Effects of the Welfare Time Limit on Welfare Use and Labor Market Outcomes: New Evidence from Welfare Reform Experiments

Seojung Oh

October 2025

For the latest version, please click here.

Abstract

The introduction of welfare time limits was arguably one of the most fundamental and striking changes of the 1996 welfare reform. This study reexamines how this policy change influenced welfare participation and labor supply among single mothers using experimental data from two welfare "waiver" experiments: Florida's Family Transition Program (FTP) and Connecticut's Jobs First Program (JFP), with a particular focus on JFP data, which has not been previously explored. I follow the identification strategy of Grogger and Michalopoulos (2003), which leverages the age of the youngest child in each household to isolate the effect of the time limit from other policy components of the 1996 welfare reform. My replication of the FTP data yields similar results with prior studies, indicating that the time limit policy reduced welfare use. However, results from the JFP data reveal novel and contrasting findings: Single mothers exposed to welfare time limits did not reduce their welfare use and may have even increased it. Importantly, more people in Connecticut began to combine both welfare and work after the time limit was introduced. I find suggestive evidence that Connecticut's unusually generous earned income disregard policy might be the reason behind these different results. This new evidence demonstrates that the effects of welfare time limits are context-dependent and should be understood within the broader context of other policies.

1 Introduction

The enforcement of a 60-month time limit on cash assistance under the 1996 welfare reform sparked a substantial and ongoing debate over whether such a policy benefits or harms disadvantaged families. A lifetime limit on welfare receipt may incentivize recipients to seek employment, thereby encouraging them to escape the poverty and achieve economic self-sufficiency. However, cutting benefits before recipients reach financial independence may result in substantial economic hardship and leave them in a poverty trap.

This paper revisits the questions of how a strict time limit policy affects welfare participation and labor supply among single mothers using a previously unexplored source of data in the time limit literature: experimental data from Connecticut's Jobs First Program (JFP). In January 1996, as part of the broader national shift from the entitlement program of Aid to Families with Dependent Children (AFDC) to the block-granted program of Temporary Assistance for Needy Families (TANF), the state of Connecticut launched the JFP as a randomized controlled trial (RCT) to test the impacts for welfare participants of strict work requirements, stronger work incentives, and strict 21-month time limits. While this dataset has been used to analyze the impacts of welfare reform on earnings, transfers, and income (Bitler et al., 2006) as well as mortality (Wilde et al., 2014), this paper is the first to use this experimental dataset to study the effects of time limits on government transfer policies.

Empirically, identifying the causal effects of time limits has three key challenges. First, it is difficult to isolate the impacts of time limits from the broader effects of the TANF reform, which includes several additional concurrent policy changes such as strengthened work requirements, enhanced work incentives (which may be exempted in the presence of a young child), and expanded childcare services. Second, it is essential to account for other confounding factors, including the strong economic conditions of the mid-1990s and the 1993 expansion of the Earned Income Tax Credit (EITC). Third, potential endogeneity in the timing and content of policy also raises concerns, as states had considerable discretion in

designing their own TANF programs. If the timing and the structure of each state's new policy package were associated with pre-reform state characteristics, and if those characteristics were systematically related to welfare caseloads or labor market conditions, before-after comparisons would yield biased estimates.

To isolate the time limits' effect from other TANF policy changes, I adopt the identification strategy of Grogger and Michalopoulos (2003, hereafter GM03), which leverages the age of the youngest child to isolate the effect of time limits from those of the broader reform. Both under AFDC and its successor, TANF, families lose eligibility for cash welfare once their youngest child turns eighteen. Therefore, only families whose youngest child was age thirteen or younger when the 60-month time limit was implemented were affected by the time limit policy. Families whose youngest child was older than thirteen serve as a counterfactual control group, as they experienced the other components of the reform but were not bound by the time limit. In addition, because the RCT design ensures the comparability of treated and control groups within the same state, using experimental data allows me to better isolate the impact of time limits from other contemporaneous confounders, such as economic conditions or the EITC expansion, and alleviates concerns related to policy endogeneity, compared to research designs that rely on variation in the timing or length of time limits across states.[1]

Assuming forward-looking behavior and wage uncertainty, GM03's model predicts that economic agents would reduce welfare use and increase labor supply even before reaching the time limit. Using experimental data from Florida's Family Transition Program (FTP), which is another welfare waiver experiment that imposed a 24-month time limit, GM03 finds that time limits reduced welfare use, consistent with their predictions. Subsequent studies using observational data support these findings, showing that anticipatory behavioral responses to

[1]Controlling for various state and policy characteristics cannot fully address the endogeneity problems because states may have unobservable characteristics that affect both welfare use and TANF policies. Including state fixed effects can control time-invariant state unobservables, but still, some of these observables may be time-varying, such as public sentiment toward welfare use or the stigma attached to cash assistance programs in each region.

time limits explain roughly 12–25 percent of the AFDC/TANF caseload decline during the late 1990s and early 2000s (Grogger, 2002, 2004; Mazzolari, 2007). Also, studies examining employment and earnings effects find an increase in labor supply after time limits took effect (Grogger, 2003; Wang, 2021).

However, analysis using Connecticut's JFP sample shows surprisingly different results. Unlike the Florida sample, whose outcomes generally align with the GM03 model and prior studies, the Connecticut sample shows puzzling patterns. In Connecticut, individuals who were exposed to the time limit due to the young age of their youngest child did not reduce—and may have even increased—their AFDC/TANF use during the first 21 months after random assignment. This finding is robust across various specifications, although it contradicts the predictions of the GM03 model and previous research. Another key difference is the positive effect on the number of months individuals in Connecticut combined welfare and work—a pattern not observed in Florida. This increased tendency to combine welfare and work is not found among those in Connecticut whose youngest child was older and thus not subject to the time limit. These results suggest that certain features of Connecticut's JFP, which were absent in Florida's FTP, influenced the unique behavioral responses of JFP participants facing the time limit. In terms of overall employment and earnings, both the FTP and JFP samples show small and statistically insignificant effects. These results indicate that the time limit did not affect the economic self-sufficiency for either the Florida or Connecticut samples.

Regarding the underlying reasons for these distinctive effects of the JFP sample, I provide suggestive evidence that the primary explanation lies in Connecticut's unusually generous policy for calculating welfare benefits for working families before reaching the time limit. Specifically, Connecticut's JFP implemented a very generous earned income disregard (EID) rule, which granted the full benefit amount for working families as long as the family's earnings remained below the federal poverty line. Unlike AFDC and FTP, which imposed a high tax rate that reduced benefits with each additional dollar of earnings, the

JFP's EID policy allowed many working families who had previously been ineligible for cash assistance to receive the full benefit without reducing their working hours. To test whether Connecticut's EID rule drives the observed results, I estimate the "simulated" number of welfare-eligible months that would have occurred if the Connecticut sample had instead been subject to Florida's less generous EID rules. When using this simulated welfare use as the dependent variable, the estimated effect of the time limit on welfare use becomes negative, consistent with the findings from the Florida sample. This simulation exercise suggests that if Connecticut's EID rule had been less generous, welfare use would likely have declined during the pre-time-limit period following the implementation of the time limit policy.

This study contributes to the literature on the impacts of the 1996 welfare reform on welfare caseloads and labor market outcomes. In particular, it adds to the body of research on welfare time limits in three key ways.[2] First, this paper presents findings from Connecticut's JFP sample that are inconsistent with the existing time-limit literature. The fact that Connecticut's earned income disregard rule led to distinct behavioral responses suggests that the effects of time limits may be heterogeneous and depend on the nature of the rule that determines benefit amounts with respect to earnings.

Second, I examine the trade-off between welfare and work more closely by analyzing whether individuals combine both, rather than only choosing one or the other, by analyzing the two outcomes separately. The findings suggest that the time limit policy in Connecticut, together with the state's generous earned income disregard rule, affected the structure of the budget set of recipients in a way that encourages them to combine work and welfare rather than substitute between them.

I provide evidence that the time limit policy, even when combined with the generous EID policy, did not enhance the economic self-sufficiency of welfare recipients. Specifically, I construct a self-sufficiency index based on welfare receipt history, employment status, total earnings, and family income following Hoynes et al. (2016), to assess the broader impact

[2]See Appendix Section 6.1 for a full literature review.

of time limits on the overall economic self-sufficiency of low-income female-headed families. While the core objective of the 1996 welfare reform was to reduce welfare dependency and promote low-income families' economic independence, the results indicate that the time limit policy did not significantly improve the overall self-sufficiency of families in either the FTP or JFP samples during the first two years after random assignment.

The remainder of the paper is organized as follows. Section 2 provides a brief background on the 1996 welfare reform and the welfare waiver experiments that preceded it. Section 3 describes the data, empirical strategies, and the construction of the main sample. Section 4 presents the results from the Florida and Connecticut samples. Section 5 discusses potential mechanisms underlying the findings from the Connecticut sample. Section 6 concludes.

2 Background and Data

2.1 The 1996 Welfare Reform and Welfare Waiver Experiments

The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), commonly known as the 1996 welfare reform, substantially reshaped the U.S. social safety net for low-income families by placing greater emphasis on employment and providing a fixed block grant to state governments. The most impactful change was the replacement of the entitlement nature of the Aid to Families with Dependent Children (AFDC) program with the block-grant structure of Temporary Assistance for Needy Families (TANF). The AFDC program is one of the earliest American means-tested safety net programs, established in 1935 as part of the New Deal. As an open-ended entitlement program, AFDC provided cash assistance to anyone who qualified and could expand flexibly in response to macroeconomic conditions. However, over the course of 60 years, there were growing concerns that the program disincentivized work, discouraging recipients from entering the labor force and making them dependent on public assistance. Additionally, as the share of single female-headed households rose, and more than 50 percent of single-mother families received AFDC

benefits O'Neill and O'Neill (1997), some critics argued that the program disincentivized marriage.

