

HANMING FANG

Yale University

MICHAEL P. KEANE

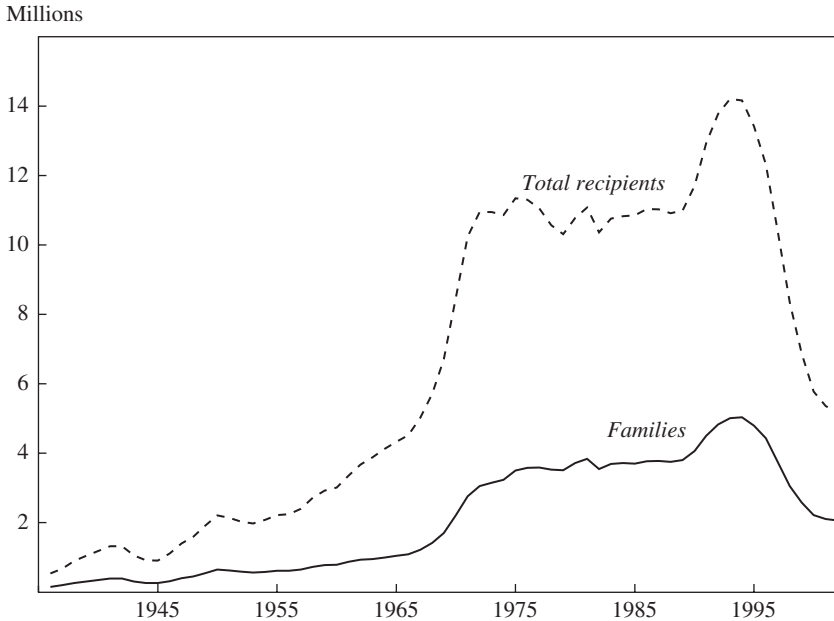
Yale University

Assessing the Impact of Welfare Reform on Single Mothers

THE PERSONAL RESPONSIBILITY and Work Opportunity Reconciliation Act (PRWORA), signed into law in 1996, transformed the U.S. welfare system. PRWORA replaced the Aid to Families with Dependent Children (AFDC) program with Temporary Assistance for Needy Families (TANF). Since its inception in 1935 as part of the Social Security Act, AFDC had been the main welfare program providing assistance to low-income single mothers. But a number of factors, particularly the rapid growth in the never-married single-mother population and a resumption of growth in caseloads in the early 1990s (following the surge of the late 1960s and early 1970s; figure 1), rendered the program unpopular.¹ Under the new TANF program, welfare participation among single mothers has dropped dramatically, from 25 percent in 1996 to 9 percent today. At the same time,

We thank our discussants for many helpful comments and suggestions. Brandon Wall provided outstanding research assistance. Raquel Bernal helped us construct some of the welfare policy variables used in the analysis. We are grateful to Nina Campbell, Karen Clairemont, and Evelyn Mills at the Administration for Children and Families of the U.S. Department of Health and Human Services, and Catherine Hine and Patrick Waldron at the U.S. Department of Agriculture, for providing additional policy data. We also thank Gretchen Rowe at the Urban Institute for clarifying some questions about its Welfare Rules database.

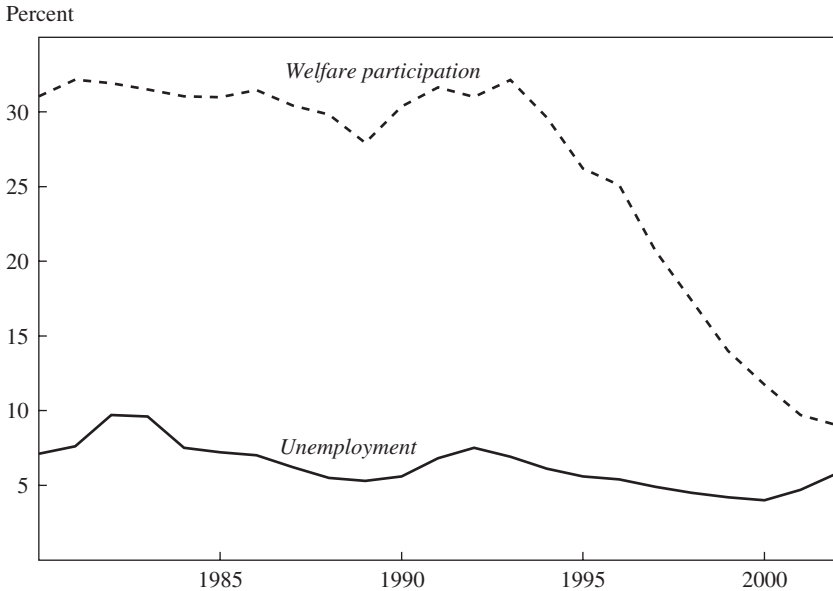
1. Before PRWORA the AFDC program had undergone a number of overhauls as well as lesser changes. For instance, in 1961 the AFDC-Unemployed Parent program (AFDC-UP; a program that provided benefits to two-parent households) was created; in 1967 the AFDC benefit reduction rate (the “tax” on wages earned while on welfare) was reduced to two-thirds from its original level of 100 percent; in 1981 the benefit reduction rate reverted to 100 percent; and in 1988 the Job Opportunities Program (JOBS) was created and AFDC-UP mandated in all states. See Garfinkel and McLanahan (1986) and Moffitt (2003) for historical accounts of the major developments in the AFDC program.

Figure 1. Welfare Caseloads, 1936–2002^a

Source: Administration for Children and Families, Department of Health and Human Services.
 a. Annual averages of monthly data on recipients of AFDC (before 1996) or TANF.

the fraction of single mothers who work has increased from 74 percent in 1996 to 79 percent today. The goal of this paper is to ascertain what features of welfare reform, if any, have been most responsible for this decline in welfare participation and increase in work among single mothers.

Two factors complicate our task. First, a key feature of PRWORA was that it reduced federal authority over welfare policy, giving the states much greater leeway in the design of their own individual TANF programs. A great deal of program heterogeneity has emerged across states, making it difficult to develop a set of variables that comprehensively characterize the different state TANF programs. Second, a number of other recent developments may also have contributed to the changes in welfare and work participation since 1996. These factors, such as the strong U.S. economy of 1996–2000 and the significant expansion of the earned income tax credit (EITC) after 1993, must be controlled for in order to isolate the impact of particular elements of state TANF policies.

Figure 2. Unemployment and Welfare Participation among Single Mothers, 1980–2002

Source: Bureau of Labor Statistics and authors' calculations based on CPS data.

One important fact lends credence to the idea that factors other than PRWORA may account for the lion's share of recent caseload declines: the dramatic drop in welfare participation (and the dramatic increase in work) among single mothers actually began in 1993–94, before PRWORA's enactment (figure 2). From 1993 to 1996 AFDC participation fell from 32 percent to 25 percent. On the other hand, beginning around 1993, many states began to obtain federal waivers allowing them to adopt TANF-like reforms of their AFDC programs. Such reforms included work requirements, time limits on benefits, sanctions for failure to meet work requirements, and family caps. These changes may have contributed substantially to caseload declines even before PRWORA.

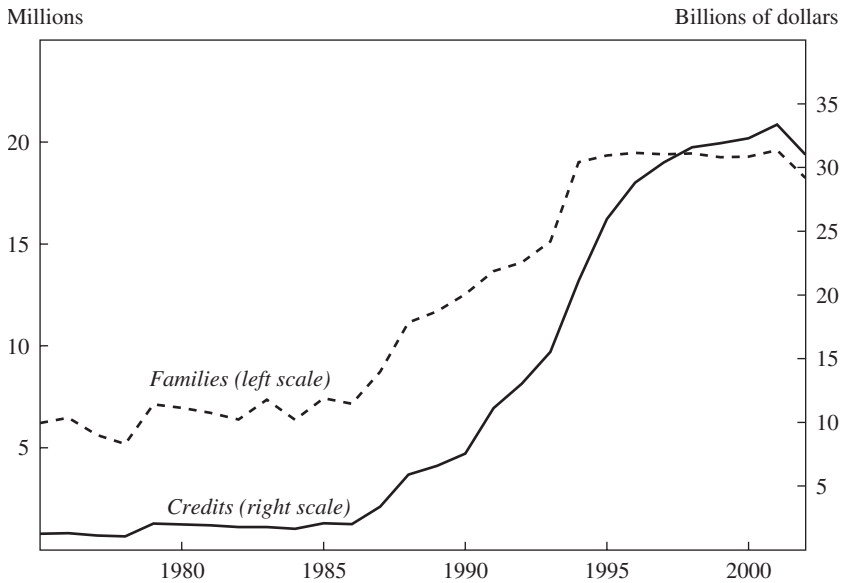
At the same time that PRWORA delegated greater control of welfare policy to the states, it also mandated nationwide many of the popular features introduced under state waivers, such as time limits and work requirements. To understand the sense in which the federal law "mandates" certain features of state TANF programs, one must understand how federal

TANF funds are distributed to the states. Under AFDC, states received federal matching funds based on their AFDC expenditures. PRWORA converted these matching funds to block grants. The block grant for a state was fixed at a level related to federal funding of AFDC benefits and other related programs in the year when that funding had been highest in that state. States were given substantial leeway in how the block grant funds could be used: for example, they may use it to support child care (an important postreform development to which we will return). However, to avoid fiscal penalties on the federal block grant, states must adhere to a “maintenance of effort” (MOE) rule: states must maintain their spending on assistance for needy families at no less than 75 to 80 percent of their pre-1996 level.²

PRWORA requires that state TANF programs set a five-year lifetime limit for any individual receiving *federally funded* aid, although states may exempt up to 20 percent of their caseload from the limit. States may elect to set shorter time limits, and many have. However, any assistance provided to recipients beyond the five-year limit must be financed solely out of state funds. Three states (Michigan, New York, and Vermont) have effectively decided not to enforce the five-year limit. And many states (such as California) do not terminate but only reduce benefits when the time limit is reached. PRWORA also requires that a specific and rising percentage of states’ TANF recipients either work or engage in work-related activities (such as job search or training), and that states impose a work requirement on any recipient who receives TANF for more than two years. Again, states may set a shorter work requirement time limit, and many have done so. States also vary greatly in the sorts of exemptions from work requirements that they allow and in the penalties they impose if work requirements are not satisfied.

Roughly contemporaneously with the changes implemented by PRWORA, the U.S. economy experienced one of its longest postwar expansions. The national unemployment rate remained below 5 percent from 1997 to 2001 and dropped as low as 4 percent in 2000 (figure 2). At about the same time, the EITC was dramatically expanded in terms of both the number of recipients and the generosity of the credit. Figure 3 shows

2. Moreover, states may carry TANF funds over from fiscal year to fiscal year without limit. Although the use of carried-over funds is, in principle, more limited than same-year funds, in practice, the restrictions do not matter.

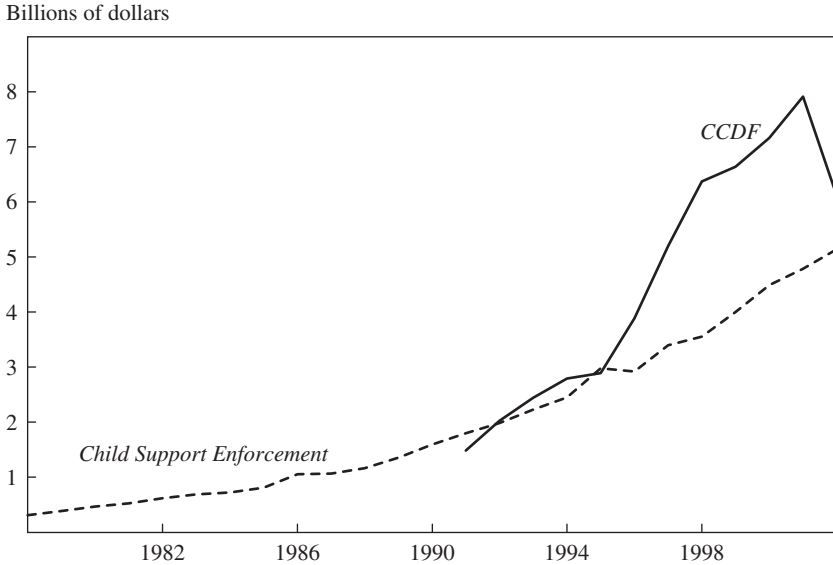
Figure 3. Families Receiving EITC and Aggregate Credits Received, 1975–2002^a

Source: Internal Revenue Service data and U.S. House of Representatives, Committee on Ways and Means (2000).
 a. Data for 2002 are projections.

that the number of federal EITC recipient families increased from about 7 million in 1980 to 19.6 million in 2001. The federal EITC phase-in rate for a single mother with one child increased from 10 percent in 1980 to 34 percent in 2002.³ Moreover, many states have enacted additional EITC programs of their own (for more details of the EITC expansion, see the discussion of the EITC under “Data” below). Other contemporaneous policy changes include the expansion of Medicaid under the Omnibus Budget Reconciliation Act of 1989 (OBRA 1989), which dramatically expanded health insurance coverage for low-income women and children who had not been receiving cash welfare benefits. Moreover, expenditure on the Child Care and Development Fund (CCDF) increased from \$1.4 billion in 1992 to \$7.9 billion in 2001 (figure 4). In fact, the value of

3. The EITC increases in proportion to earned income at the phase-in rate until the credit reaches the (fixed) maximum amount. The credit starts to decrease at the phase-out rate when earned income exceeds another fixed threshold.

Figure 4. Expenditure on Child Care and Development Fund and Child Support Enforcement, 1978–2002



Source: Department of Health and Human Services.

child care subsidies and other noncash benefits now exceeds cash assistance in total federal and state spending under TANF programs. The federal and state governments have also substantially increased expenditure for child support enforcement (figure 4). Naturally, all of these changes in the economic and policy environment could affect the incentives of single mothers to participate in welfare or work.

The changes in average yearly AFDC/TANF caseloads over the past several decades, depicted in figure 1, can be summarized as follows:

—A steep increase in AFDC caseloads occurred in the late 1960s and early 1970s, which were a time of enormous expansion in government public assistance programs, including the establishment of the food stamp and Medicaid programs. Moreover, between 1968 and 1971 the Supreme Court abolished the absent father rule, the residency requirement, and regulations that denied aid to families with “employable mothers.” These rulings increased the welfare take-up rate substantially.

—AFDC caseloads were almost flat from the early 1970s until 1990, with a mild increase in the early 1980s due to the back-to-back recessions of 1980 and 1981–82. The increase in the benefit reduction rate (the “tax” on wages earned while on welfare) from two-thirds to 100 percent during President Ronald Reagan’s first term quickly stopped that uptick.

—A dramatic increase in the caseload occurred from 1990 to 1994. This increase is puzzling because the 1990–91 recession was quite mild, and the 1988 Family Support Act had recently mandated that “work eligible” AFDC recipients participate in welfare-to-work programs. Nor did the welfare participation rate of single mothers exhibit a steep increase (figure 2). We discuss various explanations for this phenomenon in our review of the literature below.

—Welfare caseloads dropped spectacularly after the peak in 1994. The total caseload fell more than 60 percent from the peak of 1994 to 2002, a period roughly contemporaneous with the sustained economic expansion of 1992–2000. The recession that began in March 2001 did increase welfare caseloads in some states, but only slightly, and the national caseload showed a further slight decrease.

How did the different components of welfare reform and other contemporaneous economic and policy changes contribute to the spectacular drops, both in the welfare participation rate of single mothers and in welfare caseloads, that have occurred since 1993? What were the relative contributions of time limits, work requirements, the EITC, child care subsidies, and the strong macroeconomy? These are questions of immense importance for both policymakers and researchers. The answers matter for the design of improved welfare policies and for understanding how welfare policies should respond to macroeconomic conditions.

Much research has already been devoted to these questions, and we review some of the key contributions to this literature in the next section. All of these have focused on only one or a few of the policy and economic variables of interest. Thus they are unable to measure the separate contributions of each of the elements mentioned above. Furthermore, we would argue, studies that focus on only a few policy variables may yield biased estimates of the effects of the policies in question, because they exclude other important policy and environmental factors.

One of the main contributions of this paper is the construction of a detailed data set that includes measures of all the key economic and policy elements described above, on a state-by-state and year-by-year basis,

for the entire 1980–2002 period. One concern in incorporating so many features in one grand analysis was the possible *collinearity* among the policies,⁴ many of which were implemented roughly contemporaneously. We deal with this problem by exploiting both cross-state variation in the timing and form of particular policies as well as cross-sectional variation in how individuals with different characteristics are affected differently by seemingly collinear policies. We discuss in detail the sources of variation that we use to identify the effects of each variable of interest.

The individual-level data that we use, in conjunction with the economic and policy variables we compiled ourselves, are those in the Annual Demographics Supplement to the March Current Population Survey of the U.S. Bureau of the Census (March CPS).⁵ From the 1981–2003 supplements (which cover the period 1980–2002), we extracted data on all single mothers with dependent children, or, more specifically, women who were not living with a spouse at the time of the interview and who had at least one dependent child age 17 or younger. These women may be divorced, widowed, separated, or never married, and the children may be their biological, step-, or adopted children as long as the mother could claim them as her dependents. Single-mother families are not necessarily single-adult families, since single mothers may be living with other adults, including, for example, their parents or their unmarried partners or other related or unrelated individuals.⁶

We achieve two main goals in this paper. First, we show that, with a comprehensive list of control variables that include demographic, economic, and policy variables and a rich set of interaction terms, we are able to develop a model that rather successfully explains both the levels of and changes in welfare and work participation rates among single mothers across states, time, and various demographic groups for the whole 1980–2002 period. Second, using simulations of the model, we estimate the contributions of the various components of welfare reform and other

4. For instance, Grogger (2003a, p. 398) states, “Characterizing each reform is a difficult enterprise, however, which in conjunction with significant collinearity issues leads me to take a somewhat less ambitious approach here.”

5. In 2003 the Census Bureau renamed the March CPS the Annual Social and Economic Study.

6. Single women with dependent children have been the main recipients of benefits under both AFDC and TANF. Although single-parent families maintained by fathers, child-only families, and two-parent families where the primary earner is unemployed may also be eligible for benefits, single mothers account for a large majority of the caseload.

contemporaneous economic and policy changes to welfare and work participation rates. Of course, our confidence in our counterfactual decomposition relies, to a large degree, on the success of our empirical model in fitting the historical data on work and welfare participation rates.

Our main findings can be summarized as follows:

—The key economic and policy variables that contribute to the overall 23-percentage-point decrease in the welfare participation rate among single mothers from 1993 to 2002 are, in order of relative importance, work requirements (accounting for 57 percent of the decrease), the EITC (26 percent), time limits (11 percent), and changes in the macroeconomy (7 percent). This ranking holds for all years since 1997, although the contributions of the different factors differ by demographic group.

—The key economic and policy variables that contribute to the overall 11.3-percentage-point increase in the work participation rate among single mothers from 1993 to 2002 are, in order of relative importance, the EITC (33 percent), macroeconomic changes (25 percent), work requirements (17 percent), and time limits (10 percent). However, we find interesting differences in the relative importance of these variables across demographic subgroups and by time period.

These findings have important policy implications. It seems that although work requirements are highly effective at getting single mothers off welfare, they are not as effective at getting them to work. Indeed, whether single mothers work or not after leaving welfare depends crucially on conditions in the macroeconomy. One big success in public policy has been the expansion of the EITC, which contributes significantly to both getting single mothers off welfare and getting them to work. Our research highlights the crucial difference between “leaving welfare” and “working.” Indeed, we document the somewhat troubling development that nearly one-quarter of welfare leavers actually did not start work.

The paper is organized as follows. We begin with a selective critical review of some influential earlier studies. We then describe both the individual-level data from the March CPS and the economic and policy variables that we use in our empirical analysis. Next we give some descriptive statistics that emphasize the rich interactions between the economic and policy variables and the demographic characteristics of single mothers, and we use these to motivate our empirical model. Following a description of our empirical specification, we present and interpret our empirical estimates, discuss the fit of our empirical model, and use the

model to decompose the contributions of different economic and policy variables to changes in welfare and work participation rates. Finally, we draw conclusions and suggest directions for future research.

A Selective Review of the Welfare Reform Literature

In this section we discuss critically some of the key papers in the relevant literature and highlight the differences between their approaches and ours.⁷

Studies on the Effects of Time Limits

The aspect of the 1996 welfare reform that has received the greatest attention is the elimination of the entitlement status of welfare, and in particular the imposition of time limits on welfare receipt. PRWORA created a five-year lifetime limit on TANF receipt, in the sense that, except in limited special circumstances, states may not use federal funds to pay TANF benefits to any adult for more than a total of sixty months during that person's lifetime. But time limits did not originate with PRWORA. Many states had already instituted time limits on welfare receipt under federal waivers. Given the perceived centrality of time limits to the reform strategy, many studies have attempted to estimate the effects of time limits on welfare participation and other aspects of behavior.

Notable studies of time limits include those of Jeffrey Grogger and Charles Michalopoulos.⁸ These papers exploit the fact that, under both AFDC and TANF rules, only families with children under 18 are eligible for benefits. Thus time limits should have no (direct) impact on the behavior of single mothers whose children would reach the age of 18 before the limit could come into play.⁹ Therefore, in a before-and-after design, any

7. Many interesting and important papers are not discussed in this review. Grogger, Karoly, and Klerman (2002) and Blank (2002) provide extensive literature reviews.

8. Grogger (2000, 2003a) and Grogger and Michalopoulos (2003).

9. More generally, the strength of the incentive to conserve, or "bank," eligibility depends on the age of a woman's youngest child. If her youngest child is over 13, a newly imposed five-year time limit does not change her choice set at all. However, if her youngest child is under 13, then, the younger that child, the greater the option value of preserving welfare eligibility. Thus, *ceteris paribus*, time limits should enhance work incentives more for single mothers with younger children than for those with older children. Of course, time limits may also have indirect impacts. For instance, if time limits reduce welfare participa-

Table 1. Welfare Participation Rates of Single Mothers, by Age of Youngest Child

<i>Age of youngest child (years)</i>	<i>Before time limits^a (percent)</i>	<i>After time limits (percent)</i>	<i>Change</i>	
			<i>Percentage points</i>	<i>Percent</i>
0–6	41.3	23.8	–17.5	–42
7–12	23.1	13.3	–9.8	–42
13–17	16.0	11.0	–5.0	–31
All	32.0	18.8	–13.2	–41

Source: Reproduced from Grogger (2004, table 2). Data are from the March CPS from 1979 to 1999.

a. The year when time limits were introduced varies from state to state.

change in welfare participation among mothers with older children should be due solely to other time-varying factors besides the imposition of time limits (such as changes in general economic conditions or in other components of welfare reform). The change in participation rates for mothers with older children thus provides a baseline estimate of the impact of all these other factors. These mothers can therefore serve as a “control group” in estimating the effect of time limits. Under the assumption that all other time-varying factors affect the behavior of mothers with older and younger children in the same way, any incremental participation rate change among mothers with younger children isolates the effect of time limits.

Table 1, which is adapted from one of Grogger’s tables, illustrates this idea.¹⁰ A five-year time limit should not have affected the behavior of single mothers whose youngest child was between 13 and 17 years old. Thus the drop in their participation rate from 16 percent to 11 percent should be attributable entirely to other time-varying factors, such as work requirements or macroeconomic conditions. Next consider single mothers whose youngest child is 6 years old or less. These women are potentially affected by time limits, since they could use up the maximum five years of benefits long before their youngest child reaches age 18. Welfare participation dropped a much larger 17.5 percentage points among this group. Using these figures, we can estimate the impact of time limits using a difference-in-differences (DD) approach. Of the 17.5-percentage-point drop in participation for single mothers with young children, we attribute 5 percentage

tion among other groups in society (such as mothers with younger children), this may increase the stigma of welfare participation, which would indirectly impact participation rates among mothers with older children.

10. Grogger (2004).

points to the other factors besides time limits, since that is the change we observe for the control group. This leaves 12.5 percentage points as the drop in welfare participation attributable to time limits. This is a very substantial effect. It implies that 71 percent of the drop in welfare participation among mothers with young children was due to time limits.

As Grogger hastens to point out, however, this estimate relies on a number of strong assumptions.¹¹ Most critically, it supposes that all factors other than time limits have the same impact on single mothers whether their children are older or younger. This is a very strong assumption, since mothers with younger children differ from mothers with older children in important ways. To see this, note that table 1 also shows that, both before and after time limits were imposed, welfare participation rates were much higher among single mothers with younger children (41 percent before time limits) than among those with older children (16 percent). This alone illustrates the dramatic difference between the two groups and calls into serious question the assumption that they would be affected in the same way by other aspects of welfare reform or by the business cycle.

The fact that the baseline participation rates differ so greatly between the two groups creates another serious problem for the simple DD approach. Even if unmeasured time-varying factors did have a common impact across groups, to use a DD approach we need to know whether the “common impact” applies when we measure impacts in levels or in percentages. This point is also illustrated in table 1. The last column shows the percentage change in participation rates for each group following the imposition of time limits. The single mothers with older children had a 31 percent decline in welfare participation, whereas those with younger children had a 42 percent decline. So, if one assumes that the unmeasured factors have a common percentage-change effect across groups, the DD estimate of the effect of time limits on mothers with younger children is 11 percentage points. This implies that only 26 percent of the drop in welfare participation among this group of mothers was due to time limits. Thus time limits seem much less important when impacts are measured in percentages rather than levels.¹²

11. Grogger (2004).

12. To dramatize the possibility of this bias, consider the following thought experiment. Suppose that time limits had no effect on welfare participation, but that other, omitted factors (such as work requirements and work incentives) caused all single mothers to

We contend that there is only one way around this problem, and that is to do the hard work of trying to measure and control for a rich set of time-varying factors that may have affected people with different characteristics differently, and to allow for interactions between these factors and personal characteristics in constructing our model. The DD approach is *not* a panacea for dealing with unmeasured time-varying factors when the treatment and control groups are different, especially when they have different baseline participation rates.¹³

Recognizing this, Grogger extends the simple DD analysis described above to control for four specific time-varying factors that he believed might have different effects on women with younger children than on those with older children. Those time-varying factors are the unemployment rate, the minimum wage, the real level of welfare benefits (all measured at the state level), and a dummy variable for “any statewide welfare reform.”¹⁴ When these factors are controlled for, and state dummy variables and state-specific quadratic time trends are included, the estimated impact of time limits on welfare participation for single mothers with children age 6 and under drops to 8.6 percentage points.¹⁵ This is still 49 percent of the overall 17.5-percentage-point drop in participation for this group.

Thus Grogger’s results imply that time limits were a major factor driving down caseloads. His estimates of state unemployment rate effects are all insignificant, implying that the strong economy over the period did not play a significant role. His estimates do imply that falling real AFDC/TANF benefits had a significant impact on mothers with younger children. Interestingly, neither the time limit dummy nor the general reform

leave welfare. This would lead to a change of 41 percentage points for mothers with children 6 and under, and 16 percentage points for mothers with children ages 13 to 17. This would yield an estimate for the effect of time limits of 25 percentage points, when in reality the effect is zero. If instead it were known that the omitted factors operated on percentage changes rather than levels, we would get changes of –100 percent for both the first group and the second, for a (correct) difference of zero. But of course we have no way to know in advance which specification—levels or percentage changes—is the right one.

13. This criticism actually applies to many recent applications of the DD methodology, which have often involved situations where the “treatment” and “control” groups are rather different at baseline.

14. Grogger (2004).

15. We refer to the results in column 1 of table 5 in Grogger (2004), which we take to be his main results.

dummy nor the unemployment rate nor any of his other controls are significant for the single mothers with older children. Thus Grogger's results apparently attribute the 31 percent drop in welfare participation for this group to the state-specific time trends. These may be picking up the effect of the EITC expansion, a general change in "culture," or some other factor not controlled for in the model. Indeed, in a later paper that controlled for EITC expansion, Grogger found an even smaller effect of time limits on welfare participation: they now accounted for only about one-eighth of the decline in welfare use and about 7 percent of the rise in the employment rate since 1993.¹⁶ This is rather close to our own estimates, presented below, of 11 percent and 10 percent for the contributions of time limits to changes in welfare and work participation, respectively.

An important limitation of Grogger's approach is that all other aspects of welfare reform are summarized in his "any statewide welfare reform" dummy variable. This precludes him from estimating the effects of other specific policy changes. Furthermore, it will not adequately control for omitted factors if other reforms affect different demographic groups differently. As an example, one specific feature of welfare reform that Grogger omits, and which could lead to upward bias in his estimates of time limit effects, is the massive expansion of subsidized day care for low-income families that occurred largely as a result of PRWORA (figure 4). Under CCDF rules, funds may not be used to subsidize day care for children over 12 except in very rare instances (for example, for children with special needs). Hence the day care expansion should not have affected single mothers whose youngest child is 13 to 17 years old. And, obviously, subsidized day care could have a bigger effect on mothers with pre-school-aged children. That is, the effects of other contemporaneous reforms omitted from the analysis could indeed be age dependent. We note, somewhat facetiously, that if we chose to ignore time limits rather than day care, we could use table 1 to obtain a DD estimate of the effect of expanded day care spending.¹⁷

16. Grogger (2003a).

17. Using a structural model of welfare participation and labor supply estimated on data from the 1980s, Keane (1995) predicted that a policy of subsidizing single mothers' fixed costs of working (primarily day care and transport costs) would reduce their AFDC participation rate from 25 percent to 20.8 percent (a 17 percent decline) and increase their employment rate by 7 percentage points from a base rate of 60 percent. Thus our prior is that large effects of day care subsidies are plausible.

The later analysis of Grogger and Michalopoulos is less subject to these sorts of criticisms.¹⁸ They estimate the effect of time limits using data from a randomized experiment, the Florida Family Transition Program. This was a fairly small experiment in which welfare recipients in Escambia County, Florida, were randomly assigned to either a treatment group that was subject to a two- or a three-year time limit or a control group that was not.¹⁹ They estimate that the two-year time limit reduced welfare participation rates among single mothers with youngest children ages 3 to 5 by 7.4 percentage points (from a base rate of 40.3 percent) during the first two years after the time limit was imposed. This estimate implies significant effects of time limits, but it is difficult to translate it into a prediction for the aggregate welfare caseload, for two reasons: first, the estimate is based on a two-year limit, whereas most states have longer limits; and second, it conditions on a sample of women who had applied for welfare in the first place. Thus it tells us nothing about how time limits would affect entry into welfare.

Furthermore, we do not think it is possible to generalize the significant effects of time limits in the Florida context to the broader national context. Dan Bloom, Mary Farrell, Barbara Fink, and Diana Adams-Ciardullo (BFFA) provide an excellent discussion of how time limits have been implemented in practice in many states. They state that “as a relatively small pilot program . . . [the Florida program] was generously funded and heavily staffed,” and thus, “With small caseloads, workers were able to have frequent contact with participants.”²⁰ They go on to point out that “Recipients who came within six months of reaching their time limit and who were not employed were referred to specialized staff known as ‘transitional job developers,’ who worked intensively to help these individuals find jobs. The transitional job developers sometimes met with recipients several times a week, and they offered employers generous subsidies to hire their clients.” Finally, BFFA note that “. . . nearly all of those who reached the time limit had their benefits fully cancelled. Very

18. Grogger and Michalopoulos (2003).

19. A confounding feature of this experiment was that a child care subsidy was also provided to both groups. Thus the experiment does not estimate the effect of time limits alone. However, assuming no interaction between child care subsidies and time limits, the differences between the treatment and the control groups should net out the effects of child care.

