ESTIMATING THE EFFECT OF WORK REQUIREMENTS ON WELFARE RECIPIENTS:  
A SYNTHESES OF THE NATIONAL LITERATURE

Testimony of Lynn A. Karoly, Ph.D.

Senior Economist  
Director, Labor and Population Program  
RAND  
Santa Monica, CA and Arlington, VA

Before the Subcommittee on 21st Century Competitiveness  
Committee on Education and the Workforce  
U.S. House of Representatives

October 16, 2001

The opinions and conclusions expressed in this written testimony are the author’s alone and should not be interpreted as representing those of RAND or any of the sponsors of its research.
INTRODUCTION

The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), which implemented the Temporary Assistance to Needy Families (TANF) program, was enacted to further several goals: reduce dependency, increase employment, reduce unwed childbearing, promote marriage, and maintain two-parent families. With reauthorization of TANF pending, policymakers want to know how much the policy reforms embodied in the TANF legislation have been successful in achieving these objectives.

Certainly, the trends in various key indicators suggest that welfare reform may have resulted in anywhere from modest to substantial progress toward meeting these goals. In 2001, the welfare caseload is less than half of what it was at its peak level in 1994, when it was 5 million families (U.S. Department of Health and Human Services (DHHS), 2001b). Employment rates of women leaving welfare range from 62 to 90 percent (DHHS, 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, 2001). 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 been increasing, and the poverty rate has been falling (Haskins, 2001). These improvements in labor market outcomes and family incomes have been accompanied by a decline in teen fertility and an increase in two-parent families (Martin et al., 2001; Acs and Nelson, 2001; Dupree and Primus, 2001).

At the same time, other data suggest another perspective. Former welfare recipients, so-called "welfare leavers," have been monitored in over 30 state studies, including 13 funded by the Assistant Secretary for Policy and Evaluation (ASPE), U.S. Department of Health and Human Services (DHHS). A recent summary of the 13 ASPE-funded welfare leaver studies (plus two other comparable analyses) shows that most welfare leavers (62 to 90 percent), defined as those who have been off welfare for at least two months, work at least some part of the first year after leaving welfare (DHHS, 2001a). However, a smaller fraction, from one third to one half, were employed all four quarters, and the evidence is mixed about the longer-term trend in employment outcomes. Moreover, earnings and family income remain low and spells of unemployment are common. Most welfare leavers continue to receive other government support, most commonly Medicaid and, to a lesser extent, Food Stamps. Even so, rates of reported food insecurity range from 13 to 52 percent and other forms of material hardship (e.g., housing problems, issues with access to health care) have similar or somewhat lower prevalences. While work is common, 2 to 4 out of every 10 welfare leavers return to the welfare rolls in the first year after leaving. In general, these studies paint a picture of considerable diversity in the post-welfare circumstances of former recipients.

written testimony are the author’s alone and should not be interpreted as representing those of RAND or any of the sponsors of its research.
While it is tempting to infer that welfare reform explains the observed trends or the circumstances of welfare leavers, we know that these types of comparisons do not account for other changes that took place during the same time period that could have also contributed to the observed outcomes. In particular, there were other policy changes, such as increases in the Earned Income Tax Credit (EITC), a de-linking of subsidized health insurance from welfare receipt, and increases in the minimum wage. Perhaps most important, there was a long and robust economic expansion. Thus, at least some of the improvements in welfare-related outcomes resulted from changes in other policies and the improving economy rather than from changes in welfare programs. By the same reasoning, the observed outcomes of welfare leavers, while informative, may result from both welfare policy changes and other factors.

For policy purposes, it is essential to know what the effect of welfare reform has been, holding all else constant. Under a contract from the U.S. Department of Health and Human Services Administration for Children and Families (U.S. DHHS-ACF), my colleagues—Jeffrey Grogger and Jacob Klerman—and I have conducted an extensive review of recent research to synthesize the current state of knowledge about the effects of welfare reform. The primary focus of the synthesis was on the net effects of reform, 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 considered both the effect of the TANF reforms as a bundle and the effects of specific policies, including work requirements, financial incentives, and time limits.