With these concerns in mind, PRWORA fundamentally restructured the means-tested cash assistance program in 1996–1997. Unlike the AFDC program, TANF is a block grant program under which state governments receive a fixed amount of federal funding each year. TANF is specifically designed to encourage employment by providing incentives to work, imposing direct work requirements, and imposing financial sanctions against those who do not comply with work requirements. With considerable discretion given to state governments, some also adopted policies aimed at promoting two-parent families and discouraging out-of-wedlock births. Despite this latitude, the federal law imposed certain core requirements–most notably, a 60-month (or shorter) lifetime limit on benefit receipts.

Before 1996, the federal government encouraged state governments to apply for AFDC "waivers." An AFDC waiver allowed states to be exempt from certain federal requirements so they could implement experimental pilot programs to test new welfare policies. Typically, they contracted with private research organizations and randomly assigned applicants and recipients into either a treatment group, subject to the new rules, or a control group, subject to the existing AFDC rules, to compare outcomes across the two groups and observe the effects of the new policies. Between January 1993 and August 1996, 43 states received waivers to modify participation standards under AFDC, experimenting with stronger work requirements, more generous earned income disregard (EID) policies, stricter sanctions for noncompliance, family caps, and time limits. Many elements of these demonstration projects were later incorporated into the actual TANF program (Rosewater, 1997). Thirty-two out of the 43 states experimented with time limits during this period, although the length and structure of the limits varied across states (See Appendix Section A1 for more details).[3]

[3]Notably, PRWORA prohibited states from using federal funds to support families who had used up 60 months of welfare.

Table 1: AFDC, Florida's Family Transition Program, and Connecticut's Jobs First Program

		AFDC	Florida's Family Transition Program	Connecticut's Jobs First Program
Time and place		Before the welfare reform, Whole United States	Feb 1994 - 1999, Escambia County	Jan 1996 - 2000, Manchester and New Haven
Subjects		Those who meet eligibility criteria	5000+ welfare applicants and recipients	7000+ applicants and recipients
Maximum benefit payments		Same as benefits of waiver experiments	\$241, \$303, \$364 for family of 2, 3, 4	\$443, \$543, \$639 for family of 2, 3, 4
Policies:	Time limit (TL)	x	24 months in any 60-month period (36 months in any 72-months period for the least job-ready participants)	21 months, with a chance of getting renewable six-month extensions
	Work incentives (WI)	Earned income disregard (first 4 months: \$120 + 33% of earnings, months 4-12: \$120, after months 12: \$90) Asset limits \$3000, value of vehicle excluded up to \$1,500	More generous earned income disregard (first \$200 + 50% of remaining earnings) More generous asset limits (\$5,000) including the value of vehicle (\$8,150)	More generous earned income disregard (All earned income is disregarded as long as a family earns below the federal poverty level*) More generous asset limits (\$3,000) including the value of vehicle (\$9,500)
	Work requirements (WR)	Traditional welfare-to-work programs	Participation in intensive case management, employment-related activities	Participation in job searching activities including Job Search Skills Training courses
	Child care (CC)	One-years of child care assistance Parents exempt from work requirements if a child is under 3 years old in FL (2 years old in CT)	Two-years of child care assistance Parents exempt from work requirements if a child is under 6 months old	Child care assistance is provided as long as a family income < 75 % of a state median Parents exempt from work requirements if a child is under 1 year old
	Others (OT)			Medical assistance for families leaving welfare for work

Note. This table summarizes the characteristics of the AFDC program, Florida's Family Transition Program, and Connecticut's Jobs First Program. The summary is based on the information from Bloom et al. (2000) and Bloom et al. (2002).

*The 1998 Federal poverty level (FPL) was \$905, \$1138, \$1371 for family of 2, 3, 4 based on monthly earnings.

2.2 Data

To analyze the effects of time limits on welfare use and labor market outcomes, I use data from Connecticut's Jobs First Program (JFP) Evaluation, provided by the Manpower Demonstration Research Corporation (MDRC).[4] I also use data from Florida's Family Transition Program (FTP) Evaluation, which is used in GM03, to compare the results across two experiments. Although some details of the experiment differed between Connecticut and Florida, both shared the primary goal of encouraging able-bodied welfare recipients to enter the labor force and reduce welfare use, thereby promoting financial self-sufficiency.

Connecticut's JFP was implemented in Manchester and New Haven counties, Connecticut, for four years beginning in January 1996. The economies of both states were strong during the periods in which the experiments took place. The unemployment rate of the United States steadily declined after the early 1990s. Because of this trend, the local economy of both Connecticut and Florida was strong when the waiver experiments were carried out. Both FTP and JFP randomly assigned individuals who were either newly applying for AFDC benefits or renewing their applications.

Table 1 compares the characteristics of Florida's FTP and Connecticut's JFP. Among five policy components—time limits (TL), work incentives (WI), mandatory work requirements (WR), enhanced child care services (CC), and others (OT)—the key difference between FTP and JFP lies in the WI part, particularly in work incentives determined by each state's earned income disregard (EID) rule. This EID rule determines how much earned income is excluded when calculating a family's monthly benefits. Under Florida's FTP EID rule, a single parent with two children earning $300 per month would receive $253 in benefits, because the family's earned income is calculated as $50 ($253=$303-$50), as the rule disregards the first $200 out of $300, and then the half of the remaining $100. This reduces the benefit by $50. In contrast, under Connecticut's JFP EID rule, a parent in the exact same situation would receive the

[4]MDRC is one of the research institutes that participated in various state experiments during the waiver period.

full benefit amount of $543, because JFP disregards all earnings as long as a family's income remains below the federal poverty level, which was $1,138 in 1998 for a family of three. This unusually generous EID policy in Connecticut provided a strong incentive to start working or to combine work with welfare benefits.

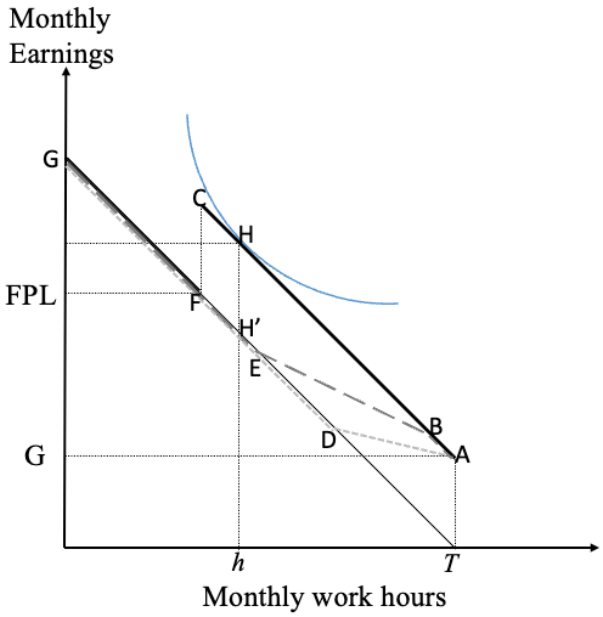

Figure 1: Graphical illustration of a budget constraint by the earned income disregard rules of Florida and Connecticut

Notes. The figure above illustrates the budget constraints under the earned income disregard rule of AFDC (light grey, short-dashed line: ADG), Florida (dark grey, long-dashed line: ABEG), and Connecticut (black, solid line: ACFG). FPL represents the federal poverty line, and G indicates the welfare benefit amount. For simplicity, I assume that the benefit amount is the same between Florida and Connecticut. For further details about the earned income disregard rules of AFDC, Florida's FTP, and Connecticut's JFP, see Table 1.

Figure 1 illustrates the budget constraints under the EID rules for AFDC (light grey, short-dashed line: ADG), Florida (dark grey, long-dashed line: ABEG), and Connecticut (black, solid line: ACFG). For example, suppose an individual's preferences are represented by the blue indifference curve. Under Connecticut's JFP rules, she chooses to work $T - h$ hours and receives the full welfare benefit since her total monthly earnings remain below the federal poverty line. Now, assume that the same person were in Florida. Without any behavioral response, she would receive no benefits if she worked $T - h$ hours, as Florida's

EID rule provides zero benefits at this level of earnings. After reaching the time limit, the budget constraint becomes TG.

There was a slight difference in their time limit as well. Florida's FTP imposed a 24-month time limit, with a 36-month limit applied to the least job-ready individuals. On the other hand, Connecticut's JFP implemented a 21-month time limit, with the possibility of receiving renewable 6-month extensions. JFP participants in Connecticut could not know in advance whether they would be granted an extension, as the application and approval process only began once a family reached the time limit.

These distinct features of FTP and JFP led to different patterns of welfare use and employment among individuals assigned to the waiver group. Panels A and B of Figure 2 illustrate trends in welfare use and employment for the waiver groups (the treated) of each state relative to the AFDC group (the controls), during the 24 months before and after randomization. Each line represents the difference in the share of people in the waiver group versus the AFDC group who fall into one of the five categories: receiving AFDC, receiving AFDC while employed, receiving AFDC without employment, employed without receiving AFDC, and neither receiving AFDC nor employed. The lines before random assignment illustrate whether pre-trends in welfare use and employment were similar between the two groups. The lines after random assignment represent the average treatment effects (ATE) of each state's waiver experiment.