20. Bloom and others (2002, p. 140).

few extensions were granted; only a handful of cases retained the child's portion of the grant; and no one was given a post-time limit subsidized job."²¹

This combination of intensive case management and strict enforcement of the time limit is wildly at variance with the norms under TANF. In fact, BFFA describe a system where, in practice, time limits are only sporadically enforced because extensions and exemptions are so common. They note that roughly 44 percent of the caseload reside in states such as Michigan, New York, and Vermont, which do not have time limits, or California, Maryland, and Washington, which only reduce (rather than terminate) benefits when the time limit is reached. Furthermore, several states, such as Oregon, stop the welfare time clock if a recipient is participating in required work or work-related activities, and many states, such as Connecticut, provide liberal extensions of the time limit if recipients have made a "good faith effort," which basically means meeting the requirements of the state TANF plan with respect to work, job search and training, and avoiding sanctions.

Thus, in many states, time limits are practically irrelevant. A typical comment is that of the U.S. General Accounting Office: "In Oregon, months count toward the time limit only if the family fails to cooperate, and the State has graduated sanctions resulting in a full family sanction for failure to participate [in required work activities]. Officials told us they do not expect any families to ever reach the State time limits in Oregon because, if families are cooperating, they can expect to receive cash assistance indefinitely (funded by the State after the waiver expires in the year 2002); if families are not cooperating, their grants will be terminated long before the time limit is reached."²² BFFA describe data on 54,148 TANF recipients who had reached the federal five-year time limit by December 2001. The bulk of these were in Michigan and New York, since these states implemented TANF relatively early on. But these states do not impose the federal limit. Of 5,143 recipients in the other states that did nominally impose time limits, BFFA report that 51 percent continued to receive TANF benefits under some sort of extension. The most common extension criteria were "good faith effort" (in Connecticut, South Carolina, and Tennessee), "disabled or caring for disabled family member" (in Georgia,

21. Bloom and others (2002, p. 142).

22. U.S. General Accounting Office (1998a, p. 55).

Louisiana, and Utah), “to complete education or training” (in Georgia), “high unemployment” (in Texas), and “other” (in Ohio).

Studies of Other TANF and TANF-Like Reforms

A number of previous studies have attempted to look more broadly at the whole range of factors that might drive caseloads. A paper by Rebecca Blank was a pioneering effort in this direction.²³ She examined the evolution of welfare caseloads by state and by year over the period 1977–95. Although her data were entirely from the pre-TANF period, a number of states had already instituted waivers in the early 1990s, making it possible to examine the impact of a number of TANF-like reforms.

The details of Blank’s specification are worth describing, because they guide much of the subsequent work in this area. Her dependent variable is the log ratio of a state’s AFDC caseload to the female population ages 15 to 44. Given that most AFDC recipients are in this age range, the dependent variable can be taken to approximate the percentage of women in this age group who participate in AFDC. This variable ranged from 6 to 8 percent over the sample period and was 7.4 percent in 1994. The policy variables include the state-specific AFDC “grant” for a family of three (that is, the benefit for a family with no earnings or outside income) and dummy variables for whether the state had been granted a waiver and, if so, whether the policies adopted under the waiver included time limits, enhanced work requirements, fewer exemptions from (or more severe sanctions for) failure to meet work requirements, or family caps. (A family cap is a policy whereby AFDC benefits are not increased by the usual per-child increment if a woman has an additional child while already on AFDC.) Controls for aggregate economic conditions were the state unemployment rate (and two lags of this variable), the median wage, and the 20th percentile wage. Blank also controlled for state demographics such as average educational attainment, the share of the population that were black, the share that were elderly, the share that were recent immigrants, and the share of households headed by single females.

Blank’s results imply that caseloads are mildly sensitive to the unemployment rate: the estimated elasticity of the welfare participation rate with respect to a sustained increase in the unemployment rate is roughly

23. Blank (2001).

0.25.²⁴ This means that a 3-percentage-point increase in the unemployment rate would raise the participation rate by about 11 percent after three years. Her results also imply that participation is quite sensitive to benefit levels: the estimated elasticity of the participation rate with respect to the benefit level is 0.56.

Blank's study has a few notable shortcomings. First, a salient feature of the data (figure 1) is that the AFDC caseload was quite flat from 1977 through 1989 (in the range of 3.5 million to 3.9 million families). But it rose sharply in the 1990–93 period (from 3.8 million in 1989 to 5.0 million in 1993), peaked in March 1994 at 5.1 million families, and then began to drop sharply in mid-1994. One might suspect that the bulge was due to the mild recession of the early 1990s. Before 1990, however, AFDC caseloads had never exhibited much cyclical sensitivity. In fact, Blank shows that half of the caseload increase in 1990–94 was due to increases in child-only and AFDC-UP cases.²⁵ Thus her dependent variable exaggerates the increase in the AFDC participation rate among single females age 15 to 44 during that period. Presumably, an ordinary least squares (OLS) estimate would attribute this exaggerated increase to the recession, leading to an overestimate of the effect of unemployment. Despite this, Blank notes that her model still does not succeed in explaining the increase in caseloads in 1990–94.

Second, Blank obtains very puzzling results for the effects of specific reform features. The coefficient on the “any major state welfare waiver” dummy implies that a waiver reduces the participation rate by roughly 11 percent. However, when this is broken down into a set of dummies for different aspects of waivers, the dummy for whether a state imposed time

24. The sum of the coefficients on the current and two lags of the unemployment rate is 0.037 (Blank, 2001, table 2). If $\log(P) = 0.038U$, where P is the participation rate and U the unemployment rate, then the elasticity of P with respect to U is $0.038U$. The mean unemployment rate in the data is 6.583 percent, so that at this mean the elasticity is 0.25.

25. The increase in AFDC caseloads during 1990–94 may have also been related to the 1986 Immigration Reform and Control Act (IRCA), which legalized 2.7 million undocumented immigrants residing in the United States since 1982, as well as certain seasonal agricultural workers, and made these legalized immigrants eligible for welfare after a five-year moratorium. Immigrants legalized under IRCA were more likely to be poor than immigrants who had entered legally, and legalization may have encouraged resident immigrants to apply for benefits for their children, even if they themselves were barred from aid receipt during the moratorium. Since most of these immigrants were legalized in 1987 and 1988, the five-year moratorium on welfare receipt ended by the beginning of 1994 (see MaCurdy, Mancuso, and O'Brien-Strain, 2000, 2002).

limits is insignificant (and has the wrong sign), and work requirements are insignificant as well. The dummy indicating that a state imposes harsher sanctions for failure to satisfy work requirements is estimated to have a significant *positive* effect on caseloads. The variables estimated to significantly reduce caseloads are dummies for reduced JOBS exemptions and for whether the state imposed a family cap. The latter policy is estimated to reduce the caseload by roughly 18 percent, which seems highly implausible. As Blank states, “the impact of family caps on the caseload in the short run should be minimal. It merely holds benefits constant for women who are already on the caseload, it does not remove anyone from the rolls.”²⁶

The Council of Economic Advisers (CEA) conducted a similar exercise using state-level data from 1976 to 1996, updated through 1998 in a second paper.²⁷ These papers use much sparser sets of controls than does Blank’s 1997 paper. The only nonwelfare factors included in the models are the current and lagged unemployment rates (along with state and year dummies). In the 1997 paper, specifications that include only a portmanteau dummy variable for “any statewide welfare waiver” imply that a waiver reduces a state’s caseload by roughly 5 percent.²⁸ When dummies for specific policies are included instead, the estimates are rather imprecise. The only clearly significant policy is stricter work requirement sanctions, which are predicted to reduce the caseload by roughly 10 percent.

It should be stressed that a fairly small amount of data underlies these estimates. For instance, according to Gil Crouse,²⁹ only five states had implemented benefit time limits by early 1996, with two more doing so in the second half of 1996. Two states implemented work requirement time limits in 1994, four more in 1995, and two more in 1996. Stricter work requirement sanctions were more common. Six states implemented these before 1995, five more in 1995, and eight more in 1996. Thus it was only in 1995–96 that a substantial number of states began to implement TANF-like policies.³⁰

26. Blank (2001).

27. CEA (1997, 1999).

28. CEA (1997, table 2, column 3).

29. Crouse (1999).

30. Schoeni and Blank (2000) use CPS data from 1977–99, thus including three years of post-TANF data. They also disaggregate state-level caseloads by age and educational attainment. They measure welfare reform using only waiver and TANF dummies, and they attempt to control for all other factors using a large set of state and time fixed effects (we

The 1997 CEA report notes that a one-year lead of the waiver dummy is significant. The estimates imply that a waiver reduces the caseload by roughly 6 percent in the year *before* it is implemented. The report points out that this could be an anticipatory effect: the knowledge that welfare policies will become stricter may deter women from welfare participation even before the waiver is implemented. But another explanation is based on policy endogeneity. It is widely accepted that the increase in welfare caseloads in 1990–93, and the increase in program costs that this induced, helped create the political momentum that led to implementation of waivers and ultimately TANF itself.³¹ However, by the time many states had implemented waiver policies in 1995–96, and certainly by the time that most had begun to implement TANF policies in 1997, a rapid decrease in the caseload had already begun.³² Any misspecified model that fails to capture the sharp decline in welfare caseloads beginning around 1995—before the implementation of most TANF-like policies—will tend to attribute these changes to the TANF and waiver dummy variables. The reason is simply that the model will produce large serially correlated residuals in the post-1995 period, and any variable that “turns on” in that period will help absorb those residuals. Thus what the CEA calls a “policy endogeneity” problem we prefer to call a misspecification or omitted variables problem.³³ The best way to deal with this problem is to look for additional

discuss their specification further later in the paper). They obtain the puzzling result that TANF had no significant effect on work participation.

31. For instance, according to the 2000 Green Book (U.S. House of Representatives, Committee on Ways and Means, 2000, p. 352), “Frustration with the character, size and cost of AFDC rolls contributed to the decision by Congress to ‘end welfare as we know it’ in 1996. Enrollment had soared to an all time peak in 1994, covering 5 million families . . . benefit costs peaked in fiscal year 1994 at \$22.8 billion,” and further, “By early 1995, many Governors pressed for a cash welfare block grant to free them from AFDC rules. The concept of a fixed block grant . . . was included in reform bills passed by Congress in 1995 and 1996; both were vetoed. But a third bill that included changes discussed during the 2 years of debate was enacted by Congress in July 1996 and was signed by President Clinton on August 22, 1996. By the time of TANF’s passage, AFDC enrollment had decreased to 4.4 million families.”

32. This can be seen quite dramatically in the state-by-state graphs of caseloads over time presented by Crouse (1999). By our count the graphs provide clear evidence that caseloads had begun to fall substantially before any implementation of waivers or TANF in at least thirty-three of the fifty states.

33. Even if policy were endogenous in the sense that increases in AFDC caseloads in 1990–93 induced the implementation of waivers and TANF policies, this would not by itself bias the estimates of policy effects. Only if the residuals are serially correlated would one get potential bias in the waiver and TANF coefficients. For instance, suppose that an

control variables that can successfully explain caseload evolution in the prereform period. This is the approach we take here.³⁴

It is interesting to note that, in a model with state fixed effects, our approach would not work. Consistency of OLS requires only that the covariates and the errors be contemporaneously uncorrelated (that is, that the policy variables be “predetermined”), whereas fixed effects estimators rely on “strict exogeneity” (that is, a lack of correlation at all leads and lags). Thus policy endogeneity would lead to inconsistent estimates in fixed effects models even if the residuals were serially independent. This is a strong argument for not including state fixed effects if we believe that policy endogeneity is present.

The CEA models certainly fail to explain both the increase in caseloads in 1990–93 and the decline beginning in 1995. Unemployment rate changes over this period—the only non-welfare-related explanatory factor in the CEA models—seem inadequate to explain the phenomenon, given the history of insensitivity of caseloads to unemployment. The 1997 CEA paper notes that “for the 1989–1993 period that saw a tremendous increase in the rate of welfare receipt . . . changes in unemployment can only explain about 30 percent of the rise . . . that leaves roughly 70 percent of the rise unexplained by this statistical analysis.”³⁵ Their model also attributes 34 percent of the decline in caseloads in 1994–96 to “other unidentified factors.” Thus a key challenge is to develop a model that can better account for caseload movements over time, particularly the pre-TANF decline in caseloads beginning in 1995. Unless a model can fit this pattern, any effects that it attributes to waiver and TANF policies may be spurious.

omitted variable was driving up caseloads in 1990–93 and then started to drive them down in 1995. The omission of this variable would generate serially correlated residuals. If one could find this variable and include it in the model, thus eliminating the serial correlation, the potential bias would vanish. The fact that the welfare policies were driven by caseload increases in the early 1990s would be irrelevant.

34. As CEA (1997) notes, another concern is that caseload increases in the early 1990s varied from state to state. If those states that had the largest caseload increases were most likely to implement waivers, then the states with the largest residuals in the early 1990s would be the ones most likely to implement waivers in 1995 and 1996. If the residuals exhibit persistence, then waivers in 1995–96 would be correlated with the 1995–96 residuals as well, inducing bias. Again, this can be thought of as a misspecification or omitted variables bias, since, if one could control for the omitted factor driving caseloads—and inducing serially correlated residuals—the bias would vanish.

35. CEA (1997, p. 8).

Robert Moffitt argues that the cyclical sensitivity of AFDC caseloads might have increased over time.³⁶ Thus, unless one takes a stand on the cyclical sensitivity of the caseload and how it has evolved over time, one cannot decide how much of the drop in welfare participation after 1994 was due to welfare reform and how much to the strong economy. If only aggregate data were available, these would leave one with a hopeless identification problem. However, Moffitt also pointed out that that cross-state variation in unemployment rates can, in principle, be used to resolve this problem. One could ask whether caseloads fell more or less in states where unemployment fell more or less, and one could even identify how the cyclical sensitivity of caseloads has varied over time, provided one assumes that it varies in the same way in all states. We today are in a much stronger position than previous researchers to identify these cyclical effects, because we can include data from the recession of 2001–02.

Studies of Non-TANF-Related Reform Policies

Other important policy changes that may have influenced the welfare and work decisions of single mothers in recent years are the expansions of Medicaid eligibility for low-income families not on AFDC and the expansion of the EITC. As Keane and Moffitt note,³⁷ the fact that single mothers would tend to lose Medicaid eligibility if they left AFDC created an important work disincentive before 1987. But a series of Medicaid eligibility expansions in 1987–2002 may have reduced this disincentive, by allowing single mothers with income above the AFDC/TANF eligibility threshold to continue to receive Medicaid benefits. Often eligibility for Medicaid expansions depended on the age of a woman's children.

Aaron Yelowitz attempted to quantify the effect of Medicaid expansions on work.³⁸ He measured the extent of eligibility expansion by a single variable, which he called GAIN%, defined as the difference between the Medicaid income eligibility threshold under the expansion and the AFDC income eligibility threshold before the expansion. Identification of Medicaid expansion effects came from the variation in GAIN% across states, over time, and across individuals. He used March CPS data from 1989 through 1992 to estimate a probit model for work participation as a

36. Moffitt (1999).

37. Keane and Moffitt (1998).

38. Yelowitz (1995).

function of GAIN%. To control for other factors that might vary across states and time, he also included year and state dummies. Yelowitz's estimates imply that the Medicaid expansion of 1989–92 led to a 1.2-percentage-point decrease in welfare participation and a 0.9-percentage-point increase in labor force participation among single mothers with at least one child under 15. However, as discussed earlier, for such a strategy to provide a consistent estimate of the effect of the policy variable in question, one has to make the strong and likely implausible assumption that all other time-varying factors, including all omitted policy variables, impact all single mothers in the same way, regardless of the ages of their children or their state of residence. Furthermore, we must know a priori whether the omitted time-varying factors affect the work participation of the “control” and “treatment” groups in terms of levels or percentages. Only then will the difference-in-differences methodology work.

Bruce Meyer and Dan Rosenbaum have undertaken a more comprehensive study of the effects of a wide range of factors on the work decisions of single mothers, but their focus is on the EITC.³⁹ They use CPS data for 1984–96 and incorporate changes in the EITC and other tax rates, AFDC and food stamp benefit levels, welfare time limits (under waivers), Medicaid expansion, and child care and training expenditures. Meyer and Rosenbaum's paper represented a significant advance over previous studies in that it controlled for a wide range of factors. Their empirical specification, however, did not control for other key TANF-like reforms under waivers, such as work requirements. Moreover, because their study used data only up to 1996, they do not address the separate contributions of various components of the 1996 welfare reform to the subsequent drop in caseloads. Meyer and Rosenbaum's estimates imply that changes in the EITC and other tax policies explain more than 60 percent of the increase in work among single mothers relative to childless single women in 1984–96. Somewhat unexpectedly, their estimates also imply that Medicaid expansions had a nonnegligible and negative effect on work participation.

We conclude with two general observations about all the studies we have described. First, they all use only dummy variables (such as whether or not a state has implemented a time limit) to capture policy effects. This is a problem because a time limit or other policy change will most likely affect rates of entry and exit from welfare, rather than simply inducing an

39. Meyer and Rosenbaum (2001).

immediate shift in the level of participation. The effect of such a policy thus builds gradually over time. In contrast, we explicitly construct measures of the time elapsed since particular policy changes might have begun to affect each single mother (based on her state of residence and demographics), thus allowing policy effects to develop gradually.

Second, all the studies we have described include state dummies to control for differences in welfare and work participation across states that the model leaves unexplained. As already mentioned, one reason for not using state fixed effects is that consistency of the fixed effect estimator requires the assumption of strict exogeneity, which we believe is invalid regarding policy changes. Furthermore, Keane and Kenneth Wolpin show how the use of state fixed effects can lead to seriously biased estimates of policy effects in a dynamic model.⁴⁰ For example, in a dynamic framework, a person decides whether to go on welfare or work or invest in human capital today based not just on benefits today but on expected future benefits as well. Suppose that each state has a typical level of benefit generosity that is persistent over time (for example, that Minnesota always has higher benefits than Alabama), but that benefits in both states fluctuate from year to year. These transitory fluctuations in benefits may have little effect on work and welfare participation decisions, which instead will be primarily driven by the permanent component of benefits. Hence a state fixed effects estimator may lead one to underestimate the effect of benefit levels. Using simulations of a dynamic model, Keane and Wolpin show that this problem can be severe.⁴¹

For these reasons we choose not to include state fixed effects in our models. Of course, this may create a problem if our control variables fail to explain the persistent differences in levels of welfare participation across states, and instead generate serially correlated residuals by state. If states with persistently negative residuals for welfare participation tended to adopt certain policies under TANF, one might falsely infer that these policies reduced participation. As we show later in the paper, our models do a reasonably good job of explaining the persistent differences in levels of welfare and work participation across states, so that we are not too concerned about this issue.

40. Keane and Wolpin (2002a, 2002b).

41. Keane and Wolpin (2002a).

To summarize, we feel that previous studies of welfare reform suffer from a number of important limitations. Typically, they examine only a subset of the many policy and economic environment variables that might affect welfare and work decisions. They often use state and time dummies to control for omitted time- and state-varying factors. This procedure is valid only under the assumption that such omitted factors affect all demographic groups equivalently and, even if this is true, that the analyst knows whether the equivalence holds in terms of levels or in terms of percentages. On the other hand, those studies that omit explicit year effects have not developed models that succeed in explaining the evolution of welfare participation over time at the national level, let alone broken down by state and demographic group.

Data

The data set used in this paper combines individual-level data from the March CPS with data on a rich set of economic and policy variables. In describing these data, we will also detail the sources of variation that we exploit to identify the effects of key economic and policy variables.

Individual Data

Our main data source is the series of March supplements to the Current Population Survey fielded between 1981 and 2003, covering activities in 1980–2002.⁴² The CPS is designed to provide a nationally representative sample by interviewing approximately 60,000 households. The sample size was increased in 2001 and 2002 to improve estimates of children's health insurance coverage by state, for the purpose of allocating federal funds under the State Children's Health Insurance Program (SCHIP) established in 1997. The CPS asks retrospective questions about demographics, work activities, and income. Questions about demographic variables, such as age, refer to the week before the interview; those about income variables refer to the previous calendar year; and those about work activity, such as hours worked and major occupation, refer to both periods.

42. Our CPS sample is extracted using the CPS Utilities produced by Unicon Research Corporation.

Our unit of analysis is families headed by single mothers. Since we condition on single-motherhood, we take marital status and the presence of children as exogenous. Of course, changes in welfare rules could affect marriage and fertility, but existing empirical work suggests that these effects are small.⁴³

For purposes of constructing a data set on single mothers, it is important to note that the CPS is organized around households defined by a unique address, for example a house or an apartment. A household may contain more than one family, with the person who rents or owns the house considered the head of the household. We select female-headed families or subfamilies as the unit of analysis.⁴⁴ We then count the number of dependents in each female-headed family or subfamily. Note that the dependent children are not necessarily the woman's biological children. Stepchildren or adopted children, grandchildren, and other unrelated children whom the woman lists as dependents are also counted.

The CPS survey asks the respondent to provide detailed demographic information (including age, race, education, and marital status) for every household member. We construct the age composition of the woman's children by counting the number of dependent children at each age. This is an important step because, as we discuss below, whether a woman is subject to particular welfare rules (such as work requirements) or eligible for particular benefits (such as child care subsidies) often depends on the precise ages of her children.

We construct our welfare utilization measures from the family's reported sources of income over the previous calendar year, and we analyze work participation decisions based on the average hours worked in that year. Specifically, we consider a single woman a welfare recipient if her income from public assistance (Unicon recode variable *incpa*) is positive.⁴⁵ The

43. See Moffitt (1992).

44. Specifically, a woman selected into our analysis must satisfy two conditions. First, she must be the head of the primary family or a subfamily, which also means that she must have dependent children. This is ensured by selecting the Unicon recode variable *_hhrel* to equal 1, 3, 5, 8, 10, 13, 15, 18, 20, 23, 26, 30, 32, 35, 38, 41, or 43. Second, her marital status, given by the Unicon recode variable *_marstat*, must be either 3 (separated), 4 (widowed), 5 (divorced), or 6 (never married).

45. The exact wording varies by year, but the essence of the question is, "How much did ___ receive in public assistance or welfare in the previous year?" and the answer is coded as *incpa*. From 1988 on, the survey also asks about the number of months in which public assistance or welfare is received. Note that *incpa* will capture cash assistance but not in-kind assistance, such as food stamps.

employment variables come directly from the CPS, which includes the “hours worked per week last year” (hrslyr). We recorded a woman as working full-time if she works for thirty-two hours or more a week, and part-time if she works between eight and thirty-two hours a week.

Policy Data

COMPONENTS OF WELFARE REFORM. An important contribution of the paper is the comprehensive documentation of the many welfare policy changes that occurred at the state level over the 1980–2002 period. We collected detailed information about states’ policies from many different sources.⁴⁶ The rest of this section describes the different policy components in detail.

Time Limits. PRWORA prohibits states from using federal TANF funds to provide benefits to adults beyond a sixty-month lifetime time limit (except that 20 percent of a state’s caseload may be exempted). Many states have opted for shorter time limits, whereas others have opted to use their own funds to provide benefits beyond the federal limit. Some states implemented their own time limits under waivers before PRWORA was enacted.⁴⁷

To understand the set of variables we use to capture the possible effects of time limits, it is useful to examine the theory of how time limits can affect behavior. A key point is that time limits may have both anticipatory and direct effects. The direct effect arises simply from the fact that a person who reaches the time limit becomes ineligible for further benefits (assuming the limit is enforced). The anticipatory effect is subtler. The basic idea is that a forward-looking person faced with time-limited

46. Sources include the State Policy Documentation Project, the U.S. General Accounting Office (1997, 1998a), Gallagher and others (1998), Johnson, Llobrera, and Zahradnik (2003), Hotz and Scholz (2003), the U.S. Department of Health and Human Services (including its Office of Family Assistance), the U.S. Department of Agriculture, the Center for Law and Social Policy, the Urban Institute, the Bureau of Labor Statistics, the National Governors Association, the Center on Budget and Policy Priorities, various issues of the U.S. House of Representatives’ Green Book, the Internal Revenue Service, and various state TANF policy handbooks.

47. A distinction is sometimes made between when a state implemented its TANF plan and when it began counting months toward time limits. Arkansas, California, Ohio, and Oregon started counting months toward time limits well after their initial TANF implementation dates. We use the actual counting date as the effective date for time limits in our analysis.

welfare benefits should try to conserve (or “bank”) her months of eligibility and use them only when truly necessary.

Consider a simple framework where a woman decides each month whether to receive welfare or go to work. A myopic person who maximizes current income would choose to participate in welfare so long as it generated one dollar more in income than she could earn by working (net of the cost of working). But a forward-looking person would choose welfare over work only if the gap between benefits and earnings were substantial. Why use up a month of welfare eligibility just to get a few extra dollars? In some future month she may confront a situation where only very low paying jobs are available, so that welfare benefits far exceed her potential earnings. It is therefore best to conserve her months of welfare eligibility for such circumstances.

Stated more formally (see appendix A), in a dynamic framework, such a woman should make welfare participation decisions by comparing the value of current-period welfare benefits with the value of current-period potential earnings *plus* the option value of conserving a month of benefit eligibility. As Grogger and Michalopoulos point out, this option value is, *ceteris paribus*, an increasing function of the time horizon over which benefits may be used (that is, the number of years until the woman’s youngest child reaches 18).⁴⁸ It is also, *ceteris paribus*, a decreasing function of the stock of remaining months of eligibility (that is, the option value of preserving a month of eligibility is greater when one has only one month left than when one has sixty).

Our empirical models include several variables designed to capture both the direct and the anticipatory effects of time limits—both those created under TANF and those created earlier under AFDC waivers. These variables and others used in the study are defined in table C1 in appendix C. Each variable has up to three subscripts: *i* for individual, *s* for state, and *t* for year. Thus the subscripts enable one to see whether each variable varies across states, across people, or both.

At the most basic level, we include a dummy variable for whether a state imposed a time limit in a given year (DTL_{st}), as well as a dummy for whether the time limit could have been binding for a particular woman (DTL_HIT_{ist}), given the ages of her children. A woman whose oldest child is *x* years old cannot have received welfare for more than *x* years.

48. Grogger and Michalopoulos (2003).

The time limit cannot bind for this woman unless x exceeds the limit, regardless of how many years ago her state implemented time limits. Thus the year in which time limits may first bind varies across women in the same state.

Note that DTL_{st} captures an anticipatory effect of time limits, and DTL_HIT_{ist} a direct effect. We also include variables that allow the anticipatory and direct effects of time limits on welfare and work decisions to develop gradually over time. First, we construct a variable called “months elapsed since the implementation of time limits” ($MONTH_SINCE_TL_START_{st}$). Second, we construct for each single mother a variable called “months elapsed since the time limits could first potentially bind” ($MONTH_SINCE_TL_HIT_{ist}$).

To evaluate the importance of the anticipatory effect of time limits, we construct two more variables motivated by the theory presented in appendix A. First, the option value of banking welfare eligibility increases with the time horizon over which a woman will be categorically eligible for benefits. This is the remaining time until her youngest child will reach age 18. We call this variable $REMAINING_CHILD_ELIG_{ist}$. Second, the option value of banking welfare eligibility decreases with the stock of eligible months that a woman currently possesses. We call this variable $REMAINING_TL_ELIG_{ist}$. To construct this measure, we first calculate the maximum number of months that a woman could have received welfare since her state started her “clock.” Subtracting this from the state time limit tells us the minimum stock of months that the woman possesses.

At this point it is worth commenting on our overall strategy in constructing covariates. We assume that a woman’s demographics, the welfare policy rules she faces, and the economic environment in her state are all exogenous. Thus, to maintain a true reduced-form specification, every covariate we use as a determinant of welfare or work participation should be a function of these demographic, policy, and economic environment variables. One can see the effect of this strategy quite clearly by looking at how we constructed covariates to measure the effects of time limits. For instance, we do not want to use a woman’s actual welfare participation history to construct the remaining months on her time limit clock, because actual participation decisions are endogenous. Similarly, in the construction of $REMAINING_CHILD_ELIG_{ist}$, we ignore the fact that a woman can always extend her months of categorical eligibility by having another

child. $\text{REMAINING_CHILD_ELIG}_{ist}$ is a function only of a woman's current demographics and state policy variables, and so it is certainly an exogenous variable driving current decisions.

A key point is that Michigan, New York, and Vermont have chosen to use state funds to provide benefits to families beyond the sixty-month federal limit.⁴⁹ In other words, these states do not have effective time limits.⁵⁰ This is a key source of variation in the data that helps identify the effect of time limits on welfare and work participation. To preview our finding that time limits have had small effects on welfare participation, we note that in Michigan the number of families on welfare dropped by 58 percent from August 1996 to June 2002, while the number of individual recipients dropped by 62 percent. Over the same period the number of families on welfare in New York dropped by 63 percent, while the number of recipients dropped 68 percent. These declines are close to the national average, suggesting that time limits are not the main factor underlying the dramatic drop in welfare participation since 1996.

Another important source of variation across states is the penalty that is imposed when a time limit is reached. Among states with effective time limits, six (Arizona, California, Indiana, Maine, Maryland, and Rhode Island) continue to provide the child portion of benefits to families even after the time limit is reached. As we discuss in appendix A, this substantially reduces the impact of time limits. Therefore we constructed a measure for each state of how benefits are reduced when the time limit is reached.

Work Requirements and Exemptions. Under PRWORA, states must require parents who receive TANF assistance to participate in "work

49. A common mistake in the literature, and in some data sources as well, is to assume that New York has a sixty-month life time limit. According to the New York State Comptroller's Office, after the TANF time limit is reached, the state will provide Safety Net Assistance (SNA) to the family in the same amount as the family's TANF grant. Twenty percent of the monthly payment standard is paid in cash for a personal needs allowance, and the rest is given on a noncash basis. Thus New York does not have a true time limit.

50. As already discussed, Oregon has a formal time limit, but it, too, is irrelevant because anyone who satisfies the work requirement for a given month does not have that month count toward the time limit, and anyone who does not satisfy the work requirement has benefits terminated immediately. Very recently, Arizona and Massachusetts have revised their TANF plans to use state funds to provide benefits to families beyond the sixty-month federal limit. This change is too recent to be relevant for our empirical work.

activities” after a maximum of twenty-four months.⁵¹ Many states have chosen to adopt shorter work requirement time limit clocks. States adopted their first TANF plans over the period from October 1996 through January 1998 and adopted revised TANF plans roughly two years later. Under the initial TANF plans, twenty states required benefit recipients to start participating in work activities immediately. Under the revised TANF plans, twenty-five states required immediate work participation. Most states that do not impose an immediate work requirement have adhered to the twenty-four-month maximum allowed under the federal law. The requirement that recipients participate in work activities may increase the disutility of welfare participation, leading to reductions in welfare caseloads and increased work among single mothers.

