on educational progress larger for boys • Favorable teacher and parent ratings of positive social behavior and teacher ratings of problem behavior larger for boys
		• Parent's employment experience pre-RA	• No differences
SSP	Cognitive/academic functioning Behavior and emotional well-being Health and safety	• Child sex	• Youngest cohort: No differences • Middle cohort: Impacts for girls more pronounced but not statistically different from boys • Oldest cohort: Impacts for girls somewhat larger and same direction as boys but not statistically different; exception is risk of depression impact for girls is unfavorable and boys it is favorable
		• Parent marital status	• No differences
		• Parent age	• No differences
		• Parent education	• No differences
		• Parent disability	• No differences
		• Family size	• No differences
		• Parent depression risk	• No differences
C. Programs that focus on mandatory work-related activities			
LA Jobs-1st GAIN	Academic functioning and schooling Behavioral and emotional adjustment Safety	• Child sex	• No differences
NEWWS Sites (COS only) Grand Rapids LFA Riverside LFA Atlanta LFA Grand Rapids HCD Riverside HCD Atlanta HCD	Behavioral adjustment School progress Health and safety	• By risk for poor development, defined separately or cumulatively by sibling structure, low maternal education, work barriers, maternal mental health	• Few impacts found within subgroups • Children at higher risk had small impacts; favorable in 2 of 3 sites for education-focused programs and unfavorable for employment-focused programs • Children at lower risk had larger impacts that tended to be unfavorable and not vary by program approach

Table A.9—Continued

Name	Child Outcome Domains	Subgroups Analyzed Defined By	Result
D. Programs that focus on financial work incentives and mandatory work-related activities			
MFIP - Recipients	Behavioral adjustment School performance	• Child sex	• Positive impacts for girls more pronounced but not statistically significantly differ from boys
		• Race	• No differences
		• AFDC recipiency history	• More favorable effects for recipients with 5+ years on aid but not statistically significantly different from short-term reci[illegible]
		• Maternal earnings history	• No differences
		• Maternal educational attainment	• No differences
		• Parents' potential barriers to work	• No differences
MFIP - Applicants	Behavioral adjustment School performance	• Child sex	• Boys had less favorable impact on school engagement
		• Race	• White children had less favorable impacts for several school performance measures
		• Maternal employment experience	• Those with recent employment experience had less favorable impacts on several school performance measures
		• Maternal educational attainment	• Those with a high school diploma had less favorable [illegible]
E. Programs that focus on other individual reforms			
F. Programs that focus on TANF-like bundle of reforms (time limits with financial incentives, work-related activities, or both)			
ABC	Maltreatment Foster care	• Years on welfare	• Neglect less favorable in 2 of 3 years for LT recipients
		• Years of education	• Neglect less favorable in 2 of 3 years for less than h.s. educ
		• Previous history of maltreatment	• Abuse and neglect less favorable in 3 of 3 years for those with prior history
		• Race	• Neglect less favorable in 2 of 3 years for nonwhites
		• Age of youngest child	• Mixed results in at most one year
		• Age of adult head	• No clear pattern of differences
FTP	School outcomes Behavior	• Risk of long-term dependence	• Less favorable outcomes on achievement and suspensions/ expulsions for those at lowest risk of LT dependence
Jobs First	School achievement Behavior problems	• Race/ethnicity	• Benefits stronger for whites but no significant group differences
		• AFDC recipiency history	• No differences
		• Level of disadvantage	• No differences

NOTES: For full program names and citations, see Table 3.4.

results for different subgroups for the experimental studies that conducted more disaggregated analyses.

Four studies consider how impacts differ for girls versus boys. While L.A. Jobs-First GAIN showed no differences by the child's gender, MFIP and SSP suggest girls in some age groups gain more than boys while the reverse is true for New Hope. SSP demonstrated impacts for girls on a range of indicators that exceed those for boys, more so in the oldest age cohort and somewhat less so in the middle age cohort. There were no differences for the youngest age group. MFIP found more pronounced effects for girls on some outcomes, particularly for recipients. In the case of New Hope, the favorable impacts on both educational progress and ratings of behavior are stronger for boys. Girls also have favorable impacts, but they are smaller than for the boys. Taken together, these studies suggest that there may well be differences between boys and girls in the impact of welfare reform, but the differences are not always consistent and may depend upon the pathways by which specific policies affect the family outcomes that determine child well-being.

Other child characteristics used to define subgroups include race and measures of developmental risk. Of the two studies that considered racial differences, MFIP found no differences for recipients, and less favorable impacts for white applicants on several school performance measures compared with blacks and a residual other ethnic group. ABC's assessment of maltreatment found nonwhites with less favorable impacts in two of three years compared with whites. The NEWWS evaluation of the COS sample available for six sites indicates that various indicators of child developmental risk may lead to differential impacts, but there was no clear pattern of variation with the orientation of the program.

Differences by family background characteristics have also been examined. SSP and ABC both find no differences by the age of the parent. ABC and SSP also found no differences by other parental background characteristics (specifically, age of the youngest child for ABC, and, for SSP, marital status, family size, disability status, and depression risk). In the case of abuse and neglect, a prior history of such behavior is associated with less favorable impacts in the ABC evaluation.

Three of the studies consider differences by measures of welfare dependency, either prior history or future risk. Long-term MFIP recipients showed more favorable impacts, but they were not statistically different from the shorter-term recipients. ABC found less favorable impacts for child neglect for long-term recipients (more than four years out of last five years), while FTP's impacts on achievement and suspensions were less favorable for those at lowest risk of long-term dependency.

Finally, two studies assess differences by prior work history of the mother or parents. Among MFIP recipients, there were no differences in child outcomes based on maternal earnings history or her potential barriers to work. The children of MFIP Applicants with recent employment experience had less favorable impacts for schooling outcomes. New Hope, despite large differences in impacts by work history, found no differences in child outcomes for those employed at random assignment versus those not employed.

Appendix B

METHODOLOGY FOR CHAPTER ELEVEN SYNTHESIS

This appendix provides additional detail regarding the synthesis of the random assignment and econometric studies presented in Chapter 11. In that chapter, we synthesize the results of the studies we reviewed in Chapters 4 to 10 according to the major policy or groups of policies they evaluate, as well as for reform as a bundle. The assignment of random assignment studies and econometric studies to the 11 policy rows shown in Table 11.1 is recorded in Table B.1.

The summary entries in each cell of Table 11.1 convey both the qualitative effect of the particular policy on the particular outcome and the depth of the knowledge base on which the entry is based. Both were determined by the signs and the significance of the underlying impact estimates reviewed in Chapters 4 to 10 and the number and quality of the studies that produced them. The procedure involved first assigning a direction (or sign) and significance indicator to each study, both experimental and econometric, that addresses each policy-outcome pair; tabulating the estimates by sign, significance, and quality of the underlying study; and then using that tabulation to assign an overall direction and knowledge-base indicator to each cell.