Before randomization and assignment for different rules, the lines are nearly flat and close to zero in both Panels A and B, indicating that the pre-trends in welfare use and employment of single mothers were generally similar between the waiver and AFDC groups in both Florida and Connecticut as expected. After random assignment, however, the trends diverge between the two groups for many of these outcomes, suggesting that the new waiver policies influenced participants' decisions regarding welfare use and employment.In addition, these changes differ between Florida's FTP group and Connecticut's JFP group. The share of individuals receiving welfare benefits increased much more substantially in Connecticut's

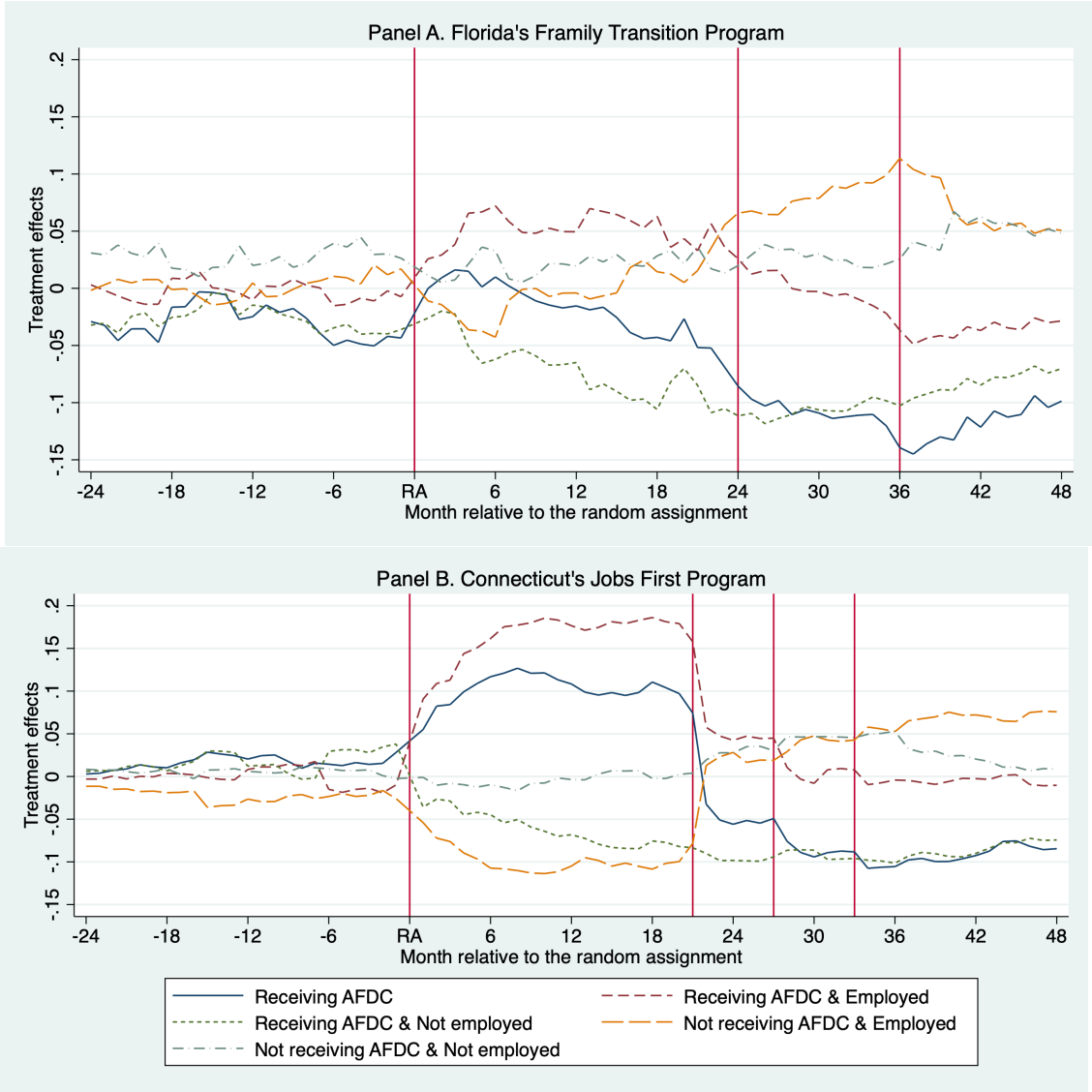

Figure 2: Difference in welfare use and employment between the waiver group and the AFDC group before and after the randomization

Notes. The above panels are based on data from Florida's FTP and Connecticut's JFP, provided by MDRC. The x-axis represents the number of months relative to the time of random assignment. The first vertical line in each panel indicates the month of random assignment. The other lines mark the months when some treated families could reach the time limit, assuming they continuously received benefits since randomization. The y-axis shows the difference in the share of individuals between the waiver and AFDC groups who fall into each of the following five categories: receiving AFDC, receiving AFDC while employed, receiving AFDC without employment, employed without receiving AFDC, and neither receiving AFDC nor employed.

waiver group compared to Florida's after random assignment. In particular, Connecticut's JFP group shows a substantial rise in the number of participants combining welfare and work, along with a decrease in those working without receiving welfare. These trends may reflect the characteristics of the JFP's policy design: a shorter but more lenient time limit with renewable 6-month extensions, combined with a generous EID rule that provided stronger incentives to work. As a result, Connecticut's waiver group had a greater incentive to combine work and welfare following randomization. Moreover, the EID policy allowed some previously ineligible working families to qualify for benefits. However, these patterns in the Connecticut group abruptly faded 21 months after randomization, when the first families reached the 21-month time limit. This suggests that the policy was actively enforced, leading to the termination of benefits for some families.

3 Empirical Strategy And Data

To identify the effects of time limits on welfare use and labor market outcomes, I follow the GM03 strategy, which leverages variation in the age of the youngest child in each household at the time when the time limit is imposed. This estimation strategy allows for the identification of the effect of time limits while accounting for policy endogeneity and other time-varying factors, under a specific identifying assumption.

Table 2 illustrates the GM03 strategy. A_0 denotes the age of the youngest child at the time of randomization, and $\bar{A}$ is the threshold age above which a welfare time limit does not affect a family's dynamic budget constraint. τ_k represents the average treatment effect (ATE) of each policy component k, where k includes time limit (TL), work requirements (WR), work incentives (WI), child care services (CC), and others (OT), as listed in Table 1. The key identification assumption in this model is that the effects of the other policy components, $\tau_{WR}+\tau_{WI}+\tau_{CC}+\tau_{OT}$, are the same regardless of the age of the youngest child. GM03 refers to this as the "age-invariance assumption." If the age-invariance assumption

holds, then the effects of the time limit, τ_{TL}, can be isolated by subtracting the ATE for the group whose youngest child is at or above the threshold age (i.e., $A_0 \geq \bar{A}$) from the ATE for the group whose youngest child is below the threshold age (i.e., $A_0 < \bar{A}$). Hereafter, I refer to this former group as the "above-threshold age group" and the latter as the "below-threshold age group." I verify that the pre-trends of the outcome variables between the below- and above-threshold age groups are parallel, as expected by random assignment, providing evidence in support of this assumption (see Appendix Figure A1-1 and A2).[5]

Table 2: Identification strategy of GM03

	Waiver Policy	**AFDC**	**Effects of Waiver**
Below-threshold age group ($A_0 < \bar{A}$)	WR, WI, CC, OT, TL	AFDC	$\tau_{WR} + \tau_{WI} + \tau_{CC} + \tau_{OT} + \tau_{TL}$
Above-threshold age group ($A_0 \geq \bar{A}$)	WR, WI, CC, OT	AFDC	$\tau_{WR} + \tau_{WI} + \tau_{CC} + \tau_{OT}$

For FTP families whose youngest child is 0.5-3 years old, I cannot identify the effect the time limit because FTP parents are exempted from the mandatory work requirements if thier youngest child is under 3 years old, while the AFDC (or control) group was exempted if their youndest child is under six months old. Therefore, I drop families with youngest child aged 0-3 from the main sample. In Connecticut, AFDC families were exempt if their youngest child was under 2 years old, while JFP parents were exempt if their youngest child was under 1 year old. Although this would allow identification of the time limit effects for Connecticut families whose youngest child is over 2, I instead drop families with a youngest child aged 0–3 (rather than 0-2) to maintain consistency with the Florida sample.

[5]GM03 tests the age-invariance assumption using the FTP data by examining welfare reform initiatives that implemented experiments similar to Florida's FTP but did not include the time limit, such as National Evaluation of Welfare-to-Work Services (NEWWS), the Vermont Welfare Restructuring Project (WRP), and the Minnesota Family Investment Program (MFIP). The point estimates show little variation between groups with younger versus older youngest children, offering empirical support for the age-invariance assumption.

I restrict the main sample to families whose youngest child was between 13 and 17 years old at the time of randomization, as they are least likely to violate the age-invariance assumption. Families with infants or toddlers may respond differently to waiver policies than families with teenagers, due to differences in child care needs and parental availability for work. This range is narrow enough to minimize systematic differences in responses to work incentives and child care policies across age groups, while still providing plenty of variation in time limit exposure.

Using experimental data additionally offers advantages in isolating the effects of time limits, compared to relying on cross-state variation in the timing and duration of time limits. Since the 1996 reform gave state governments substantial latitude in designing their TANF programs based on local needs, program content was likely influenced by the socioeconomic conditions of each state and other contemporaneous changes, which may also be related to welfare caseloads and employment decisions among the AFDC/TANF target population. Moreover, evaluating the effects of individual TANF policy components is vulnerable to policy endogeneity. Because states were granted the authority to design their own programs, they made decisions not only about the timing and duration of time limits but also about other elements, such as the maximum benefit amounts, types of assistance (e.g., cash vs. in-kind benefits), work incentive structures, and the content and strictness of mandatory work requirements.[6] If states with longer time limits also tend to offer more generous benefits in other areas, estimates of the time limit's effect may be upwardly biased. Including state fixed effects does not fully address this issue, as there may still be time-varying unobservables that are simultaneously correlated with time limit policies and the key outcome variables.[7]

[6]Substantial heterogeneity in cash assistance amounts has existed across states since the AFDC era and continues under TANF. For example, as of July 2018, the maximum TANF cash benefit for a family of three ranged from $170 in Mississippi to $1,039 in New Hampshire (Congressional Research Service, 2022).