Section 407, paragraph (b)(5), of PRWORA gives states the option to exempt single parents with a child up to 1 year of age from work requirements. However, many states, such as California, have chosen to exempt only those single mothers with children under 3 or 6 months of age, and a few have granted longer exemptions. Thus there is considerable variation in the variable we call “age of child exemption from work requirements” ($CHILD_EXEMPT_AGE_{st}$). We use this variable, in conjunction with the state-specific work requirement time limit and the age of the woman’s youngest child, to construct an indicator for whether a woman could be subject to a work requirement. We call this variable SWR_{ist} .

Thus we have two key sources for the identification of the effects of work requirement time limits. First, because of the variation in when states implemented their TANF plans and in the length of their work requirement time limit clocks, there is substantial variation across states in how early a single mother could have been subject to binding work requirements. For instance, under AFDC waivers, work requirements could have come into force as early as mid-1994 in Iowa, October 1995 in Michigan, and mid-1996 in Wisconsin. TANF work requirements could have been binding as early as the fall of 1996 in Alabama, Connecticut,

51. Work activities as defined in PRWORA include “(1) unsubsidized employment; (2) subsidized private sector employment; (3) subsidized public sector employment; (4) work experience (including work associated with refurbishing of publicly assisted housing); (5) on-the-job training; (6) job search and job readiness assistance; (7) community service programs; (8) vocational educational training; . . . and (12) the provision of child care services to an individual who is participating in a community service program.”

Florida, Indiana, Kansas, Nebraska, New Hampshire, Oklahoma, Oregon, and Utah. On the other hand, work requirements were not binding until December 1998 in New York, January 1999 in Louisiana, February 1999 in New Jersey, March 1999 in Pennsylvania, and July 1999 in Illinois, Minnesota, and Missouri.⁵²

Second, as already noted, we can exploit individual variation based on children's ages. For example, assume that two otherwise similar women living in different states have both been on TANF long enough to have reached their state's work requirement time limit. Suppose that each has a youngest child who is 9 months old. Suppose further that their states have similar policies, except that one state exempts women with children under 12 months old and the other exempts only women with children under 6 months old. Then only the woman in the first state is exempt from the work requirement, and any difference in welfare participation and work behavior between these women will provide additional evidence on the effects of work requirements. Similarly, take two otherwise similar women living in different states, each of whom has just one child, who is 18 months old. Suppose their states have similar policies, except that one imposes an immediate work requirement whereas the other imposes a work requirement only after twenty-four months on welfare. The woman in the first state may be subject to a work requirement, but the woman in the second cannot be. Since her only child is only 18 months old, she cannot yet have been on welfare for twenty-four months.⁵³

Besides the exemption based on age of youngest child, many states allow other exemptions from work requirements under TANF. These include exemptions for single parents with children under age 6 who are unable to obtain child care, and for recipients who are disabled or have a disabled household member.⁵⁴ We call the total number of these exemptions

52. We calculate that the fraction of women who were potentially subject to a work requirement ($SWR = 1$) was 5 percent in 1995 and then rose to 16 percent in 1996, 46 percent in 1997, 62 percent in 1998, 85 percent in 1999, and 91 percent in 2000. It then stabilized at about 91 percent in 2001–02.

53. It is important to understand how the exemption for age of youngest child interacts with the work requirement time limit clock. Suppose that a state has a twenty-four-month time limit and that mothers whose youngest child is less than 12 months are exempt. If a woman is on welfare starting from the time the child is born, then when the child reaches 1 year of age she will have just 1 year left on the clock.

54. States must maintain certain work participation rates among TANF recipients in order to avoid penalties to their TANF block grants. Originally, 25 percent of all families receiving assistance were required to participate in work activities for at least twenty hours

$N_WR_EXEMPTION_{st}$. States also differ as to whether they impose a full or a partial benefit sanction if a recipient does not satisfy the work requirement. A “partial” sanction generally means that only the adult portion of benefits, and not the children’s portion, is denied. In 1996 nine states imposed a full sanction. That number increased to twenty-three in 1997 and stayed close to thirty from 1998 onward. We call the dummy variable indicating imposition of a full sanction $DFULLSANCTION_{st}$.⁵⁵ We view both the sanction variable and the exemption variable as indicators of the strictness with which a state enforces its work requirement time limit, and we interact the work requirement variables with these measures of strictness.⁵⁶

Finally, work requirements can, in theory, have anticipatory effects just as time limits do. If a state adopts a work requirement with a twenty-four-month time limit before the requirement is triggered, this creates an incentive to avoid welfare participation even before the twenty-four months are used up. One reason is to conserve time on the clock. Another reason is that, since the time limit reduces expected future welfare participation, it increases the value of human capital investment today. Thus we also include in our models a dummy for whether a state has a time limit in effect ($DWORKREQ_{st}$).

Benefit Reduction Rates and Earnings Disregards. The AFDC program always imposed a “tax” on a recipient’s earnings while on welfare, called the benefit reduction rate (BRR). Allowance was made for deductions for work and child care expenses, and over the history of the AFDC program the amounts of these work expense deductions were changed several times, as was the BRR itself. Notably, the BRR was decreased from 100 percent to 67 percent in 1967, but it was raised back

a week. The required rate was gradually raised to 50 percent in 2002, and the hours requirement was raised to twenty-five hours in 1999 and thirty hours in 2000. However, these requirements were relaxed for states that achieved substantial caseload reductions. Because caseloads fell so dramatically after 1996, states were rarely subject to significant participation rate requirements.

55. States often have a more lenient sanction policy for first-time violators of work requirements. Although we have information about these first-time sanction rules, we use only the “ultimate” sanction rule in our analysis. There is a high correlation between the first-time and ultimate sanctions.

56. Pavetti and Bloom (2001) classify twenty-five states as “strict” and thirteen as “lenient” in terms of the benefits denied to families of noncompliant individuals. Their classification is roughly consistent with our direct classification of states with full versus partial sanctions.

to 100 percent in 1982. Starting that year the work expense deduction was set at \$90 a month, and an additional child care expense deduction was introduced.

In addition, in an effort to encourage work among participants, the AFDC program at various times in its history included “earnings disregards.” That is, for a specified time after an AFDC recipient started a job, a part of her earnings (above and beyond the work and child care expense deductions) would not be subject to the BRR. In general, this earnings disregard consisted of a fixed component (for example, the first \$30 of monthly earnings) and a variable component (for example, one-third of earnings beyond the first \$30) and applied only during the first several months of work.⁵⁷ Starting in late 1992, again in an effort to encourage work, many states used waivers to enhance their earnings disregards. PRWORA did not mandate specific disregard policies, and, as a result, substantial heterogeneity has emerged in how states set disregards. Many states have expanded disregards and allowed them to apply indefinitely. For instance, under its TANF plan implemented in January 1998, California set the fixed portion of the monthly disregard at \$225 and the variable portion at 50 percent, with no phase-out over time. Since the variable part of the disregard is not phased out, it acts just like a BRR of 50 percent, and this is in fact how we code it. Across states, as of 2002, fixed disregard amounts varied from zero to \$252, and variable disregards ranged from zero to 100 percent.

Obviously, earnings disregards, the BRR, and work expense deductions directly affect a woman’s incentive to work by altering her effective after-tax wage rate. A lower effective tax rate makes welfare receipt more attractive. Furthermore, as we discuss in appendix A, effective tax rates also affect the incentive to bank months of eligibility when time limits are present. The higher the effective tax rate, the greater the incentive to forgo participating in welfare in a month when work can be found.

Diversion Programs. Under TANF many states have developed “diversion” programs under which new TANF applicants can receive a

57. After 1982 the rule was as follows: For each of the first four months of work, the first \$30 of earned income, plus one-third of the remainder, was disregarded when calculating the monthly benefit. After four months and until one year, only the \$30 monthly disregard continued. After one year there was no earnings disregard. This means that, after one year, a recipient’s grant amount was reduced by one dollar for every dollar she earned above the \$90 work expense deduction.

few months' worth of benefits up front if they agree not to participate in TANF for some stated period of time. A typical program may offer three months of benefits up front to a person who agrees to stay off "welfare" for three months. We view this as largely an accounting device to make TANF caseloads appear smaller, and so we code such diversion payment recipients as welfare recipients. Eight states, however, have introduced what we regard as genuine diversion programs, whereby TANF applicants agree to stay off welfare for an extended period in return for short-term cash payments (or loans) whose value is well below the maximum value of the forgone benefits.⁵⁸ In the empirical analysis we simply introduce a dummy variable to indicate whether the woman lives in a state with a genuine diversion program.

Child Support Enforcement and Treatment of Child Support Income. Although nonpayment is widespread, child support is an important source of income for single women with dependent children (see table 4 below). Under AFDC, recipients were required to assign child support collections to the welfare agency. States were then required to pass through the first \$50 of monthly child support payments to the family. This pass-through income was disregarded for purposes of benefit calculation. Between January 1993 and August 1996, states requested and received waivers of a number of AFDC provisions related to child support enforcement. These waivers sometimes involved changing the pass-through amount or allowing single mothers to keep child support payments, in which case they would be subject to certain disregards just like earned income. Under TANF, all states have discretion to set their own policy in terms of pass-through or disregard of child support payments.

The Child Support Enforcement and Paternity Establishment (CSE) program was enacted in 1975 to address the problem of nonpayment of child support owed by noncustodial parents. CSE has programs to help locate absent parents and establish paternity. The CSE administrative expenditure is an important indication of how likely it is that a single woman will be able to collect child support. Figure 4 showed the large increase in CSE expenditure, from \$2.92 billion in 1996 to \$5.14 billion in 2002, a 76 percent jump. To measure state-level CSE activity, we take state-level CSE expenditure and divide it by the state population of single

58. These eight states are Arkansas, Florida, Kentucky, Idaho, Montana, Texas, Washington, and Wisconsin.

mothers (excluding widows).⁵⁹ This, combined with variation in CSE spending across states and over time, provides the three key sources of variation that identify the effect of child support enforcement expenditure on welfare and work participation.

In terms of the incentives created, there are important interactions between CSE expenditure and the pass-through and disregard rules. Since child support payments are heavily taxed under TANF rules in many states, enhanced child support collections make welfare less appealing. On the other hand, enhanced pass-throughs or disregards may reduce this incentive.

Child Care Subsidies and the Child Care and Development Fund. In the late 1980s several new programs expanded federal support for child care. The Family Support Act of 1988 created two programs, AFDC Child Care and Transitional Child Care. AFDC Child Care was designed as an entitlement for single parents on AFDC who were working or enrolled in job training or education programs. Transitional Child Care provided a temporary child care subsidy to single parents with young children for twelve months after they left AFDC to start working. Both programs used AFDC participation as an eligibility criterion. The Omnibus Budget Reconciliation Act of 1990 (OBRA 1990) created the Child Care and Development Block Grant and the At-Risk Child Care program. These programs gave states funds with which to subsidize child care for low-income working families who were not on AFDC. However, unlike AFDC Child Care and Transitional Child Care, these benefits were not an entitlement. PRWORA consolidated these four preexisting programs into the Child Care and Development Fund. The CCDF provides federal funds to the states to use in providing child care subsidies to low-income working families, whether or not these families are current or former TANF recipients. Under the CCDF a great deal of heterogeneity has emerged in the design of states' child care subsidy programs. In particular, many states ration benefits, and states differ in terms of whether they give priority to low-income families who are on TANF or to those just transitioning off TANF.

We use state CCDF expenditure per single mother as a measure of the availability and generosity of child care subsidies in a state. A key factor

59. CSE expenditure should not impact the work or welfare decisions of widows, who do not have ex-husbands from whom to collect alimony or child support.

identifying the effect of these subsidies is that they are essentially irrelevant for women whose children are older than 12, since they are not eligible for subsidies except in rare instances (for example, for children with special needs). Also, the effect of child care subsidies is presumably stronger for women whose children are not yet of school age.

As we discuss in appendix B, an important aspect of PRWORA is the maintenance-of-effort requirement, which requires each state to maintain spending on assistance for needy families at a minimum of 75 percent of its pre-1996 level in order to receive the full TANF block grant. The MOE requirement interacts with the CCDF in an important way. The CCDF funding system is rather complex, consisting of federal funds to which states are entitled, plus federal matching funds that require state contributions, plus discretionary state contributions, including a certain level of funds that states are allowed to transfer out of the TANF block grant. But the key point is that the state part of CCDF spending counts as MOE spending. Thus, when welfare caseloads began to drop unexpectedly rapidly after 1996, causing state spending on TANF cash assistance to drop, the states shifted substantial resources into the CCDF as one way to achieve the MOE requirement. This dynamic was partly responsible for the rapid growth in total expenditure on CCDF from 1996 to 2002 (figure 4).

An alternative to using CCDF expenditure per single mother as a measure of the generosity of a state's child care program would be to use detailed program parameters, such as the monthly income limit for eligibility and the co-payment rate, which are state-specific and have varied over time within states. We choose not to use this approach because of the problems created by rationing. A state with a seemingly generous program (for example, a high income eligibility threshold and a low co-payment) will tend to have a longer waiting list. Thus program generosity is more accurately measured by the state's actual expenditure per case than by the income eligibility threshold and co-payment rates.

CONTEMPORANEOUS POLICY CHANGES. Our data set also contains detailed information about state policies other than those directly related to AFDC and TANF.

Earned Income Tax Credit. The EITC, enacted in 1975, is a refundable federal income tax credit that supplements wages for low-income working families. Major expansions of the EITC occurred in 1986, 1991, and 1994–96. Because of these expansions, the number of families receiving EITC increased from 6.2 million in 1975 to 19.5 million in 2000, and

total EITC payments increased from \$1.25 billion to more than \$31 billion (figure 3).⁶⁰

The EITC rules specify four parameters: a phase-in rate, a phase-out rate, a phase-in income range, and a phase-out income range. These parameters depend on family size. After the expansions of the mid-1990s, the EITC became a sizable wage subsidy to low- and moderate-income families. Thus it may provide an important work incentive. For example, in 2003 the phase-in and phase-out rates for a family with one child were 34 percent and 15.98 percent, respectively. The phase-in annual income range is from zero to \$7,490, and the phase-out range is from \$13,730 to \$29,666. Thus a single mother with one child with taxable income between \$7,490 and \$13,730 would receive an EITC of \$2,547. The EITC phase-in rate is even higher (40 percent in 2003) for families with two or more children.

As of 2003, seventeen states had enacted their own EITCs that supplement the federal credit. Most of these were enacted in the 1990s. Our econometric analysis combines the federal and state EITC programs and characterizes them by two parameters: the phase-in rate and the maximum credit amount.⁶¹ Many sources of variation help identify the effects of the combined EITC. One source is variation across time. For example, the federal EITC phase-in rate for families with one child increased from 10 percent in 1980–84 to 14 percent in 1987–90, 16.7 percent in 1991, 17.6 percent in 1992, 18.5 percent in 1993, 26.3 percent in 1994, and 34 percent in 1995, where it has remained since. Second, since 1991 a different EITC phase-in rate and maximum credit have applied to families with one child than to families with two or more children, thus introducing variation across individuals. Third, the implementation of state EITC programs at different times and with different parameters has introduced variation across states.

Food Stamps. The food stamp program provides coupons that can be exchanged for food at participating stores. The value of the coupons to which a family is entitled depends on a grant level, which depends on family size, and a benefit reduction rate, which is applied to income. Unlike AFDC/TANF benefits, food stamp benefit levels are set at the federal level, and the same rules apply in all states except Alaska and Hawaii.

60. See also U.S. House of Representatives, Committee on Ways and Means (2000, p. 813).

61. We collect state EITC information from Johnson, Llobrera, and Zahradnik (2003), Hotz and Scholz (2003), and state government websites.

We collect the food stamp program parameters directly from the U.S. Department of Agriculture. Currently, the food stamp benefit reduction rate is 30 percent.

Medicaid and SCHIP. AFDC/TANF participants have had health insurance coverage provided by Medicaid since the inception of the Medicaid program in 1965. Since 1987 a number of expansions of Medicaid eligibility have enabled single mothers, under various circumstances, to leave AFDC/TANF while maintaining Medicaid coverage. Between 1987 and 1990 several legislative options and mandates were enacted to expand Medicaid eligibility for pregnant women, infants, and children. OBRA 1989 required states to cover all pregnant women, as well as all children below age 6, living in families with income at or below 133 percent of the federal poverty line. OBRA 1990 required states to phase in coverage of children born after September 30, 1983, and living in families with income below the poverty line, until all children through age 18 were covered. As of October 1, 1997, children 14 years of age and younger were covered in all states, and the upper age limit of 18 was reached in all states in October 2002.

The States Children's Health Insurance Program (SCHIP), established under the Balanced Budget Act of 1997, appropriated roughly \$24 billion in federal grants over five years for states to use to provide health insurance to uninsured children under age 19 in families with incomes below 200 percent of the federal poverty line. SCHIP covers approximately 5.3 million children nationwide. States are using this new grant money to expand Medicaid, to develop new programs or expand existing programs that provide health insurance, or both.

We collected Medicaid rules for each state since 1987 (and SCHIP rules since 1997) from the annual Maternal and Child Health updates of the National Governors Association.⁶² These updates provide detailed information on the age limits of children covered by Medicaid (independent of welfare status) and the age-specific income eligibility thresholds (as a percentage of the poverty line). We combine these rules with the ages of the children of each single mother to obtain the variable $MEDICID_PCT_{ist}$, which measures the percentage of children who would be covered by Medicaid if their mother left welfare but earned less than the income threshold for Medicaid eligibility, which is coded by the

62. The updates from 1990 to 2002 can be found at www.nga.org.

variable $MEDICAID_FPL_{ist}$. Since the income threshold varies by age of the child, we used the threshold applicable to the woman's youngest eligible child as a percentage of the federal poverty line in constructing $MEDICID_FPL_{ist}$.

MACROECONOMIC VARIABLES. We include several variables in our model to control for state and national economic conditions. We obtain state unemployment rates from the Bureau of Labor Statistics. From the Urban Institute–Brookings Tax Policy Center we obtain information on personal and standard income tax deductions (deflated by the consumer price index) and the federal income tax rate for the lowest bracket. Data on minimum wage rates are collected from the Department of Labor website. Finally, we construct the 20th percentile wage rate for each state (deflated by the consumer price index) from CPS data.

Descriptive Statistics on Single Mothers

Our data set contains 127,119 observations on single mothers 18 years and older over 1980–2002. Here we provide descriptive statistics about the single-mother population and their welfare and work participation over that period.

Demographics

Table 2 summarizes basic demographic information about single mothers. The racial composition of single mothers has been very stable over time, with about 62 to 65 percent white and 32 to 35 percent black. On the other hand, there has been a dramatic and steady increase in the proportion of never-married single mothers, from 15.6 percent in 1980 to 41.3 percent in 2002. In fact, in 1997 “never married” overtook “divorced” as the most common marital status among single mothers. The fact that the proportion of never-married single mothers continued to increase after 1996 is interesting, as an explicit objective of PRWORA was to lower the incidence of out-of-wedlock births.

Table 2 also shows a slow downward trend in the average size of families headed by single mothers. The proportion of single mothers with only one child increased from 48.3 percent in 1980 to 54.5 percent in 2002. The share of single mothers with four or more children decreased

Table 2. Demographic Characteristics of Single Mothers, 1980–2002^a

Year	Race				Marital status				No. of children			Educational attainment			
	White	Black	Other	Separated	Widowed	Divorced	Never married	One	Two	Three or more	High school dropout	Finished high school	Some college ^b	Finished college	
1980	64.6	33.3	2.1	24.7	13.8	46.0	15.6	48.3	31.7	20.0	34.5	39.1	19.5	7.0	
1981	64.9	33.0	2.1	22.8	12.0	43.9	21.3	49.1	31.3	19.6	32.8	40.3	19.2	7.7	
1982	62.3	35.3	2.4	21.3	10.9	43.6	24.2	50.3	30.5	19.2	32.2	40.8	19.4	7.6	
1983	63.0	34.6	2.5	20.6	10.9	44.5	24.0	52.7	29.7	17.6	31.8	41.4	19.4	7.5	
1984	63.6	34.1	2.3	20.0	10.6	44.3	25.1	52.4	30.8	16.9	30.6	40.1	21.4	7.9	
1985	64.7	33.2	2.0	20.1	10.0	44.4	25.4	51.0	31.3	17.7	29.0	40.5	22.3	8.2	
1986	64.1	33.6	2.3	19.5	9.5	43.4	27.6	53.0	30.5	16.5	29.8	40.8	20.9	8.4	
1987	62.8	34.3	2.9	19.4	10.0	41.4	29.2	52.0	30.1	17.9	29.1	41.0	21.2	8.8	
1988	62.5	34.7	2.9	18.7	9.3	41.9	30.0	52.5	30.2	17.3	28.0	40.7	21.7	9.7	
1989	62.8	34.7	2.5	19.3	9.9	41.1	29.7	52.8	30.8	16.4	26.9	41.1	23.3	8.7	
1990	62.3	34.9	2.8	19.0	8.3	40.3	32.4	52.1	30.5	17.4	27.8	40.2	22.8	9.1	
1991	62.8	34.3	3.0	19.4	7.2	39.8	33.6	51.3	30.5	18.2	26.7	39.6	24.7	9.1	
1992	62.6	34.2	3.2	18.7	6.9	40.5	33.9	52.9	30.0	17.1	24.6	39.5	26.5	9.4	
1993	61.8	34.7	3.6	17.9	7.0	39.4	35.7	52.2	30.3	17.6	24.8	36.7	29.4	9.1	
1994	62.8	33.6	3.7	18.2	7.4	40.4	34.0	50.5	31.7	17.8	22.9	37.0	30.3	9.8	
1995	64.2	32.6	3.3	18.5	6.6	40.2	34.7	51.7	30.5	17.8	22.3	36.4	30.7	10.6	
1996	62.8	33.7	3.5	17.2	6.0	38.9	38.0	51.6	31.4	16.9	22.0	37.6	30.3	10.1	
1997	63.9	33.3	2.8	15.9	6.9	38.2	38.9	52.7	29.6	17.8	21.3	37.0	30.7	11.0	
1998	63.8	33.1	3.1	15.4	6.6	38.7	39.3	52.2	30.8	17.0	20.4	35.7	31.7	12.2	
1999	63.6	32.5	3.9	14.4	6.4	39.1	40.2	53.7	30.4	15.9	20.7	36.8	30.8	11.7	
2000	64.4	31.6	4.0	13.9	6.5	38.8	40.7	52.3	31.1	16.6	19.6	36.1	32.3	12.0	
2001	64.1	31.5	4.5	14.4	6.4	38.1	41.1	53.4	30.1	16.4	18.7	36.7	31.6	13.0	
2002	63.1	31.8	5.1	14.4	6.5	37.8	41.3	54.5	29.0	16.6	19.3	35.4	32.7	12.6	

Source: Data from the annual March Current Population Survey.

a. Percentages may not sum to 100 because of rounding.

b. Includes those earning associate degrees.

from 7.7 percent in 1980 to 4.7 percent in 2002 (not shown). On average, single mothers have about 1.7 to 1.8 children.

Finally, table 2 summarizes single mothers' educational attainment. The share of single mothers who are high school dropouts declined from 34.5 percent in 1980 to 19.3 percent in 2002. At the same time, the share with at least some college increased from 26.5 percent to 45.3 percent. However, the bulk of this rather substantial increase in educational attainment occurred before 1996.

An important message of table 2 is that shifts in the demographics of single mothers since 1996 have been rather gradual. The largest shift over this period was the increase in never-married single mothers, and this shift is not favorable regarding work. Thus demographic shifts alone will be unable to account for much of the drop in welfare caseloads since 1996.

Welfare Participation Rates

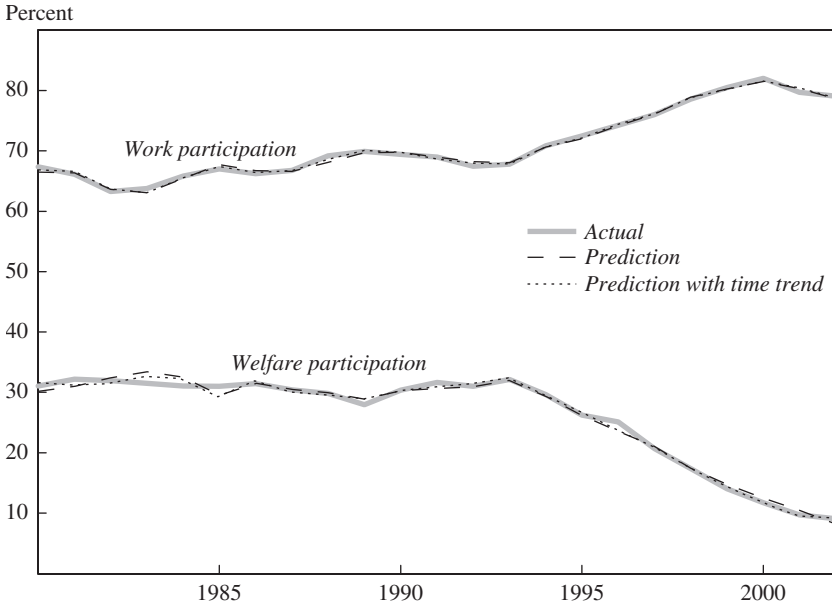
The solid lines in figure 5 show welfare and work participation rates for single mothers from 1980 to 2002.⁶³ In contrast to the trend in the total AFDC/TANF caseload (figure 1), the welfare participation *rate* was much more stable before 1994, hovering around 30 percent, with a peak of 32.2 percent in 1993. Since 1993, however, welfare participation has dropped spectacularly, all the way to 9.0 percent in 2002, or by 72 percent.⁶⁴

Figure 6 reports welfare participation rates for eight large states. Clearly, both levels and trends in participation rates differ substantially from state to state. The participation rate peaked in California in 1993, and in Texas and Florida in 1992; all these observations are roughly consistent with the national trend. But in Michigan the participation rate has trended down since 1983, and in Illinois it has trended down since 1987. The peak year in Pennsylvania was 1984, but a second run-up followed, which peaked in 1992. Peak years in New York and North Carolina were 1990 and 1991, respectively—a bit earlier than the national peak.

63. Recall that we define "welfare participation" as receipt of cash public assistance.

64. Since the March CPS consists of repeated cross-sectional data (with only a small panel component), we cannot determine the extent to which the decrease in welfare participation is due to an increase in exit from or a decrease in entry into welfare. Grogger, Haider, and Klerman (2003) used data from the Census Bureau's Survey of Income and Program Participation to examine the importance of entry in explaining the drop in welfare caseloads.

Figure 5. Actual and Predicted Welfare and Work Participation Rates among Single Mothers, 1980–2002

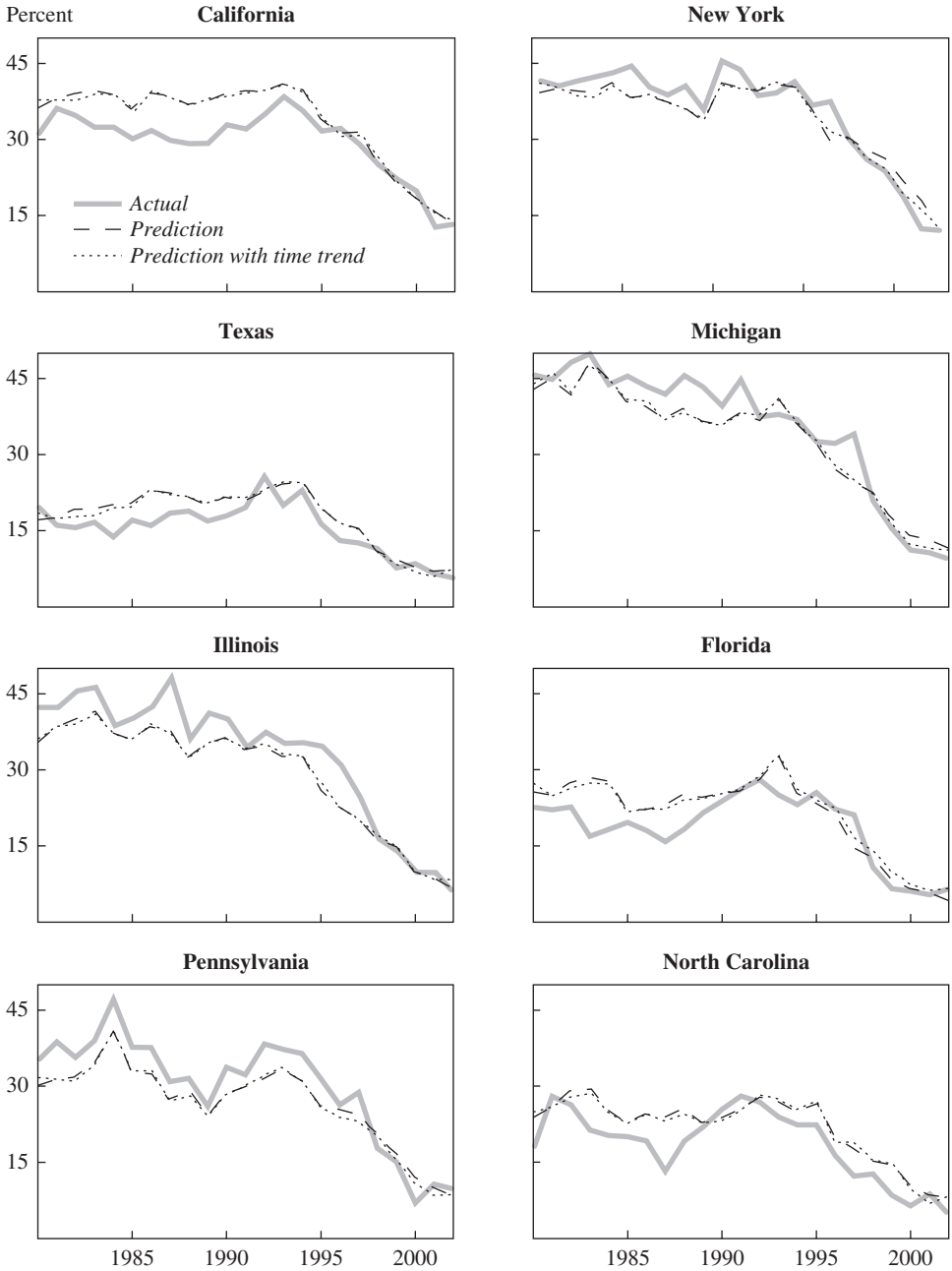


Sources: CPS data and authors' calculations.