With a very small number of exceptions, which we discuss below, we first assigned a single direction and significance level to each study that addressed a particular policy-outcome pair. This was to avoid double-counting the studies for which the tables present multiple estimates, which would arbitrarily give them more weight than the others. If all the estimates for a study agreed in direction, then that direction was assigned to that study. Most of the studies fell into this category. If any of the estimates were significant, then the study was coded as having a significant effect. Among studies with mixed estimates, three approaches were taken. If the only significant estimates all had the same sign, then the study was assigned that direction and coded as significant. If none of the estimates were significant, then the direction for that study was assigned based on the majority of the estimates and coded as insignificant. If there was no majority, then no direction was assigned. If the study had significant estimates that were both positive and negative, then the study contributed two significant entries, one in each direction, to the second-stage tabulation.

Once each study had been scored, the results were tabulated for all studies within a policy-outcome cell using the assignments shown in Table B.1. Studies were categorized according to direction (positive or negative), significance (significant or not), and quality (high or moderate). The summary direction for the cell and the indicator of the depth of the knowledge base for the cell were based on this tabulation.

For the most part, the direction for the cell was easy to assess, since most studies within a cell tended to have the same direction. However, there were some cases where the underlying studies yielded results that varied enough that it was difficult to reasonably classify the direction of the impact. For example, in the case of Medicaid use (column (E)) for policy row (5) (mandatory work-related activities and strong financial work incentives), MFIP leads to a statistically significant positive impact, while FIP finds a statistically significant negative impact. In this case and others like it, the cells in Table 11.1 are labeled as “mixed.” In two cases, we label a cell as “no change.” These occur in cells F3 and G3 for the impact of mandatory work-related activities on marriage and fertility, where a dozen or so high-quality studies all find a statistically insignificant impact, thus providing a relatively high level of confidence that the null hypothesis of no impact is true.

Assessing the depth of the knowledge base within each cell is more subjective and necessarily involves some judgment. The extremes are reasonably clear. Some cells are populated by a substantial number of high-quality studies. When nearly all their results point in the same direction, and nearly all are significant, that clearly provides a deep knowledge base about the effect of that policy on those outcomes. In other cases, there is only a single high-quality study that addresses a particular policy-outcome combination and provides a statistically significant impact. That cell has a more shallow knowledge base. Other cells are empty or nearly empty because the few studies in that cell provide no statistically significant results.

Although some of the cases in between are hard to classify, we adopt a four-level scale to relate how much is known about the effect of a particular policy on a particular outcome. The criteria for the levels and their representation on Table 11.1 are as follows:

- No evidence (blank cells): No studies at all.
- Little evidence (unshaded cells): A single high-quality study that yielded a significant result or two moderate-quality studies that yielded significant results of the same sign. Similar combinations also fall into this category, such as two high-quality studies that yield results of the same sign, only one of which is significant. We also use this shading to denote cases when results are mixed because there are two high-quality studies with statistically significant impacts of opposite sign or there are three or more high-quality studies with mixed signs but only one is statistically significant.
- Some evidence (lightly shaded cells): All nonempty (or nearly empty) cells that fall neither in the “little” or “much” (defined below) categories reflect cells about which we know “some.”
- Much evidence (darkly shaded cells): At least 4 high-quality studies that yield significant results of the same sign.[113] This shading is also used in two cases signed as “no change” where 12 high-quality studies are nearly evenly split between positive and negative insignificant impacts with magnitudes close to zero, and in one other case signed as “mixed” where 13 high-quality studies are nearly evenly

[113]The number four is arbitrary, but we note there were only four studies of female-family heads from the Negative Income Tax experiments, the results from which have been widely regarded as conclusive (Burtless, 1986).

split between positive and negative impacts and only one impact of each sign is significant.

In addition to the blank cells in the first category above, we also leave some cells unsigned even though there may be one or more studies in the cell. These cases, denoted by an asterisk, occur when there are up to three moderate and/or high-quality studies with no significant impacts or a single moderate-quality study with a significant impact. We felt these cells provided too little evidence to assign a direction of impact with even a minimal level of confidence.

Table B.1—Assignment of Studies for Synthesis

Policy or Policy Bundle	Random Assignment Studies	Econometric Studies	
1. Financial Work Incentives	CWPDP MFIP-IO WRP-IO	CEA (1997) CEA (1999) Hofferth, Stanhope and Harris (2000a) Hofferth, Stanhope and Harris (2000b)	Horvath and Peters (1999) Meyer and Rosenbaum (2001) Moffitt (1999) Ziliak et al. (2000)
2. Financial Work Incentives Tied to Hours Worked	New Hope SSP SSP-Plus SSP-A		
3. Mandatory Work-Related Activities	L.A. Jobs-First GAIN NEWWS Programs: Atlanta LFA Grand Rapids LFA Riverside LFA Portland Atlanta HCD Grand Rapids HCD Riverside HCD Columbus Integrated Columbus Traditional Detroit Oklahoma City IMPACT-Basic Track	CEA (1997) CEA (1999 Hofferth, Stanhope and Harris (2000a) Hofferth, Stanhope and Harris (2000b) Horvath and Peters (1999)	MaCurdy, Mancuso, and O'Brien-Strain (2000) Moffitt (1999) Paxson and Waldfogel (2001) Rector and Youssef (1999) Ziliak et al. (2000)
4. Sanctions for Noncompliance		CEA (1997) CEA (1999) Hofferth, Stanhope and Harris (2000a) Hofferth, Stanhope and Harris (2000b) Levine and Whitmore (1998)	MaCurdy, Mancuso, and O'Brien-Strain (2000) Mead (2001) Moffitt (1999) Paxson and Waldfogel (2001) Rector and Youssef (1999)

Table B.1—Continued

Policy or Policy Bundle	Random Assignment Studies	Econometric Studies	
5. Mandatory Work-Related Activities and Strong Financial Work Incentives	MFIP FIP		
6. Mandatory Work-Related Activities and Weak Financial Work Incentives	WRP TSMF		
7. Time Limits (Before Recipients Reach Limit)		CEA (1997) CEA (1999) Grogger (2000) Grogger (2002) Grogger (forthcoming) Grogger and Michalopoulos (forthcoming) Hofferth, Stanhope and Harris (2000a) Hofferth, Stanhope and Harris (2000b) Horvath and Peters (1999)	Kaushal and Kaestner (2001) Kearny (2001) MaCurdy, Mancuso, and O'Brien-Strain (2000) Meyer and Rosenbaum (2001) Moffitt (1999) Paxson and Waldfogel (2001) Ziliak et al. (2000)
8. Time Limits (After Recipients Reach Limit)	FTP Jobs First	Meyer and Rosenbaum (2001)	
9. Family Cap	AAWDP FDP	CEA (1997) CEA (1999) Hofferth, Stanhope and Harris (2000a) Hofferth, Stanhope and Harris (2000b) Horvath and Peters (1999) Kaushal and Kaestner (2001)	Kearny (2001) Levine (2001) Moffitt (1999) Paxson and Waldfogel (2001) Ziliak et al. (2000)
10. Parental Responsibility	PPI PIP	Horvath and Peters (1999)	

Table B.1—Continued

Policy or Policy Bundle	Random Assignment Studies	Econometric Studies	
11. Reform as a Bundle (Before Recipients Reach Time Limits)	EMPOWER	Bartik and Eberts (1999)	Kaushal and Kaestner (2001)
	IMPACT	Bitler, Gelbach and Hoynes (2001)	Kearny (2001)
	VIP/VIEW	Blank (2000)	Ku and Garrett (2000)
	ABC	CEA (1997)	Levine (2001)
	FTP	CEA (1999)	Levine and Whitmore (1998)
	Jobs First	Currie and Grogger (2001)	Moffitt (1999)
		Figlio, Gunderson and Ziliak (2000)	Mueser et al. (2000)
		Figlio and Ziliak (1999)	O'Neill and Hill (2001)
		Grogger (2000)	Paxson and Waldfogel (2001)
		Grogger (forthcoming)	Schoeni and Blank (2000)
		Horvath and Peters (1999)	Wallace and Blank (1999)
		Huang et al. (2000)	Wilde et al. (2000)

NOTES: The eleven rows for individual reform policies or groups of policies correspond to the rows in Table 11.1. For full names of and citations for the random assignment programs, see Table 3.4.