[7]The rapid nationwide rollout of TANF—including its time limit provisions—further limits the usefulness of cross-state variation. The reform was implemented at the state level and spread across states over roughly five years. In contrast, the Food Stamp Program was rolled out more gradually—from 1961 to 1975—at the county level, offering more variation across time and region for identifying policy effects (Hoynes et al., 2016).

To identify the effects of a time limit, τ_{TL}, I estimate the following equation:

$$y_i = \beta_0 + \beta_{TL}\ E_i \cdot 1\{A_{i0} < \bar{A}\} + \beta_E\ E_i + \beta_A\ 1\{A_{i0} < \bar{A}\} + X_i \cdot \Gamma + \epsilon_i \tag{3.1}$$

where E_i is a dummy variable indicating whether a family i is assigned to the treatment group; A_{i0} is the age of the youngest child in a family i at the time of randomization; $\bar{A}$ is the threshold age above which a family is not subject to time limits. Therefore, the indicator $1\{A_{i0} < \bar{A}\}$ equals 1 if the time limit binds the family's dynamic budget constraint. The matrix of covariates X_i includes race dummies for whether a mother is Black or Hispanic, the mother's years of education, a dummy for whether a mother is under age 25 at the time of randomization, the number of children, the age of the youngest child, the number of months the family received AFDC and Food Stamps in the 24 months before randomization, the number of quarters with employment in the 8 quarters prior to randomization, and total earnings over the same period. For the Florida sample, I additionally control for whether the individual was subject to the 36-month time limit rather than the 24-month limit.

In Equation (3.1), β_{TL} is the coefficient of interest that captures the effect of time limits under the age-invariance assumption. β_E captures the effect of waiver policy components other than time limits, i.e., $\tau_{WR} + \tau_{WI} + \tau_{CC} + \tau_{OT}$, and β_A represents the effect of having a youngest child below the threshold age, $\bar{A}$. The GM03 model predicts that β_{TL} should be negative when the outcome variable is welfare use, and positive when the outcome is employment.

Another key prediction of the GM03 model is that the effects should be larger for families with younger youngest children. The following equation captures the heterogeneous effects based on the age of the youngest child:

$$y_i = \lambda_0 + \lambda_{TL} \cdot E_i \cdot 1\{A_{i0} < \bar{A}\} \cdot (A_{i0} - \bar{A}) + \lambda_E \cdot E_i + \lambda_A \cdot 1\{A_{i0} < \bar{A}\} + X_i \cdot \Pi + u_i \tag{3.2}$$

where $(A_{i0} - \bar{A})$ represents the difference between the youngest child's age and the threshold

age. Assuming a linear age effect on y_i, λ_{TL} captures the impact of the time limit based on the child's age. For example, under the 24-month time limit, families with a 10-year-old youngest child would experience a change in their outcome variables of $-6 \cdot \lambda_{TL}$, relative to the changes experienced by families whose youngest child is older than the threshold age. The theory predicts λ_{TL} to be positive, suggesting that the effects of time limits diminish as the child's age increases and approaches the threshold. I later relax the assumption of a linear age effect and examine whether alternative functional forms, such as a quadratic specification, alter the overall results.

The threshold age $\bar{A}$ is 16 for individuals in Florida who faced the 24-month time limit and 15 for those who were granted the 36-month time limit. For the Connecticut sample, the threshold age is 16.25, as the time limit was 21 months. However, due to data limitations, I use age 16 as the threshold age for the Connecticut sample. The JFP experiment data only report the age of the youngest child on a yearly basis, unlike Florida's data, which provides monthly age information. Although this choice may introduce measurement error and lead to attenuation bias, using age 16 as the baseline threshold for Connecticut maximizes the number of observations in the counterfactual group, where the youngest child is older than the 16.25 threshold at the time of random assignment.[8] I also provide results using age 17 as the threshold age for the Connecticut sample, excluding families whose youngest child was 16 years old to avoid ambiguity in identifying whether their youngest child was above the 16.25 threshold.

For the outcome variables, I use various measures of welfare use and employment over the 24 months following randomization for the Florida sample and 21 months for the Connecticut sample. Specifically, I examine the number of months of welfare use, quarterly employment, the number of months an individual received welfare benefits while employed, the number

[8]The number of observations for families whose youngest child was 17 at the time of random assignment is only 72. This small number is due to two factors. First, the JFP experiment had a relatively small scale, with approximately 7,000 families participating, and the public data contains information on only about 4,000 families. Second, families with older youngest children were less common than those with younger children, as the sampling overrepresented families with younger children.

of months an individual received benefits without employment, the number of months an individual was employed but did not receive benefits, and the number of months an individual neither worked nor received benefits. I also explore the effects on monthly Food Stamp use, total earnings, and total family income to see how the time limits affect the overall economic activities and resources of single mothers.

Additionally, I constructed a self-sufficiency index to measure the time limit's effect on the overall self-sufficiency of single mothers. The self-sufficiency index follows the method of Hoynes et al. (2016), using z-scores for (1) the number of months without receiving AFDC benefits, (2) the number of months without receiving Food Stamp benefits, (3) quarters worked, (4) total earnings, and (5) total family income. I sum the z-scores for each variable and then calculate the z-score of the summed value. As a result, the mean and standard deviation of the score are zero and one, respectively. The higher the z-score of a family, the more "self-sufficient" the family is perceived to be.

My main sample includes individuals who are female, single-parents, have at least one child older than three, and have non-missing information about the youngest child and all control variables I use in the analysis for both Florida's FTP and Connecticut's JFP samples. These criteria leave me 1,423 individuals (out of 2,508) for Florida and 2,104 individuals (out of 4,000) for Connecticut. The reduction occurs mainly because I drop families whose youngest child is below three, while the sampling method of both waiver experiments put larger weight on families with young children. A few additional observations are dropped because of missing information on demographic characteristics, prior welfare use, and prior employment.

Appendix Table A2 presents the summary statistics for the Florida and Connecticut samples. Both samples consist of disadvantaged families. For example, the average years of schooling for female household heads in both samples is 11 years, which is below the level of a high school graduate. The racial/ethnic composition differs slightly between the samples: 50 percent of the Florida sample is Black, with very few Hispanics, while 39 percent of the

Connecticut sample is Black and 22 percent is Hispanic. The Connecticut sample seems to have received more AFDC and Food Stamp benefits prior to randomization than the Florida sample. However, the Connecticut sample also has, on average, 0.24 more quarters of employment experience than the Florida sample, a difference that is statistically significant at the 1 percent level. The fourth and eighth columns show that the randomization was successful in both Florida and Connecticut. I observe well-balanced explanatory variables, including demographic characteristics, pre-randomization welfare use, and pre-randomization labor market outcomes, except for the number of months of welfare use during the 24 months prior to random assignment in the Connecticut sample.

4 Results

4.1 Effects on AFDC Use, Labor Supply, and Combining Welfare and Employment

Table 3 presents the estimation results of Equation (3.1). The treatment effect for the below-threshold age group (individuals whose youngest child is less than the threshold age, above which the time limit does not apply) is $\beta_{TL} + \beta_E$, while the effect for the above-threshold age group is β_E. The difference between the two ATEs, β_{TL}, captures the effect of the time limit on each outcome variable. Each column in each panel reports results from a separate regression. I report only β_{TL} and β_E and their standard errors in Table 3. Coefficients for all other explanatory variables are provided in Appendix Tables A3-A, A3-B, and A3-C.

Column 1 of Panel A shows that in Florida, the waiver experiment increased AFDC use among the above-threshold age group by 2.4 months ($= \hat{\beta}_E$), but had almost no effect for the below-threshold age group ($= \hat{\beta}_{TL} + \hat{\beta}_E$). The estimated effect of the time limit, identified by the difference in ATEs between the two groups, is -2.5 months ($= \hat{\beta}_{TL}$), implying that the time limit decreased welfare use by 2.5 months during the first 24 months after randomization. Although we cannot reject the null hypothesis, it is narrowly insignificant

Table 3: Effects of the time limit policy on welfare use and employment, by state

Outcome variable	(1) AFDC	(2) Empl.	(3) AFDC & Empl.	(4) AFDC & No empl.	(5) No AFDC & Empl.	(6) No AFDC & No empl.
Panel A. Florida						
Interaction (β_{TL})	-2.477	-0.422	-1.290	-1.187	0.0238	2.453
	(1.61)	(0.63)	(0.88)	(1.56)	(1.70)	(1.69)
E (β_E)	2.350	0.638	2.248***	0.102	-0.335	-2.015
	(1.56)	(0.61)	(0.83)	(1.52)	(1.66)	(1.65)
Pre-treatment mean	12.38	2.38	2.73	9.66	4.41	7.21
Panel B. Connecticut (16+ as a control group)						
Interaction (β_{TL})	0.373	0.194	1.507*	-1.134	-0.924	0.551
	(1.20)	(0.45)	(0.78)	(1.19)	(1.04)	(1.16)
E (β_E)	1.696	0.446	2.364***	-0.668	-1.026	-0.670
	(1.16)	(0.43)	(0.73)	(1.15)	(1.00)	(1.14)
Pre-treatment mean	15.93	2.62	3.85	12.09	4.01	4.05
Panel C. Connecticut (17+ as a control group)						
Interaction (β_{TL})	2.565*	0.183	2.924***	-0.359	-2.375	-0.190
	(1.41)	(0.72)	(0.72)	(1.58)	(1.82)	(1.76)
E (β_E)	-0.502	0.457	0.942	-1.444	0.430	0.0719
	(1.38)	(0.72)	(0.66)	(1.55)	(1.80)	(1.74)
Pre-treatment mean	15.93	2.62	3.85	12.09	4.01	4.05
Controls	O	O	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (1). Standard errors are in parentheses. The observations used for the analyses in Panel A, B, and C are 1,423, 2,104, and 2,031. The effect of overall waiver policies for the above-threshold age group is captured by β_E, while the effect for the below-threshold age group is captured by $\beta_{TL} + \beta_E$. Thus, β_{TL} is a DD estimator that identifies the effect of the time limit under the age-invariance assumption. For each column, I provide the mean of each outcome variable during the 24 months prior to randomization. All variables are measured at the monthly level except the employment variable, which is measured at the quarterly level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(p-value = 0.12), and the sign is consistent with the model's prediction and empirical findings from prior studies. Relative to the pre-treatment mean, the reduction can be translated into a 20 percent reduction. Panel A of Appendix Figure A2 plots the time limit's effects on AFDC use by quarter, and shows that the effects were generally negative over the eight

quarters following randomization. Regarding labor supply response, the effect of the time limit on months of employment is close to zero (Column 2).