The left-hand panels of figures 7 through 11 show how the welfare participation rates of single mothers vary with their demographic characteristics. Of course, such differences are not surprising. What is more interesting is that the *trends* in participation rates also differ in important ways across demographic groups. For instance, the left-hand panels of figure 7 show that welfare participation rates differ substantially by educational attainment, as one would expect. In 1994 the participation rate was 47.7 percent among single mothers who were high school dropouts, 26.9 percent among those who were high school graduates without a college degree, and 5.8 percent among those with at least a college degree.⁶⁵ More interesting, however, is the fact that, as a percentage, participation has dropped less (62 percent) for the least educated group; the participa-

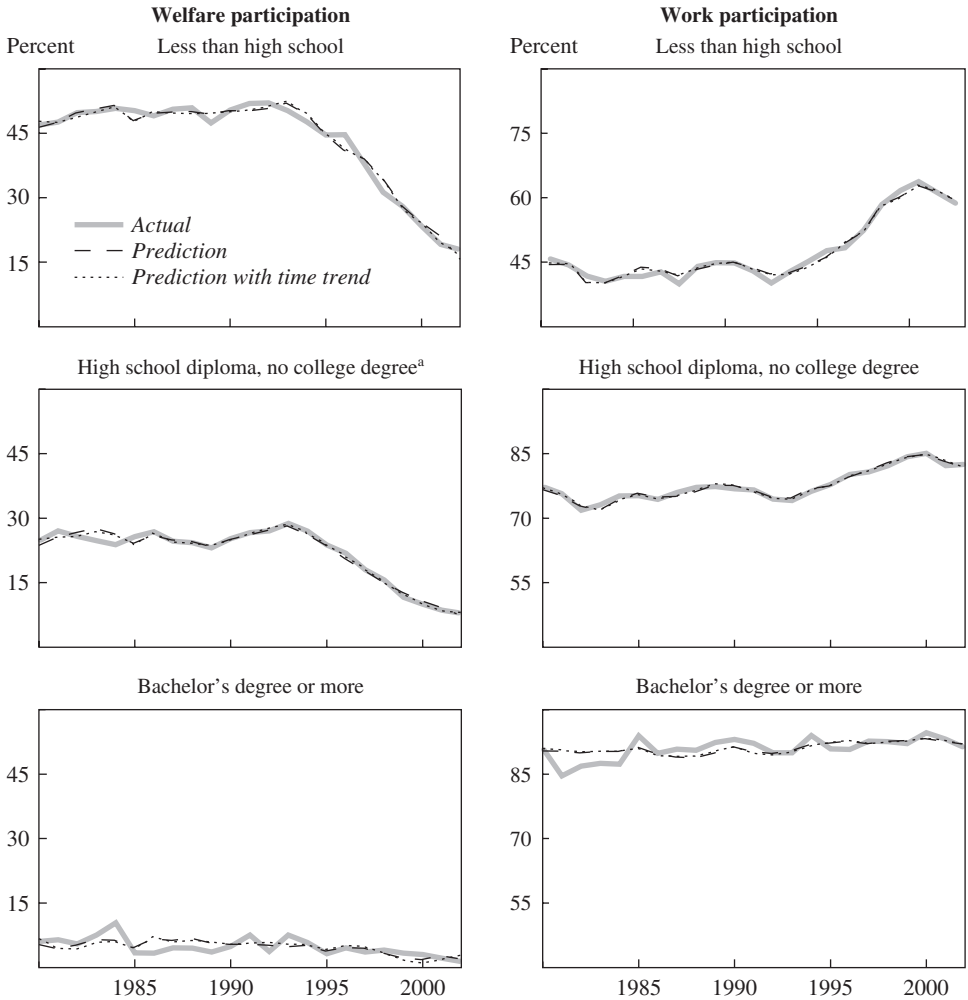
65. The second group combines those single mothers who had only a college degree with those who had some college (and possibly an associate degree) but not a bachelor's. The participation rate among single mothers with a bachelor's degree and no further education was 7.1 percent.

Figure 6. Actual and Predicted Welfare Participation Rates among Single Mothers, Selected States, 1980–2002



Sources: CPS data and authors' calculations.

Figure 7. Welfare and Work Participation Rates among Single Mothers by Educational Attainment, 1980–2002



Sources: CPS data and authors' calculations.
a. Combined data for "High school diploma only" and "Some college."

tion rate declines since 1994 for the other two groups were 71 and 80 percent, respectively.

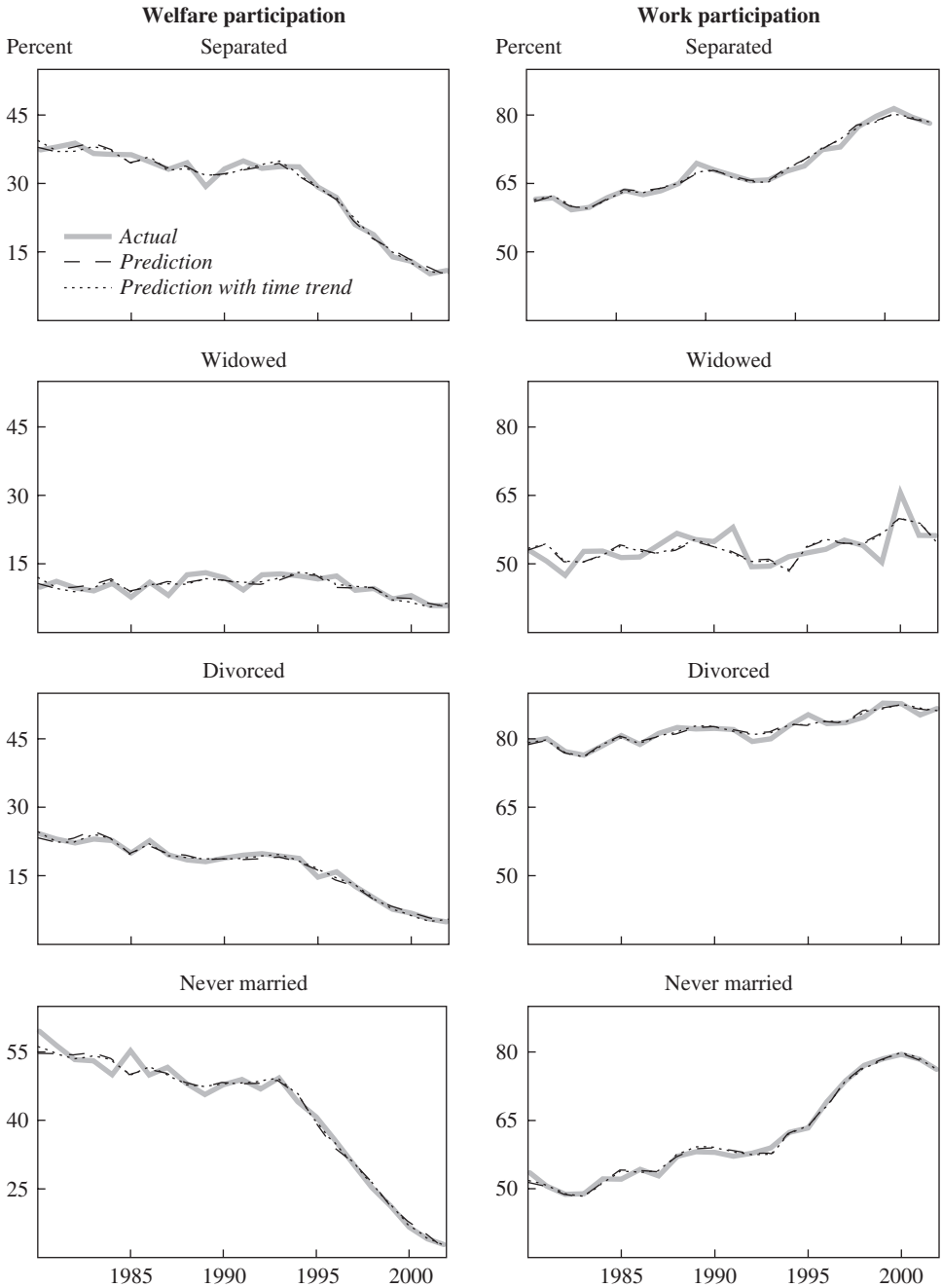
The left-hand panels of figure 8 show that the welfare participation rates of single mothers also differ substantially by marital status. The participation rate of the never-married mothers has historically been the highest (44.1 percent in 1994), followed in that same year by separated (33.7 percent), divorced (18.8 percent), and widowed mothers (12.3 percent). Interestingly, the percentage drops since 1994 for these four groups also differed, at 71, 67, 74, and 52 percent, respectively. Because of the relatively large drop in their participation rate, divorced single mothers are now the least likely to be on welfare.

As the left-hand panels of figure 9 show, welfare participation rates have historically been much higher for black than for white single mothers. However, the participation rate for whites was fairly stable at roughly 25 percent from 1980 to 1994, while the rate for blacks fell from 42.6 percent to 37.0 percent. Thus in 1994 the participation rate for blacks was 47 percent higher than that for whites. Since the welfare reform of 1996, racial differences in participation rates have narrowed further: in 2002 the rates were 8 and 10.5 percent for whites and blacks, respectively, so that the rate for blacks was only 31 percent higher. Thus the decline in welfare participation rates has been much greater for blacks than for whites and started much earlier.

The left-hand panels of figure 10 show that participation rates are much higher for single mothers with younger children, as already discussed. Interestingly, the drop in participation from 1994 to 2002 is larger for mothers whose youngest child is 6 to 12 years old (70 percent) than for those whose youngest child is less than 6 years old (68 percent) or those whose youngest child is 13 to 17 years old (63 percent). The same pattern is found for specific ages at the low end of these ranges: 76, 62, and 47 percent for mothers whose youngest child is 6, 1, and 13 years old, respectively; not shown. Thus the notion of a pure anticipatory time limit effect implies a monotonically decreasing rate of decline as the age of the youngest child increases, *ceteris paribus*. These figures seem somewhat inconsistent with that story.

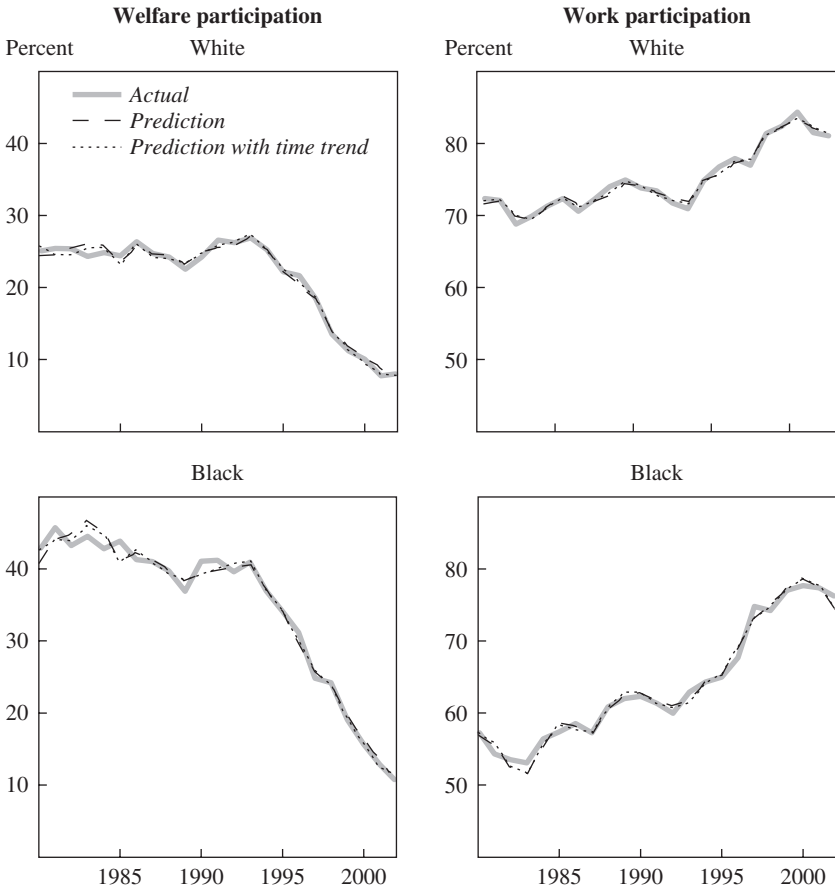
Finally, the left-hand panels of figure 11 show that single mothers with more than one child are more likely to be on welfare than are single mothers with only one child. However, the percentage drop in welfare participation from 1994 to 2002 was similar for single women with one,

Figure 8. Welfare and Work Participation Rates among Single Mothers by Marital Status, 1980–2002



Source: CPS data and authors' calculations.

Figure 9. Welfare and Work Participation Rates among Single Mothers by Race, 1980–2002



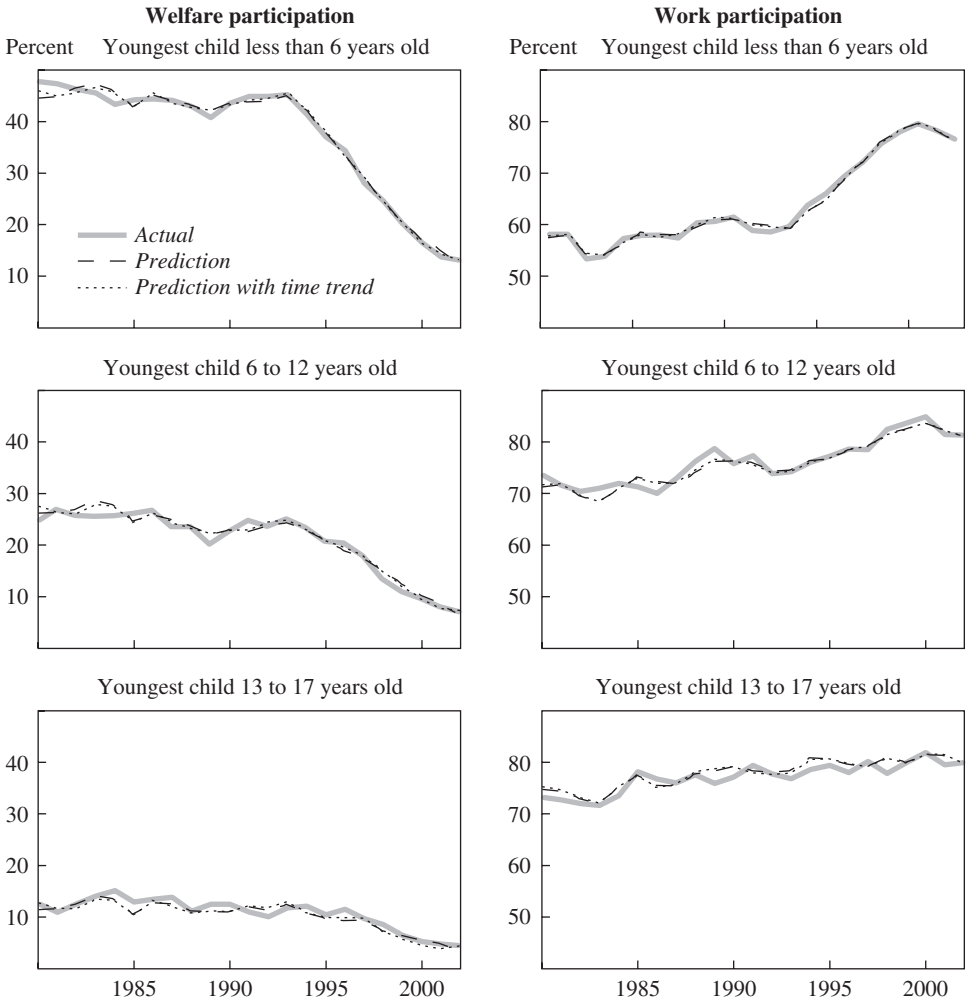
Source: CPS data and authors' calculations.

two, three, or four or more children (69, 71, 65, and 66 percent, respectively; not shown).

Work Participation Rates

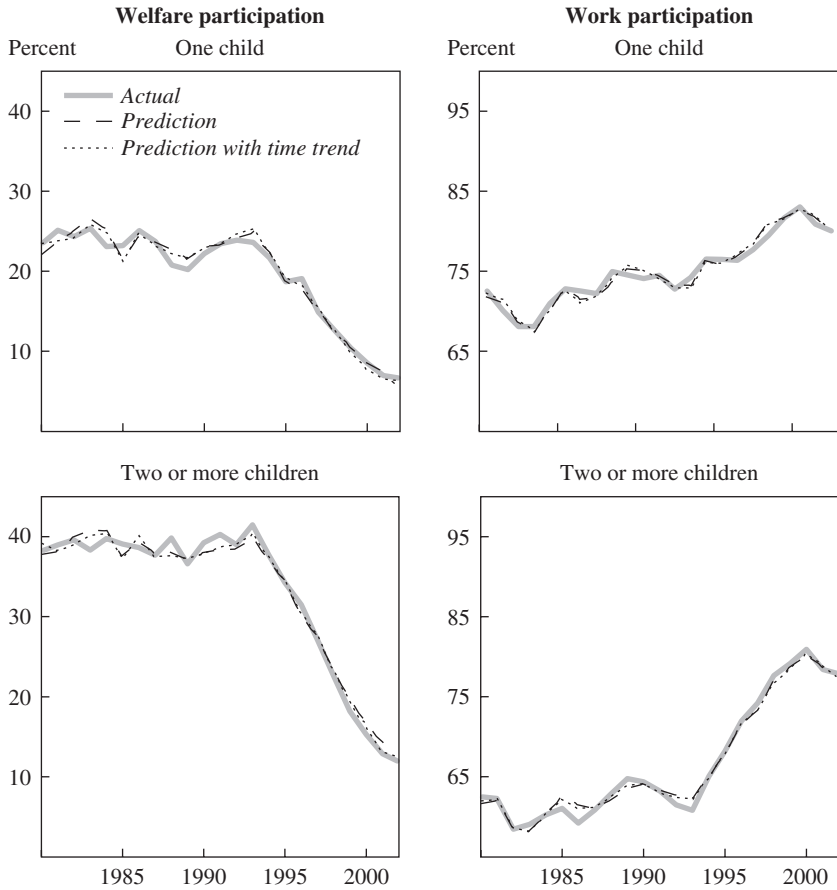
In summarizing trends in work participation rates for single mothers from 1980 to 2002, we combine part-time work (defined above as from eight to

Figure 10. Welfare and Work Participation Rates among Single Mothers by Age of Youngest Child, 1980–2002



Source: CPS data and authors' calculations.

Figure 11. Welfare and Work Participation Rates among Single Mothers by Number of Children, 1980–2002



Source: CPS data and authors' calculations.

thirty-two hours a week) and full-time work (more than thirty-two hours a week) into a single “working” category. The general patterns we describe here are robust to plausible changes in these definitions.

Figure 5 shows that the share of single mothers who work increased from 67.8 percent in 1993 to 82.0 percent in 2000. With the onset of the recession, the working share slipped back, to 79.1 percent in 2002. It is interesting that the upward trend in work participation began a year earlier

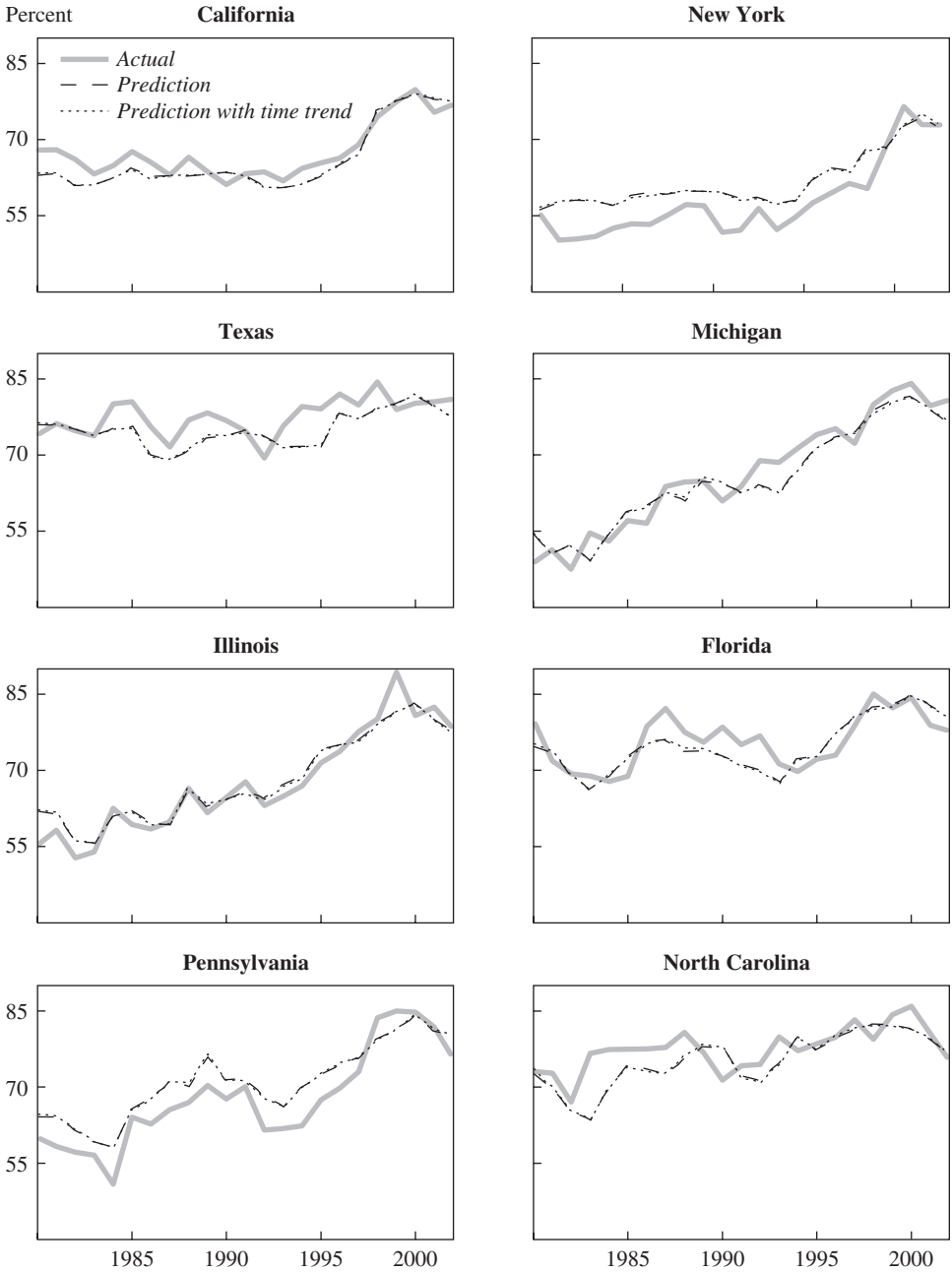
than the dramatic drop in welfare participation. Welfare participation rose rather noticeably in 1993, whereas work participation also increased that year, but only slightly. It is plausible that this occurred because the expansion of the EITC provided a substantial enhancement of work incentives in 1993, whereas regulations that made AFDC less attractive, such as work requirements under waivers, were not widely introduced until 1994. Both the share of single mothers not working and the share on welfare start to trend down strongly together in 1994. Not shown in figure 5 but also notable is that almost all the increase in work activity took the form of increased full-time work. The share of single mothers working full time increased from 53.3 percent in 1993 to 67.3 percent in 2000, while the share working part time stayed fairly flat (in the 14 to 15 percent range).

Figure 12 reports work participation rates for eight large states. Clearly, both levels and trends in work differ substantially by state. In California work participation is rather stable except for a dramatic increase in 1996–2000. In contrast, in Michigan participation trends up over the whole 1980–2002 period. Florida and Pennsylvania show clear cyclical patterns, but the participation rate is rather flat in Texas and North Carolina throughout our sample period. New York shows a slight upward trend in the mid-1990s followed by a sharp increase in 1998–2000. Illinois has an upward trend from 1980 through 1999, followed by a decline.

The right-hand panels of figures 7 through 11 show how the work participation rates of single mothers vary with their demographic characteristics. The right-hand panels of figure 7 show, not surprisingly, that work is much more prevalent among the more educated. Since 1993, however, the share of single mothers not working has declined at all education levels. For single mothers with less than high school, those who had completed high school but not college, and those who had at least completed college, the declines in the share not working were 28, 33, and 13 percent, respectively. (The declines were 31 percent for those with a high school diploma only, 35 percent for those with some college, and 29 percent for those who had a bachelor's degree but no further college; not shown.)⁶⁶

66. We prefer to report percentage declines in the share not working, rather than percentage increases in the share working, because the former can always range from zero to 100 percent regardless of the baseline. Thus the percentage decrease in the share not working should be more comparable across groups with different baseline rates of work participation.

Figure 12. Actual and Predicted Work Participation Rates among Single Mothers, Selected States, 1980–2002



Sources: CPS data and authors' calculations.

For certain other demographic characteristics, however, the trends in work participation differ across groups in important ways. For instance, the right-hand panels of figure 8 show the work participation rates of single mothers of different marital status. Divorced single mothers are the most likely to work, and widowed single mothers the least. In 1993 the shares of widowed, never-married, separated, and divorced single mothers at work were 49.5, 58.9, 65.8, and 80.0 percent, respectively. In 2002 these percentages had risen to 56.2, 75.9, 78.2, and 86.8 percent, respectively. Thus the decrease in the not-working share is greater for never-married single mothers (41 percent) than for the other groups. It is interesting that the never-married group and the separated group show slight upward trends in work participation in the pre-1993 period, whereas the divorced and widowed groups do not.

As the right-hand panels of figure 9 show, work participation rates for white single mothers have historically been higher than those for black single mothers. The work participation rate for whites held stable at roughly 72 percent from 1980 to 1994, while that for blacks rose from 57.5 percent to 64.3 percent; these patterns roughly mirror those of the welfare participation rates for both races. Since the welfare reform of 1996, racial differences in work participation have narrowed further. In 2002 the work participation rates for whites and blacks were 81 percent and 76 percent, respectively, a difference of only 5 percentage points (or 6.6 percent).

The right-hand panels of figure 10 summarize the work participation rates of single mothers according to the age of their youngest child. In 1993 only 59.6 percent of single mothers with children ages 0 to 5 worked. By 2000 this rate had increased to 79.6 percent, but with the recession it dropped back down, to 76.6 percent in 2002. In contrast, 74.2 percent of single mothers with children ages 6 to 12 were at work in 1993, as were a slightly larger fraction of those with older children. The overall decrease in the not-working share from 1993 to 2002 was 42 percent for single mothers with children ages 0 to 5, 28 percent for those with children 6 to 12, and 14 percent for those with children 13 to 17. Comparable figures for women with children of specific ages are 41 percent for those with infant children, 35 percent for those with 6-year-old children, and 36 percent for those with 13-year-old children (not shown). Meanwhile the working share of women with 17-year-old children stayed fairly flat at about 75 percent.

The right-hand panels of figure 11 show that work participation rates tend to be lower for single women with more than one child. What is more interesting is that the increase in work from 1993 to 2002 was much greater for women with two or more children than for those with only one child. The not-working share of mothers with one child declined from 25.8 percent in 1993 to 20.0 percent in 2002—only a 22 percent decrease. But among those with two children (not shown), the share not working fell from 32.0 percent to 18.7 percent, a 42 percent drop. For women with more children the percentage decreases were slightly larger (not shown).

One plausible explanation for this pattern would attribute it to the EITC, since the EITC phase-in rate for women with one child increased by only 5.8 percentage points from 1993 to 2002, while that for women with two or more children increased by 13 percentage points. Of course, it is also possible that child care or Medicaid expansions, or both, were more important for women with more children, or that work requirements had a greater effect on women with more children.

In general, the key fact that these discussions bring home is that there are important differences across states and demographic groups in how work and welfare participation have changed over time. A successful model should therefore explain changes in work and welfare participation among single mothers not just at the national level, but also at the state level and at the level of particular demographic groups. We will allow for interactions of our policy measures with the key demographic measures discussed here (education, marital status, race, age, age of children, and number of children) in order to accommodate the fact that different policies may affect different groups differently.

Differences in Welfare and Work Participation Rate Changes

As noted above, the welfare participation rate among single mothers overall dropped from 32.2 percent in 1993 to 9.0 percent in 2002, a 23.2-percentage-point decrease. At the same time, the work participation rate increased from 67.8 percent to 79.1 percent, an 11.3-percentage-point increase. The gap between the drop in welfare and the increase in work is thus a full 11.9 percentage points. What explains this discrepancy?

One factor is that that work and welfare are not mutually exclusive. If a single mother who is working while on welfare then leaves welfare but continues to work, overall welfare participation falls but work participa-

Table 3. Accounting for the Discrepancy between Falling Welfare Participation and Rising Work Participation among Single Mothers, 1993–2002

Percent of all single mothers except where stated otherwise

	1993			2002		
	<i>Not on welfare</i>	<i>On welfare</i>	<i>Total</i>	<i>Not on welfare</i>	<i>On welfare</i>	<i>Total</i>
Not working	11.35	20.86	32.21	16.43	4.51	20.94
Working	56.49	11.30	67.79	74.54	4.52	79.06
Total	67.84	32.16	100.00	90.97	9.03	100.00
Fall in welfare participation rate			= 32.16 – 9.03	= 23.13	percentage points	
Rise in work participation rate			= 79.06 – 67.79	= 11.27		
			Difference	= 11.86		
Fall in share working, on welfare			= 11.30 – 4.52	= 6.78	percentage points	
Rise in share not working, not on welfare			= 16.43 – 11.35	= 5.08		
			Sum	= 11.86		

Source: Authors' calculations using data from the March CPS.

tion does not increase. By the same token, women may leave welfare without finding work. Table 3 decomposes the discrepancy between the changes in the overall welfare and work participation rates. It shows that the fraction of single mothers who both work and receive welfare dropped from 11.3 percent in 1993 to 4.5 percent in 2002, a decrease of 6.8 percentage points. Meanwhile the fraction of single mothers who neither worked nor collected welfare increased from 11.4 percent to 16.4 percent, an increase of 5.1 percentage points. (The small difference is due to rounding.) Together these components exactly account for the gap between the increase in work and the decrease in welfare participation. The fact that the share of single mothers who neither work nor receive welfare increased by 5 percentage points is troublesome, because this may be a vulnerable group.⁶⁷ We return to this issue later in the paper.

Income and Other Quality-of-Life Measures

Table 4 summarizes trends in the incomes of single mothers over 1980–2002. Table 5 does the same for several other life quality measures, namely, housing arrangements, number of hours worked per week, and average hourly wages.

67. Moffitt (1983) proposed and estimated a model of welfare stigma to explain why a large fraction of welfare-eligible single mothers did not participate in welfare.