BIBLIOGRAPHY

Acs, Gregory, and Pamela Loprest, *Initial Synthesis Report of the Findings from ASPE's Leavers Grants*: *Report to the US Department of Human Services*, Washington, DC: Urban Institute, December 2000.

Acs, Gregory, and Sandi Nelson, "'Honey I'm Home.' Changes in Living Arrangements in the Late 1990s," Series B, No. B-38, Washington, DC: Urban Institute, June 2001.

Acs, Gregory, et al., "Comings and Goings: The Changing Dynamics of Welfare in the 1990s," unpublished manuscript, Washington, DC: Urban Institute, 2001.

Adams, John, and V. Joseph Hotz, "The Statistical Power of National Data to Evaluate Welfare Reform," in Robert A. Moffitt and Michele Ver Ploeg, eds., *Evaluating Welfare Reform in an Era of Transition*, Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, Washington, DC: National Academy Press, Committee on National Statistics, Committee on Behavioral and Social Sciences and Education, and the National Research Council, 2001.

Ahn, Jay, Shon Kraley, Debra Fogarty, Faith Lai, and Laurie Deppman, *A Study of Washington State's TANF Leavers and TANF Recipients: Welfare Reform and Findings from Administrative Data. Final Report*, Washington Department of Social and Health Services, Offices of Planning and Research, Economic Service Administration, February 2000.

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

Bartik, Timothy J., and Randall W. Eberts, "Examining the Effect of Industry Trends and Structure on Welfare Caseloads," in Sheldon H. Danziger, ed., *Economic Conditions and Welfare Reform,* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1999, pp. 119-157.

Becerra, Rosina M., Vivian Lew, Michael N. Mitchell, and Hiromi Ono, *California Work Pays Demonstration Project: Report of First Forty-Two Months (Final Report)*, Los Angeles, CA:

School of Public Policy and Social Research, University of California, Los Angeles, October 1998.

Becker, Gary, "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, Vol. 76, No. 2, March-April 1968, pp. 169-217.

Bell, Stephen H., "Why Are Welfare Caseloads Falling?" Washington, DC: Urban Institute, Working Paper #01-02, March 2001.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes, "The Impact of Welfare Reform on Living Arrangements," unpublished manuscript, July 2001.

Blank, Rebecca M., "Analyzing the Length of Welfare Spells," *Journal of Public Economics*, Vol. 39, No. 3, August 1989, pp. 245-273.

Blank, Rebecca M., "What Causes Public Assistance Caseloads to Grow?" Chicago, IL: Joint Center for Poverty Research, Working Paper #18, June 2000.

Blank, Rebecca M., "Evaluating Welfare Reform in the United States," *Journal of Economic Literature*, forthcoming.

Blank, Rebecca M., David Card, and Philip K. Robins, "Financial Incentives for Increasing Work and Income Among Low-Income Families," in David E. Card and Rebecca M. Blank, eds., *Finding Jobs: Work and Welfare Reform*, New York, NY: Russell Sage Foundation, 2000, pp. 373-419.

Blank, Rebecca M., and Patricia Ruggles, "When Do Women Use AFDC and Food Stamps? The Dynamics of Eligibility versus Participation," *Journal of Human Resources*, Vol. 31, No. 1, Winter 1996, pp. 57-89.

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

Bloom, Dan, Richard Hendra, and Charles Michalopoulos, *Vermont's Welfare Restructuring Project: Key Findings from the Forty-Two-Month Client Survey*, New York, NY: Manpower Demonstration Research Corporation, June 2000.

Bloom, Dan, James J. Kemple, Pamela Morris, Susan Scrivener, Nandita Verma, and Richard Hendra, *The Family Transition Program: Final Report on Florida's Initial Time-Limited Welfare Program*, New York, NY: Manpower Demonstration Research Corporation, December 2000a.

Bloom, Dan, Laura Melton, Charles Michalopoulos, Susan Scrivener, and Johanna Walter, *Jobs First: Implementation and Early Impacts of Connecticut's Welfare Reform Initiative*, New York, NY: Manpower Demonstration Research Corporation, February 2000b.

Bloom, Dan, and Charles Michalopoulos, *How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research*, New York, NY: Manpower Demonstration Research Corporation, January 2001.

Bloom, Dan, Charles Michalopoulos, Johanna Walter, and Patricia Auspos, *Implementation and Early Impacts of Vermont's Welfare Restructuring Project*, New York, NY: Manpower Demonstration Research Corporation, October 1998.

Bloom, Dan, Susan Scrivener, Charles Michalopoulos, Pamela Morris, Richard Hendra, Diana Adams-Ciardullo, Johanna Walter, with Wanda Vargas, *Jobs First: Final Report on Connecticut's Welfare Reform Initiative*, New York, NY: Manpower Demonstration Research Corporation, February 2002.

Borjas, George J., "Food Insecurity and Public Assistance," Chicago, IL: Joint Center for Poverty Research, Working Paper No. 243, May 2001a.

Borjas, George J., "Welfare Reform and Immigration," in Rebecca M. Blank and Ron Haskins, eds., *The New World of Welfare*, Washington, DC: Brookings Institution Press, 2001b, pp. 369-390.

Bos, Johannes M., Aletha C. Huston, Robert C. Granger, Greg J. Duncan, Thomas W. Brock, and Vonnie C. McLoyd, *New Hope for People with Low Incomes: Two-Year Results of a Program to Reduce Poverty and Reform Welfare*, New York, NY: Manpower Demonstration Research Corporation, August 1999.

Bos, Johannes M., and Wanda G. Vargas, "Maternal Employment and Changes in Adolescent Outcomes: Evidence from Two Evaluations of Programs to Promote Work," draft paper prepared for the Biannual Research Conference of the Society for Research on Child Development, April 2001.

Burke, Vee, and Melinda Gish, *Welfare Reform: Work Trigger Time Limits, Exemptions, and Sanctions under TANF*, Washington, DC: Congressional Research Service, 98-697 EPW, August 6, 1998.

Burtless, Gary, "The Work Response to a Guaranteed Income: A Survey of Experimental Evidence," in Alicia Munnell, ed., *Lessons from the Income Maintenance Experiments*, Boston, MA: Federal Reserve Bank of Boston, 1986.

Burtless, Gary, "The Case for Randomized Field Trails in Economic and Policy Research," *Journal of Economic Perspectives*, Vol. 92, No. 2, Spring 1995, pp. 63-84.