Given the focus of this hearing on work requirements, I devote the bulk of my testimony today to a discussion of what we know about the impact of this particular policy component. However, before doing so, I briefly review our approach to determining the causal impact of
TANF on various outcomes of interest. At the end of this testimony, I turn to a discussion of some issues for TANF reauthorization suggested by the research.

**HOW DO WE DETERMINE THE CAUSAL IMPACT OF TANF?**

Distinguishing the effects of welfare reform from the effects of other policies and the economy is not a simple matter. For example, we want to know how much of the decline in the welfare caseload or rise in employment of single mothers is the result of the 1996 PRWORA legislation and the TANF programs it implemented. To answer this question, we need to know what would have happened to the welfare caseload or employment in the absence of welfare reform; in other words, we want to know what would have happened if the status quo prior to August 1996 had continued. However, we do not actually observe that outcome: that outcome—referred to as a counterfactual—does not exist. Instead, we observe a world where the policy change occurred but other policies and other factors like the economy changed as well. Since we do not observe the counterfactual, we must design and implement a research strategy that holds everything else constant.

There are two primary research methodologies available to measure causal impacts, in other words, to hold everything else constant. The first is to implement a random assignment experiment. To test a new program or policy, a study population is chosen and individuals are randomly assigned to either the control group, which is subject to the baseline policy environment, or the treatment group, which is subject to the new policy environment. If randomization is executed properly, there should be no systematic differences between the treatment and control groups other than those attributable to the different policy environments and chance. As the sample becomes larger, the likely effect of chance shrinks so that 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 outcomes between the two groups.

Such random assignment experiments can be a powerful evaluation tool, either for evaluating specific policy changes such as work requirements or combined (or bundled) policy changes such as work requirements packaged with financial incentives and time limits.1 Recognizing the analytical benefit of this approach, DHHS-ACF required—in the pre-TANF period—that random assignment evaluations be a component of granting section 1115 waivers. Under these waivers, states began to implement reforms that, for the most part, were later incorporated into their TANF plans. The body of experimental studies from these waiver evaluations represents a major accumulation of knowledge about policy effects and, thus, comprises a substantial part of the evidence base we draw on for our synthesis. More specifically, we review the published findings from 28 major experimental evaluations conducted during the 1990s prior to TANF implementation.

The second research approach to estimating causal impacts is to analyze observational data using statistical methods to hold constant as many potential confounding factors as possible. This approach allows researchers to evaluate of policies not subject to the random assignment studies and to capture the entry effects that the random assignment studies miss. In our synthesis, we review over two dozen high-quality studies that utilize this approach.2

__________

1Despite their advantages, however, random assignment studies have several drawbacks. For a variety of reasons, they are not always feasible, their findings may not generalize to a universally implemented program, and they can be implemented poorly such that the impact estimates may be biased. In addition, in the context of welfare reform, random assignment studies only capture the effect of the new policy environment for those already on welfare. The impact on entry onto welfare is generally not captured in these studies.

2Like the experimental studies, there are potential methodological issues with observational studies and the statistical methodologies they employ.
Drawing on both types of studies, our synthesis considers the following welfare-related outcomes:

- Welfare use and the caseload;
- Employment and earnings;
- Utilization of other government programs (e.g., Food Stamps and Medicaid);
- Family structure, specifically marriage and fertility;
- Income and poverty;
- Consumption and other material well-being (e.g., food security, housing security, and health insurance coverage);
- Child well-being (e.g., child development and school progress).

We examine the effect of specific welfare policies on these outcomes. The specific policies we consider are limited by those considered in the literature and include:

- Financial work incentives, including earnings disregards and benefit reduction rates;
- Requirements to work or participate in work-related activities;
- Time limits;
- Family caps and minor residence requirements;
- Family responsibility requirements.

In the case of the policies related to work participation mandates, we also consider variation in program content or approach, such as the human-capital development model (that emphasizes basic skills and education) and the work-first model (that emphasizes job search and employment).
The best way to view our synthesis is as a matrix that addresses the intersection between specific policy and welfare-related outcome cells in that matrix and seeks to fill them in based on the evidence in the literature. While there is a considerable body of research to draw on, our review documents that we do not know the impact of each policy on every welfare-related outcome.