Table 4. Sources of Single Mothers' Real Incomes, 1980–2002
Constant (2001) dollars

Year	Total income ^a	Wages	Public assistance	Food stamps	Alimony ^b	Child support	Unempl. compensa- tion ^c	Worker's compensa- tion	SSP ^d	Social Security	Other	EITC ^e
1980	18,468.2	12,530.4	1,982.6	910.5	1,543.7	...	331.0	...	153.3	954.0	62.7	216.77
1981	17,360.4	11,650.5	1,903.2	812.6	1,618.1	...	302.8	...	123.8	894.6	54.8	189.49
1982	17,028.6	11,503.2	1,839.7	843.8	1,472.7	...	329.0	...	124.3	887.0	28.9	156.74
1983	17,006.8	11,630.5	1,789.6	837.6	1,414.2	...	337.5	...	121.5	833.3	42.6	138.77
1984	17,486.4	12,199.8	1,767.9	794.5	1,431.5	...	250.6	...	122.3	699.0	220.8	135.84
1985	18,278.3	12,720.6	1,752.7	774.6	1,592.2	...	233.3	...	130.9	719.3	387.7	155.73
1986	18,002.6	12,745.6	1,794.2	794.0	1,407.9	...	190.5	...	152.9	643.7	273.8	147.61
1987	18,956.2	13,330.9	1,679.9	818.0	1,159.7	...	79.1	...	194.5	699.4	750.5	266.13
1988	18,676.8	13,384.9	1,569.7	768.1	256.1	1,081.0	85.0	46.4	161.7	648.2	676.8	356.07
1989	19,442.1	13,871.4	1,470.9	757.6	229.9	1,155.2	100.9	58.9	143.8	703.5	950.0	359.61
1990	18,700.8	13,301.6	1,536.4	868.7	178.3	1,139.7	111.6	95.7	193.6	608.5	666.7	354.34
1991	18,522.4	13,353.6	1,583.0	927.9	148.2	1,240.1	136.2	74.8	188.1	564.4	306.1	449.26
1992	18,396.0	13,326.8	1,447.0	908.5	112.0	1,253.6	172.7	62.8	222.8	637.4	252.4	481.35
1993	18,498.4	13,065.5	1,528.9	921.3	118.1	1,208.0	159.7	51.5	296.3	727.3	421.8	519.03
1994	19,271.7	13,932.0	1,400.6	947.6	124.3	1,253.2	152.9	80.0	292.4	773.2	315.5	796.39
1995	20,162.6	14,627.5	1,212.1	850.3	103.2	1,318.4	128.7	77.5	360.7	735.4	748.8	950.12
1996	20,280.0	15,153.1	1,036.0	801.7	132.7	1,334.4	120.9	41.7	338.0	720.5	601.0	1,074.46
1997	20,324.9	15,210.4	802.5	719.9	157.9	1,272.8	106.5	41.8	331.5	794.0	887.6	1,116.82
1998	21,468.2	16,535.1	647.1	612.0	134.9	1,331.8	103.1	41.5	297.0	636.0	1,129.7	1,175.87
1999	22,440.9	17,285.2	509.8	505.9	203.7	1,472.1	91.2	58.7	260.7	686.6	1,367.0	1,179.53
2000	23,164.4	18,338.7	400.9	449.3	130.1	1,494.5	99.5	56.3	285.7	597.3	1,312.1	1,151.16
2001	23,885.1	18,739.5	306.8	442.8	170.3	1,531.5	166.0	46.6	277.7	668.1	1,535.8	1,104.65
2002	23,068.3	18,225.9	301.5	499.1	154.8	1,475.4	277.9	55.2	273.6	676.6	1,128.3	1,088.48

Source: Authors' calculations using data from the March CPS, 1981–2003.

a. Data are averages and include imputed cash value of food stamps; does not include EITC (earned income tax credit).

b. Before 1987, includes child support.

c. Before 1987, includes worker's compensation, veterans' payments, and pensions to government workers.

d. Supplemental Security Income.

e. Simulated from federal and state EITC rules.

Table 4 reports single mothers' average real incomes (in 2001 constant dollars) and the main sources of that income. From 1980 through 1993 the mean real income of single mothers was basically flat, except for a brief decline in 1981–83. However, from 1993 to 2002 their mean real income increased from \$18,498 to \$23,068, or by 25 percent. Their mean real wage earnings increased by \$5,161, or 39.5 percent, over the same period. Much less important sources of income that showed some gain in this period were child support and alimony, perhaps reflecting the increases in state expenditure on child support enforcement. At the same time, mean income from public assistance and food stamps dropped substantially, from \$2,450 in 1993 to \$800 in 2002.

The source of the real wage earnings increase can be decomposed into the fraction due to the increased work participation rate, that due to increases in average hours worked per week (conditional on employment), and that due to increases in real hourly wages. Recall that the work participation rate of single mothers increased from 67.8 percent in 1993 to 79.1 percent in 2002, a 16.7 percent increase. According to table 5, the mean hourly wage rate increased from \$11.16 to \$12.88 over that period, a 15.4 percent increase.⁶⁸ Average hours worked per week increased from 37.6 to 38.3, a 1.7 percent increase. Thus the hourly wage increase together with the increased work participation rate explains almost all of the 39.5 percent increase in real wage earnings experienced by single mothers in this period. The last column in table 4 shows the average (simulated) value of federal and state EITC payments. (The CPS imputes these EITC payments rather than querying for them directly, and so we do not include them in the total real income measures.) Note that from 1993 to 2002 the average real EITC more than doubled.

Table 5 also shows that the share of single mothers living in public housing declined from 11.4 percent in 1993 to 9.4 percent in 2002. Somewhat surprisingly, the rate of cohabitation increased only slightly over the same period, from 30.5 percent to 31.9 percent.⁶⁹ We have already seen that the share of single mothers who do not work and are not on welfare increased by 5 percentage points over that period. In future work we plan

68. To obtain an hourly wage measure for each woman, we divide total wage income (Unicon recode variable *incwage*) by total hours worked, which is the product of hours worked per week last year (*hrslyr*) and total weeks worked last year (*wkslyr*).

69. We say a single mother cohabits if she lives in a house or apartment where another person pays the rent.

Table 5. Living Arrangements, Working Hours, and Wages of Single Mothers, 1980–2002

Year	Living arrangements (percent of all single mothers)		Average hourly wage (constant 2001 dollars)		Average hours worked per week	
	Share living in public housing	Share cohabiting ^a	Full- and part- time workers	Full-time workers only	Full- and part- time workers	Full-time workers only
1980	8.7	13.7	10.93	11.34	37.69	40.95
1981	9.0	19.0	10.33	10.80	37.59	40.82
1982	10.3	22.4	10.75	11.07	37.60	40.90
1983	10.2	22.6	10.90	11.52	37.36	41.05
1984	11.2	22.6	10.79	11.30	37.49	40.98
1985	10.9	21.9	11.02	11.52	37.49	41.20
1986	11.1	22.4	11.14	11.54	37.58	41.19
1987	10.8	28.5	11.20	11.71	38.09	41.24
1988	10.1	27.7	10.96	11.57	37.67	41.31
1989	10.0	28.6	11.29	11.85	38.30	41.60
1990	10.2	29.2	11.46	11.61	37.98	41.37
1991	10.4	31.2	11.09	11.64	37.78	41.47
1992	10.2	31.1	11.19	12.00	37.84	41.39
1993	11.4	30.5	11.16	11.53	37.64	41.49
1994	9.8	31.9	11.28	11.94	37.86	41.56
1995	10.0	31.8	11.48	11.89	38.09	41.82
1996	11.1	31.2	11.73	12.20	37.87	41.53
1997	10.3	32.1	11.62	11.97	37.89	41.53
1998	10.4	31.4	11.93	12.37	38.41	41.59
1999	10.0	32.5	12.26	12.69	38.45	41.44
2000	9.7	33.2	12.47	12.80	38.46	41.58
2001	9.4	31.2	13.00	13.43	38.49	41.45
2002	9.4	31.9	12.88	13.23	38.27	41.45

Source: Authors' calculations using data from the March CPS, 1981–2003.

a. Defined as living in a house or apartment where another person pays the rent.

to investigate more carefully the income sources of these women. Preliminary results suggest that they are more likely to reside in public housing and to cohabit, and that they tend to receive above-average benefits from Social Security and Supplemental Security Income, but that their incomes are still well below the average for single mothers.

The Empirical Specification

Demographics and Identification

Table C1 in appendix C describes the variables used in our empirical analysis. Our dependent variable is either welfare reciprocity status ($WEL_RECEIPT_{ist}$) or work participation ($WORK_{ist}$), both of which are categorical, zero-or-one variables. The individual-level demographic variables included in the model are age, which is continuous, and several categorical variables: race (three categories), educational attainment (four categories), marital status (four categories), state of residence (fifty-one categories), and urban or rural residence (two categories). We also categorize family composition by five variables: numbers of children in age groups 0–5, 6–12, and 13 and above, and the ages of the youngest and the oldest child.

A completely saturated model would include a separate dummy variable for each demographic group in each state in each time period. But because six of the demographic variables are continuous, the number of demographic “groups” would be enormous. For the purpose of understanding identification, it is useful to think of a simpler situation in which the data on age and number of children are discretized. Suppose that all we observed in the data was that the woman is in one of four age intervals, that her youngest child is in one of three age intervals, that her oldest child is likewise in one of three age intervals, and that she has either one child, two children, or three children or more. We would then have $4 \times 3 \times 3 \times 3 = 108$ categories of family age composition. In addition, we have $3 \times 4 \times 4 \times 2 = 96$ types of mothers in terms of race, education, marital status, and urban or rural residence, giving $108 \times 96 = 10,368$ demographic cells. A fully saturated model that interacted demographics \times states \times time would then include $10,368 \times 51 \times 23 = 12,161,664$ parameters and would fit the data on welfare and work participation (by cell) perfectly.

Of course, a fully saturated model does not permit the identification of policy effects. Since the model fits the data perfectly using demographic \times state \times time dummies, it is impossible to identify the effect of any *particular* time-varying factor, such as a policy variable.⁷⁰ If we wish to identify a policy effect, we must exclude certain interactions. The source of identification depends on which interactions are allowed and which are excluded.

As already discussed in our review of the literature, much of the previous literature in this area has relied on specifications that include state dummies, year dummies, and state-specific quadratic time trends. This gives a model with $51 + 23 + 100 = 174$ parameters. A typical procedure is then to include a measure of a *single* policy, such as a time-varying dummy variable for whether a state has yet imposed time limits.

It is important to understand the assumptions that underlie identification in such a specification. One is assuming that any omitted time-varying factors, including policy variables other than the one being investigated, either have common effects across all states (picked up by the year dummies), or, if they do have differential effects by state, that these are captured by the smoothly varying state-specific quadratic time trends. Both these assumptions would be violated by an omitted policy variable that “turned on” discretely in a particular year (say, 1993) and that had differential effects across states (say, because it affects different demographic groups differently and the demographic composition of states differs).⁷¹

To avoid these problems, one could use a more flexible specification that included state \times year interactions (that is, state-specific time dummies rather than state-specific time trends). Such a specification would have $51 \times 23 = 1,173$ parameters, plus the additional parameters characterizing the single policy under study. In that case identification of the policy effects would rely on how the policy affects different demographic groups within a state differently over time. The key assumption is that any omitted

70. Note that the lack of identification has nothing to do with the fact that the number of parameters in a saturated model would exceed the number of available data points (127,119 in our case). This is a finite-sample problem, whereas identification analysis proceeds under the hypothesis that one has unlimited data. Rather, the lack of identification comes from the fact that, if the model is saturated, then all time-varying factors are controlled for.

71. They would also be violated by state-specific policy variables that turn on discretely at particular points in time (so long as timing differs across states or, if timing is synchronized, effects differ across states).

policy variables have common effects across all demographic groups. We have already argued at length that this assumption is implausible.

One could try to deal with this problem by including demographics \times time interactions. Given that we have 10,368 demographic cells, this would generate $10,368 \times 23 = 238,464$ parameters, although one could perhaps reduce this by defining groups much more coarsely. Even here, however, one is continuing to assume that any omitted policies that affect different demographic groups differently are national policies and, as such, cannot be implemented at different times in different states. This is obviously false in the case of welfare reform. But relaxing this assumption brings one back to the saturated specification.⁷²

Given that, over the 1980–2002 period, states pursued an array of different policies that clearly have different effects on different demographic groups, and given that these policies were implemented at different times in different states, we feel it is not possible, under reasonable identifying assumptions, to identify the effect of any single policy (or small set of policies) while using an array of state and time dummies to control for all other aspects of policy. Therefore we take a very different course. We have attempted, as best we can, to include in our model measures of the entire range of policy changes that occurred at the state level over the whole 1980–2002 period. We also interact these policy variables with a range of demographic controls to allow for the fact that policies affect different groups differently.

72. Schoeni and Blank (2000) adopt a hybrid approach by including some demographics \times time interactions and state-specific time trends. Rather than use individual data, they use March CPS data from 1977 through 1999 to calculate welfare participation rates by state and by year for each of three education and four age categories, giving $23 \times 3 \times 4 \times 51 = 14,076$ cells to be fit. Their model includes 234 parameters, since it includes 50 state dummies, 9 demographic dummies (age, education, and race), 50 state-specific time trends, 23 aggregate time dummies, 69 aggregate time dummies interacted with education, and a total of 27 interactions of the three education group dummies with current and lagged unemployment and employment growth rates, the AFDC grant for a family of three, and four age dummies. They then include just six parameters to capture the effects of welfare reform. These are dummies for whether a state had a waiver or had implemented TANF, both of which are interacted with the three education categories. This model is identified because it assumes that any omitted policies that affect different education groups differently are purely national rather than state-specific, and because it rules out omitted variables that affect different age groups differently. We view such assumptions as untenable, given the great heterogeneity of policies across states, and because many policies might affect women of different ages differently. For example, older mothers are likely to have higher wages and therefore may be more affected by the EITC; they will also tend to have older children and therefore are likely to be less affected by the CCDF, time limits, and work requirements.

Thus, in estimating the effect of any particular policy, we are in effect controlling for other time-varying factors not through dummies, but rather by including those other policy and economic environment factors explicitly in the model, and by including policy \times demographics interactions that allow those other factors to affect different groups differently. The key identifying assumption is that we have adequately controlled for all the important time-varying factors that influenced the welfare and work participation decisions of single mothers over the 1980–2002 period. Of course, this is a very strong assumption, but it should at least in principle be achievable (if not by us, then at least by others who can improve on our specification). In contrast, the approach of using portmanteau dummies to control for all other aspects of policy seems to rely necessarily on assumptions that are clearly untenable.

Our model that includes demographics, policy variables, and economic environment variables, along with a rich set of interactions among these groups of variables, contains a total of 245 parameters. This is remarkably parsimonious relative to the portmanteau dummy variable specifications described above. It is also a small parameter set relative to our sample size ($N = 127,119$).⁷³ Despite this parsimony, as we shall demonstrate, our model does quite a good job of explaining differences in welfare participation and work both across states and demographic groups and over time within states and demographic groups.

Policy and Economic Environment Variables

The third panel of table C1 lists our individual-level policy measures. These are constructed from the individual-level demographic variables in conjunction with relevant policy variables. For example, the variable WELFARE_BEN_{ist} (the AFDC/TANF benefit level for each individual single mother) is constructed using the state payment standard for the corresponding family size of the single mother. Similarly, the variable EITC_RATE_{ist} , which denotes the EITC phase-in rate, is constructed by combining information on federal and state EITC phase-in rates with

73. Schoeni and Blank (2000) provide an interesting point of comparison, as theirs is fairly typical of work that uses a small set of variables to measure policy and attempts to control for omitted policy variables using state and time dummies. Their model has 234 parameters, and, since they fit it to 14,076 state \times year \times age \times education cells, they have a smaller ratio of data points to parameters than we do (60:1 versus 519:1).

information on family size. In general, since we have individual-level data, we have exploited every opportunity to tailor policy variables to the individuals based on their demographics, which we again assume are exogenous.

Another example is the variable $\text{MONTH_SINCE_WR_HIT}_{ist}$, which measures the cumulative time (in months) elapsed since the woman may *potentially* have been subject to work requirements. In actual implementations of work requirement time limits, a woman who fails to satisfy work activity requirements is not typically denied benefits as soon as the time limit is reached. Rather, she becomes subject to a series of sanctions and remedial measures, which may eventually result in benefit termination if she fails to make a “good faith” effort to comply. Thus we hypothesize that the effect of a binding time limit on behavior is likely to be increasing in the time that has elapsed since the time limit was reached.

Construction of individual-level work requirement measures is rather involved. Recall that states often exempt a woman with children below some threshold age (typically around 12 months) from work requirements. Thus we must examine the ages of all the woman’s children and ask for each child whether that child would have exempted her from the work requirement when he or she was born, and, if so, for how long. (This is complicated because, over the years, many states have changed the exemption for very young children.) We then add up all the possible child age exemptions from work requirements and use this information to calculate how long the woman may potentially have been subject to work requirements.⁷⁴ In contrast to this duration measure, the variable SWR_{ist} (“subject to work requirement,” as defined previously) is simply a measure of the fraction of the year t that a woman may be subject to work requirements. This depends on whether her state of residence has a work requirement, on whether she has potentially reached her work requirement time limit, and on the fraction of the year that she may be exempted from work requirements if she has a young child.

The fourth panel of table C1 lists our state-level policy and economic variables. These vary both across states and across time but do not vary

74. In constructing this variable we do *not* measure whether a woman is actually subject to a work requirement, or for how long a woman has actually been up against a binding time limit. The reason we use “potential” rather than “actual” measures is that the actual measures would be endogenous (dependent on the woman’s work or welfare participation decisions) and therefore do not belong in a reduced-form equation.

across individuals in the same state and year. For instance, this set includes the length of the state's time limit, TL_LENGTH_{st} ; the time elapsed since the state's time limit clock started (under either waivers or TANF), $MONTH_SINCE_TL-START_{st}$; and whether the child's portion of TANF benefits continues after the exhaustion of the time limit, $DCHILD BEN_{st}$. The last panel of table C1 lists our federal-level policy variables. These variables, which vary only across time, are the federal minimum wage (in 2001 constant dollars an hour), MIN_WAGE_t , and the lowest-bracket federal income tax rate, $INCTAX_RATE_t$.

The Empirical Specification

In our regression models, the dependent variable, either $WEL_RECEIPT_{ist}$ or $WORK_{ist}$, is regressed on the full set of individual-level demographic variables, individual-level policy variables, and state and federal policy variables listed in table C1. We also include a wide range of terms that interact the policy variables with the demographic characteristics of respondents (table C2 in appendix C).

The main rationale that led to most of the interaction terms in our model is the notion that welfare policy variables should have different effects on women with different labor market opportunities (that is, different offer wage rates), different nonlabor incomes (for example, differing access to alimony or child support), and different fixed costs of working (for example, depending on whether they have young children). These three characteristics are, in turn, determined by the woman's age, race, education, marital status, and children's ages. Thus our basic strategy was to interact this set of demographic variables with each major policy variable.

From the descriptive statistics cited earlier, we know that welfare and work participation, and how they have changed over time, differ substantially across these demographic groups. Thus we expect that these interaction terms will be crucial in fitting the data. There could also be important interactions between policy variables. For example, single mothers may be more or less responsive to work requirements if the EITC is more generous. Our model thus includes a number of policy interactions as well.

We stress, however, that our specification was not chosen as the result of a specification search. That is, we neither added variables in an attempt to fit the data better nor deleted variables that proved insignificant. Instead

we specified our list of demographic, policy, and economic environment measures, as well as the list of interaction terms, a priori.

Empirical Results

Evaluating the Fit of the Model

Before we can take seriously the implications of our model regarding the impact of welfare policy on behavior, it is important to verify that the model provides a good fit to the data. Figure 5 above shows that the model accurately tracks both the welfare and work participation rates of single mothers at the national level over the 1980–2002 period and the changes in those rates. This accomplishment may seem trivial, but, as noted in our review of the literature, previous models that omitted time dummies have failed to achieve this result. Because we have no time dummies, our model explains changes in welfare participation over time based on changes in demographics, policy, and the economic environment alone.

On the other hand, the earlier models that included time effects attributed much of the change in welfare participation to the time effects, which is in effect an admission of ignorance. As figure 5 shows, inclusion of a fifth-order time polynomial in our model leads to essentially no improvement in fit,⁷⁵ and to almost no change in the model's predictions regarding various policy changes. In other words, the model assigns no significant role to unmeasured time-varying factors at the national level.

Figure 6 shows the model's fit to welfare participation rates in eight large states. It is not surprising that our model does not fit the changes in welfare participation over time as well at the state level as at the national level, since at the state \times year level the sample sizes are much smaller, generating much more noise. Nevertheless, our model replicates quite well both the differences in levels across states and the changes in participation rates within states over time.

For example, in the early 1980s welfare participation in Texas was around 20 percent, while that in Michigan was around 45 percent. Our model is able to generate these cross-state differences quite accurately using demographic and policy differences alone, without state dummies.

75. Our model produces an R^2 of 0.2403 for welfare participation and 0.2064 for work participation. Adding a fifth-order polynomial in time increases the R^2 by 0.0003 for both specifications, which is trivial.

The main failure of the model is that it consistently overestimates welfare participation in California by about 5 percentage points in the 1980–93 period. But, on the whole, the fit at the state level seems quite good.⁷⁶ Most strikingly, the model correctly predicts the downward trend in welfare participation in Illinois, Michigan, and Pennsylvania that began back in the mid-1980s, well before the national downtrend began.

Figure 12 shows the model's fit to work participation rates in the same eight large states. Here the fit is excellent. For instance, the model correctly predicts the steady upward trend in work participation in Michigan over the whole 1983–2000 period. It also correctly predicts that work participation was flat in California from 1980 to 1995, jumped up rapidly in 1996–98, and then flattened again. And it predicts the several turning points in work in Florida and Pennsylvania quite well. Bear in mind that this is all done without using any state or national time effects.

Figures 7 through 11 show how the model fits the behavior of various demographic groups. All the figures convey the message that our model fits the differences in levels across demographic groups, as well as changes over time within groups, very well. All these group differences are explained without the use of any group-specific time effects.

Our model fits equally well when we further narrow down the demographic groups to, for example, combinations of race and marital status, and when we apply the model to other states, as well as to various demographic groups within states (results not shown). That the model fits quite well in all these dimensions is comforting, as it suggests that we have successfully included most of the key time-varying factors driving behavior over this period.

One might argue that it is not surprising that the model fits the data so well given that we have 245 terms in our regression. However, we see such criticism as misguided, for two reasons. First, as we have pointed out, an alternative empirical model that included state \times year interactions would have 1,173 parameters plus any policy variables. Such a model would not be able to explain differences across demographic groups unless it also included demographics \times policy interactions, leading to a vastly expanded version of the model. On the other hand, inclusion of demographics \times year interactions would lead to many thousands of additional

76. Adding state dummies increases the R^2 to 0.2489 for welfare participation and to 0.2095 for work participation. These changes are significant but seem quantitatively small.

parameters. Thus the model is actually quite parsimonious compared with such alternatives.

Second, we require that our model fit not only the national work and welfare participation trends, but also the variation in participation rates over time by state and demographic group. This is a very stringent test. For example, although, as noted above, a simple fifth-order polynomial in time fits national rates quite well, it completely fails to capture how changes in welfare and work participation rates have differed across states and demographic groups. A model with state \times year effects would fit changes over time by state while failing to fit changes over time by demographic group, yet it would have many more parameters than our model. Thus one can easily envisage specifications with many more parameters than ours that would nevertheless fail to fit well in all the dimensions we examine.

Interpreting the Estimates

In models with many interaction terms, individual coefficient estimates become difficult to interpret. Thus, instead of presenting our parameter estimates, we try to give an intuition of what the estimates mean by presenting predicted probabilities of welfare participation for a set of single mothers with different demographic characteristics under a variety of policy regimes. We focus on the model's implications regarding the different impacts of work requirements, time limits, and the unemployment rate.

Table 6 reports the probability of welfare participation as predicted by the model for sixteen different types of single mothers under two policy regimes: one without any work requirement or time limit, the other with both a work requirement and a time limit. To obtain our sixteen representative types, we vary the mother's race, education, marital status, and age of youngest child while holding other characteristics fixed. For each dimension we consider only two settings: black versus white, high school dropout versus college graduate, never married versus divorced, and youngest child age 2 versus youngest child age 13. Regarding the other characteristics, it is assumed that each woman has two children, with the older child age 15; that the woman herself is age 35; that they live in a state where monthly welfare benefits are \$500, and so forth. We also vary the economic environment by setting the unemployment rate at either 4 percent, 6 percent, or 8 percent.

Table 6. Probability of Welfare Participation among Single Mothers of Differing Demographic Characteristics in Response to Work Requirements and Time Limits
Percent

Demographic characteristic ^a			Unemployment rate											
			4 percent				6 percent				8 percent			
Race	Education	Marital status	Age of youngest child	No WR or TL ^b	WR and TL	Difference	No WR or TL	WR and TL	Difference	No WR or TL	WR and TL	Difference		
White	HS dropout	Never married	2	46.03	32.35	13.68	46.43	31.97	14.46	46.59	31.95	14.64		
Black	HS dropout	Never married	2	51.93	37.59	14.34	54.12	38.90	15.12	55.66	39.76	15.90		
White	College graduate	Never married	2	9.59	7.36	2.23	10.25	7.9	2.36	11.33	8.20	3.13		
Black	College graduate	Never married	2	17.43	14.80	2.63	17.94	14.93	3.01	18.80	15.01	3.79		
White	HS dropout	Divorced	2	38.92	27.22	11.71	40.26	27.57	12.69	42.95	29.48	13.47		
Black	HS dropout	Divorced	2	45.82	33.26	12.56	47.94	37.46	10.48	49.42	38.29	11.13		
White	College graduate	Divorced	2	4.03	3.26	0.77	4.08	3.50	0.58	4.09	3.51	0.58		
Black	College graduate	Divorced	2	11.33	10.87	0.45	11.76	10.53	1.24	11.86	10.44	1.42		
White	HS dropout	Never married	13	39.81	30.39	9.43	44.41	34.20	10.21	44.95	34.73	10.21		
Black	HS dropout	Never married	13	43.27	33.18	10.09	49.37	38.50	10.87	49.39	38.20	11.19		
White	College graduate	Never married	13	10.78	11.47	-0.68	11.29	13.19	-1.90	12.93	13.91	-0.98		
Black	College graduate	Never married	13	12.24	14.26	-2.03	16.24	17.48	-1.24	16.95	17.41	-0.46		
White	HS dropout	Divorced	13	33.56	25.91	7.65	34.76	26.27	8.49	35.93	26.72	9.21		
Black	HS dropout	Divorced	13	37.01	28.70	8.31	39.66	30.57	9.09	42.95	33.08	9.87		
White	College graduate	Divorced	13	1.53	4.29	-2.76	1.58	5.26	-3.68	1.59	5.43	-3.84		
Black	College graduate	Divorced	13	5.98	9.38	-3.40	6.54	9.55	-3.02	8.56	10.79	-2.24		

Source: Calculated from results of authors' regressions.

a. It is assumed that each woman has two children, with the older child age 15; that the woman herself is age 35; and that they live in a state where the monthly welfare benefit is \$500.
b. WR, work requirements; TL, time limit.

Table 6 shows that our model yields plausible response patterns. In all cases the more educated women have much lower predicted rates of welfare participation. The drop in welfare participation (in percentage points) in response to the imposition of time limits and work requirements is much greater for high school dropouts than for college graduates (who should be relatively insensitive to welfare policy). Typically, blacks respond more to work requirements and time limits than do whites. And women with younger children respond more than women with older children. The model also predicts that welfare participation rates are higher, and the welfare-reducing effects of work requirements and time limits slightly greater, when unemployment is higher.

Explaining the Drop in Welfare Participation and the Increase in Work

Here we present the central element of our analysis, which uses the model to decompose the contributions of various welfare reform components and of other economic as well as policy variables to both the drop in the welfare participation rate and the increase in the work participation rate from 1993 to 2002. Our approach is as follows. We conduct six counterfactual experiments, which are detailed below. In each experiment we use the model to simulate what welfare and work participation would have been from 1994 through 2002 under the hypothesis that a specific economic or policy variable of interest stayed fixed at its 1993 level, while all other policy and economic variables followed their actual post-1993 paths. The difference between the predicted welfare (or work) participation rate under the experiment and that observed when the variable in question is allowed to take its actual historical path is then said to be the contribution of that variable to the change in welfare (or work) participation from 1994 through 2002. The six experiments are as follows:

—*No time limit.* We assume that no states implement time limits. The counterfactual data are generated by setting DTL_{st} (and thus all terms interacting with DTL_{st}) to zero for all years from 1993 onward.

—*No work requirement.* We assume that no states implement work requirements. The counterfactual data are generated by setting $DWORKREQ_{st}$ (and its interaction terms) to zero from 1993 onward.

—*No EITC expansion.* We assume that the federal and state EITC phase-in rates of 1993 are maintained through 2002, and that the real

value of the maximum EITC credit stays fixed at the 1993 level. Recall that $EITC_RATE_{ist}$ and $EITC_MAX_{ist}$ are both individual-level policy variables. Thus we hold the way they vary with family size fixed as per the 1993 rules as well.

—*No unemployment rate change.* We assume that state unemployment rates do not change from 1993 onward.

—*No CCDF expenditure.* We assume that states do not have child care subsidy programs. The counterfactual data are generated by setting $CHILDCARE_{st}$ (and its interaction terms) to zero.

—*No Medicaid expansion.* We assume that Medicaid does not expand from 1993 onward. We construct counterfactual values of $MEDICAID_PCT_{ist}$ and $MEDICAID_FPL_{ist}$ for all individuals observed after $t > 1993$ using the Medicaid rules used in state s in 1993.

DECOMPOSING THE CONTRIBUTIONS TO THE WELFARE PARTICIPATION RATE DROP. Table 7 summarizes our results on the effects of various welfare reform policies on welfare participation, by year from 1997 through 2002. The first data column reports the percentage-point change in welfare participation from 1993 until that year, as predicted by our model. The remaining columns of the table correspond to various policy changes. In each case we report how many percentage points less the welfare participation rate would have dropped if that policy change had not been implemented.

For example, in the row for 2002 in the top panel of table 7, the first data column indicates that our model predicts a welfare participation rate drop of 23.8 percentage points from 1993 to 2002. The next two columns indicate that, had time limits not been implemented in any state, the drop in welfare participation would have been 2.5 percentage points less, which is equal to 10.6 percent of the overall 23.8-percentage-point drop in participation. Thus our model implies that time limits were a relatively small factor in generating the overall caseload decline.

In contrast, the next two columns of table 7 show that, according to our model, the drop in the welfare participation rate from 1993 to 2002 would have been 13.6 percentage points less if no states had implemented work requirements, and thus that work requirements accounted for 57 percent of the decline in welfare participation among single mothers from 1993 to 2002.