Camasso, Michael J., Carol Harvey, and Radha Jagannathan, *An Interim Report on the Impact of New Jersey's Family Development Program*, New Brunswick, NJ: Public Affairs, 1996.

Camasso, Michael J., Carol Harvey, Radha Jagannathan, and Mark Killingsworth, *New Jersey's Family Development Program: Results on Program Impacts, Experimental Control Group Analysis*, Trenton, NJ: New Jersey Department of Family Services, October 1998.

Camasso, Michael, Carol Harvey, Mark Killingsworth, and Radha Jagannathan, "New Jersey's Family Cap and Family Size Decisions: Some Findings from a 5-Year Evaluation," unpublished manuscript, April 1999.

Cancian, Maria, Robert Haveman, Thomas Kaplan, and Barbara Wolfe, *Post-Exit Earnings and Benefit Receipt Among Those Who Left AFDC in Wisconsin*, Madison, WI: Institute for Research on Poverty, January 1999a.

Cancian, Maria, Robert Haveman, Thomas Kaplan, Daniel Meyer, and Barbara Wolfe, "Work, Earnings, and Well-Being After Welfare: What Do We Know?" in Sheldon H. Danziger, ed., *Economic Conditions and Welfare Reform,* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1999b, pp. 161-186.

Cancian, Maria, Robert Haveman, Daniel R. Meyer, Barbara Wolfe, *Before and After TANF: The Economic Well-Being of Women Leaving Welfare*, Madison, WI: Institute for Research on Poverty, Special Report #77, May 2000.

Center on Budget and Policy Priorities (CBPP), "State Time Limits," Unpublished data, 2001.

Citro, Constance F., and Robert T. Michael, eds., *Measuring Poverty: A New Approach*, Washington, DC: National Academy of Sciences Press, 1995.

Committee on Ways and Means, *Green Book*, Washington, DC, 2000.

Congressional Research Service, *Medicaid Source Book: Background and Analysis*, Washington, DC: U.S. Government Printing Office, 1993.

Coulton, Claudia, and Nandita Verma, *Employment and Return to Public Assistance Among Single, Female Headed Families Leaving AFDC in Third Quarter, 1996*, Washington, DC: Cuyahoga Work and Training, Manpower Demonstration Research Corporation, Case Western Reserve University, September 2000.

Council of Economic Advisers (CEA), *Explaining the Decline in Welfare Receipt, 1993-1996*, Washington, DC, May 1997.

Council of Economic Advisers (CEA), *The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update*, Washington, DC, August 1999.

Currie, Janet, and Jeffrey Grogger, "Explaining Recent Declines in Food Stamp Program Participation," *Brookings-Wharton Papers on Urban Affairs*, Vol. 2, 2001, pp. 203-229.

DeMarzo, Peter M., Michael J. Fishman, and Kathleen M. Hagerty, "The Optimal Enforcement of Insider Trading Regulations," *Journal of Political Economy*, Vol. 106, No. 3, June 1998, pp. 602-632.

Dickens, William T., Lawrence F. Katz, Kevin Lang, and Lawrence H. Summers, "Employee Crime and the Monitoring Puzzle," *Journal of Labor Economics*, Vol. 7, No. 3, July 1989, pp. 331-347.

Dion, M. Robin, and LaDonna Pavetti, *Access to and Participation in Medicaid and the Food Stamp Program: A Review of the Recent Literature*, Princeton, NJ: Mathematica Policy Research, Inc., March 2000.

Du, Jean, Debra Fogarty, Devin Hopps, and James Hu, *A Study of Washington State's TANF Leavers and TANF Recipients: Findings from the April-June 1999 Telephone Survey. Final Report*, Washington Department of Social and Health Services, Offices of Planning and Research, Economic Service Administration, February 2000.

Duncan, Greg J., and Jeanne Brooks-Gunn, eds., *Consequences of Growing Up Poor*, New York, NY: Russell Sage Foundation, 1997.

Duncan, Greg J., and P. Lindsay Chase-Lansdale, "Welfare Reform and Children's Well-Being," in Rebecca M. Blank and Ron Haskins, eds., *The New World of Welfare*, Washington, DC: Brookings Institution Press, 2001, pp. 391-417.

Dupree, Allen, and Wendell Primus, *Declining Share of Children Lived With Single Mothers in the Late 1990s*, Washington, DC: Center on Budget and Policy Priorities, 2001.

Edin, Kathryn, and Laura Lein, *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work*, New York, NY: Russell Sage Foundation, 1997.

Ellwood, David T., "The Impact of the Earned Income Tax Credit and Social Policy Reforms on Work, Marriage, and Living Arrangements," *National Tax Journal*, Vol. 53, No. 4, Part 2, December 2000, pp. 1063-1106.

Fein, David J., *Will Welfare Reform Influence Marriage and Fertility? Early Evidence from the ABC Demonstration*, Cambridge, MA: Abt Associates, Inc., June 1999.

Fein, David J., Erik Beecroft, William L. Hamilton, Wang S. Lee, Pamela A. Holcomb, Terri S. Thompson, and Caroline E. Ratcliffe, *The Indiana Welfare Reform Evaluation: Program Implementation and Economic Impacts After Two Years*, Cambridge, MA: Abt Associates, Inc., November 1998.

Fein, David J., and Jennifer Karweit, *The ABC Evaluation. The Early Economic Impacts of Delaware's A Better Chance Welfare Reform Program*, Cambridge, MA: Abt Associates, Inc., December 1997.

Fein, David J., and Wang S. Lee, *The ABC Evaluation. Impacts of Welfare Reform on Child Maltreatment*, Cambridge, MA: Abt Associates, Inc., December 2000.

Fein, David J., Wang S. Lee, and E. Christina Schofield, *The ABC Evaluation: Do Welfare Recipients' Children Have a School Attendance Problem?* Cambridge, MA: Abt Associates, Inc., August 1999.

Fein, David J., David A. Long, Joy M. Behrens, and Wang S. Lee, *The ABC Evaluation. Turning the Corner: Delaware's A Better Chance Welfare Reform Program at Four Years*, Cambridge, MA: Abt Associates, Inc., January 2001.

Figlio, David N., Craig Gunderson, and James P. Ziliak, "The Effects of the Macroeconomy and Welfare Reform on Food Stamp Caseloads," *American Journal of Agricultural Economics*, Vol. 82, No. 3, November 2000, pp. 635-641.

Figlio, David N., and James P. Ziliak, "Welfare Reform, The Business Cycle, and the Decline in AFDC Caseloads," in Sheldon H. Danziger, ed., *Economic Conditions and Welfare Reform*, Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1999, pp. 17-48.

Fogarty, Debra, and Shon Kraley, *A Study of Washington State's TANF Leavers and TANF Recipients: Findings from Administrative Data and the Telephone Survey. Summary Report*, Washington Department of Social and Health Services, Offices of Planning and Research, Economic Service Administration, March 2000.

Foster, E. M., *Amended Quarterly Progress Report: Outcomes for Single Parent Leavers by Cohort Quarter*, Georgia State University, April 1999.

Fraker, Thomas M., and Jonathan E. Jacobson, *Iowa's Family Investment Program: Impacts During the First 3 1/2 Years of Welfare Reform*, Washington, DC: Mathematica Policy Research, Inc., May 2000.