THE IMPACT OF REQUIREMENTS FOR MANDATORY WORK-RELATED ACTIVITIES

One specific policy we consider in our synthesis is requirements for mandatory work-related activities, also known as work requirements. Starting with the Job Opportunities and Basic Skills (JOBS) training program in 1988, nonexempt welfare recipients were required to spend 20 hours a week in work-related activities, including work, education, or training. Those who failed to participate were subject to sanctions, which involved forfeiting the adult’s portion of the Aid to Families with Dependent Children (AFDC) benefit. First under waivers and then under TANF, requirements for mandatory work-related activities were strengthened to include higher hours requirements, more restrictive definitions of permissible work-related activities, and lower age-of-youngest-child exemptions. Many states also reoriented their welfare-to-work programs as well, emphasizing job search and employment (so-called work-first programs) over basic skills and education (so-called human-capital development programs). Stiffer sanctions policies (i.e., penalties for noncompliance) were also implemented.

Of the various welfare reform policies, requirements for mandatory work-related activities are the best studied. Eleven random assignment experiments, collectively known as the National Evaluation of Welfare-to-Work Strategies (NEWWS), evaluated a reform that primarily consisted of stronger requirements for mandatory work-related activities, with approaches that
included the Human Capital Development (HCD) (e.g., education and training) model and the Labor Force Attachment (LFA) (i.e., work first) model. The sites included Atlanta (Georgia), Grand Rapids (Michigan), Riverside (California), Portland (Oregon), Columbus (Ohio), Detroit (Michigan), and Oklahoma City (Oklahoma). Two similar evaluations in Los Angeles (Jobs-First GAIN) and Indiana (the basic track of the Indiana Manpower Placement and Comprehensive Training Program (IMPACT)) also primarily evaluated mandatory work-related activities. Each of these studies have published impact estimates for a two-year follow-up period (Freedman, Knab et al., 2000; Freedman, Friedlander et al., 2000, McGroder et al., 2000; Fein et al., 1998).

The results from these 13 studies support the following conclusions about the impact of requirements for mandatory work-related activities:

- Employment rates and earnings increase, while welfare use and the use of Food Stamps and Medicaid decrease.
- While earnings rise, welfare payments fall, so that there appears to be no effect on income (i.e., the sum of earnings and welfare), although poverty may improve.
- Like income, there appears to be no change in marriage or fertility when work requirements are implemented. The evidence base is weaker with respect to the impact of work requirements on broader measures of well-being, and on child development and school progress.

We discuss each of these conclusions in more detail below.

---

3These programs included some other reforms such as extended transitional child care or a family cap (Indiana only) but unlike TANF, they did not include time limits or financial incentives to make work pay (e.g., more generous earned income disregards or benefit reduction rates).
Employment Rates and Earnings Increase, While Welfare Use and the Use of Food Stamps and Medicaid Decrease

Figure 1 shows the impact of work requirements on employment, as measured by the experimental (random assignment) evaluations that hold all else constant. Employment is measured by whether the participant was ever employed over the two-year follow-up period. With one exception, employment is higher in the treatment group than in the control group, and nine of the twelve positive impact estimates are statistically significant (as indicated by the asterisk on the program name in the figure). On average, these programs increase 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, result in larger average employment gains than the HCD programs, which emphasize skill-building and generally require the participant 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. Part of this difference may arise because LFA participants generally start looking for work right away, while HCD participants spend part of the follow-up period in classes or training and therefore on average do not start work until later.
Figure 1: Impact of Work Requirements on Employment, 13 Experimental Studies

NOTES: * indicates difference in employment rates between treatment and control groups is statistically significant at the 5 percent level or better.
The earnings results (not shown in the figure) from these programs are also positive, although the earnings impacts are small. Twelve of thirteen programs produced positive effects on earnings, nine of which were statistically significant. The one negative effect was insignificant. The average earnings impact over the two-year follow-up exceeded $700; only four of the programs failed to produce earnings gains of at least $400. Again, the gains were greater for the search-oriented programs than for the skills-oriented programs. Among the four work-first programs, earnings impacts averaged about $1,200. Among the human-capital programs, earnings impacts were smaller, averaging just under $400.