The number of months receiving AFDC benefits can be decomposed into months during which individuals received benefits while working and while not working (Columns 3 and 4). Similarly, total months of employment can be broken down into months worked with and without receiving welfare benefits (Columns 3 and 5). The estimation result in Column 3 shows that the time limit reduced the number of months combining welfare and work by 1.3, although the standard error is too large to reject the null hypothesis (p-value = 0.14). The magnitude of the effect on months receiving welfare without employment (Column 4) is similar, but the larger standard error results in an even higher p-value. Interestingly, Column 6 shows that the time limit had a positive, though narrowly insignificant, effect on the number of months during which individuals neither worked nor received welfare (p-value = 0.15). Overall, the results in Panel A suggest that the time limit slightly reduced AFDC use among FTP participants but did not necessarily lead to an increase in labor supply.

The results from the Connecticut sample in Panel B and C shows a different result. In Columns 1 and 2 of Panel B, the effects of the time limit on AFDC use and employment are both close to zero. However, when I use age 17 as the threshold while dropping observations with children aged 16, the point estimate of the effect on AFDC use in Panel C becomes positive and statistically significant at the 10 percent level (p-value = 0.07). The estimated effect on employment remains largely unchanged. Overall, the point estimates in Column 1 of Panels B and C suggest that families in the below-threshold age group in Connecticut did not reduce–and may have even increased–their AFDC use during the first 21 months after random assignment.

Another distinctive finding from the Connecticut sample is the positive effect of the time limit on combining welfare and work. Results in Column 3 of both Panels B and C indicate that the time limit increased the number of months individuals simultaneously received welfare benefits and were employed by approximately 1.5-2.9 months, which can be translated

into 39-75.3% increase compared to the pre-treatment mean. In contrast, Column 4 shows a negative (although statistically insignificant) effect on the number of months receiving AFDC benefits without employment. These results suggest that, when faced with the 21-month time limit, families in Connecticut were more likely to combine work and welfare and less likely to rely on benefits while unemployed. This is not surprising given the generous EID rule that the Connecticut sample faced. Appendix Figure A3 breaks down the time limit's effects on combining welfare and work on a quarterly basis. Panel B shows that the effects were consistently positive from quarter 1 through 7 following randomization, with the impact gradually increasing over time.

To summarize, the results in Table 3 reveal that individuals in Florida and Connecticut responded differently to the introduction of the welfare time limit. Unlike in Florida, the Connecticut sample did not actively reduce the welfare use in response to the policy. The time limit in Connecticut also led single mothers to combine work and welfare, a pattern not observed in Florida. In contrast, the signs and magnitudes of β_E across specifications are broadly similar between the two states, suggesting that the effects of waiver policies other than the time limit were generally consistent. Thus, it is specifically the time limit policy that had differential impacts in Florida and Connecticut.

As shown earlier in Figure 2, Table 3 shows that the Florida and Connecticut samples differ in how they combine welfare and work. This difference is likely influenced by Connecticut's much more generous earned income disregard (EID) policy, which makes it easier for JFP participants to simultaneously receive welfare benefits and engage in employment. In other words, the choice set for families in Florida is largely limited to either welfare or work, whereas families in Connecticut can choose welfare, work, or a combination of both. I will examine this mechanism in greater detail in Section 5.

4.2 Heterogeneity by the Age of the Youngest Child

A key prediction of the GM03 model is that the effect of the time limit should be smaller when the youngest child is older. Table 4 presents the estimation results of Equation (3.2), which allows for heterogeneity in the time limit effect by the age of the youngest child. Under this specification, the effect of the time limit for a family whose youngest child is A_0 years old is given by $(A_0 - \bar{A}) \cdot \lambda_{TL}$ when $A_0 < \bar{A}$.[9] Note that the age effect is not defined when $A_0 \geq \bar{A}$.

The results from the Florida and Connecticut samples again show different results. When using Connecticut's JFP sample, the age effect on the monthly AFDC use is close to zero (Panel B). The effect in Panel C is -0.15, which is negative and statistically significant at the 0.1 level (p-value = 0.07). This negative estimate of the age effect implies that if the age of a child among the below-threshold age group increases by one year, then the number of months receiving AFDC benefits decreases by -0.15 months. Again, the result for the AFDC use of the Connecticut sample is not consistent with the GM03 model's prediction that the (negative) effect of the time limit on welfare use decreases as children get older.

I find a notable difference between the Connecticut and Florida samples in how families combine welfare and work depending on the age of the youngest child, again likely because of the different benefit reduction rates. In Column (3), both specifications in Panels B and C yield negative and statistically significant point estimates of -0.12 and -0.15, respectively. These results imply that families were more likely to combine welfare and work when the age of their youngest child was farther below the threshold age. For example, based on the estimate from Panel B, Connecticut families whose youngest child was six years old, which is ten years younger than the threshold age of sixteen, spent an additional 1.2 months receiving AFDC benefits while working during the first 21 months after random assignment, compared

[9] Appendix Table A4 replicates the estimation in Table 4, but assumes a quadratic functional form for age effects. The estimated age effects of the time limit on the six outcome variables are generally linear in both Panel A and B, as the coefficients on the squared age terms are close to zero. In Panel C, some quadratic terms are statistically significant, but the overall results remain qualitatively similar.

to families in the above-threshold age group. This represents a 31.2 percent increase relative to the pre-treatment average of 3.85 months combining work and welfare. Since the estimated effect of the time limit on months combining welfare and work was positive in Table 3, the negative coefficient on the age variable in Panel B of Table 4 can be interpreted that families with younger children were more strongly impacted by the time limit.

Table 4: Heterogeneous effects of the time limit policy on welfare use and employment by the age of the youngest child

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome variable	AFDC	Empl.	AFDC & Empl.	AFDC & No empl.	No AFDC & Empl.	No AFDC & No empl.
Panel A. Florida						
Age effects (λ_{TL})	0.114	0.00410	0.0202	0.0939	-0.00792	-0.106
	(0.10)	(0.04)	(0.06)	(0.09)	(0.10)	(0.09)
E (λ_E)	0.910	0.272	1.190**	-0.279	-0.375	-0.536
	(0.87)	(0.33)	(0.55)	(0.84)	(0.87)	(0.80)
Pre-treatment mean	12.38	2.38	2.73	9.66	4.41	7.21
Panel B. Connecticut (16+ as a control group)						
Age effects (λ_{TL})	-0.0776	-0.0185	-0.121*	0.0434	0.0654	0.0122
	(0.08)	(0.03)	(0.07)	(0.08)	(0.07)	(0.06)
E (λ_E)	1.463**	0.489**	2.863***	-1.401**	-1.398**	-0.0650
	(0.66)	(0.24)	(0.55)	(0.67)	(0.55)	(0.57)
Pre-treatment mean	15.93	2.62	3.85	12.09	4.01	4.05
Panel C. Connecticut (17+ as a control group)						
Age effects (λ_{TL})	-0.146*	-0.0169	-0.145**	-0.00164	0.0938	0.0525
	(0.08)	(0.03)	(0.07)	(0.08)	(0.08)	(0.07)
E (β_E)	0.701	0.486	2.507***	-1.806**	-1.047	0.347
	(0.77)	(0.30)	(0.66)	(0.79)	(0.71)	(0.67)
Pre-treatment mean	15.93	2.62	3.85	12.09	4.01	4.05
Controls	O	O	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (2). Standard errors are in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.3 Effects on Food Stamp Use, Earnings, Income, and Self-Sufficiency Index

In this section, I examine the effects of the AFDC waiver's time limit policy on Food Stamp use, earnings, family income, and the self-sufficiency index to provide a more comprehensive understanding of its impact on the overall economic self-sufficiency of welfare recipients.

Table 5 presents the estimated effects of the time limit policy on additional outcome variables, including the number of months receiving Food Stamp benefits, total earnings, total family income, and the self-sufficiency index during the first 24 months following randomization for Florida and the first 21 months for Connecticut. Panel A and Panel B report results from Equation (3.1) (labeled as "*Baseline DD*") and Equation (3.2) (labeled as "*Age heterogeneity*"), respectively.

In Panel A, results from the Florida sample suggest that the time limit reduced Food Stamp use by 3.7 months. The age effect, $\lambda_{TL,1}$, is positive and statistically significant, indicating that the negative impact of the time limit becomes weaker as the age of the youngest child approaches the threshold. This pattern mirrors the diminishing age effect observed for AFDC use in Table 4. These consistent results likely stem from the close linkage between AFDC and Food Stamp benefits. For example, individuals who apply for AFDC/TANF benefits are typically automatically enrolled in Food Stamps if they meet the eligibility criteria for both programs. Conversely, when AFDC benefits are terminated, Food Stamp benefits often end as well—unless recipients take steps to reapply for them. The estimates show minimal impact on total earnings and family income, suggesting that the quality of jobs did not change significantly in response to the time limit or other waiver policies. Overall, the time limit policy does not appear to have improved the self-sufficiency index of participants in Florida's FTP experiment.