According to our model, the second-largest factor driving down welfare participation was EITC expansion, as shown in the next two columns

Table 7. Factors Contributing to the Cumulative Decline in Welfare Participation among Single Mothers after 1993, by Race of Mother and Year^a

Year	Decrease in participation rate from 1993 (percentage points) ^b	Factor																	
		Time limits			Work requirements			EITC			Unemployment rate			CCDF			Medicaid		
		Contribution (pct. of total)	Percentage	Per cent of total	Contribution (pct. of total)	Percentage	Per cent of total	Contribution (pct. of total)	Percentage	Per cent of total	Contribution (pct. of total)	Percentage	Per cent of total	Contribution (pct. of total)	Percentage	Per cent of total	Contribution (pct. of total)	Percentage	Per cent of total
<i>All single mothers</i>																			
1997	11.09	1.68	15.15	4.73	42.69	4.46	40.25	2.38	21.49	-0.36	-3.23	0.05	0.46						
1998	14.43	3.06	21.19	6.74	46.68	4.67	32.33	2.69	18.60	-0.63	-4.35	0.25	1.76						
1999	17.23	3.21	18.61	8.77	50.93	4.68	27.14	2.79	16.20	-0.85	-4.93	-0.04	-0.23						
2000	19.58	3.11	15.89	11.07	56.54	4.81	24.56	2.86	14.61	-0.88	-4.50	-0.12	-0.61						
2001	21.44	2.33	10.85	11.86	55.29	5.38	25.07	2.23	10.39	-0.43	-2.02	-0.13	-0.59						
2002	23.77	2.52	10.60	13.61	57.26	6.21	26.14	1.58	6.63	-0.11	-0.44	-0.09	-0.39						
<i>Whites</i>																			
1997	8.85	1.64	18.54	4.50	50.82	3.26	36.87	1.80	20.30	-0.19	-2.15	0.17	1.94						
1998	12.96	3.20	24.70	6.36	49.11	3.60	27.75	2.00	15.40	-0.34	-2.59	0.42	3.26						
1999	15.11	3.16	20.92	8.35	55.21	3.64	24.07	2.01	13.32	-0.44	-2.88	0.09	0.59						
2000	16.81	2.89	17.18	10.66	63.40	3.77	22.45	2.03	12.08	-0.52	-3.07	0.02	0.11						
2001	18.18	1.93	10.59	11.38	62.57	4.31	23.69	1.41	7.75	-0.09	-0.52	0.02	0.12						
2002	20.20	1.88	9.30	13.02	64.44	5.31	26.28	0.99	4.88	0.07	0.36	0.05	0.26						
<i>Blacks</i>																			
1997	14.73	1.84	12.53	5.09	34.54	6.95	47.21	3.74	25.38	-0.71	-4.84	-0.27	-1.80						
1998	16.85	3.02	17.93	7.17	42.54	6.88	40.83	4.33	25.69	-1.29	-7.66	-0.16	-0.92						
1999	20.97	3.80	18.12	9.16	43.66	6.95	33.15	4.71	22.45	-1.75	-8.37	-0.38	-1.82						
2000	24.33	4.23	17.40	11.22	46.14	7.14	29.34	5.00	20.54	-1.85	-7.59	-0.52	-2.13						
2001	27.12	4.13	15.21	11.89	43.83	7.90	29.14	4.41	16.28	-1.33	-4.91	-0.55	-2.02						
2002	30.09	4.88	16.20	13.56	45.05	8.38	27.85	3.18	10.58	-0.62	-2.07	-0.55	-1.81						

Source: Authors' calculations.

a. Contributions may not sum to total decrease, and percentages of total may not sum to 100, because of possible interactions among factors.

b. As predicted by the model described in the text.

of table 7. Our estimates imply that this factor accounted for 6.2 percentage points of the drop in welfare participation from 1993 through 2002, or 26 percent.

The next two columns of table 7 report the effect of the unemployment rate. Interestingly, according to the model, from 1993 through 1997 the unemployment rate accounts for a 2.4-percentage-point drop in the welfare participation rate, which was 21 percent of the overall decline up until that time. However, in the recession of 2001–02, the impact of unemployment is lessened, because the unemployment rate in 2002 was no longer so much lower than it had been in 1993. Thus, for the whole 1993–2002 period, our model says that macroeconomic conditions account for only 1.6 percentage points, or 7 percent, of the total decline in the welfare participation rate.

Aside from work requirements, the EITC, time limits, and the macroeconomy, no other variables seemed to have a major effect on the evolution of welfare caseloads.⁷⁷ Table 7 also reports results for CCDF expansion and Medicaid expansion, both of which had very small predicted effects. In fact, these effects are of the “wrong” sign relative to our expectations, but they are so close to zero that we doubt they are significant. These findings could have several explanations. For example, many states give preference to TANF recipients or to women transitioning off

77. Note that total shares do not necessarily sum to less than one. The reason is that, in the actual model, we included interaction terms among various combinations of the policy variables. Our method of decomposition, however, assumes that in each counterfactual only one variable deviates from the actual. Previous research, such as CEA (1997, 1999), suggested that the strictness of sanctions for failure to satisfy work requirements is a key factor. A related variable is the ease with which one can obtain exemptions from work requirements. Our variables capturing these aspects of policy are whether a state has a full or partial (ultimate) benefit sanction for failure to satisfy work requirements, and the number of work requirement exemptions allowed (maximum = 3). To examine the importance of these variables, we conducted two counterfactual experiments. In the first, all economic and policy variables were kept at their actual values, except that all states are assumed to be “lenient” (with only partial sanctions and three exemptions). In the other, all states are assumed to be “strict” (with full sanctions and no exemptions). Our model predicts that welfare participation would have been 1.5 percentage points higher in 2002 under the lenient regime than under the strict regime, and that work participation would have been 0.5 percentage point higher under the strict regime. Thus the strictness of work requirements does have a noticeable effect (about half as large as the effect of time limits), but it is far less important than work requirements per se.

TANF in the allocation of limited CCDF funds. This could actually create an incentive for TANF participation.⁷⁸

The bottom two panels of table 7 examine the determinants of the fall in the welfare participation rate separately by race. According to our model, macroeconomic conditions played a larger role in the decline for black single mothers than for whites. This is consistent with the notion that employment opportunities are more sensitive to macroeconomic conditions for blacks than for whites. In fact, our results in table 9 below confirm this. (Table 9 is similar to table 7, except that it examines the increase in work participation rates, rather than the decrease in welfare participation rates.) According to table 9, changes in the macroeconomy led to a 4.4-percentage-point increase in the work participation rate for black single mothers over the 1993–2002 period, but only a 1.9-percentage-point increase in the work participation rate for whites. Our model also implies that work requirements are relatively more important in explaining the rise in the work participation rate for whites than for blacks, whereas time limits played a relatively larger role for blacks.

Table 8 examines the determinants of the drop in welfare participation from 1993 to 2002 separately by demographic group. The first panel breaks down the effects of different policies by age of the single mother's youngest child. Regardless of the youngest child's age, the importance of time limits is small compared with that of work requirements and the EITC. There is evidence that time limits are more important for single mothers with younger children.⁷⁹ However, consistent with our earlier discussion, the difference is much less apparent if one looks at percentage changes, since single mothers with younger children start from a much higher base participation rate.

The second panel in table 8 shows that time limits were a much more important factor for single mothers who are high school dropouts than for those with a high school but not a college diploma. This is true both in percentage-point terms (7 percentage points versus 2) and in percentage terms (19 percent of the drop in welfare participation versus 9 percent). This is what we would expect in a dynamic model, since mothers who are

78. See U.S. General Accounting Office (1998b).

79. This is consistent with results in Grogger (2004) and Grogger and Michalopoulos (2003).

Table 8. Factors Contributing to the Cumulative Decline in Welfare Participation among Single Mothers after 1993, by Demographic Group^a

Demographic group	Factor																		
	Decrease in participation rate from 1993 (percentage points) ^b	Time limits			Work requirements			EITC			Unemployment rate			CCDF			Medicaid		
		Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total
<i>Age of youngest child</i>																			
0-5 years	32.85	4.84	14.72	19.06	58.03	7.72	23.50	0.66	2.02	-0.87	-2.63	-0.60	-1.82						
6-12 years	18.01	1.81	10.04	11.47	63.70	5.76	31.96	2.79	15.47	0.74	4.10	-0.18	-1.02						
<i>Educational attainment</i>																			
Less than high school	36.16	7.03	19.44	19.85	54.90	9.97	27.57	2.16	5.97	-0.24	-0.65	-0.16	-0.46						
High school but not college degree ^c	20.97	1.99	9.48	13.82	65.93	6.10	29.11	1.75	8.34	-0.33	-1.57	-0.64	-3.05						
<i>Marital status</i>																			
Never married	37.37	6.42	17.17	16.73	44.75	10.25	27.42	1.86	4.97	-0.93	-2.49	-0.60	-1.61						
Separated	24.99	1.95	7.79	16.66	66.67	3.71	14.86	1.34	5.37	-0.18	-0.70	-0.04	-0.15						
Divorced	14.58	-1.01	-6.96	10.20	69.94	3.87	26.55	1.71	11.74	0.80	5.50	0.30	2.05						
<i>No. of children</i>																			
One	19.55	0.63	3.22	12.54	64.13	5.83	29.84	1.23	6.31	-0.20	-1.00	0.12	0.60						
Two or more	28.06	4.78	17.03	14.89	53.05	6.66	23.74	1.97	7.04	-0.01	-0.05	-0.34	-1.23						

Source: Authors' calculations.

a. Contributions may not sum to total decrease, and percentages of total may not sum to 100, because of possible interactions among factors.

b. As predicted by the model described in the text.

c. Combines those with a high school diploma only and those with some college but not a bachelor's degree.

high school dropouts have higher rates of unemployment and therefore a greater incentive to bank eligibility under time limits.

The third and fourth panels of table 8 show that time limits are relatively important for never-married single mothers and for single mothers with two or more children. This is again consistent with these groups having relatively high baseline unemployment rates, implying that they have a greater incentive to conserve their eligibility.

DECOMPOSING THE CONTRIBUTIONS TO THE WORK PARTICIPATION RATE INCREASE. Table 9 summarizes our results on the effects of various welfare reform policies on work participation. According to the top panel, out of the overall 10.8-percentage-point predicted increase in work from 1993 to 2002,⁸⁰ the model implies that 3.6 percentage points (33 percent) was due to the EITC expansion, 2.7 percentage points (25 percent) to macroeconomic conditions, 1.8 percentage points (17 percent) to work requirements, and 1.1 percentage points (10 percent) to time limits.

Thus the ranking of the policy variables in terms of their impact on work participation is drastically different from that for welfare participation. Macroeconomic conditions, as captured by local unemployment rates, were until 2001 the most important contributor to the increase in work participation. For the 1993–2000 period, the macroeconomy accounts for more than 40 percent of the total increase in work participation. But its contribution has recently dropped, reflecting the recession that began in March 2001. By 2002 the EITC had become the most important factor.⁸¹

The top panel of table 10 examines the determinants of the increase in work in 1993–2002 separately by age of the youngest child. A key result

80. Recall that the work participation rate of single mothers actually increased by 11.3 percentage points from 1993 to 2002 in the data.

81. According to the top panel of table 7, our model implies that the EITC generated 6.21 percentage points of the drop in welfare participation from 1993 to 2002. Thus it may seem puzzling that, according to table 9, the EITC accounts for only a 3.6-percentage-point increase in the work participation rate. Presumably, the EITC gets women off welfare by getting them to work, and therefore one might expect that these effects should be roughly equal. The discrepancy arises because, as discussed earlier, work and welfare are not mutually exclusive. Expansion of the EITC encourages some single mothers on welfare who were working to leave welfare and continue to work. This reduces welfare participation while not increasing work participation. Thus the number of single mothers who leave welfare because of EITC expansion should be larger than the number who start working because of the expansion, and this is consistent with what we find. Recall that, in general, the total decrease in welfare participation (23 percentage points) was more than twice as great as the total increase in work participation (11.3 percentage points).

Table 9. Factors Contributing to the Rise in Work Participation among Single Mothers after 1993, by Race of Mother and Year^a

Year	Increase in participation rate from 1993 (percentage points) ^b	Factor																							
		Time limits			Work requirements			EITC			Unemployment rate			CCDF			Medicaid								
		Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total	Contribution (pct. of total points)	Per cent of total	Per cent of total						
<i>All single mothers</i>																									
1997	7.99	0.30	3.81	1.83	22.92	0.67	8.39	4.73	59.18	0.51	6.34	0.18	2.23	0.30	3.81	1.83	22.92	0.67	8.39	4.73	59.18	0.51	6.34	0.18	2.23
1998	10.96	1.25	11.43	1.82	16.60	1.05	9.55	5.21	47.51	0.81	7.38	0.30	2.77	0.30	3.81	1.82	16.60	1.05	9.55	5.21	47.51	0.81	7.38	0.30	2.77
1999	12.35	0.98	7.90	1.61	13.01	2.07	16.77	5.63	45.58	1.14	9.20	0.32	2.57	0.32	4.06	1.61	13.01	2.07	16.77	5.63	45.58	1.14	9.20	0.32	2.57
2000	13.48	0.85	6.32	1.73	12.83	2.71	20.10	5.42	40.18	1.28	9.48	0.32	2.36	0.32	4.06	1.73	12.83	2.71	20.10	5.42	40.18	1.28	9.48	0.32	2.36
2001	12.26	0.91	7.46	0.79	6.46	3.33	27.14	4.06	33.14	0.75	6.08	0.32	2.58	0.32	4.06	0.91	7.46	3.33	27.14	4.06	33.14	0.75	6.08	0.32	2.58
2002	10.82	1.11	10.25	1.79	16.54	3.61	33.31	2.66	24.55	0.06	0.59	0.33	3.00	0.33	4.06	1.11	10.25	3.61	33.31	2.66	24.55	0.06	0.59	0.33	3.00
<i>Whites</i>																									
1997	6.16	0.10	1.54	1.90	30.78	0.74	11.98	3.85	62.49	0.38	6.21	0.09	1.38	0.09	1.54	1.90	30.78	0.74	11.98	3.85	62.49	0.38	6.21	0.09	1.38
1998	9.63	0.92	9.51	2.22	23.08	1.17	12.13	4.23	43.87	0.50	5.15	0.22	2.27	0.22	9.51	2.22	23.08	1.17	12.13	4.23	43.87	0.50	5.15	0.22	2.27
1999	10.56	0.58	5.48	2.19	20.69	2.14	20.28	4.34	41.06	0.71	6.69	0.21	2.01	0.21	5.48	2.19	20.69	2.14	20.28	4.34	41.06	0.71	6.69	0.21	2.01
2000	11.52	0.28	2.46	2.64	22.87	2.71	23.56	4.17	36.17	0.83	7.24	0.22	1.87	0.22	2.46	2.64	22.87	2.71	23.56	4.17	36.17	0.83	7.24	0.22	1.87
2001	10.21	0.29	2.84	2.04	20.01	3.34	32.72	2.96	28.98	0.31	3.07	0.21	2.06	0.21	2.84	2.04	20.01	3.34	32.72	2.96	28.98	0.31	3.07	0.21	2.06
2002	9.58	0.24	2.46	3.45	35.95	3.70	38.64	1.86	19.44	-0.17	-1.78	0.22	2.29	0.22	2.46	3.45	35.95	3.70	38.64	1.86	19.44	-0.17	-1.78	0.22	2.29
<i>Blacks</i>																									
1997	10.61	0.61	5.76	1.70	15.98	0.40	3.78	6.63	62.52	0.77	7.23	0.43	4.02	0.43	5.76	1.70	15.98	0.40	3.78	6.63	62.52	0.77	7.23	0.43	4.02
1998	12.90	1.74	13.45	1.14	8.84	0.60	4.63	7.42	57.53	1.41	10.93	0.57	4.41	0.57	13.45	1.14	8.84	0.60	4.63	7.42	57.53	1.41	10.93	0.57	4.41
1999	15.45	1.51	9.79	0.70	4.55	1.64	10.60	8.61	55.70	1.97	12.77	0.63	4.05	0.63	9.79	0.70	4.55	1.64	10.60	8.61	55.70	1.97	12.77	0.63	4.05
2000	16.83	1.75	10.42	0.18	1.09	2.44	14.47	8.44	50.16	2.15	12.74	0.65	3.85	0.65	10.42	0.18	1.09	2.44	14.47	8.44	50.16	2.15	12.74	0.65	3.85
2001	15.65	1.87	11.97	-1.17	-7.48	3.08	19.66	6.72	42.92	1.62	10.33	0.68	4.33	0.68	11.97	-1.17	-7.48	3.08	19.66	6.72	42.92	1.62	10.33	0.68	4.33
2002	12.93	2.49	19.23	-0.55	-4.27	3.06	23.64	4.40	34.06	0.49	3.80	0.69	5.33	0.69	19.23	-0.55	-4.27	3.06	23.64	4.40	34.06	0.49	3.80	0.69	5.33

Source: Authors' calculations.

a. Contributions may not sum to total decrease, and percentages of total may not sum to 100, because of possible interactions among factors.

b. As predicted by the model described in the text.

Table 10. Factors Contributing to the Cumulative Rise in Work Participation among Single Mothers after 1993, by Demographic Group^a

Demographic group	Increase in participation rate from 1993 (percentage points) ^b	Factor											
		Time limits		Work requirements		EITC		Unemployment rate		CCDF		Medicaid	
		Contri- bution (pct. of total points)	Per- cent of total	Contri- bution (pct. of total points)	Per- cent of total	Contri- bution (pct. of total points)	Per- cent of total	Contri- bution (pct. of total points)	Per- cent of total	Contri- bution (pct. of total points)	Per- cent of total	Contri- bution (pct. of total points)	Per- cent of total
<i>Age of youngest child</i>													
0-5 years	16.79	2.21	13.15	3.50	20.86	3.91	23.30	3.02	17.97	0.26	1.55	0.03	0.16
6-12 years	6.48	0.97	14.99	1.25	19.25	3.70	57.10	2.57	39.64	0.00	-0.01	0.29	4.48
<i>Educational attainment</i>													
Less than high school	17.16	2.59	15.08	7.13	41.52	2.70	15.73	3.12	18.19	0.91	5.32	0.53	3.10
High school but not college degree ^c	7.24	1.09	15.11	0.50	6.89	4.01	55.39	2.89	39.97	-0.24	-3.26	-0.22	-3.00
<i>Marital status</i>													
Never married	18.15	2.51	13.80	0.86	4.75	5.13	28.26	4.06	22.34	0.39	2.16	-0.07	-0.37
Separated	12.84	1.93	15.02	2.86	22.30	3.16	24.59	2.25	17.55	0.00	-0.02	0.20	1.53
<i>No. of children</i>													
One	6.98	0.15	2.15	1.20	17.25	2.83	40.62	2.53	36.26	0.07	1.02	0.42	6.08
Two or more	14.89	2.26	15.16	2.49	16.74	4.53	30.41	2.80	18.82	0.05	0.36	0.21	1.38

Source: Authors' calculations.

a. Contributions may not sum to total decrease, and percentages of total may not sum to 100, because of possible interactions among factors.

b. As predicted by the model described in the text.

c. Combines those with a high school diploma only and those with some college but not a bachelor's degree.

is that the macroeconomy has been much less important relative to other factors for mothers with young children (those age 5 and under). For this group our model says that EITC expansion and work requirements were the largest factors increasing work over the 1993–2002 period, accounting for 44 percent of the increase. This is not surprising, because mothers with young children are traditionally much less likely to participate in the labor market, and therefore they should be less sensitive to macroeconomic conditions.

The fact that single mothers with young children have, more than other groups, been forced into employment by work requirements does raise concerns about the impact of a mother's work on her young children. This is an important topic for future research.

The second panel of table 10 contains some interesting results on how various policies have different effects on single mothers who are high school dropouts than on those with a high school education or more. It is striking that work requirements account for 42 percent of the 17-percentage-point increase in work participation among high school dropouts, but a negligible part of the 7-percentage-point increase in work participation among the more educated single mothers. The increase in work for the more educated group is driven almost entirely by the EITC (55 percent) and by the macroeconomy (40 percent). Yet these results are exactly what one would expect. The more educated women have higher offer wage rates and are therefore more likely to have been close to the margin where they would be better off working than on welfare. A stronger macroeconomy (and, for some at least, the EITC wage subsidy) could easily push them over that margin. On the other hand, the high school dropouts have poorer labor market opportunities, so that a work requirement may be necessary to push them off welfare and into market work. The third and fourth panels of the table show the decomposition by marital status and number of children.

This also suggests that the impact of welfare reform on “welfare” in the economic sense for these two groups of women may be radically different. Women who choose to work because an improved economy and enhanced EITC have driven up their effective wage rates must be better off. Women who entered the labor market because of a binding constraint induced by work requirements must be worse off. This is an important topic for future research.

To sum up: our simulations seem to have a great deal of face validity. When we predict that different policies have had different effects on different groups, the differences are in line with what one would expect in light of economic theory.

ROBUSTNESS CHECKS. Given that our model contains 245 variables, to allay concerns that our results might be sensitive to possible collinearity problems, we also estimated a “sparse” specification that eliminated eighty-eight of the interaction terms (those indicated in brackets in table C2). We dropped these terms because we judged, a priori, that they represented interaction effects that would be relatively weak or subtle.⁸² This simpler model fit nearly as well as the full model, and it produced very similar predictions. We take this as evidence that collinearity is not a concern.

Some critics have suggested that the model succeeds in explaining the recent dramatic drop in caseloads because it includes the variable *MONTH_SINCE_WR_START_{st}*, which plays a role similar to a linear time trend that starts around 1996–97 in most states. These critics argue that even if we had randomly assigned the state-specific policy variables to the individual women, regardless of their true state of residence, the model would still fit the data well. To address this concern, we constructed an artificial data set in which welfare policy variables were indeed randomly assigned to each woman. Specifically, we assigned to each woman in the CPS data, with equal probability, the set of policy variables appropriate for one of the fifty states or the District of Columbia.

The resulting model fit the aggregate patterns of welfare and work participation and the patterns by demographic group rather well. But it fit the state data quite poorly. Not surprisingly, it largely missed the important cross-state differences in both levels of and changes in welfare and work participation that we discussed in detail above. In particular, it completely misses the fact that welfare participation peaked much earlier than 1993–94 in many states. This model also produced some very odd predictions of policy effects. For example, it implied that work requirements accounted for almost the entire drop in welfare participation, that the

82. More specifically, the basic rationale was as follows: In our model each of several policy areas, such as time limits and work requirements, is characterized by several variables. In the full model the demographics are interacted with each of the variables within each policy area. In the sparse model we interact the demographics with only the one or two variables within each policy area that we judged a priori to be the most important.

macroeconomy played a negligible role in the 1993–99 period (and, in fact, that it slightly *reduced* employment), and that time limits slightly *increased* welfare participation (while nevertheless accounting for a large part of the increase in work). We take these very odd results as confirmation that it is important to carefully code policy variables at the state and the individual level in order to avoid collinearity and provide plausible estimates of policy effects.

Conclusion

It has been a decade since states began implementing welfare reform under AFDC waivers, and seven years since the overhaul of the U.S. welfare system under PRWORA. Judging from the more than 60 percent drop in welfare caseloads and welfare participation rates, and the close to 20 percent increase in work participation rates among single mothers, these policies would seem to have been a major success. However, this success was achieved during one of the greatest economic expansions since World War II and amid a wide range of other economic and policy changes, most notably a dramatic expansion in the EITC, Medicaid, and CCDF expenditure. To make better policy in the future, it is important to understand what exactly each of these various policy components contributed to the behavioral changes observed over the past decade.

Whereas many researchers have studied the impact of particular policies or subsets of policies, this paper is more ambitious in that we have made a major effort to compile, at the state level, measures of all the key policy and economic environment variables that we think may have substantially influenced the behavior of single mothers over the 1980–2002 period. We then merged these policy data with individual-level data from the March CPS from 1981 to 2003. Using these data, we developed and tested a model that successfully explains both the levels of and changes in welfare and work participation rates across states, across time, and across demographic groups—all without using state and year dummy variables—for the 1980–2002 period.

We then used our estimated model to disentangle the contributions of various components of the welfare reforms, as well as other contemporaneous economic and policy changes, to the changes in welfare and work

participation rates of single mothers from 1993 to 2002. Our main findings are that the key economic and policy variables that account for the 23-percentage-point decrease in the welfare participation rate were work requirements (57 percent of the decrease), the EITC (26 percent), time limits (11 percent), and macroeconomic conditions (7 percent). The main factors contributing to the 11-percentage-point increase in the work participation rate of single mothers during 1993–2002 were the EITC (33 percent), macroeconomic conditions (25 percent), work requirements (17 percent), and time limits (10 percent).

The results of the model imply some important differences across demographic groups in the impact of these policies. For instance, whereas economic conditions and the EITC largely explain the increase in work among relatively well educated single mothers, work requirements were a much more important factor for high school dropouts. This is not surprising: since more-educated mothers presumably command higher wages, an enhanced EITC wage subsidy plus a stronger economy could easily push them over the margin from choosing welfare to choosing to work. On the other hand, if women who have dropped out of high school enter the labor market because of a binding constraint induced by work requirements, they are presumably made worse off. Thus how welfare reform has affected the well-being of high school dropout single mothers and their children is an important topic for future research.

Our research also highlights the crucial difference between leaving welfare and going to work. A troubling development is that about one-quarter of the welfare leavers actually did not enter the work force. What are the characteristics of these people? What happens to their children's well-being and to their own? These, too, are important questions for future research. In this regard the EITC seems to be a particularly attractive policy because it scores high both as a factor reducing welfare participation and as a factor increasing work. Work requirements, on the other hand, seem to be much more effective at getting single mothers to exit welfare than at getting them to enter market work.

This paper has used a simple methodology to address some important policy questions. But we assumed the exogeneity of educational attainment, marital status, and number of children. In a life-cycle model with forward-looking mothers, these demographic characteristics will certainly be endogenous. Thus yet another important topic for further

research is how welfare policy affects education, marriage, and fertility decisions.⁸³

Another issue is that, in a dynamic framework, not just current policy measures but also expected future policy measures affect current decisions. A fundamental tension in previous research on the impact of welfare time limits (both benefit eligibility and work requirement time limits), including our own work reported here, is that the incentive to bank time that is estimated in this work exists only if women are forward looking. But if women are indeed forward looking, any model that fails to include expected future benefits as a determinant of current choices is misspecified, except under strong assumptions about expectations and the process generating future benefits.

More concretely, it is entirely possible that some part of the welfare participation decline that began in the mid-1990s occurred because single mothers were forward looking and anticipated that welfare participation would become more difficult in the future, because of some combination of work requirements, time limits, and reduced benefits. Anyone who thinks that his or her future welfare participation has become less likely, and future work more likely, will have a greater incentive to invest more in human capital today, including by working. Our modeling framework cannot address this point.

Finally, the reliability of our decomposition of the roles of various factors in reducing welfare and increasing work hinges crucially on the assumption that we have successfully measured and included in our analysis all the key factors involved. If we have omitted any important factor, its effect may be spuriously transferred to the factors that we have included. After months of intensive data collection, we were unable to identify other aspects of the policy or economic environment that we felt could plausibly be considered important. Indeed, we attempted to include a reasonable measure of every aspect of welfare reform and the economic environment that Blank's 2002 survey hypothesized as potentially important.⁸⁴ Of course, it is quite likely that some of our policy *measures* could be improved, but it is difficult to think of any key policy variables we have completely omitted.

83. See Keane and Wolpin (2002a, 2002b, 2003) for some work on this topic.

84. Blank (2002).

Perhaps our most plausible omission is a change in “culture”—an intrinsically difficult-to-quantify concept. A change in culture could take two forms: either a change in the culture of welfare offices toward a “welfare-to-work” emphasis, or a change in the general culture that makes welfare receipt somehow seem less desirable. However, we do not understand how such changes in culture could be generated except through such measurable things as the imposition of work requirements, stronger sanctions for violating work requirements, and the imposition of time limits, all of which we have measured and included. In that case it is quite appropriate, in a reduced-form specification, for the coefficients on these measurable policy instruments to pick up how they affect culture.

APPENDIX A

The Impact of Time Limits

HERE WE LAY out a simple model of welfare participation decisions by forward-looking, wealth-maximizing agents when there are time limits. Suppose that a single mother has only two choices in a given month: welfare participation only (choice 0) and work only (choice 1). The value of each choice is given by

$$W_{0t}(S, T) = B_t + dV_{t+1}(S-1, T-1)$$

$$W_{1t}(S, T) = E_t + dV_{t+1}(S, T-1).$$

Here B_t is the welfare benefit in month t , and E_t is the earnings (determined by her wage offer net of the cost of working) the woman can obtain if she works. The term $V_{t+1}(S, T)$ denotes the expected present value of lifetime wealth at time $t + 1$ given that the woman has S months of eligibility that may be spread over a T -month horizon. The term $W_{0t}(S, T)$ denotes the value of participating in welfare (choice 0) at time t , given that the woman has S months of eligibility that may be spread over a T -month horizon. This equals the current welfare benefit the woman will receive, B_t , plus the discount factor $d \in (0, 1)$ times $V_{t+1}(S-1, T-1)$, which is the woman’s expected present value of lifetime wealth at time $t + 1$ given that she has chosen to be on welfare at t . Note that the first

argument of this function is $S - 1$, since if the woman accepts welfare today, she will have only $S - 1$ periods of eligibility left when she gets to the next period. Similarly, $W_{1t}(S, T)$ denotes the value of working only (choice 1). This equals the current earnings the woman will obtain from working, E_t , plus the discount factor d times $V_{t+1}(S, T - 1)$, the woman's expected present value of lifetime wealth at time $t + 1$ given that she has chosen to work and not be on welfare at t . Note that the first argument of this function is S , since if the woman does not go on welfare today, she will still have S periods of eligibility left when she gets to the next period.

A key point is that $V_{t+1}(S, T - 1) > V_{t+1}(S - 1, T - 1)$ as long as $S < T$. That is, one is better off if one gets to time $t + 1$ with more available months of eligibility remaining. Optimal behavior in this model is to try to time the use of one's S periods of potential welfare participation eligibility to coincide with those periods when the realization of E_t is relatively low. Define $D_{t+1}(S, T) = d[V_{t+1}(S, T - 1) - V_{t+1}(S - 1, T - 1)] \geq 0$ as the "option value" of preserving a month of welfare eligibility. The optimal decision rule for working in this model is to work if

$$W_{1t}(S, T) - W_{0t}(S, T) = E_t - B_t + D_{t+1}(S, T) > 0.$$

Or, more intuitively,

$$E_t + D_{t+1}(S, T) > B_t.$$

The main point is that it is not optimal to choose welfare over work just because $B_t > E_t$. In fact, B_t must exceed E_t by an increment at least as great as the option value of saving a month of eligibility, $D_{t+1}(S, T)$, in order for it to be optimal to choose welfare. This is the basic intuition for why time limits would be expected to reduce welfare participation, even before people have reached the limit.

Things get more complex if we add the option of working and participating in welfare at the same time (choice 2). The value of this option is

$$W_{2t}(S, T) = B_t + (1 - \tau)E_t + dV_{t+1}(S - 1, T - 1),$$

where τ is the rate at which earnings are taxed in the welfare benefit formula. Now, in order for it to be optimal not to participate in welfare, a second condition must hold. It must also be the case that

$$W_1(S, T) - W_{2t}(S, T) = \tau E_t - B_t + D_{t+1}(S, T) > 0.$$

By working only (choice 1) rather than working *and* going on welfare (choice 2), a woman gains τE_t and loses B_t . Equivalently but more intuitively,

$$\tau E_t + D_{t+1}(S, T) > B_t.$$

As the benefit tax rate approaches 0, it becomes less likely that this condition will be satisfied. In fact, as $\tau \rightarrow 0$, the condition approaches $D_{t+1}(S, T) > B_t$, and the woman would always choose welfare. To see this, note that the largest possible value of $D_{t+1}(S, T)$ occurs when $S = 1$ and the woman is certain that she will choose to participate in welfare at $t + 1$. In that case $D_{t+1}(S, T) = dB_t$, since, by using up the month of eligibility, she loses B_t with certainty next month. Thus the condition becomes $dB_t \geq D_{t+1}(S, T) > B_t$, which is impossible for $d < 1$. This further implies that there exists some $\bar{\tau} > 0$ such that it is never optimal to “bank” years of eligibility if $\tau \leq \bar{\tau}$.

We also point out that $D_{t+1}(S, T)$ will be *decreasing* in the pool of extensions, *decreasing* in the fraction of benefits one continues to receive after reaching the time limit, *increasing* in the likelihood of future unemployment, *increasing* in the level of benefits, and *decreasing* in the probability of marriage (or any other event or change in variable that would reduce the probability of participating in welfare in the future). Thus we should interact indicators for time limits with any variables that help determine the above quantities (for example, education could affect the probability of unemployment).