In twelve of the thirteen evaluations, compared with the control group, welfare use was lower after two years for the treatment group subject to stricter work requirements and all impacts were statistically significant. This is consistent with economists’ expectations that work requirements should make welfare less attractive and therefore lower welfare use. The average reduction in welfare use was 5.1 percentage points. In relative terms, the average reduction was 8.7 percent. Across the programs, there is evidence that the jobs-first model generated somewhat greater reductions in welfare use than the skills-oriented programs.

Food Stamps and Medicaid use (also not shown in the figure) generally follow the impacts for welfare. Nine of the thirteen studies find a statistically significant negative impact on Food Stamp use, while the one study that measures Medicaid use also finds a negative impact. The negative impacts on Food Stamp and Medicaid participation are generally not as large as those for welfare use, indicating that some who leave welfare continue to receive these benefits.
While Earnings Rise, Welfare Payments Fall, So That There Appears to Be No Effect on Income, Although Poverty May Improve

Twelve of the thirteen welfare to work programs found no significant impact of mandatory work requirements on income (the sum of recipient earnings and welfare plus the Earned Income Tax Credit) at the two-year follow-up. The sole exception was Los Angeles Jobs-First GAIN, where a broader measure of monthly household income was about $90 higher for the treatment group compared with the control group, a statistically significant impact. Impacts on poverty were almost all insignificant as well, although more impact estimates were negative, suggesting that these programs may be somewhat more effective at raising incomes near the poverty threshold than at the bottom of the income scale. At the same time, many of the programs increased the fraction with incomes below 50 percent of the poverty line.

The small or nonexistent impacts on income and poverty are consistent with the combination of the positive earnings gains produced by these programs and high benefit reduction rates under the old AFDC rules that led to a significant reduction in welfare payments as earnings rose. Since income 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 share of income from earnings. Even so, the average earnings share for treatment group members never exceeds 50 percent.

Like Income, There Appears to Be No Change in Marriage and Fertility When Work Requirements Are Implemented—Evidence Base Weaker for Other Well-Being Outcomes

With one exception, the twelve programs that evaluate the impact of work requirements on marriage and fertility show no significant impacts on either outcome as of the two-year follow-up. In terms of other measures of well-being, measures of health care coverage for adults and children are available for 12 of the 13 studies. Almost all the impact estimates indicate 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 a few are statistically significant. The reductions in health insurance coverage are consistent with the move off welfare to employment associated with mandatory work requirements. The loss of automatic Medicaid coverage is not entirely made up by transitional Medicaid coverage, coverage under the poverty-related Medicaid expansions, or transitions to employment-based coverage.

Twelve of the thirteen studies also examine two-year impacts on a range of outcome domains for children, including behavioral adjustments, school progress, and health and safety. Overall, the child outcome results from these studies show no clear pattern of beneficial or harmful effects for children. Both favorable and unfavorable effects are found across all the domains, sometimes for the same program. A comparison across program models does not reveal any strong patterns for employment-focused or education-focused programs.

THE IMPACT OF WORK REQUIREMENTS COMBINED WITH OTHER POLICIES

Under states’ TANF programs, work requirements have typically not been implemented in isolation; rather, they are combined with other major reforms. Our synthesis draws on another set of experimental evaluations that assess impacts of work requirements combined with other policy reforms, namely financial incentives and time limits. Compared to the 13 studies that focus on work requirements by themselves, there are fewer studies that analyze work requirements in conjunction with other reforms. Thus, our conclusions about such reforms are less definitive than our conclusions about work requirements implemented by themselves.