In Panel B, for the Connecticut sample, the effect of the time limit on Food Stamp use is positive, though narrowly insignificant (p-value = 0.13). There is little evidence

Table 5: Effects of the time limit policy on monthly Food Stamp use, total earnings, total family income, and overall self-sufficiency

Outcome variable	(1) Food Stamp use	(2) Total earnings	(3) Total family income	(4) Self-sufficiency index
Panel A. Florida				
Baseline DD				
Interaction (β_{TL})	-3.744**	-352.8	-1472.0	0.130
	(1.58)	(2498.32)	(2432.16)	(0.24)
E (β_E)	3.165**	1026.3	1597.6	-0.0601
	(1.53)	(2466.01)	(2399.89)	(0.23)
Age heterogeneity				
Age effects ($\lambda_{TL,1}$)	0.196**	-47.70	8.257	-0.0120
	(0.10)	(120.44)	(118.50)	(0.01)
E (β_E	1.177	317.9	273.4	-0.0312
	(0.84)	(1122.66)	(1086.60)	(0.11)
Panel B. Connecticut (16+ as a control group)				
Baseline DD				
Interaction (β_{TL})	1.696	1525.8	2339.4*	0.0549
	(1.12)	(1390.28)	(1352.94)	(0.15)
E (β_E)	-0.701	-643.2	399.5	0.00322
	(1.08)	(1320.04)	(1288.05)	(0.14)
Age heterogeneity				
Age effects ($\lambda_{TL,1}$)	-0.0656	-46.96	-165.8	-0.00279
	(0.07)	(103.20)	(101.10)	(0.01)
E (β_E)	0.389	427.6	1339.0	0.0335
	(0.62)	(847.41)	(838.01)	(0.08)
Controls	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (1) and (2). Results from Equation (1) are presented under *Baseline DD* section and results from Equation (2) are presented under *Age heterogeneity* section.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

of an effect on total earnings. However, a positive and statistically significant effect is observed on total family income (p-value = 0.08), but only in the Connecticut sample. This increase likely stems from individuals in the below-threshold age group in Connecticut increasingly combining welfare and work after encountering the 21-month time limit, driven by the incentives created by the state's generous EID rule and lower benefit reduction rate. Since the effect on total earnings is negligible, the rise in family income is likely driven primarily by higher AFDC take-up rates and benefit amounts. Lastly, as in the Florida

sample, the introduction of the time limit does not appear to improve the self-sufficiency index of participants in the JFP experiment.

4.4 Heterogeneity by Welfare Use and Employment History before Randomization

To examine how these changes in cash assistance policy and the introduction of time limits affected the strength of the safety net and individuals' self-sufficiency, I examine heterogeneity in impacts between single mothers who are more job-ready and those who are less job-ready. To do this, I assess the effect of the time limit based on individuals' welfare use and employment history prior to randomization. On average, individuals in the Florida and Connecticut samples received AFDC benefits for 12.4 and 15.9 months, respectively. The average number of quarters worked prior to randomization was 2.4 in Florida and 2.6 in Connecticut. I classify individuals as "less job-ready" if they received welfare for more than the sample average and were employed for fewer than the average number of quarters at the time of random assignment.

Table 3.6.1 shows that the results from the Florida sample are primarily driven by the more job-ready subsample, which includes individuals who used welfare less than average and were employed more than average prior to randomization. In particular, the estimated effect of the time limit is -5.7 among individuals with longer employment histories, while the effect is -1.4 and statistically insignificant among those with shorter employment histories. More job-ready individuals also increased their labor supply by 1.9 quarters in response to the welfare time limit. In contrast, less job-ready individuals decreased their labor supply by 1.5 quarters, a result inconsistent with the predictions of the GM03 model. These findings suggest that the behavioral responses predicted by the GM03 model appear only among individuals with some prior work experience. Although the estimated effects of the time limit on the self-sufficiency index are not statistically significant, the point estimate is larger for the more job-ready group than for the less job-ready group. Overall, the results imply

that individuals with limited work experience may struggle to adjust their welfare use and labor supply decisions in response to the additional constraint imposed by the time limit policy.

Table 3.6.1: Heterogeneous effects of the time limit policy by prior welfare use and employment history —Results from the Florida sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome variable	AFDC	Empl.	AFDC & Empl.	AFDC & No empl.	FS	Earn.	Inc.	Self-suff.
Panel A. By welfare use of 2 years prior to the random assignment								
Lower welfare use								
Interaction (β_{TL})	-1.908	0.0169	-0.816	-1.093	-1.922	801.6	159.3	0.168
	(2.46)	(0.88)	(1.44)	(2.74)	(2.12)	(1791.54)	(1867.80)	(0.24)
E (β_E)	2.030	0.346	2.401*	-0.371	2.157	-364.9	81.78	-0.124
	(2.40)	(0.86)	(1.36)	(2.67)	(2.07)	(1731.74)	(1809.93)	(0.23)
Higher welfare use								
Interaction (β_{TL})	-2.829	-0.915	-1.950**	-0.878	-5.629**	-1691.2	-3255.9	0.0730
	(2.07)	(0.88)	(0.96)	(1.66)	(2.30)	(4482.37)	(4213.26)	(0.40)
E (β_E)	2.500	0.959	2.245**	0.255	4.236*	2538.0	3212.2	0.0172
	(1.99)	(0.85)	(0.91)	(1.59)	(2.21)	(4432.92)	(4161.76)	(0.39)
Panel B. Subgroup by the employment histor								
Shorter employment history								
Interaction (β_{TL})	-1.383	-1.483**	-1.439	0.0562	-4.186**	-2463.4	-3362.7	-0.114
	(1.94)	(0.71)	(1.00)	(1.99)	(1.90)	(2562.07)	(2571.73)	(0.25)
E (β_E)	0.853	1.665**	2.128**	-1.276	3.262*	2933.9	3209.5	0.194
	(1.87)	(0.68)	(0.94)	(1.93)	(1.83)	(2518.64)	(2524.88)	(0.24)
Longer employment history								
Interaction (β_{TL})	-5.710**	1.926*	-1.636	-4.074*	-2.945	5295.3	3117.0	0.746
	(2.73)	(1.16)	(1.63)	(2.20)	(2.77)	(5217.39)	(4860.61)	(0.50)
E (β_E)	6.043**	-1.632	2.919*	3.124	2.900	-4145.7	-2440.5	-0.674
	(2.65)	(1.14)	(1.55)	(2.14)	(2.69)	(5152.63)	(4803.21)	(0.49)
Controls	O	O	O	O	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (1). Standard errors are in parentheses. I present results separately by individuals' welfare use and employment history prior to randomization. The four subsamples are defined based on whether an individuals' pre-randomization welfare use or employment duration was above or below the sample average.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.6.2: Heterogeneous effects of the time limit policy by prior welfare use and employment history —Results from the Connecticut sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome variable	AFDC	Empl	AFDC & Empl	AFDC & No empl	FS	Earn	Inc	Self-suff
Panel A. By welfare use of 2 years prior to the random assignment								
Higher								
Interaction (β_{TL})	-0.796	0.0141	1.394	-2.190	2.218*	-101.8	369.0	-0.0439
	(1.54)	(0.52)	(0.98)	(1.69)	(1.30)	(1437.37)	(1410.52)	(0.17)
E (β_E)	2.522*	0.753	2.673***	-0.150	-1.577	1303.9	2470.2*	0.152
	(1.50)	(0.50)	(0.90)	(1.64)	(1.27)	(1357.40)	(1338.79)	(0.16)
Lower welfare use								
Interaction (β_{TL})	1.821	0.502	1.516	0.305	2.094	4174.9	5660.7**	0.191
	(1.92)	(0.75)	(1.23)	(1.63)	(1.93)	(2560.11)	(2443.72)	(0.25)
E (β_E)	0.975	-0.0771	2.080*	-1.104	-0.224	-3902.9	-2989.3	-0.232
	(1.83)	(0.73)	(1.14)	(1.55)	(1.84)	(2410.53)	(2292.40)	(0.24)
Panel B. Subgroup by the employment history								
Shorter employment history								
Interaction (β_{TL})	0.451	0.404	2.606***	-2.155	1.642	373.3	750.0	-0.00336
	(1.47)	(0.54)	(0.92)	(1.53)	(1.31)	(1374.41)	(1435.81)	(0.16)
E (β_E)	0.716	0.613	1.351	-0.635	-1.166	1436.6	2176.4	0.182
	(1.42)	(0.52)	(0.84)	(1.46)	(1.26)	(1286.18)	(1347.46)	(0.15)
Longer employment history								
Interaction (β_{TL})	-0.678	0.378	0.0892	-0.767	0.981	5398.5*	6380.6**	0.364
	(1.97)	(0.83)	(1.55)	(1.80)	(2.08)	(3128.60)	(2874.10)	(0.31)
E (β_E)	4.034**	-0.244	3.772**	0.262	0.715	-5731.6*	-3804.3	-0.467
	(1.90)	(0.81)	(1.48)	(1.76)	(2.02)	(3022.61)	(2784.61)	(0.30)
Controls	O	O	O	O	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (1). Standard errors are in parentheses. I present results separately by individuals' welfare use and employment history prior to randomization. The four subsamples are defined based on whether an individuals' pre-randomization welfare use or employment duration was above or below the sample average.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

On the other hand, the results from the Connecticut sample in Panel B of Table 3.6.2 show that the positive effect of the time limit on the months combining welfare and work, found in Table 3, is driven by the less job-ready group in terms of employment history. However, the time limit policy decreased the number of months receiving welfare without employment (Column 4), although the effect is statistically insignificant (p-value = 0.16). The positive effect on the months combining welfare and work is offset by the negative impact on the months receiving welfare without employment, which masks the overall effect on AFDC use, bringing it close to zero. This pattern is not observed among individuals

who were employed more than the average for the total Connecticut sample. Interestingly, the below-threshold age group affected by the time limit experienced an increase in total earnings by $5,400 compared to the above-threshold age group. Since the effect of the time limit on employed quarters is close to zero (0.378), it can be inferred that individuals in this group were more likely to secure higher-paying jobs than those with shorter employment histories. This increase in earnings translated into a $6,400 rise in total family income for the group with longer working experience, but no such effect was found in the group with less employment history. The increased family income may also stem from higher AFDC benefit amounts due to the EID rule in the JFP policy, which provides the full amount of benefits to families whose earnings are below the federal poverty line.