The idea that time limits could have played a major role in the decline of welfare caseloads in the early TANF period of 1996–2001 rests on the presumption that the anticipatory effect is substantial, since few people were subject to binding time limits before 2002. But it strikes us as implausible that the effect of time limits could have been substantial, given how TANF was actually implemented. The very simple analysis of the anticipatory effect presented above ignores several crucial features of actual state TANF plans. Most obviously, we have noted that a large percentage of the caseload resided in states that did not enforce a time limit or that stopped the clock for working participants. Other features of many actual state plans that reduce the option value of banking months of eligibility include generous earnings disregards for employed TANF participants, and rather modest partial benefit reductions when the time limit is reached.

The ways in which partial benefit reductions and generous treatment of earned income reduce incentives to bank time can be clarified using some simple numerical examples, which also help elucidate how the AFDC/TANF benefit formulas work. For example, in Illinois the monthly TANF benefit for a family of three with no income is \$377, and the benefit reduction rate is 33 percent of earnings. A woman working 130 hours a month at \$6.00 an hour would be taxed \$257 (ignoring work expense deductions), leaving \$120 per month in TANF benefits if she decided to participate. In principle, there might be an incentive to pass up the \$120 in order to preserve eligibility to get the full \$377 in some future month, if the woman thought it likely that she would face some future protracted unemployment spell. But in Illinois, if a woman works while receiving welfare, that time is not counted against the clock, so there is no such incentive.

Even if work did not stop the clock, it is not at all clear that banking the month would be optimal in this situation. It could be optimal to pass up a certain \$120 today in favor of \$377 at some hypothetical future point only if the probability of future unemployment were quite high. For example, a just-divorced woman with an 8-year-old child facing this choice in a state with a five-year time limit should only begin to consider passing up the \$120 today if she feels there is a nonnegligible probability that she will be unable to find work for five years out of the next ten (her time horizon until the child reaches 18). Otherwise there is almost no chance she will ever be able to use the banked time. Accounting for discounting or for the fact that she might remarry in the future would further diminish the option value to preserving months of eligibility.

Some states have even more generous earnings disregards. For example, Connecticut exempts 100 percent of earnings up to the point where total income from work and benefits reaches the poverty line. Under this system there is no incentive whatsoever to bank months, so long as the person can save and the sum of earnings plus benefits does not exceed the poverty line. As long as she discounts future income, a woman in the circumstances described above and living in Connecticut should prefer to take her TANF benefits now rather than later. Apparently, the participants and caseworkers in Connecticut realized this. According to Bloom and others (2002, p. 133), "Surveys of recipients and staff [in Connecticut] found that workers did not actively encourage most recipients to leave welfare quickly in order to bank their months of eligibility. In fact, such a

message would not have been credible. . . . Individuals who found employment would usually continue to receive their full welfare grant. . . . Thus, in order to bank months, a recipient would need to give up \$543 per month in benefits.”

As another example, California has a five-year time limit, but the only penalty for reaching the limit is loss of the adult portion of benefits. In 2000 the maximum TANF benefit for a family of three was \$626 a month, whereas that for a family of two was \$505 a month. So the penalty is the loss of only \$121 a month. No one would pass up \$626 today just to preserve eligibility for an additional \$121 in some future month.

Does the option to work while on welfare change the calculation? California disregards the first \$225 of monthly earnings, plus 50 percent of earnings beyond that. Thus, if a woman could earn \$780 a month, her benefit reduction would be \$278. This gives a three-person-family benefit of $\$626 - \$278 = \$348$, and a two-person-family benefit of $\$505 - \$278 = \$227$. Could it ever be optimal to give up \$348 today to preserve eligibility for a benefit of \$348 rather than \$227 in some future month? That would mean reducing this month’s income from \$1,128 to \$780, in order to have an income of \$348 instead of \$227 in the event of some future month of unemployment. One could devise a numerical example in which such a choice would be optimal, by ruling out saving, making marginal utility diminishing extremely strongly in income, making the risk of future unemployment very high, and ruling out any other sources of support when unemployed. But such a scenario seems quite implausible.

To summarize, time limits may make the option of working (and staying off welfare) more attractive relative to welfare by adding an extra element to the value of working, namely, the option value of preserving a future month of welfare eligibility. But, in states with generous earnings disregards and in states that only reduce (rather than eliminate) benefits when the limit is reached, this option value is likely to be small relative to the current TANF benefits that one would have to pass up in order to bank a month of eligibility. Thus, as a practical matter, we expect that any anticipatory effects of time limits in most existing state TANF plans should have been small.

APPENDIX B

Effects of Maintenance-of-Effort Requirements

THE MAINTENANCE-OF-EFFORT provision in Section 409 of PRWORA stipulates that the Department of Health and Human Services may reduce a state's federal TANF block grant if the state fails to maintain its level of assistance for needy families at 75 percent of the historical level.⁸⁵ The "historical level" was defined as peak-year (usually 1994) spending on the whole range of programs replaced by TANF (such as AFDC and AFDC-related child care). This feature was designed to prevent a feared "race to the bottom," in which many states might start to cut assistance once federal AFDC matching funds vanished.

But the MOE requirement has had some dramatic and unexpected consequences. The critical feature of the MOE requirement was that "qualified" expenditures were defined as including not just cash assistance paid through TANF, but a range of non-TANF spending as well. These alternatives included child care assistance and educational and job training activities. Critically, such benefits could be paid to any low-income family, even if they were not TANF recipients. As welfare caseloads dropped dramatically after 1996, causing expenditure on TANF cash assistance to fall, the states were essentially forced by the MOE requirement to redirect money into other qualified programs. To a great extent, the states responded by funneling substantial resources into subsidized day care for low-income families (U.S. General Accounting Office, 1998c). The effect can be seen in figure 4 in the text, which shows the increase in CCDF expenditure from roughly \$3.0 billion in 1995 to \$8.0 billion in 2001.

Since child care is obviously one of the most important costs of working for single mothers with young children, the increase in child care subsidies after 1996 should have provided enhanced work incentives for this group. Interestingly, the MOE requirement can thus create a feedback loop that perpetuates the impact of welfare reform. That is, welfare reforms that reduce caseloads and encourage work cause state spending

85. If a state failed to achieve a required work participation rate for its welfare participants, the MOE requirement could be raised to 85 percent. The work requirement was reduced if a state achieved certain caseload reduction targets. Since caseloads fell so dramatically, these caseload reduction credits rendered the work requirement targets essentially irrelevant until recently.

on cash assistance to fall. This in turn induces states to spend more on day care and other work expense subsidies, which causes caseloads to drop further, in a virtuous cycle. Also interesting is that the MOE rule can create a situation of multiple equilibria, with high state welfare caseloads and low work expense subsidies in one equilibrium, and low caseloads and high subsidies in the other. Moreover, the high-subsidy equilibrium is fiscally sustainable because welfare spending is low. We formalize this argument below and show that, under plausible assumptions about the dynamics of states' budget processes, the equilibrium with high child care subsidies and low welfare participation is the only stable equilibrium.

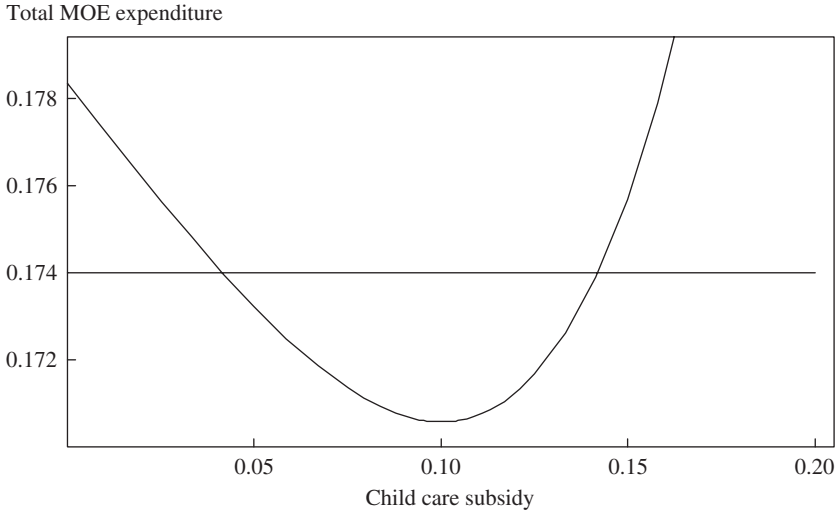
Our model of the effect of the MOE clause on welfare caseloads can be described as follows: Suppose that there is a continuum of single mothers with measure 1 in the population. In every period, single mothers receive a job offer with wages (net of the cost of working) independently drawn from a distribution F , and each decides whether or not to work. If a woman works, she obtains her net wage draw, and she may receive a child care subsidy $s \geq 0$ from the state government. If she chooses not to work, she receives welfare payment $z > 0$. Thus a woman will work if and only if $w + s > z$, or $w > z - s$. Given the policy variable pair (z, s) , the total measure of women participating in welfare is $F(z - s)$.

Following the spirit of the MOE requirement of PRWORA, we assume that the state is required to spend a total of $B > 0$ on assistance to single mothers. We assume that the welfare assistance level z is fixed through time. As the law stipulates, the state government's expenditure on both cash welfare assistance and child care subsidies to low-income women both qualify as MOE expenditure. Thus, for a fixed z , any level of s that satisfies

$$zF(z-s) + s[1-F(z-s)] = B$$

will constitute an equilibrium. Depending on the level of B , multiple levels of s may be consistent with equilibrium (see figure B1 for an illustration).

Now we assume that a state's fiscal allocation is determined in an adaptive fashion as follows. Suppose that, in period t , the state's welfare caseload is given by $F(z - s_t)$, so that the cash welfare expenditure is $zF(z - s_t)$. Then, in period $t + 1$, the government will adjust its child care subsidy s_{t+1} according to

Figure B1. Possible Multiple Equilibria

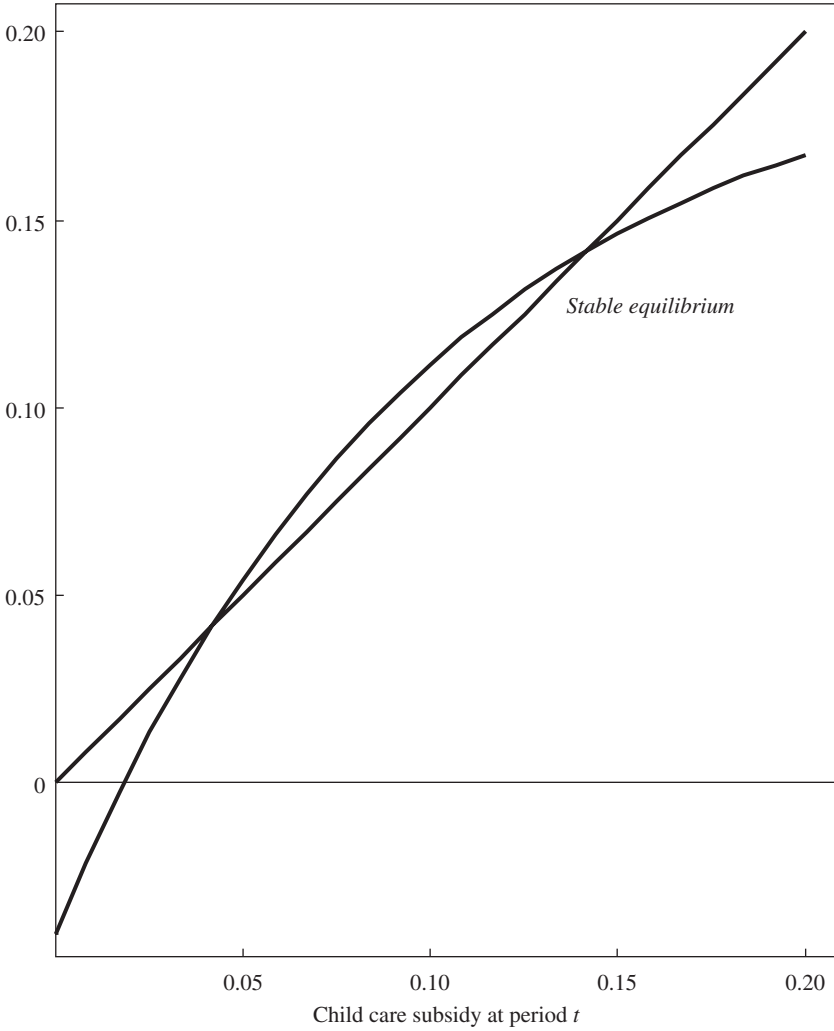
Source: Authors' model described in the text.

$$s_{t+1} = \frac{B - zF(z - s_t)}{1 - F(z - s_t)}.$$

That is, we assume that the state sets the child care subsidy for period $t + 1$ to ensure that the MOE spending amount B is satisfied under the myopic assumption that the number of people who work in period $t + 1$ will remain the same as in period t . Although we do not have direct evidence for this particular specification of the fiscal dynamics, it is certainly plausible. Under reasonable assumptions about the shape of F , the system will have two equilibria, and only the one with the higher level of child care subsidy is stable. Therefore, under the MOE requirement clause, any initial shock experienced by the economy, such as the booming macroeconomy beginning in 1996, will lead the system to converge to an equilibrium with a high child care subsidy and low welfare participation. This equilibrium Pareto-dominates the other equilibrium with a low child care subsidy and high welfare participation (figure B2).

Figure B2. Stable Equilibrium

Child care subsidy at period $t + 1$



Source: Authors' model described in the text.

APPENDIX C

*Variables Used in the Analysis***Table C1. Description of Variables**

<i>Variable</i>	<i>Description</i>
<i>Dependent variables</i>	
WEL_RECEIPT _{ist}	Categorical variable indicating whether woman <i>i</i> received welfare during period <i>t</i>
WORK _{ist}	Categorical variable indicating whether woman <i>i</i> worked at least part time during period <i>t</i>
<i>Individual-level demographic variables</i>	
AGE _{ist}	Continuous variable indicating age of woman <i>i</i> in years
RACE _{ist}	Categorical variable indicating white, black, or other
EDU _{ist}	Categorical variable indicating educational attainment (less than high school, finished high school, some college, or finished college)
MARITAL _{ist}	Categorical variable indicating never married, divorced, separated, or widowed
URBAN _{ist}	Categorical variable indicating urban or rural residence
NCHILD05 _{ist} , NCHILD612 _{ist} , NCHILD1317 _{ist}	Number of children of woman <i>i</i> younger than 6, between 6 and 12, or between 13 and 17, respectively
DCHILD05 _{ist} , DCHIL612 _{ist}	Dummy variable indicating whether any child of woman <i>i</i> is younger than 6 or between 6 and 12, respectively
YOUNGEST _{ist} , OLDEST _{ist}	Age of youngest or oldest child of woman <i>i</i> , respectively
<i>Individual-level policy variables</i>	
WELFARE_BEN _{ist}	Real AFDC or TANF maximum benefit (assuming zero earnings) received by woman <i>i</i> , calculated using her state's benefit rule and her family composition
FOOD_STAMP _{ist}	Real food stamp benefits, taking into account that welfare benefits count toward income
EITC_RATE _{ist} , EITC_MAX _{ist}	EITC phase-in rate and real maximum EITC amount, constructed from both federal and state EITC rules, together with family composition
INC_EXEMPTION _{ist}	Exemption amount for federal income tax; constructed from personal exemption times the number of people in the family, plus the standard deduction
SWR _{ist}	Fraction of year <i>t</i> woman <i>i</i> may be subject to the state's work requirement, constructed from state's work requirement time limits, child age exemptions, and family composition
DTL_HIT _{ist}	Dummy variable indicating whether woman <i>i</i> would have hit her time limit for welfare receipt
MONTH_SINCE_TL_HIT _{ist}	Time elapsed since woman <i>i</i> may potentially have become subject to time limit

(continued)

Table C1. Description of Variables (continued)

<i>Variable</i>	<i>Description</i>
MONTH_SINCE_WR_HIT _{ist}	Cumulative time elapsed since woman <i>i</i> may potentially have become subject to work requirement
REMAINING_TL_ELIG _{ist}	Maximum potential remaining length of woman <i>i</i> 's time limit, constructed by $TL_LENGTH_{st} - \min\{MONTH_SINCE_TL_START_{st}, OLDEST_{ist} \cdot 12\}$
REMAINING_CHILD_ELIG _{ist}	Remaining length of time woman <i>i</i> may be categorically eligible for welfare benefits, constructed from $18 - YOUNGEST_{ist}$
MEDICAID_PCT _{ist}	Percentage of woman <i>i</i> 's children covered by Medicaid expansion, constructed from her family composition and her state's Medicaid and/or SCHIP expansion
MEDICAID_FPL _{ist}	Medicaid expansion eligibility income threshold for woman <i>i</i> 's youngest child
<i>State-level policy and economic variables</i>	
DTL _{st}	Dummy variable indicating whether state <i>s</i> had time limit in place at year <i>t</i> , under either waiver
TL_LENGTH _{st}	Length of time limit in state <i>s</i>
MONTH_SINCE_TL_START _{st}	Months elapsed since implementation of time limits, under either waiver or TANF
DCHILDBEN _{st}	Dummy variable indicating whether child portion of welfare benefit continues after time limits are exhausted
FIXED_DISREGARD _{st}	Fixed amount of earnings disregarded in calculating AFDC/TANF benefit levels
BBR _{st}	Benefit reduction rate (incorporates "permanent" percent income disregard)
DWORKREQ _{st}	Dummy variable indicating whether state <i>s</i> has work requirement in place at year <i>t</i> , under either waiver or TANF
LENGTH_WR_TL _{st}	Length of work requirement time limit
MONTH_SINCE_WR_START _{st}	Months elapsed since implementation of work requirement time limits, under either waiver or TANF
CHILD_EXEMP_AGE _{st}	Age of youngest child below which the mother will be exempted from work requirement
N_WR_EXEMPTION _{st}	Number of work requirement exemptions
DFULLSANCTION _{st}	Dummy variable indicating whether state <i>s</i> has an ultimate full sanction when work requirement is not satisfied
CHILDSUPPORT_ENFORCE _{st}	Child support enforcement expenditure in year <i>t</i> per single mother
CHILDCARE _{st}	Child care subsidy expenditure in year <i>t</i> per single mother

(continued)

Table C1. Description of Variables (continued)

<i>Variable</i>	<i>Description</i>
WAGE20 _{st}	Real hourly wage rate at the 20th percentile of the wage distribution
UNEMPLOYMENT _{st}	State unemployment rate
DIVERSION _{st}	Dummy indicating whether state <i>s</i> has a diversion program
<i>Federal-level policy variables</i>	
MIN_WAGE _t	Real federal minimum hourly wage
INCTAX_RATE _t	Federal income tax rate applying to lowest bracket

Sources: See text for a description of the data sources.

Table C2. Interaction Terms Used in the Empirical Specification

<i>Policy variable</i>	<i>Interaction terms</i>
<i>Variables related to time limits</i>	
DTL _{st}	TL_LENGTH _{st} , DTL_HIT _{ist} , DCHILDBEN _{st}
REMAINING_TL_ELIG _{ist}	WELFARE_BEN _{ist} , WELFARE_BEN _{ist} × EDU _{ist} AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist} REMAINING_CHILD_ELIG _{ist} [REMAINING_CHILD_ELIG _{ist} × WELFARE_BEN _{ist}] [REMAINING_CHILD_ELIG _{ist} × WELFARE_BEN _{ist} × EDU _{ist}] [REMAINING_CHILD_ELIG _{ist} × AGE _{ist}] [REMAINING_CHILD_ELIG _{ist} × MARITAL _{ist}] [REMAINING_CHILD_ELIG _{ist} × EDU _{ist}] [REMAINING_CHILD_ELIG _{ist} × RACE _{ist}] [REMAINING_CHILD_ELIG _{ist} × UNEMPLOYMENT _{st}] [REMAINING_CHILD_ELIG _{ist} × UNEMPLOYMENT _{st(t-1)}]
MONTH_SINCE_TL_START _{st}	AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}
MONTH_SINCE_TL_HIT _{ist}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}]
<i>Variables related to work requirements</i>	
DWORKREQ _{st}	LENGTH_WR_TL _{st}
SWR _{ist}	AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist} [NCHILD05 _{ist} , NCHILD612 _{ist}] [WELFARE_BEN _{ist} , WELFARE_BEN _{ist} × EDU _{ist}] DFULLSANCTION _{st} , N_WR_EXEMPTION _{st} [UNEMPLOYMENT _{st} , UNEMPLOYMENT _{st(t-1)}]
MONTH_SINCE_WR_START _{st}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}]
MONTH_SINCE_WR_HIT _{ist}	AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}

(continued)

Table C2. Interaction Terms Used in the Empirical Specification (continued)

<i>Policy variable</i>	<i>Interaction terms</i>
<i>Variables related to EITC</i>	
EITC_RATE _{ist}	[WAGE20 _{st} , UNEMPLOYMENT _{st} , UNEMPLOYMENT _{s(t-1)}]
EITC_MAX _{ist}	[SWR _{ist} , DTL _{ist}] AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}
<i>Variables related to child care</i>	
CHILDCARE _{st}	NCHILD05 _{ist} , NCHILD612 _{ist} , AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist} [UNEMPLOYMENT _{st} , UNEMPLOYMENT _{s(t-1)}] DCHILD05 _{ist} × AGE _{ist} , DCHILD05 _{ist} × MARITAL _{ist} DCHILD05 _{ist} × EDU _{ist} , DCHILD05 _{ist} × RACE _{ist} [DCHILD05 _{ist} × UNEMPLOYMENT _{st}] [DCHILD05 _{ist} × UNEMPLOYMENT _{s(t-1)}] DCHILD612 _{ist} , AGE _{ist} , DCHILD612 _{ist} × MARITAL _{ist} , DCHILD612 _{ist} × EDU _{ist} DCHILD612 _{ist} × RACE _{ist} [DCHILD612 _{ist} × UNEMPLOYMENT _{st} , DCHILD612 _{ist} × UNEMPLOYMENT _{s(t-1)}]
<i>Variables related to AFDC/TANF benefit levels</i>	
WELFARE_BEN _{ist}	[WAGE20 _{st} , UNEMPLOYMENT _{st} , UNEMPLOYMENT _{s(t-1)}] [SWR _{ist} , DTL _{st}] AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}
BBR _{st}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}]
<i>Variables related to food stamps</i>	
FOOD_STAMP _{ist}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}]
<i>Variables related to child support enforcement expenditure</i>	
CHILDSUPPORT_ENFORCE _{st}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}] [NCHILD05 _{ist} , NCHILD612 _{ist}]
<i>Variables related to Medicaid</i>	
MEDICAID_PCT _{ist}	[AGE _{ist} , MARITAL _{ist} , EDU _{ist} , RACE _{ist}] [NCHILD05 _{ist} , NCHILD612 _{ist}]

Source: See text for a description of the data sources.

a. Bracketed terms are omitted from the “sparse” specification.

Comments and Discussion

Rebecca M. Blank: This paper adds to the growing literature that attempts to assess the impact of the welfare reforms of the mid-1990s. A key problem in this literature is how to parameterize the reforms in a way that allows them to be clearly distinguished from other changes that occurred at the same time, including a major economic boom and changes in closely related policies (such as the minimum wage and the earned income tax credit). Hanming Fang and Michael Keane claim to do a better job than previous papers by using a more complete specification to measure the impact of policy changes, and their results suggest they are fitting the data very well. In these comments I will primarily focus on this specification and its interpretation.

Before turning to more critical comments, however, let me stress the things that this paper does well. First, Fang and Keane have done an excellent job of coding detailed state-level policy changes, with much more sophistication than previous papers. For instance, rather than use a dummy variable to indicate whether a time limit is in effect, they try to find variables that differentiate between the anticipatory behavioral effects of time limits (those that occur before the limits actually hit) and the effects that occur once enough time has passed that some women are facing benefit losses due to time limits.

My own experience suggests that coding specific state policies is difficult and that the information available for some states and some policies is open to multiple interpretations. Different observers code these policies in different ways, because there is no standardized interpretation (and often there is not even a wholly reliable source of information about what policies are in effect each year). I suspect future researchers will make

some modifications to Fang and Keane's coding, but I suspect that they will borrow some of their ideas as well. I hope that the authors plan at some point to make information available on the coding they use and their sources of information, so that future researchers can both learn from and improve upon their data.

Second, Fang and Keane study a longer time period than do most previous papers, and they estimate the effects of reform through 2002. The addition of 2001 and 2002 to the data is potentially important, since these are years of more economic variation and thus provide information on how reformed welfare policies operate in a more sluggish economic environment.

Third, the paper provides a very useful description of trends in welfare and work participation. Most researchers have focused on aggregate changes in welfare participation and labor force involvement among single mothers. Fang and Keane expand this and look at these trends by the mother's educational attainment, race, marital status, age of children, and number of children. There is much to be learned by looking at the comparative changes among these different groups.

One minor concern about these descriptive trends (and about the data more generally), however, is whether these trends on program participation correlate well with the administrative data from government agencies on actual program reciprocity. The CPS data have always undercounted welfare participants, but my understanding is that this undercount appears to have increased in the late 1990s.¹ This is a problem for all researchers who use CPS data to study welfare reform and deserves more attention from the research community. The paper's analysis of specific trends by demographic group led me to wonder whether there is any way to compare the administrative data by demographic group with these data to see whether the undercount trend is stronger among some groups than others.

Let me focus the remainder of my comments on my concerns about how to interpret the paper's results. Fang and Keane make a very strong claim about their specification and their results. They note that, ". . . the key identifying assumption is that we have adequately controlled for all the important time-varying factors that influenced the welfare participation and work decisions of single mothers over the 1980–2002 period." They argue that they have therefore avoided problems with

1. Bavier (2000).

omitted variables and have no need to include year-specific effects. They buttress this claim by showing that they are able to fit the data extremely well, in the aggregate, by state, and by demographic group. I have some questions, however, about the specification choices the authors have made, about their claim of completeness for the specification, and about these well-fitted results.

I will start by admitting to a strong bias against the authors' claim (or, indeed, any researcher's claim) of a complete specification. Almost never do we in the social sciences have the data or the knowledge necessary to fully specify a behavioral process, and this paper is no exception.

In fact, many would argue that Fang and Keane's approach to this problem is wrongly specified from the start, since they are trying to estimate a final outcome (welfare or work participation) with panel data created from sequential cross sections. Jacob Klerman and Steven Haider have argued that the appropriate specification to use to understand case-load changes (and presumably labor market participation changes as well) is one that looks at flows, since it is decisions about entry into and exit from welfare that are actually being affected.² Others have argued that longitudinal data on individual entries and exits provide better information.³ I think most people would agree that it is hard to estimate well-specified behavioral models from panel data created from sequential CPS cross-sections. At best, Fang and Keane are estimating a type of reduced-form model, which must have some underlying behavioral parameters that determine the entry and exit decisions of individual women.

But many researchers (including myself) have estimated models using data similar to Fang and Keane's because these data are more readily available. Even within the parameters of what the authors are doing, I have doubts about their completeness claim. It is not hard to think of potential missing variables in their specification. For instance, they have data from 1980–2002, yet they include no variables that provide information about the 1986 tax reforms that significantly reduced federal tax rates on lower-income families. There is no information on the substantial changes in child support laws over this period, which should affect other income sources available to women. There is no information on the extent to which states are able to enforce or implement the parameters and regu-

2. Klerman and Haider (2001).

3. Grogger (forthcoming).

lations they have put on the books. These other variables may or may not be important to welfare and work participation decisions, but until they are included in the model, one must worry about omitted variables and about coefficients that include the effects of omitted factors in their estimated values.

Many readers will think I am beating a dead horse—don't we all know that omitted variables exist everywhere? Can't we just acknowledge that and move on to build the best specification possible? I feel the need to make the point nevertheless, partly because the completeness claim is made more strongly in this paper than I have seen in most research, and I want to uphold my reputation for skepticism by saying, "I don't believe it!" Furthermore, this claim is crucial to the authors' decision to omit year effects from the regression. Their identification of policy effects relies on the omission of year effects, since some of their key variables would be collinear with such effects.

A major problem in earlier research on the effects of welfare reform is that certain specific policy parameters (time limits, work exemptions, and others) have been difficult to identify. Fang and Keane solve the identification problem largely by coding these parameters in a way that interacts the policy with demographic information on individual women. For instance, a single mother with one 2-year-old child can have been eligible for welfare for at most two years, and hence can only have been subject to time limits for two years even if the state has had time limits in effect for four years. This demographic interaction with the policies is creative but has potential problems. As the authors themselves note, one must assume that fertility and marital status are exogenous. A growing body of literature is investigating the effects of welfare reform on these outcomes, and some of this research shows significant effects.⁴

One way of thinking about this identification strategy is to ask whether one views these policy changes as one-time state effects that are known to all and "turn on" in the state at the same time everywhere, or as person-specific (or, more accurately in this paper, demographic group-specific) effects that turn on for individual women only when they become single mothers and are directly affected by them. If one believes in full-information decisionmaking, the decision to have children and go

4. For summaries of this literature see Blank (2002) and Grogger, Karoly, and Klerman (2002); see also Gennetian (2003) for some more recent references.

on welfare will take into account the presence of time limits or work exemptions after 1996, and the state-level variables are the more accurate ones to include in the regression, since everybody's calculations change once the new state law is enacted. If instead one's view is that people are more shortsighted and do not think about these things until the law affects them personally, then the policy should be interacted with individual demographic circumstances. I do not know what the right model is; certainly, the more evidence that becomes available about significant effects of welfare reform on fertility and marriage behavior, the more hesitant I am about this identification approach.

But if the proof is in the pudding, then Fang and Keane's results must indicate that they are doing it right. Unlike all other research on this topic, which can only partly explain welfare and labor force participation trends since 1996, Fang and Keane's models fit these data almost exactly. Is this evidence that their claims must be true?

I am not persuaded. Fitting the data may or may not mean that one has the right model. My first hesitation about this paper is one that I have to put on record, again, simply to keep my reputation as a skeptic. The paper does not report any of the actual regression results. Since I do not know what the coefficients of this model look like, I cannot know whether their individual values make sense. I see only the aggregate simulations that the authors provide. I know there are lots of coefficients, but I still would like to see them in an appendix table at least.

My primary concern, however, is that the authors' specification overfits the data. Not only do they enter dummy variables for the implementation of time limits or work exemptions (both interacted with individual demographic variables); they also enter a host of variables that are, in effect, time trends. They include a variable that counts forward as soon as a woman is potentially affected by a ticking time limit (the number of months since she could have first been on welfare in her state of residence after time limits were enacted). This means that these time trends start just when TANF programs are being implemented and then count forward from 1997 for many single mothers. Another variable counts how long it has been since a woman could have first had an actual time limit imposed (the number of months since she could have first been on welfare in her state after time limits were enacted, minus the time on her time limit clock.) Similar time trends count the months since the work requirements clock might have started ticking and since sanctions might have been

imposed. The result is a large number of trend variables (which are also interacted with all of the demographic variables; see the authors' appendix table C2). There are strong trends within the data on caseloads and work participation, and I suspect that putting in these trend variables helps the regression fit the data. The question is whether they tell us anything about behavior at the same time.