Freedman, Stephen, Daniel Friedlander, Gayle Hamilton, JoAnn Rock, Marisa Mitchell, Jodi Nudelman, Amanda Schweder, and Laura Storto, *National Evaluation of Welfare-to-Work Strategies: Evaluating Alternative Welfare-to-Work Approaches: Two-Year Impacts of Eleven*

Programs, New York, NY: Manpower Demonstration Research Corporation, U.S. Department of Health and Human Services, and U.S. Department of Education, June 2000a.

Freedman, Stephen, Jean Tansey Knab, Lisa A. Gennetian, and David Navarro, *The Los Angeles Jobs-First GAIN Evaluation: Final Report on a Work First Program in a Major Urban Center*, New York, NY: Manpower Demonstration Research Corporation, June 2000b.

Gais, Thomas L., Richard P. Nathan, Irene Lurie, and Thomas Kaplan, *The Implementation of the Personal Responsibility Act of 1996: Commonalities, Variations, and the Challenge of Complexity*, in Rebecca M. Blank and Ron Haskins, eds., *The New World of Welfare*, Washington, DC: Brookings Institution Press, 2001, pp. 35-69.

Garfinkel, Irwin, and Sara McLanahan, *Single Mothers and Their Children: A New American Dilemma,* Changing Domestic Priorities Series, Washington, DC: Urban Institute Press, 1986.

Garrett, Bowen, and John Holahan, *Welfare Leavers, Medicaid Coverage, and Private Health Insurance*, Urban Institute, B-13, March 2000.

General Accounting Office (GAO), *Food Stamp Program: Various Factors Have Led to Declining Participation,* Washington, DC: U.S. General Accounting Office, GAO/RCED-99-185, July 1999a.

General Accounting Office (GAO), *Medicaid Enrollment: Amid Declines, State Efforts to Ensure Coverage After Welfare Reform Vary,* Washington, DC: U.S. General Accounting Office, GAO/HEHS-99-163, September 1999b.

General Accounting Office (GAO), *Welfare Reform: Information on Former Recipients' Status,* Washington, DC: U.S. General Accounting Office, GAO/HEHS-99-48, 1999c.

General Accounting Office (GAO), *Welfare Reform: State Sanction Policies and Number of Families Affected*, Washington, DC: U.S. General Accounting Office, GAO/HEHS-00-44, March 2000.

General Accounting Office (GAO), *More Research Needed on TANF Family Caps and Other Policies for Reducing Out-of-Wedlock Births*, Washington, DC: U.S. General Accounting Office, GAO-01-924, 2001.

Gennetian, Lisa A., and Cynthia Miller, *Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program (Vol. 2: Effects on Children)*, New York, NY: Manpower Demonstration Research Corporation, September 2000.

Gladden, Tricia, and Christopher Taber, "Wage Progression Among Less-Skilled Workers," in David E. Card and Rebecca M. Blank, eds., *Finding Jobs: Work and Welfare Reform,* New York, NY: Russell Sage Foundation, 2000, pp. 160-192.

Goldberg, Heidi, and Liz Schott, *A Compliance-Oriented Approach to Sanctions in State and County TANF Programs*, Washington, DC: Center on Budget and Policy Priorities, October 1, 2000.

Gordon, Anne, and Roberto Agodini, *Early Impacts of the Virginia Independence Program: Final Report*, Princeton, NJ: Mathematica Policy Research, Inc., November 1999.

Grogger, Jeffrey, "Time Limits and Welfare Use," Cambridge, MA: National Bureau of Economic Research, Working Paper #7709, May 2000.

Grogger, Jeffrey, "The Behavioral Effects of Welfare Time Limits," *American Economic Review*, Vol. 92, No. 2, May 2002.

Grogger, Jeffrey, "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income Among Female-Headed Families," *Review of Economics and Statistics*, forthcoming.

Grogger, Jeffrey, and Charles Michalopoulos, "Welfare Dynamics under Time Limits," Cambridge, MA: National Bureau of Economic Research Working Paper #7353, September 1999.

Grogger, Jeffrey, and Charles Michalopoulos, "Welfare Dynamics under Time Limits," *Journal of Political Economy*, forthcoming.

Gruber, Jonathan, "Medicaid," Cambridge, MA: National Bureau of Economic Research, Working Paper #7829, August 2000.

Guyer, Jocelyn, *Health Care After Welfare: An Update of Findings from State-Level Leaver Studies*, Washington, DC: Center on Budget and Policy Priorities, August 2000.

Haider, Steven, Alison Jacknowitz, and Robert F. Schoeni, "The Impact of Welfare Work Requirements on Breastfeeding," unpublished manuscript, Santa Monica, CA: RAND, January 2002.

Haider, Steven, Jacob Alex Klerman, and Elizabeth Roth, "The Relationship Between the Economy and the Welfare Caseload: A Dynamic Approach," unpublished manuscript, Santa Monica, CA: RAND, April 2001.

Haider, Steven, Robert F. Schoeni, Yuhua Bao, and Caroline Danielson, Immigrants, "Welfare Reform, and the Economy in the 1990s," Santa Monica, CA: RAND, DRU-2681, 2001.

Hamilton, Gayle, Stephen Freedman, Lisa Gennetian, Charles Michalopoulos, et al., *National Evaluation of Welfare-to-Work Strategies: How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs*, New York, NY: Manpower Demonstration Research Corporation, U.S. Department of Health and Human Services, and U.S. Department of Education, December 2001.

Hamilton, Gayle, Stephen Freedman, and Sharon M. McGroder, *Do Mandatory Welfare-to-Work Programs Affect the Well-Being of Children? A Synthesis of Child Research Conducted as Part of the National Evaluation of Welfare-to-Work Strategies*, New York, NY: Manpower Demonstration Research Corporation, U.S. Department of Health and Human Services, and U.S. Department of Education, June 2000.

Harvey, Carol, Michael J. Camasso, and Radha Jagannathan, "Evaluating Welfare Reform Waivers Under Section 1115," *Journal of Economic Perspectives*, Vol. 14, No. 4, Fall 2000.

Haskins, Ron, "Effects of Welfare Reform on Family Income and Poverty," in Rebecca M. Blank and Ron Haskins, eds., *The New World of Welfare*, Washington, DC: Brookings Institution Press, 2001, pp. 103-136.

Haveman, Robert, and Barbara Wolfe, "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, Vol. XXXIII, December 1995, pp. 1829-1878.

Health Care Financing Administration (HCFA), *A Profile of Medicaid: Chart Book 2000*, Washington, DC: Department of Health and Human Services, 2000.

Heckman, James J., and Jeffrey A. Smith, "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, Vol. 92, No. 2, Spring 1995, pp. 85-110.

Heckman, James J., Jeffrey A. Smith, and Christopher Taber, "Accounting for Drop-Outs in Evaluations of Social Programs," *Review of Economics and Statistics*, Vol. 80, No. 1, February 1998, pp. 1-14.

Heim, Bradley T., "Does Child Support Enforcement Reduce Divorce Rates? A Reexamination," unpublished manuscript, Northwestern University, September 2001.

Hendra, Richard, and Charles Michalopoulos, *Forty-Two-Month Impacts of Vermont's Welfare Restructuring Project*, New York, NY: Manpower Demonstration Research Corporation, September 1999.