5 Mechanism: Unusually Generous Earned Income Disregard Rules in Connecticut

In this section, I examine the hypothesis that Connecticut's unusually generous EID rules contributed to its distinct results. In Connecticut, individuals can receive 100 percent of the maximum cash assistance benefits as long as their earnings are below the federal poverty line. In contrast, Florida's EID rule disregards the first $200 of income and then 50 percent of any remaining earnings when calculating benefit amounts. To investigate whether the difference in EID rules explains Connecticut's different results, I seek to answer the following question: If treated individuals in Connecticut had been subject to Florida's less generous EID rule, would the time limit's effects have aligned more closely with the GM03 model's prediction?

To explore this possibility, I calculate the "simulated" monthly benefit receipt that individuals in one state's treatment group would have experienced had they faced the other state's EID rules. For simplicity, I ignore other requirements such as asset limits and behavioral responses due to a change in the budget constraint. To isolate the income effects attributable to the different maximum benefit amounts between Florida and Connecticut

(e.g., Florida provides $303 for a family of three, while Connecticut provides $543), I replace only the EID rule without altering the maximum benefit amount. Appendix Table A6 shows how the benefit amounts change under Florida and Connecticut's EID rules. The simulation method is explained in more detail in Appendix Section A6, where I also discuss the potential bias resulting from ignoring the behavioral response.

Table 8 presents the results for all simulated variables related to the months of AFDC use. The variables in Columns 2 and 3 represent the upper and lower bounds of the simulated number of months using AFDC, while the variables in Columns 4 and 5 correspond to the upper and lower bounds of the number of quarters during which AFDC benefits were received. Unlike the zero effect of the time limit on actual AFDC use reported in Column 1, the negative effects of the time limit on AFDC use become evident when examining all four simulated variables. These results align with the predictions of the GM03 model.

Table 8: Effects of welfare time limit on "simulated" welfare use: Connecticut

Outcome variable	(1) Original AFDC (Col.1 of Table 3)	(2) Sim. AFDC (Upper bound)	(3) Sim. AFDC (Lower bound)	(4) Sim. ever used AFDC (Upper bound)	(5) Sim. ever used AFDC (Lower bound)
Interaction (β_{TL})	0.373	-1.980***	-7.930***	-1.980***	-4.655***
	(1.20)	(0.34)	(0.77)	(0.34)	(0.43)
E (β_E)	1.696	17.47***	13.29***	3.468***	1.605***
	(1.16)	(0.32)	(0.75)	(0.32)	(0.42)
Controls	O	O	O	O	O

Note. The table presents the estimated effects of the time limit on actual and simulated measures of AFDC use for the Connecticut sample. Column 1 reports the effect on actual observed months of AFDC receipt, while Columns 2–5 show the effects on simulated measures that estimate how long individuals would have received benefits had Connecticut implemented Florida's less generous Earned Income Disregard (EID) rule. The simulated outcomes include upper and lower bounds for both monthly and quarterly measures of AFDC use, as described in Appendix Section A6.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

This exercise suggests that the generous EID rule of Connecticut's JFP played a critical role in driving the unique patterns observed in the Connecticut sample. Due to the surpris-

ingly generous EID rule, participants of the JFP were able to work while still receiving the maximum welfare benefits. Additionally, the time limit policy further increased the tendency to combine welfare and work.

6 Robustness Checks

To test the robustness of the findings, I examine three additional factors that might potentially explain the distinctive results from the Connecticut sample: (1) the less strict time limit policy of Connecticut's JFP compared to Florida's FTP, (2) the identifying assumptions of the empirical model, and (3) the assumptions of the GM03 model. The results suggest that none of these factors are the primary drivers of the Connecticut findings.

Less Strict Enforcement of the Time Limit and Potential 6-Month Extensions in Connecticut

One possible explanation for the effects of the time limit in Connecticut is that participants may not have fully understood or believed that benefits would terminate due to the time limit. The lack of awareness or skepticism about enforcement could weaken the policy's impact.

However, results from the Interim Client Survey indicate that most participants in the JFP experiment understood the existence of the 21-month time limit and the potential for 6-month extensions. Only about 9 percent believed that they would not lose benefits even after reaching the time limit Bloom et al. (2002). These survey results suggest that the Connecticut sample had a clear understanding of the time limit structure and the possibility of benefit loss after reaching the time limit.

To further explore whether beliefs about enforcement shaped different behavioral responses, I exploit a unique institutional difference between Manchester and New Haven Counties. In Manchester, caseworkers tended to be less definitive about the possibility of extensions, emphasizing the need for recipients to find a job and prepare for the potential

loss of benefits. In contrast, staff in New Haven were more likely to reassure recipients that they could receive extensions if they complied with all program requirements.[10] Therefore, if disbelief about the enforcement of the time limit were a key driver of Connecticut's divergent results, then we would expect to see the most distinctive patterns in the New Haven subsample. In contrast, outcomes in the Manchester subsample should more closely resemble those observed in Florida.

Table 7: Effects of the time limit policy—Manchester and New Haven

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome variable	AFDC	Empl	AFDC & Empl	AFDC & No empl	FS	Earn	Inc	Self-suff
Panel A. Manchester								
Interaction (β_{TL})	1.312	-0.223	1.909	-0.597	1.003	-1479.0	-709.7	-0.169
	(2.45)	(1.00)	(1.50)	(2.32)	(2.47)	(3282.01)	(2801.85)	(0.36)
E (β_E)	0.812	1.186	2.922**	-2.110	0.211	2477.1	3546.9	0.255
	(2.36)	(0.97)	(1.38)	(2.23)	(2.39)	(3129.44)	(2643.68)	(0.34)
Panel B. New Haven								
Interaction (β_{TL})	0.0571	0.359	1.538*	-1.481	1.897	2524.5*	3403.9**	0.134
	(1.35)	(0.49)	(0.90)	(1.38)	(1.22)	(1487.36)	(1511.58)	(0.16)
E (β_E)	1.998	0.169	2.018**	-0.0201	-0.883	-1764.3	-738.3	-0.0935
	(1.31)	(0.48)	(0.84)	(1.33)	(1.18)	(1414.66)	(1443.40)	(0.15)
Controls	O	O	O	O	O	O	O	O

Note. This table contains the estimation results for selected variables from Equation (1), separately for the Manchester and New Haven subsamples. Standard errors are in parentheses. The observations of the Manchester and New Hanve samples are 501 and 1,603.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Results in Column 1 of Table 7 suggest that neither the Manchester nor the New Haven subsample exhibits the expected negative effect of the time limit on AFDC use. In New Haven, where recipients were more confident about the possibility of extensions, the positive effects of the time limit on combining welfare and work, total earnings, and family income

[10]Bloom et al. (2002) describes this difference as follows: "Some workers said that recipients who cooperated with the program's mandates would likely receive an extension if they could not find a job; they emphasized that recipients should comply with all program rules to ensure that they did not make themselves ineligible for an extension. Other workers were much less definite, saying they did not know which recipients would receive an extension and that clients thus needed to make every effort to find a job to prepare for the possible loss of benefits. The staff survey results indicate that New Haven staff were much more likely to adopt the former approach, while Manchester staff adopted the latter (page 59)."

are stronger and statistically significant. This suggests that belief in the availability of extensions may have encouraged recipients to continue working while receiving benefits, rather than saving their eligibility for future periods. In contrast, no significant effects are observed in the Manchester subsample, nor the negative effects predicted by the GM03 model. Instead, the Manchester sample shows similar patterns to New Haven, with a positive (albeit imprecise) effect on the number of months combining welfare and work. Given that recipients in Manchester, like those in Florida, were less certain about receiving extensions, these results suggest that disbelief in benefit loss alone does not explain the distinctive findings in Connecticut.

Identifying Assumptions of the Empirical Model

Second, I revisit the validity of the age-invariance assumption for the Connecticut sample, which is a key identifying assumption of the empirical strategy. Specifically, I check whether the effect of the waiver policies other than the time limit, $\tau_{WR} + \tau_{WI} + \tau_{CC} + \tau_{OT}$, was the same across the above- and below-threshold age groups. I conclude that it is highly likely to be so, as the 24-month and 21-month time limits were somewhat arbitrarily set, making the threshold age above which families are not affected by the time limit effectively random. Additional evidence supporting this is provided by Appendix Figures A1-1 and A2, which clearly show that the pre-trends of welfare use and employment variables are quite similar between the above- and below-threshold age groups. While the graphical illustration is not a sufficient condition to guarantee the age-invariance assumption, it provides favorable evidence in support of it. To further improve the precision of the estimation, I control for demographic variables, pre-randomization welfare use, and pre-randomization labor supply.

I also consider the possibility that the effects of the time limit and other programs are not additive in Connecticut. While there is no straightforward way to test this, it is an important consideration because if the effects of the time limit and the other programs are not additive, the GM strategy would not be able to isolate the impact of the time limit policy.

Assumptions of the GM03 Model

I next examine whether the assumptions of the GM03 model—forward-looking utility-maximizing individuals, wage uncertainty, and credit constraints—hold in the context of 1990s Connecticut. The following evidence suggests that none of these assumptions are violated in the case of the Connecticut sample.