Combining mandatory work-related activities with strong financial incentives and/or time limits generally still results in positive impacts on employment and earnings. At the same time,
financial incentives, by allowing recipients to keep more of their welfare grant as their earnings increase, typically produce smaller decreases in transfer payments, which lead to higher income and a reduction in poverty. Welfare reforms that combine work requirements, financial incentives, and time limits appear to initially increase welfare use. Welfare participation then declines, first as more individuals leave the rolls to preserve benefit eligibility for the future, and then as recipients reach their time limit. There is also evidence that programs that include time limits along with work requirements and financial incentives produce initial increases in income that eventually fade as welfare use declines, a pattern that is consistent with the changes in employment, earnings, and welfare use.

There is a more limited evidence base to suggest that strong financial incentives can increase marriage, but this is not always the case. The evidence is mixed with respect to child outcomes, with some random assignment studies finding some improvements in child development when programs include financial incentives, while others show little or no impact. Negative impacts on child schooling and behavior, particularly for adolescents, have been found in one evaluation that includes time limits as part of the reform package.

In sum, the available results suggest that work requirements when combined with other reforms generally increase work and earnings, just as they do in isolation. Both welfare use and income can increase when strong financial incentive are part of the reform package, or they can both decrease when time limits are part of the reforms and those time limits become binding. The impact on other outcomes, such as income, fertility and marriage, and child well-being, can differ from what is observed when work requirements are the only reform.

ISSUES FOR REAUTHORIZATION

Our synthesis suggests several issues as Congress debates the reauthorization of TANF.
First, policymakers need to recognize the trade-offs between the various goals of welfare reform and the ability of different policy components to affect the goals of reform.

While the primary reforms to the welfare system, such as work requirements, financial incentives, and time limits, generally serve to raise employment and earnings, it is not the case that each of these policies simultaneously reduces welfare use, increases income, reduces poverty, lowers fertility, increases marriage, and promotes child well-being. For example, as discussed above, while work requirements reduce welfare use, they do not have much of an impact on income and poverty. Consequently, they also do not have much of an impact on marriage, fertility, or child well-being. In contrast, rather than reducing dependency, programs with generous financial incentives generally increase welfare receipt and total transfer payments, at least in the short-run. Since financial incentives allow families to keep more of their welfare benefits as their earnings rise, they also increase income, decrease poverty, and improve material well-being. Moreover, when incomes increase, there is more of a tendency for child outcomes to improve as well, or at least to not become worse.

Second, the federal government should continue to coordinate and fund evaluation research of welfare reform, including longer-term follow-up of existing random assignment populations and new studies that evaluate specific reform components.

Understanding the causal impact of welfare reform as a whole and specific policies in particular requires a solid research base of high-quality random assignment and observational studies of the kinds we review in our synthesis. While the knowledge base is quite strong in some areas, such as assessing the impact of work requirements, in other areas, it is weak or nonexistent. For example, we know very little about the impact of various sanction policies on welfare-related outcomes. As another example, time limits have been evaluated almost
exclusively as part of a package of reforms, so it is difficult to isolate the separate impact of time limits on behavior. There is also little basis for knowing what will happen under policies that might be adopted in the future, such as different forms of time limits.

Finally, much of our knowledge base tells us about the short-term impacts of welfare reform, typically over a two- to four-year horizon. Some impacts, such as those on marriage, fertility, and child well-being, may take longer to respond to a new policy environment. Hence, longer-term follow-up of experimental populations is vital for assessing the longer-run consequences of these policies.

Both to better understand the impacts of policies already implemented and to gauge the impact of policies that might be implemented in the future, it is imperative that the federal government continue to coordinate and fund new research to augment what we already know. Our knowledge base in 2001 is stronger because of research programs put in place in the late 1980s and early 1990s, and that increase in knowledge occurred only as a result of major expenditures on program development and evaluation. The random assignment studies in particular were conducted to satisfy the requirement that waiver-era reforms be evaluated.

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, serious consideration should be given to mechanisms to reinvigorate the evaluation of program reforms. Even given TANF’s devolution of welfare policy to the states, a strong federal role in research and evaluation remains appropriate and necessary. This is especially true because research conducted in one state can be informative for policymakers throughout the nation. Such investments in the knowledge base will ensure that
policymakers better understand the trade-offs embodied in different reform policies at the next reauthorization of PRWORA.

REFERENCES