Hendra, Richard, Charles Michalopoulos, and Dan Bloom, *Three-Year Impacts of Connecticut's Jobs First Welfare Reform Initiative*, New York, NY: Manpower Demonstration Research Corporation, April 2001.

Henry, Mark, Willis Lewis, Lynn Reinschmiedt, and Darren Hudson, "Reducing Food Stamp and Welfare Caseloads in the South: Are Rural Areas Less Likely to Succeed Than Urban Centers?" Chicago, IL: Joint Center for Poverty Research, Working Paper #188, June 2000.

Hofferth, Sandra L., Stephen Stanhope, and Kathleen Mullan Harris, "Remaining Off Welfare in the 1990s: The Influence of Public Policy and Economic Conditions," Ann Arbor, MI: Institute for Social Research, October 2000a.

Hofferth, Sandra L., Stephen Stanhope, and Kathleen Mullan Harris, "Exiting Welfare in the 1990s: Did Public Policy Influence Recipients' Behavior," Ann Arbor, MI: Institute for Social Research, October 2000b.

Horvath-Rose, Ann, and H. Elizabeth Peters, "Welfare Waivers and Non-Marital Childbearing," Chicago, IL: Joint Center for Poverty Research, Working Paper #128, September 1999.

Hotz, V. Joseph, Guido W. Imbens, and Jacob A. Klerman, "The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program," Cambridge, MA: National Bureau of Economic Research, Working Paper #8007, November 2000.

Hotz, V. Joseph, Charles Mullin, and John Karl Scholz, "The Earned Income Tax Credit and Labor Market Participation of Families on Welfare," unpublished manuscript, 2001.

Hotz, V. Joseph, Charles Mullin, and John Karl Scholz, "Welfare, Employment and Income: Evidence on the Effects of Benefit Reductions from California," *American Economic Review*, Vol. 92, No. 2, May 2002.

Hu, Wei-Yin, "Marriage and Economic Incentives: Evidence from a Welfare Experiment," unpublished manuscript, University of California, Los Angeles, 2000.

Huang, Chien-Chung, Irwin Garfinkel, and Jane Waldfogel, "Child Support and Welfare Caseloads," Madison, WI: Institute for Research on Poverty, Discussion Paper 1218-00, 2000.

Hurst, Erik, and James P. Ziliak, "Welfare Reform and Household Saving," unpublished manuscript, July 2001.

Isaacs, Julia B., and Matthew R. Lyon, *A Cross-State Examination of Families Leaving Welfare: Findings from the ASPE-Funded Leavers Studies*, presented at the National Association for Welfare Research and Statistics 40th Annual Workshop, Scottsdale, AZ, August 1, 2000, Division of Data and Technical Analysis, Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, November 2000.

Julnes, George, Anthony Halter, Steven Anderson, Lee Frost-Kumpf, Richard Schuldt, Francis Staskon, and Barbara Ferrara, *Illinois Study of Former TANF Clients: Final Report*, Springfield, IL: Institute for Public Affairs, University of Illinois at Springfield, School of Social Work, University of Illinois at Urbana-Champaign, August 2000.

Kaushal, Neeray, and Robert Kaestner, "From Welfare to Work: Has Welfare Reform Worked?" *Journal of Policy Analysis and Management*, Vol. 20, No. 4, 2001, pp. 699-719.

Kearney, Melissa Schettini, "Is There an Effect of Incremental Welfare Benefits on Fertility Behavior? A Look at the Family Cap," unpublished manuscript, Cambridge, Massachusetts, 2001.

Kenney, Genevieve, and Jennifer Haley, "Why Aren't More Uninsured Children Enrolled in Medicaid or SCHIP?" Washington, DC: Urban Institute, Series B, No. B-35, May 2001.

Kerpelman, Larry C., David B. Connell, and Walter J. Gunn, "Effect of a Monetary Sanction on Immunization Rates of Recipients of Aid to Families with Dependent Children," *Journal of the American Medical Association*, Vol. 284, No. 1, July 2000.

Klerman, Jacob A., and Steven Haider, "A Stock-Flow Analysis of the Welfare Caseload: Insights from California Economic Conditions," Santa Monica, CA: RAND, March 2000.

Kornfeld, Robert, Laura Peck, Diane Porcari, John Straubinger, Zachary Johnson, and Clemintina Cabral, *Evaluation of the Arizona EMPOWER Welfare Reform Demonstration: Impact Study Interim Report*, Cambridge, MA: Abt Associates, Inc., May 1999.

Ku, Leighton, and Brian Bruen, "The Continuing Decline in Medicaid Coverage," Washington, DC: Urban Institute New Federalism: Issues and Options for States Series Number A-37, 1999.

Ku, Leighton, and Bowen Garrett, "How Welfare Reform and Economic Factors Affected Medicaid Participation: 1984-96," Washington, DC: Urban Institute, Discussion Paper 00-01, February 2000.

Levine, Philip B., "The Impact of Social Policy and Economic Activity Throughout the Fertility Decision Tree," unpublished manuscript, Wellesley College, 2001.

Levine, Phillip B., and Diane M. Whitmore, "The Impact of Welfare Reform on the AFDC Caseload," *National Tax Association Proceedings–1997*, Washington, DC: National Tax Association, 1998.

Loeb, Susana, and Mary Corcoran, “Welfare, Work Experience, and Economic Self-Sufficiency,” *Journal of Policy Analysis and Management*, Vol. 20, Winter 2001, pp. 1-20.

Lofstrom, Magnus, and Frank D. Bean, “Labor Market Conditions and Post-Reform Declines in Welfare Receipt Among Immigrants,” unpublished manuscript, July 2001.

Loprest, Pamela, *Families Who Left Welfare: Who Are They and How Are They Doing?* Washington, DC: Urban Institute, Discussion Paper # 99-02, 1999.

Loprest, Pamela, *How Are Families Who Left Welfare Doing Over Time? A Comparison of Two Cohorts of Welfare Leavers*, Washington, DC: Urban Institute, 2000.

Loprest, Pamela, *How Are Families that Left Welfare Doing? A Comparison of Early and Recent Welfare Leavers*, Washington, DC: Urban Institute, Series B, No. B-36, April 2001.

Loprest, Pamela, and Gregory Acs, *The Status of TANF Leavers in the District of Columbia: Interim Report*, Washington, DC: Urban Institute, February 2000.

MaCurdy, Thomas, David Mancuso, and Margaret O’Brien-Strain, “How Much Does California’s Welfare Policy Explain the Slower Decline of Its Caseload,” unpublished manuscript, December 2000.

Martin, Joyce A., Brady E. Hamilton, and Stephanie J. Ventura, *Births: Preliminary Data for 2000*, Hyattsville, MD: National Center for Health Statistics, National Vital Statistics Reports, Vol. 49, No. 5, www.cdc.gov/nchs/data/nvsr/nvsr49/nvsr49_05.pdf, July 24, 2001.

Mayer, Susan E., *What Money Can’t Buy: Family Income and Children’s Life Chances*, Cambridge, MA: Harvard University Press, 1997.