I am dubious about the interpretation of these trend variables for several reasons. The authors interpret the results as "anticipatory" or "reactive" effects: women are responding differently over time to the threat of time limits or to their actual implementation. But nothing suggests to me that this should be modeled by a linear time trend. I agree that implementation and learning effects are surely going on as these new welfare reforms are being put in place. But why should I believe they are captured by a linear trend? I would expect anticipatory effects to grow nonlinearly as women approach the imposition of a policy change. And I would expect learning effects to fade over time as information is gained.

In discussing their robustness tests, the authors note that, "Critics argue that, even if we had randomly assigned the state-specific policy variables to women, regardless of their true state of residence, the model would still fit the data well." I will out myself: I offered this argument. I believe that these time trends cause the good fit and that, even if they were included for the wrong state, they would still provide a good fit. (The reason is that all states see turning points within a year or two of each other and then experience almost straight-line declines in welfare participation.) The authors go on to say that, when states were randomly assigned, the model fit "rather well" the aggregate patterns of welfare and work participation, as well as the patterns by demographic group. But because these estimates did not fit the state-level data nearly as well as their model, they suggest that this means the time-trend specification is not causing the good fit.

I would interpret this finding very differently. If the same specification applied in a meaningless way (that is, by randomly assigning states and state policies to individuals) produces almost as good a fit as the original specification, both within the aggregate data and within demographic groups, it suggests that it is something in the nature of the specification (I suspect the time-trend variables) that is causing the good fit. In short, the specification is not correct in any model-based sense; it simply happens to be designed to fit time-trending data very well. True, the results of this randomized misspecification do not fit the data as well at the state level—

the states did implement different policies at different times, and the randomization ignores these differences entirely and probably misses the exact turning points by state. But if the randomization works for the aggregate data and the demographic groups, I believe that my concerns are justified.

I will end with two more minor points. First, nowhere do Fang and Keane show standard errors on any of their descriptive trend data or their simulations. Although many of the changes they report are large, the CPS samples of single mothers by state are smaller than any of us would ideally like (and this is even more true once the data are disaggregated by demographic group). It would be useful to see the standard error bounds on the simulations in the authors' tables 7 through 10.

Second, despite an extensive set of interactions, the authors' model includes only two variables that describe the macroeconomy: the state unemployment rate and the 20th-percentile wage in a state. Surprisingly, these two variables, unlike the policy variables, are not interacted with demographic characteristics. This is surprising because research has shown that women with different skill levels and different demographic characteristics respond differently to changes in economic conditions.⁵ One effect of this omission is to give less scope for the specification to find economic effects as opposed to policy effects.

In short, I am not convinced that this paper has solved the problem of how to "explain" the impact of welfare reform on the behavior of single mothers. I am not entirely sure that the authors have found the best identification strategy for welfare reform effects, or that the results can be interpreted as cleanly as the authors suggest. All that said, this is a paper that others will want to read. The authors have performed a very useful service in carefully coding state-specific changes in the components of welfare reform and are creative in trying to use this information to identify policy effects better than the existing literature has. I suspect that this paper will spawn other research and drive ongoing interest in the impact of these detailed policy changes.

Jeffrey Grogger: Hanming Fang and Michael Keane have produced a useful study that corroborates much of what has been learned about the effects of time limits and work requirements on welfare participation.

5. Hoynes (2000).

Their study also could shed valuable light on the effects of child care subsidies and of sanctions for failure to meet work requirements, on which there has been less previous research. My comments will focus on the context of the study; on the authors' estimates of the effects of time limits, work requirements, sanctions, and child care subsidies; and on some specification issues. I will close with some parting questions.

CONTEXT. Fang and Keane seek to understand the effects of welfare reform on single mothers' welfare and work behavior. These are important issues. PRWORA represents the most sweeping change in U.S. welfare policy in sixty years. It ended the entitlement status of welfare, imposed time limits and work requirements on welfare recipients, and gave the states new latitude to formulate policy in other aspects of their welfare programs. It was highly controversial at the time of its enactment in 1996.

Not surprisingly, the effects of welfare reform have been the subject of a large body of research.¹ That research falls into two broad categories: First, a number of social experiments have been conducted to evaluate the effects of pre-PRWORA reforms carried out in individual states during the early 1990s. These experiments have primarily sought to estimate the effects of incremental reforms. They generally involved changes to a single policy or a small bundle of policies. Work requirements, in particular, have been the subject of extensive experimental study. Only a few experiments have sought to evaluate TANF-like bundles involving multiple policy reforms. Second, a number of observational studies have estimated the effects of both pre- and post-PRWORA reforms. In contrast to the experimental studies, the observational studies have been most successful in estimating the effect of reform as a whole, or of separate reforms considered as a bundle. Most of these studies have been less successful in isolating the effects of specific reforms.

The reason for this is easy to understand. The first statewide welfare reforms were carried out in 1992 under waivers from the AFDC program. By the beginning of 1998, all fifty states had implemented their TANF plans. Within that period, some states moved earlier and some later, but by the end of the six years, all states had moved from AFDC to TANF. The bundles of reforms that the states implemented did vary, but many

1. Blank (2002) surveyed a selection of roughly twenty studies. Grogger, Karoly, and Klerman (2002), who had a broader mandate, covered sixty-seven studies in their review of the literature. Additional studies have been released since those reviews were completed.

states made only one change: from the old AFDC program, under which states differed primarily in terms of the generosity of the benefit payment, to the new TANF program, under which state programs varied along multiple dimensions.

Adding to the evaluation problem is the fact that other changes that independently affected welfare behavior took place at the same time. The economy experienced the longest unbroken expansion in postwar history. The EITC became considerably more generous. Public health insurance was expanded, especially for children. Disentangling the effects of numerous specific welfare reforms, the economy, and other policy changes has proved to be an inordinately difficult task for analyses that rely solely on differences in the timing and nature of state welfare reforms as the basis for estimation.

A few studies have gone beyond differences in the timing of reforms to exploit the demographic dimensions of policy variation as well. Some of these derive from theory. For example, theory suggests that welfare families should reduce their welfare participation in anticipation of time limits, and that such reductions should be greater among families with young children than among families without young children. The reason is that families with younger children have longer eligibility horizons than families with older children, so that the value of preserving their eligibility for the future is greater.² Evidence in support of this prediction has come from a Florida welfare reform experiment, from the Current Population Survey, and from the Survey of Income and Program Participation (SIPP).³

In other cases the demographic dimensions of policy variation come from program rules. For example, most states exempt women with very young children from work requirements, but the definition of “very young” varies from state to state.⁴ Fang and Keane provide several other examples of program rules whose impact varies according to the demographic composition of welfare households.

It is fair to say that Fang and Keane push these demographic dimensions of policy variation further than any previous study. Thus their study represents the most ambitious observational analysis of multiple welfare reforms to date.

2. Grogger and Michalopoulos (2003).

3. Grogger and Michalopoulos (2003); Grogger (2002, 2003a, 2004).

4. Rowe (2000).

Table 1. Qualitative Comparison of Results on Effects of Time Limits

<i>Age of youngest child</i>	<i>Decline in welfare participation rate due to time limits</i>		
	<i>Fang and Keane (this volume), for period 1993–2002</i>		<i>Grogger (2003), for period through 1999^a (percentage points)</i>
	<i>Percentage points</i>	<i>Percent of total decline</i>	
0 to 5 years	4.84	14.72	4.57
6 to 12 years	1.81	10.04	0.50

Sources: Fang and Keane (this volume, table 8); Grogger (2003a, table 2).

a. Evaluated at age-group midpoints.

TIME LIMITS. Fang and Keane discuss the effects of time limits at some length. This emphasis is well warranted. Time limits were arguably the sharpest break from past policy and the most controversial of the reforms implemented under PRWORA. They have also been the subject of previous study.

Indeed, Fang and Keane are fairly critical of previous studies of time limits. They make two arguments as to why previous estimates may have overstated the effects of time limits. First, they argue that other features of state time-limit policies, and of state welfare reform plans more generally, may limit the impact of the “pure” time limits studied theoretically by Charles Michalopoulos and myself.⁵ Second, they argue that the relatively lenient treatment by caseworkers and welfare agencies of recipients who have reached their time limits should have further mitigated the anticipatory effects.

Those critiques notwithstanding, Fang and Keane’s results corroborate both qualitatively and quantitatively much of what we have already learned about the anticipatory effects of time limits. Qualitatively, their results follow the age pattern predicted by the Grogger-Michalopoulos model: the effects of time limits are larger for families with younger children. Quantitatively, they credit time limits with an even greater share of the caseload decline than do previous studies.

Table 1 above addresses the qualitative comparison. The first two data columns are drawn directly from the authors’ table 8. They show how much of the 1993–2002 decline in welfare receipt can be attributed to time limits, by the age of the youngest child in the family. In both

5. Grogger and Michalopoulos (2003).

Table 2. Quantitative Comparison of Results on Effects of Time Limits

Year	<i>Decrease in welfare participation rate between 1993 and indicated year (percentage points)</i>	<i>Portion of decline due to time limits</i>		
		<i>Fang and Keane (this volume)</i>		<i>Grogger (2003) (percent of total)</i>
		<i>Percentage points</i>	<i>Percent of total</i>	
1997	11.09	1.68	15.15	...
1998	14.43	3.06	21.19	...
1999	17.23	3.21	18.61	11.9
2000	19.58	3.11	15.89	...
2001	21.44	2.33	10.85	...
2002	23.77	2.52	10.60	...

Sources: Fang and Keane (this volume, table 7); Grogger (2003a, table 4).

absolute and relative terms, their results are consistent with Grogger and Michalopoulos's prediction. The numbers in the third column come from calculations based on regression results reported in my own recent paper.⁶ They show the same age pattern. Indeed, comparing the results in the first and third columns of table 1 shows that Fang and Keane's estimates are not just qualitatively similar to mine, but quantitatively similar as well.

Table 2 provides further evidence on the quantitative comparison. The first three data columns are taken directly from Fang and Keane's table 7. The second and third columns show the amount and the percentage share, respectively, of the cumulative caseload decline between 1993 and the indicated year that is explained by time limits. Fang and Keane focus on the last row in the third column, which shows that time limits account for about 11 percent of the caseload decline. The number in the last column of table 2 is again taken directly from my 2003 paper.⁷ It shows that time limits account for 12 percent of the caseload decline.

Because these estimates cover different sample periods, they are less comparable than they seem at first glance. Fang and Keane's estimate pertains to the 1993–2002 period, whereas mine pertains to 1993–99. When the results of both papers over the latter period are compared, Fang and Keane attribute 19 percent of the decline to time limits, an amount that exceeds my estimate by more than half.

6. Grogger (2003a, table 2).

7. Grogger (2003a, table 4).

This presents a puzzle. In light of Fang and Keane's critique of earlier work, how could their estimates be so similar to the earlier estimates? Even more puzzling, how can they attribute an even greater share of the caseload decline to time limits, when their critique suggests that time limits should have little effect at all?

To solve the puzzle, consider again their first argument, which notes that "real-world" time limits differ from pure time limits because other features of state welfare programs, such as adult-only time limits and financial incentives to work, may lessen their impact. If anything, this sounds like an argument for why the earlier studies, which did not control for such policies explicitly, should understate the effects of time limits. If adult-only time limits have smaller effects than conventional time limits, and enhanced earnings disregards pull recipients in the direction of more, rather than less, welfare receipt, then estimates that fail to control for those policies should understate the extent to which time limits reduce the welfare rolls. Put differently, the broader set of policy controls employed by Fang and Keane should allow them to better isolate the effects of pure time limits, all else equal.

What about the argument based on caseworker and welfare agency behavior? If welfare offices were so lenient in granting exemptions and extensions, how could time limits have had any effect at all? There are at least two possible answers.

The first involves differences in the timing of recipient and agency behavior. When time limits were imposed, the important question facing recipients was, What would happen to them in the future if and when they exhausted their benefits? Only after recipients started reaching their limits could they have learned that the welfare office might go easy on them, or at least go easy on others in similar circumstances. If recipients in the most lenient states knew in advance what to expect, they would have had no reason to leave the rolls in order to preserve their eligibility. The question is how much they responded before they knew what the welfare office would do.

Table 2 provides a hint of evidence. It shows that the effect of time limits peaked in 1998 or 1999. This was one to three years after the typical state imposed time limits, in 1996 or 1997, and two to four years before recipients in the typical state could have exhausted their benefits. After 1999 the effect of time limits fell. Thus it appears that much of the

response to time limits took place well before recipients could have known how the welfare office would handle time-limited cases.

Furthermore, even if one takes the caseworker argument at face value, time limits still could have reduced welfare participation, because the caseworker argument applies only to welfare recipients. Nonrecipients at risk of entering welfare do not have caseworkers, and so it is difficult to see how the caseworker argument would apply to them. Yet the same theory that predicts that *recipients* should *leave* the welfare rolls to preserve eligibility also predicts that *nonrecipients* should *avoid* the rolls to preserve eligibility. Moreover, nonentry on the part of nonrecipients accounts for a substantial portion of the caseload decline. Based on an analysis of SIPP data, Steven Haider, Jacob Klerman, and I estimated that 40 percent of the caseload decline was due to reductions in the entry rate.⁸ Administrative data from California put the figure at 60 percent and show that the decline in the entry rate was overwhelmingly concentrated among would-be first-time entrants. Even if recipients fully anticipated caseworker and agency behavior, the responses of the would-be entrants seem large enough to account for the 12 percent, or possibly even 19 percent, share of the reduction in the caseload that has been attributed to time limits.

As a whole, therefore, Fang and Keane's evidence corroborates much of what is already known about the anticipatory effects of time limits. They show that time limits had greater effects on families with younger children, consistent with the prediction of Grogger and Michalopoulos. They show that, if anything, previous analyses of time limits may have understated the contribution of time limits to the recent decline in welfare caseloads. An alternative interpretation is that Fang and Keane may overstate the contribution. I return to this point below.

WORK REQUIREMENTS. Of all the reforms implemented after PRWORA, work requirements have been analyzed the most extensively. In a forthcoming book, Lynn Karoly and I review thirteen experimental studies designed to isolate the effects of time limits that were implemented under pre-PRWORA waivers from the AFDC program.⁹ In all but one of these studies, work requirements had a significant negative effect on welfare receipt. Many of the effects were sizable. Fang and Keane are

8. Grogger, Haider, and Klerman (2003).

9. Grogger and Karoly (forthcoming).

the first to quantify the importance of work requirements in explaining the caseload decline, but their general effects have been known for some time.

SANCTIONS. Compared with the effects of work requirements themselves, the effects of sanctions for failure to meet work requirements have been little studied. No experiments have been conducted to isolate the effects of sanctions, and none of the experimental variation found in policy bundles that included sanctions involved anything as severe as the full-benefit sanctions imposed in many states' TANF plans. Although only four previous observational studies estimated the effects of sanctions, all reported results that can only be regarded as sensational. The smallest estimate indicates that full-benefit sanctions reduce welfare caseloads by 16 percent,¹⁰ and the largest finds a reduction of 39 percent.¹¹ The other two yield estimates of 20 percent and 25 percent.¹² In the period immediately following PRWORA's enactment, this led some analysts to conclude that sanctions were the most important contributor to the fall in the caseload. One spreadsheet calculation indicated that full-benefit sanctions caused 540,000 families to leave the welfare rolls between 1997 and 1999.¹³

Fang and Keane, however, report very different results. Their estimates suggest that full-benefit sanctions reduced the welfare caseload by 1.5 percentage points, or by roughly 10 percent of its 2002 level. To my way of thinking, this is a much more plausible estimate than those of the earlier studies. Nevertheless, the discrepancy calls for much greater elaboration than mere mention in a footnote. Why are the numbers so different? Which estimates are right? Given the popular attention that sanctions have received, it is disappointing that Fang and Keane chose to give their own estimates such short shrift.

CHILD CARE SUBSIDIES. Even less research has been done on the effects of child care subsidies than on those of sanctions. No welfare reform experiments have sought to isolate these effects, nor, to the best of my knowledge, have any observational studies undertaken during the reform era done so. Fang and Keane make a forceful argument that child care

10. Mead (2000).

11. Council of Economic Advisers (1999).

12. MaCurdy, Mancuso, and O'Brien-Strain (2002), and Rector and Youssef (1999), respectively.

13. Goldberg and Schott (2000).

subsidies should affect welfare participation: pre-PRWORA simulations suggested that they should have large effects, and child care expenditure did grow rapidly in the post-PRWORA period. Yet none of their estimates suggest that child care subsidies affect either welfare participation or employment very much, and to the extent that they do affect welfare participation, the estimates show they increase it rather than decrease it.

It would be helpful to hear the authors' explanation for these results. Do they stem from data quality problems arising from the consolidation of four programs into one? or from specification issues? Or should the reader interpret them to mean that the recent explosion in child care spending has had no effect on these two important outcomes? One reason why child care spending may have little effect is that many low-skill women take advantage of free care provided by relatives. Thus subsidies may affect the mode of care without affecting the availability of care, employment, or welfare participation. If that is the case, child care subsidies may merely provide economic rents rather than affect behavior (unless, of course, paid care improves child outcomes). Given the potential policy importance of these results, it is regrettable that Fang and Keane spend so little time on them.

SPECIFICATION ISSUES. Observational studies of welfare reform have typically included state and year dummy variables in the regression model. State dummies control for unobservable, time-invariant characteristics of states, and year dummies control for unobservable, nationwide trends in factors that may influence welfare participation. Fang and Keane include neither. Rather, they make the “. . . key identifying assumption . . . that we have adequately controlled for all important time-varying factors that influenced the welfare participation and work decisions of single mothers over the 1980–2002 period.”

This is an unconventional approach to identification that practically invites the reader to think of counterexamples. In fact, one variable that the authors do omit is the “message” that recipients hear about welfare. By passing PRWORA, policymakers sought to change that message and to change the culture of welfare offices.¹⁴ Richard Bavier has argued that such changes have had important effects on the caseload.¹⁵ If he is right,

14. Lurie and Riccucci (2003).

15. Bavier (2003).

Fang and Keane's model suffers from omitted variables bias, since the message is correlated with the observed policy changes in their model.

Furthermore, there is some evidence to suggest that the omission may be important. A comparison of the authors' table 7 with table 4 of my 2003 paper shows that Fang and Keane attribute larger shares of the 1993–99 caseload decline not just to time limits, but also to the EITC and macroeconomic conditions. It is possible that they get larger estimates because they control more precisely for the policy environment. However, it is also possible that their larger estimates stem from omitted year dummies, which would help absorb the effects of the message. Fang and Keane note that adding year dummies to their model has little effect on their R^2 , but the real question is whether the year dummies affect the other coefficients in the model. The authors could resolve this issue by presenting estimates from models that include year dummies.

PARTING QUESTIONS. I close with two questions. The first is, How do the authors' estimates square with those in the rest of the literature? A problem in the experimental literature is that subgroup variation in the effects of work requirements differs across experiments.¹⁶ Can the authors shed any light on why this is so? Considering the interest in subgroup effects, this would be a valuable contribution, and a way for the authors' model to provide a bridge between the experiments and the observational studies.

The second question is, Does Fang and Keane's model explain too much? This again relates to the message of welfare reform. My guess (which admittedly is only that) is that the message made a difference to the decline in welfare receipt. If so, a model that does not account for the message should not be able to fully explain that decline. Yet, to judge from their figure 5, the authors' model tracks welfare receipt closely.

Furthermore, although I expect that much of the message was nationwide in nature—stemming from the rhetoric of the 1992 presidential campaign, the 1994 Contract with America, and the 1996 passage of PRWORA itself—I would also expect some variation among the states. California and Michigan provide points in contrast. California was a reluctant participant in welfare reform. It was the last state to implement its TANF plan, and the plan it implemented involved adult-only time

16. Grogger, Karoly, and Klerman (2002, appendix A).

limits, relatively lenient sanctions, and generous financial incentives to work. It seems doubtful that California's welfare recipients heard much of a message beyond what they got through the national media. In Michigan, in contrast, a charismatic governor staked his reputation as a policy innovator on the success of welfare reform in his state. If recipients anywhere got a state-specific message about welfare reform, it was in Michigan (and perhaps Wisconsin).

If the message matters, it should be easier to track the caseload in California than in Michigan, since the model does not account for the message. Yet the authors' figure 6 shows that their model tracks postreform welfare caseloads equally well in both states. This leaves the reader with a choice between two conclusions: either the model explains too much, or the message does not matter, despite the best intentions of Congress, the president who signed the legislation, and welfare officials throughout the country.

Fang and Keane argue that the message should somehow work through the measurable policies rather than have an independent effect. They conclude that "it is quite appropriate . . . for the coefficients on these measurable policy instruments to pick up how they affect culture." But if one's goal is to predict the effect of future policy changes, one needs estimates of the effects of the various policy instruments themselves, unless one believes that future policy changes are likely to be accompanied by similar changes in the message. More likely, at some point in the future, policymakers will want to change time limits, work requirements, or some other relatively isolated aspect of the welfare program. If the regression model fails to distinguish the effects of the policy change per se from the effects of the message, it will not accurately predict the effects of such future policy changes.

General discussion: Christopher Sims complimented the authors for their enterprise in obtaining such detailed data on the timing and features of welfare reform, which allowed them to estimate a rich model of behavior. He noted that the rich specification, ironically, was the primary reason that several panelists found it difficult to know how impressed they should be by the model's close fit of state and demographic data on welfare and work participation. One complaint was that some potentially useful explanatory variables had been omitted, whose effects may therefore be spuriously attributed to the variables that were included. This concern led to the suggestion to check whether the inclusion of fixed effects, for a

time trend or for states, affected the coefficient estimates. Another concern was overfitting, because there were so many free parameters. Including fixed effects would only amplify this problem, increasing the likelihood that multicollinearity would result in unreasonable or statistically insignificant coefficients.

Sims argued that in these circumstances Bayesian methods provide a practical solution. As an example, he suggested including terms interacting fixed effects with various variables with a prior mean of zero. Varying the weights on this prior would provide a check on the robustness of the results. This method could also be used to obtain measures of fit and to compare specifications in a continuous fashion. The systematic use of priors on the coefficients on the variables would offer protection from spurious correlation resulting in coefficient estimates that make no sense from the point of view of theory.

Olivier Blanchard remarked that Europe is now struggling with reform of its social welfare programs, including unemployment insurance. The issues there—the effectiveness of time limits, work requirements, monitoring, and others—are the same as those studied in this paper, and it therefore may help inform the European debate. But Blanchard wondered whether cultural differences made some of the conclusions inapplicable. William Nordhaus commended the authors for their careful attention to differences in timing and enforcement across states and for using information about who was directly affected by specific welfare changes. But he observed that cultural change had been one of the motivations for welfare reform, and he wondered whether the authors' specification accurately captured the resulting dynamics. He noted that changes in the welfare "culture" are likely to be gradual. There are substantial "setup" costs, psychological as well as economic, of going on welfare or going to work. Hence the response to a discontinuous change in policy may take place over many years, as people slowly change their attitudes and the way they organize their lives.

Nordhaus also found striking the substantial discrepancy between the dramatic drop in welfare participation among single mothers and the increase in the fraction that work. It is hard to know how to evaluate welfare reform without knowing the characteristics of those who have moved off welfare yet are not working, and the factors affecting their behavior. He agreed with the authors that this was an important topic for future research.

References

- Bavier, Richard. 2000. "Accounting for Increases in the Failure to Report AFDC/TANF Receipt." Office of Management and Budget (March).
- . 2003. "Non-economic Factors in Early Welfare Caseload Declines." Office of Management and Budget.
- Blank, Rebecca M. 2001. "What Causes Public Assistance Caseloads to Grow?" *Journal of Human Resources* 36, no. 1: 85–118.
- . 2002. "Evaluating Welfare Reform in the United States." *Journal of Economic Literature* 40, no. 4: 1105–66.
- Bloom, Dan, Mary Farrell, and Barbara Fink, with Diana Adams-Ciardullo. 2002. "Welfare Time Limits: State Policies, Implementation, and Effects on Families." U.S. Department of Health and Human Services. Available at www.acf.dhhs.gov/programs/opre/welfare_timelimits/welfare_timelimits.pdf.
- Council of Economic Advisers. 1997. "Explaining the Decline in Welfare Receipt 1993-1996." Washington (May).
- . 1999. "The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update." Washington (August).
- Crouse, Gil. 1999. "State Implementation of Major Changes to Welfare Policies, 1992–1998." U.S. Department of Health and Human Services, Office of Human Services Policy. Available at aspe.hhs.gov/hsp/Waiver-Policies99/policy_CEA.htm.
- Gallagher, L. Jerome, and others. 1998. "One Year After Federal Welfare Reform: A Description of State Temporary Assistance to Needy Families (TANF) Decisions as of October 1997." Discussion paper. Urban Institute (May).
- Garfinkel, Irwin, and Sara S. McLanahan. 1986. *Single Mothers and Their Children*. Washington: Urban Institute Press.
- Gennetian, Lisa A. 2003. "The Long-term Effects of the Minnesota Family Investment Program on Marriage and Divorce Among Two-Parent Families." New York: MDRC (October).
- Goldberg, Heidi, and Liz Schott. 2000. "A Compliance-Oriented Approach to Sanctions in State and County TANF Programs." Washington: Center on Budget and Policy Priorities (October 1).
- Grogger, Jeffrey. 2002. "The Behavioral Effects of Welfare Time Limits." *American Economic Review* 92, no. 2: 385–89.
- . 2003a. "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* 85, no. 2: 394–408.
- . 2004. "Time Limits and Welfare Use." *Journal of Human Resources* 39, no. 2: 405–24.
- . Forthcoming. "Welfare Transitions in the 1990s: The Economy, Welfare Policy and the EITC." *Journal of Policy Analysis and Management*.
- Grogger, Jeffrey, and Lynn A. Karoly. Forthcoming. *Welfare Reform: Effects of a Decade of Change*. Harvard University Press.

- Grogger, Jeffrey, and Charles Michalopoulos. 2003. "Welfare Dynamics under Time Limits." *Journal of Political Economy* 111, no. 3: 530–54.
- Grogger, Jeffrey, Steven J. Haider, and Jacob Klerman. 2003. "Why Did the Welfare Rolls Fall During the 1990s? The Importance of Entry." *American Economic Review Papers and Proceedings* 93, no. 2: 288–92.
- Grogger, Jeffrey, Lynn A. Karoly, and Jacob A. Klerman. 2002. *Consequences of Welfare Reform: A Research Synthesis*. Document DRU-2676-DHHS. Santa Monica, Calif.: RAND. Available at www.acf.hhs.gov/programs/opre/welfare_reform/reform_cover.html.
- Hotz, V. Joseph, and John Karl Scholz. 2003. "The Earned Income Tax Credit." In *Means-Tested Transfer Programs in the United States*, edited by Robert A. Moffitt. University of Chicago Press.
- Hoynes, Hilary W. 2000. "The Employment, Earnings, and Income of Less Skilled Workers Over the Business Cycle." In *Finding Jobs: Work and Welfare Reform*, edited by David E. Card and Rebecca M. Blank. New York: Russell Sage.
- Johnson, Nicholas, Joseph Llobrera, and Bob Zahradnik. 2003. "A Hand Up: How State Earned Income Tax Credits Help Working Families Escape Poverty in 2003." Washington: Center on Budget and Policy Priorities (March).
- Keane, Michael P. 1995. "A New Idea for Welfare Reform." *Federal Reserve Bank Minneapolis Quarterly Review* 19, no. 2: 2–28.
- Keane, Michael P., and Robert Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review* 39, no. 3: 553–89.
- Keane, Michael P., and Kenneth I. Wolpin. 2002a. "Estimating Welfare Effects Consistent with Forward Looking Behavior. Part I: Lessons from a Simulation Exercise." *Journal of Human Resources* 37, no. 3: 570–99.
- . 2002b. "Estimating Welfare Effects Consistent with Forward Looking Behavior. Part II: Empirical Results." *Journal of Human Resources* 37, no. 3: 600–22.
- . 2003. "Public Welfare and Life Cycle Decisions of Young Women." Yale University, Department of Economics.
- Klerman, Jacob A., and Steven J. Haider. 2001. "A Stock-Flow Analysis of the Welfare Caseload: Insights from California Economic Conditions." Labor and Population Program Working Paper Series 01-02. Santa Monica, Calif.: RAND (January).
- Lurie, Irene, and Norma M. Riccucci. 2003. "Changing the 'Culture' of Welfare Offices—From Vision to the Front Lines." *Administration and Society* 34, no. 6: 653–77.
- MaCurdy, Thomas E., David C. Mancuso, and Margaret O'Brien-Strain. 2000. "The Rise and Fall of California's Welfare Caseload: 1980–1999." San Francisco: Public Policy Institute of California.
- . 2002. "Does California's Welfare Policy Explain the Slower Decline of Its Caseload?" San Francisco: Public Policy Institute of California.

- Mead, Lawrence M. 2000. "Caseload Change: An Exploratory Study." *Journal of Policy Analysis and Management* 19, no. 3: 463–72.
- Meyer, Bruce D., and Dan T. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 116, no. 3: 1063–14.
- Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *American Economic Review* 73, no. 5: 1023–35.
- . 1992. "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature* 30, no. 1: 1–61.
- . 1999. "The Effect of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior." In *Economic Conditions and Welfare Reform*, edited by S. Danziger. Kalamazoo, Mich.: W. E. Upjohn Institute for Employment Research.
- . 2003. "The Temporary Assistance for Needy Families Program." In *Means-Tested Transfer Programs in the United States*, edited by Robert A. Moffitt. University of Chicago Press.
- Pavetti, LaDonna, and Dan Bloom. 2001. "Sanctions and Time Limits." In *The New World of Welfare*, edited by Rebecca M. Blank and Ron Haskins. Brookings.
- Rector, Robert E., and Sarah E. Youssef. 1999. "The Determinants of Welfare Caseload Decline." Report CDA99-04. Washington: Heritage Foundation Center for Data Analysis (May).
- Rowe, Gretchen. 2000. *Welfare Rules Databook: State TANF Policies as of July 1999*. Washington: Urban Institute.
- Schoeni, Robert F., and Rebecca M. Blank. 2000. "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Income, Poverty and Family Structure." Working Paper 7627. Cambridge, Mass.: National Bureau of Economic Research (March).
- U.S. General Accounting Office. 1997. "Welfare Reform: States' Early Experience with Benefit Termination." GAO/HEHS-97-74.
- . 1998a. "Welfare Reform: States Are Restructuring Programs to Reduce Welfare Dependence." GAO/HEHS-98-109 (January).
- . 1998b. "Welfare Reform: Child Support an Uncertain Income Supplement for Families Leaving Welfare." GAO/HEHS-98-168 (August).
- . 1998c. "Welfare Reform: Early Fiscal Effects of the TANF Block Grant." GAO/AIMD-98-137 (August).
- U.S. House of Representatives, Committee on Ways and Means. 2000. *The 2000 Green Book: Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means*.
- Yelowitz, Aaron S. 1995. "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *Quarterly Journal of Economics* 110, no. 4: 909–39.