McGroder, S., M. Zaslow, K. Moore, and S. LeMenestrel, *National Evaluation of Welfare-to-Work Strategies: Impacts on Young Children and Their Families Two Years After Enrollment: Findings from the Child Outcomes Study*, Washington, DC: 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, and Office of Vocational and Adult Education, 2000.

Mead, Lawrence M., “The Decline of Welfare in Wisconsin,” *Journal of Public Administration Research and Theory*, Vol. 9, No. 4, 1999.

Mead, Lawrence M., “Caseload Change: An Exploratory Study,” *Journal of Policy Analysis and Management*, Vol. 19, No. 3, 2000.

Mead, Lawrence M., "Governmental Quality and Welfare Reform," unpublished manuscript, New York University Department of Politics, 2001.

Meyer, Bruce, "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, Vol. 13, No. 2, April 1995, pp. 151-161.

Meyer, Bruce D., and Dan T. Rosenbaum, *Making Single Mothers Work: Recent Tax and Welfare Policy and Its Effects*, Chicago, IL: Joint Center for Poverty Research, Working Paper #152, January 2000.

Meyer, Bruce D., and Dan T. Rosenbaum, "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers," *Quarterly Journal of Economics*, Vol. CXVI, August 2001, pp. 1063-1114.

Meyer, Bruce D., and James X. Sullivan, "The Effects of Welfare and Tax Reform: The Material Well-Being of Single Mothers in the 1980s and 1990s," Cambridge, MA: National Bureau of Economic Research, Working Paper #8298, May 2001.

Meyers, Marcia K., Bonnie Glaser, and Karin MacDonald, "On the Front Lines of Welfare Delivery: Are Workers Implementing Policy Reforms?" *Journal of Policy Analysis and Management*, Vol. 17, No. 1, 1998, pp. 1–22.

Michalopoulos, Charles, and Gordon Berlin, "Financial Work Incentives for Low-Wage Workers: Encouraging Work, Reducing Poverty, and Benefiting Families," in Bruce Meyer and Greg Duncan, eds., *The Incentives of Government Programs and the Well-Being of Families*, Chicago, IL: Joint Center for Poverty Research, 2001.

Michalopoulos, Charles, David E. Card, Lisa A. Gennetian, Kristen Harknett, and Philip K. Robins, *The Self-Sufficiency Project at 36 Months: Effects of a Financial Work Incentive on Employment and Income*, Ottawa, Ontario, Canada: Social Research and Demonstration Corporation, June 2000.

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

Michalopoulos, Charles, and Christine Schwartz, with Diana Adams-Ciardullo, *What Works Best for Whom: Impacts of 20 Welfare-to-Work Programs by Subgroup*, New York, NY: Manpower Demonstration Research Corporation, U.S. Department of Health and Human Services, and U.S. Department of Education, August 2000.

Midwest Research Institute, *Chapter 1–Employment and Earnings of Former AFDC Recipients in Missouri, Chapter 2–Household Income and Poverty, Chapter 3–The Continuing Use of*

Assistance by Former Missouri AFDC Recipients, and Chapter 4–Insecurity Among Former AFDC Recipients, interim reports for the Missouri Department of Social Services, June 2000.

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

Miller, Cynthia, Virginia Knox, Lisa A. Gennetian, Martey Dodoo, Jo Anna Hunter, and Cindy Redcross, *Reforming Welfare and Rewarding Work: Final Report on the Minnesota Family Investment Program (Vol. 1: Effects on Adults)*, New York, NY: Manpower Demonstration Research Corporation, September 2000.

Minkovitz, Cynthia, Elizabeth Holt, Nancy Hughart, William Hou, Larry Thomas, Eugene Dini, and Bernard Guyer, "The Effect of Parental Monetary Sanctions on the Vaccination Status of Young Children," *Archives of Pediatric Adolescent Medicine*, Vol. 153, December 1999.

Moffitt, Robert A., "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature*, Vol. XXX, March 1992, pp. 1-61.

Moffitt, Robert A., "The Effect of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior," in Sheldon H. Danziger, ed., *Economic Conditions and Welfare Reform,* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1999, pp. 91-118.

Moffitt, Robert A., and Michele Ver Ploeg, eds., *Data and Methodological Issues for Tracking Former Welfare Recipients: A Workshop Summary*, Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, Committee on National Statistics, National Research Council, Washington, DC: National Academy Press, 1999.

Moffitt, Robert A., and Michele Ver Ploeg, *Evaluating Welfare Reform in an Era of Transition*, Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, Washington, DC: National Academy Press, Committee on National Statistics, Committee on Behavioral and Social Sciences and Education, and the National Research Council, 2001.

Morris, Pamela A., Aletha C. Huston, Greg J. Duncan, Danielle A. Crosby, and Johannes Bos, *How Welfare and Work Policies Affect Children: A Synthesis of Research*, New York, NY: Manpower Demonstration Research Corporation, March 2001.

Morris, Pamela A., and Charles Michalopoulos, *The Self-Sufficiency Project at 36 Months: Effects on Children of a Program that Increased Parental Employment and Income*, Ottawa, Ontario, Canada: Social Research and Demonstration Corporation, June 2000.

Moses, Anne, and D. C. Macuso, *Examining Circumstances of Individuals and Families Who Leave TANF: Assessing the Validity of Administrative Data*, SPHERE Institute, May 1999.

Mueser, Peter R., Julie L. Hotchkiss, Christopher T. King, Philip S. Rokicki, and David W. Stevens, *The Welfare Caseload, Economic Growth and Welfare-to-Work Policies: An Analysis of Five Urban Areas*, July 2000.

Nixon, Lucia A., "The Effect of Child Support Enforcement on Marital Disolution," *Journal of Human Resources*, Vol. 32, No. 1, Winter 1997, pp. 159-181.

Olsen, Edgar O., "Housing Programs for Low-Income Households," unpublished manuscript, University of Virginia, February 2001.

O'Neill, June E., and M. Anne Hill, *Gaining Ground? Measuring the Impact of Welfare Reform on Welfare and Work,* New York, NY: Manhattan Institute for Policy Research, Civic Report #17, July 2001.

Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects, *Econometrica*, Vol. 49, November 1981, pp. 1417-1426.

Pavetti, LaDonna, and Dan Bloom, *State Sanctions and Time Limits*, in Rebecca M. Blank and Ron Haskins, eds., *The New World of Welfare*, Washington, DC: Brookings Institution Press, 2001, pp. 245-269.

Paxson, Christina, and Jane Waldfogel, "Parental Resources and Child Abuse and Neglect," *American Economic Review*, Vol. 89, No. 2, May 1999, pp. 239-244.

Paxson, Christina, and Jane Waldfogel, "Public Policies, Family Resources, and Child Maltreatment," in Bruce Meyer and Greg Duncan, eds., *The Incentives of Government Programs and the Well-Being of Families,* Chicago, IL: Joint Center for Poverty Research, 2001.

Polit, Denise F., Andrew S. London, and John M. Martinez, *Food Security and Hunger in Poor, Mother-Headed Families in Four U.S. Cities*, New York, NY: Manpower Demonstration Research Corporation, May 2000.

Polit, Denise F., Andrew S. London, and John M. Martinez, *The Health of Poor Urban Women: Findings from the Project on Devolution and Urban Change*, New York, NY: Manpower Demonstration Research Corporation, May 2001.

Primus, Wendell, L. Rawlings, K. Larin, and K. Porter, *The Initial Impacts of Welfare Reform on the Incomes of Single Mother Families*, Washington, DC: Center on Budget and Policy Priorities, 1999.

Quets, Gail, Phillip K. Robins, Elsie C. Pan, Charles Michalopoulos, and David Card, *Does SSP Plus Increase Employment? The Effect of Adding Services to the Self-Sufficiency Project's Financial Incentives*, Ottawa, Ontario, Canada: Social Research and Demonstration Corporation, May 1999.

Rector, Robert E., and Sarah E. Youssef, *The Determinants of Welfare Caseload Decline*, Washington, DC: The Heritage Foundation, Paper # CDA99-04, May 1999.

Rockefeller Institute, New York State Office of Temporary and Disability Assistance, and the New York State Department of Labor, *After Welfare: A Study of Work and Benefit Use After Case Closing: Revised Interim Report*, December 1999.

Rosenbaum, Dan T., *Taxes, The Earned Income Tax Credit, and Marital Status*, Chicago, IL: Joint Center for Poverty Research, Working Paper #177, August 2000.

Ryan, S., M. Theilbar, S. Choi, J. Qu, M. Deng, and L. Ellebracht, *Preliminary Outcomes for 1996 Fourth Quarter AFDC Leavers: Revised Interim Report*, University of Missouri-Columbia, September 1999.

Schoeni, Robert F., and Rebecca M. Blank, "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Income, Poverty, and Family Structure," Cambridge, MA: National Bureau of Economic Research, Working Paper #7627, March 2000.

Turturro, Carolyn, Brent Benda, and Howard Turney, *Arkansas Welfare Waiver Demonstration Project. Final Report*, Little Rock, AR: University of Arkansas at Little Rock, School of Social Work, June 1997.

U.S. Bureau of the Census, *Historical Poverty Tables*, 2001, www.census.gov/hhes/poverty/histpov/histpov1.html.

U.S. Bureau of the Census, *Living Arrangements of Children Under 18 Years Old: 1960 to Present*, June 29, 2001.

U.S. Bureau of Labor Statistics (USBLS), *Employment and Wages Annual Averages*, Washington, DC: U.S. Government Printing Office, 1989.

U.S. Bureau of Labor Statistics (USBLS), *Labor Force Statistics from the Current Population Survey*, 2002, www.bls.gov/cps/home.htm.

U.S. Department of Agriculture (USDA), Food and Nutrition Service (FNS), *Who is Leaving the Food Stamp Program? An Analysis of Caseload Changes from 1994 to 1997*, www.fns.usda.gov/oane/MENU/Published/FSP/FILES/Participation/CDR.html, March 1999.

U.S. Department of Agriculture (USDA), Food and Nutrition Service (FNS), Office of Analysis, Nutrition, and Evaluation. *Characteristics of Food Stamp Household: Fiscal Year 2000*, 2001a.

U.S. Department of Agriculture (USDA), Food and Nutrition Service (FNS), Office of Analysis, Nutrition, and Evaluation, *The Decline in Food Stamp Participation: A Report to Congress*, www.fns.usda.gov/oane/MENU/Published/FSP/FILES/Participation/PartDecline.htm, July 2001b.

U.S. Department of Agriculture (USDA), Food and Nutrition Service (FNS), *Food Stamp Program Participation and Costs*, www.fns.usda.gov/pd/fssummar.htm, May 2001c.

U.S. Department of Agriculture (USDA), Food and Nutrition Service (FNS), *Program Data: Child Nutrition Tables*, www.fns.usda.gov/pd/cnpmain.html, May 2001d.

U.S. Department of Health and Human Services (USDHHS), *Report to Congress on Out-of-Wedlock Childbearing*, Washington, DC, September 1985.

U.S. Department of Health and Human Services (USDHHS), Administration for Children and Families (ACF), Office of Planning, Research and Evaluation, *Temporary Assistance for Needy Families (TANF Program), Third Annual Report to Congress*, Washington, DC, August 2000.

U.S. Department of Health and Human Services (USDHHS), *Status Report on Research on the Outcomes of Welfare Reform*, Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, Administration for Children and Families, aspe.hhs.gov/hsp/welf-ref-outcomes01/index.htm, July 2001a.

U.S. Department of Health and Human Services (USDHHS), *Total Number of Families, Temporary Assistance for Needy Families*, Washington, DC: Administration for Children and Families, www.acf.dhhs.gov/news/families.htm, 2001b.

U.S. National Center for Health Statistics, *Births: Final Data for 2000*, Hyattsville, MD: National Vital Statistics Reports, February 12, 2002.

Verma, Nandita, and B. Goldman, *Los Angeles County Post-TANF Tracking Project: Quarterly Progress Report*, New York, NY: Manpower Demonstration Research Corporation, January 2000.

Wallace, Geoffrey, and Rebecca M. Blank, "What Goes Up Must Come Down? Explaining Recent Changes in Public Assistance Caseloads," in Sheldon H. Danziger, ed., *Economic Conditions and Welfare Reform,* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1999, pp. 49-89.

Werner, Alan, and Robert Kornfeld, *The Evaluation of "To Strengthen Michigan Families": Final Impact Report*, Cambridge, MA: Abt Associates, Inc., September 1997.

Westra, Karen L., and John Routley, *Arizona Cash Assistance Exit Study: Cases Exiting Fourth Quarter 1996*, Arizona Department of Economic Security, July 1999.

Westra, Karen L., and John Routley, *Arizona Cash Assistance Exit Study: First Quarter 1998 Final Report*, Washington, DC: Arizona Department of Economic Security, U.S. Department of Health and Human Services, January 2000.

Wilde, Parke, Peggy Cook, Craig Gundersen, Mark Nord, and Laura Tiehen, *The Decline in Food Stamp Program Participation in the 1990s*, Washington, DC: U.S. Department of Agriculture, Economic Research Service, Food and Rural Economics Division, Food Assistance and Nutrition Research Report No. 7, June 2000.

Wilson, William Julius, and Kathryn Neckerman, "Poverty and Family Structure: The Widening Gap between Evidence and Policy Issues," in William J. Wilson, ed., *The Truly Disadvantaged*, Chicago: The University of Chicago Press, 1987.

Zedlewski, Sheila R., and Sarah Brauner, "Are the Steep Declines in Food Stamp Participation Linked to Falling Welfare Caseloads?", Washington, DC: Urban Institute, Series B, No. B-03, November 1999.

Zedlewski, Sheila R., and Amelia Gruber, "Former Welfare Families and the Food Stamp Program: The Exodus Continues," Washington, DC: Urban Institute, Series B, No. B-33, March 2001.

Ziliak, James P., and David N. Figlio, "Geographic Differences in AFDC and Food Stamp Caseloads in the Welfare Reform Era," Chicago, IL: Joint Center for Poverty Research, Working Paper # 180, June 2000.

Ziliak, James P., David N. Figlio, Elizabeth E. Davis, and Laura S. Connolly, "Accounting for the Decline in AFDC Caseloads: Welfare Reform or the Economy?" *The Journal of Human Resources*, Vol. 35, No. 3, January 2000.