First, regarding the "forward-looking" and "utility-maximizing" assumptions, there is no reason to believe that only the Connecticut sample was not forward-looking or unable to optimize its behavior to maximize lifetime utility. However, one possibility worth exploring is recent research on how poverty can impair mental capacities or "bandwidth" through factors like malnutrition, sleep deprivation, exposure to risky and unhealthy environments, or easy access to substance use (Schilbach et al., 2016). If certain characteristics of Connecticut's JFP limited participants' mental bandwidth, leaving them with less cognitive capacity to reoptimize decisions on welfare use and labor supply when faced with the time limit, then this could help explain the puzzling patterns observed in the Connecticut sample.

However, the results from Tables 3.6.1 and 3.6.2 suggest that this is unlikely. In Florida, it was the less disadvantaged individuals with lower prior welfare use and more working experience who adjusted in line with the model's predictions. This aligns with recent findings suggesting that more vulnerable populations may have limited mental bandwidth. In contrast, in Connecticut, both the more and less job-ready groups did not react as expected based on the model. Instead, the group with shorter employment experience that appeared to face greater barriers prior to randomization responded more actively by combining welfare and work in response to the time limit. Based on these observations, I conclude that there is insufficient evidence to argue that the forward-looking and utility-maximizing assumptions were violated for Connecticut's JFP sample.

I also examine the validity of the assumptions regarding wage uncertainty and credit constraints. During the evaluation period, both Florida and Connecticut experienced relatively strong local economies (Bloom et al., 2000, 2002), with decreases in unemployment rates

according to the Local Area Unemployment Statistics from the Bureau of Labor Statistics. Therefore, it is unlikely that the wage uncertainty assumption was violated only for the Connecticut sample. Additionally, survey results suggest that the credit constraint assumption is not severely violated. Seventy-five percent of the Florida sample reported having no savings, and 57 percent had borrowed less than $5,000.[11] Similarly, 79 percent of the Connecticut sample had no savings, and 62 percent had borrowed less than $5,000. These findings suggest that the credit constraint assumption holds reasonably well for both the Florida and Connecticut samples.

7 Conclusion

In this paper, I examine the effects of the time limit policy on welfare use and labor supply. I find that, contrary to expectations, the time limit did not reduce welfare use among Connecticut's JFP sample. Instead, it encouraged individuals to combine welfare and work more. This positive effect on the number of months combining welfare and work diminishes as the age of the youngest child increases. Additionally, I observe a statistically significant increase in total household income. The increased tendency to combine welfare and work is primarily driven by the less job-ready subgroup, characterized by lower levels of prior work experience.

I present suggestive evidence that Connecticut's unusually generous earned income disregard rule is the key factor driving these distinct patterns. By simulating the application of Florida's less generous earned income disregard rule to Connecticut's sample, I demonstrate that the time limit would have led to decreased AFDC use under Florida's rules. These findings suggest that the effects of welfare time limit policies are context-dependent and must be considered alongside other policies, particularly those related to earned income disregard rules.

[11]The question about borrowing is as follows: "Not counting any money you might owe on a mortgage or a car, do you (or your (husband/wife/partner)) now owe any money? That is, do you owe any money to stores, banks, hospitals or doctors, credit unions, credit card companies, friends or relatives, and so on?"

References

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006). What mean impacts miss: Distributional effects of welfare reform experiments. *American Economic Review 96*(4), 988–1012.

Bloom, D., J. Kemple, P. Morris, S. Scrivener, N. Verma, R. Hendra, D. Adams-Ciardullo, D. Seith, and J. Walter (2000). The family transition program: Final report on florida's initial time-limited welfare program. *Manpower Demonstration Research Corporation*.

Bloom, D., S. Scrivener, C. Michalopoulos, P. Morris, R. Hendra, D. Adams-Ciardullo, J. Walter, and W. Vargas (2002). Jobs first: Final report on connecticut's welfare reform initiative. *Manpower Demonstration Research Corporation*.

Chan, M. K. (2013). A dynamic model of welfare reform. *Econometrica 81*(3), 941–1001.

Chan, M. K. (2017). Welfare dependence and self-control: An empirical analysis. *The Review of Economic Studies 84*(4), 1379–1423.

Chan, M. K. (2018). Measuring the effects of welfare time limits. *Journal of Human Resources 53*(1), 232–271.

Congressional Research Service (2022). The temporary assistance for needy families (tanf) block grant: Responses to frequently asked questions. *Congressional Research Service Report*.

Council of Economic Advisers (1997). Explaining the decline in welfare receipt, 1993–1996.

Fang, H. and M. P. Keane (2004). Assessing the impact of welfare reform on single mothers. *Brookings Papers on Economic Activity 2004*(1), 1–95.

Fang, H. and D. Silverman (2004). On the compassion of time-limited welfare programs. *Journal of Public Economics 88*(7-8), 1445–1470.

Fang, H. and D. Silverman (2009). Time-inconsistency and welfare program participation: Evidence from the nlsy. *International Economic Review 50*(4), 1043–1077.

Grogger, J. (2002). The behavioral effects of welfare time limits. *American Economic Review 92*(2), 385–389.

Grogger, J. (2003). 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*(2), 394–408.

Grogger, J. (2004). Time limits and welfare use. *Journal of Human Resources 39*(2), 405–424.

Grogger, J. and C. Michalopoulos (1999). Welfare dynamics under time limits. *NBER Working Paper* (w7353).

Grogger, J. and C. Michalopoulos (2003). Welfare dynamics under time limits. *Journal of Political Economy 111*(3), 530–554.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-run impacts of childhood access to the safety net. *American Economic Review 106*(4), 903–34.

Mazzolari, F. (2007). Welfare use when approaching the time limit. *Journal of Human Resources 42*(3), 596–618.

Moffitt, R. A. et al. (1999). The effect of pre-prwora waivers on afdc caseloads and female earnings, income, and labor force behavior. *Economic conditions and welfare reform*, 91–118.

O'Neill, D. M. and J. O'Neill (1997). Lessons for welfare reform: An analysis of the afdc caseload and past welfare-to-work programs. *WE Upjohn Institute for Employment Research*.

Rosewater, A. (1997). Setting the baseline: A report on state welfare waivers. *Office of the Assistant Secretary for Planning and Evaluation*.

Schilbach, F., H. Schofield, and S. Mullainathan (2016). The psychological lives of the poor. *American Economic Review 106*(5), 435–40.

Swann, C. A. (2005). Welfare reform when recipients are forward-looking. *Journal of Human Resources 40*(1), 31–56.

Wang, J. S.-H. (2021). State tanf time limit and work sanction stringencies and long-term trajectories of welfare use, labor supply, and income. *Journal of Family and Economic Issues 42*(4), 650–696.

Wilde, E. T., Z. Rosen, K. Couch, and P. A. Muennig (2014). Impact of welfare reform on mortality: an evaluation of the connecticut jobs first program, a randomized controlled trial. *American Journal of Public Health 104*(3), 534–538.

Ziliak, J. P., D. N. Figlio, E. E. Davis, and L. S. Connolly (2000). Accounting for the decline in afdc caseloads: Welfare reform or the economy? *Journal of Human Resources*, 570–586.

Appendix

A1. Three Types of Time Limits Applied during the Waiver Period

Formats of time limits imposed during this waiver period can be categorized into three groups according to the classification of the Center for Law and Social Policy: termination time limits, work requirement time limits, and reduction time limits (Rosewater, 1997). Termination time limits cut the full benefit of a family who reached the time limits. Work requirement time limits require a parent to meet mandatory work requirements after using up all time-limited benefits. The benefits are granted, however, as long as a family complies with the requirements. Finally, reduction time limits cut only a certain amount or percentage of benefits for a family who hits the time limit. Appendix Table A1 provides the list of states classified into each category. This study mainly focuses on the effects of the termination time limits using Florida and Connecticut data, one of the earliest states that adopted the termination time limit.

Table A1: Types of Time limits

Termination	Work Requirement	Reduction
Connecticut	California	
Delaware	Delaware	
Florida	Colorado	
Hawaii	Georgia	
Illinois	Illinois	
Iowa	Maryland	
Louisiana	Massachusetts	Arizona
Nebraska	Michigan	Indiana
North Carolina	Missouri	Texas
Ohio	Montana	Washington
Oregon	New Hampshire	
South Carolina	North Dakota	
Tennessee	Oklahoma	
Virginia	South Dakota	
Wisconsin	Vermont	

Note. Source: Table II.A of Rosewater (1997).

A2. Literature Review

This section provides the literature review on the studies about welfare time limits. I first describe the theoretical framework constructed by GM03. Then I summarize the results of reduced-form studies, which provide empirical evidence on the behavioral responses of welfare recipients upon facing time limits. Finally, I outline the findings from theory-based studies.

GM03's theoretical prediction regarding the effects of time limits

Studies on the effects of time limits mostly focus on its impact on welfare use to explain the rapid caseload reduction around and after the 1996 reform. GM03 is the key paper that contributes to the literature on the effects of time limits on welfare use. The study provides a theoretical framework to understand the behavior response that time-limited welfare use makes people reduce welfare use even before they reach the time limit.[12]

The outline of the model is as follows. Assume that a consumer is forward-looking. Assume that she faces wage uncertainty and credit constraint, which is not uncommon for disadvantaged families. Let us say a time limit is started at time $t = 0$. A consumer is endowed with an initial stock of "options" to receive a welfare benefit, N. A consumer's problem is not a lifetime utility maximization problem but a problem limited to a time horizon H, a time span until the youngest child of a consumer becomes 18 and loses eligibility.

At each period $t > 0$, a consumer compares the utility of using an option now and the option value of saving it for later. Based on the decision at each period, her stock of remaining options of welfare eligibility at time t can be represented by $S_t = N - \sum_{j=0}^{t-1} p_j$, where p_j is a welfare use indicator at time j. If $S_t \geq H$, then a consumer's problem becomes static because she cannot use the number of remaining options of welfare use, S_t, before her youngest child turns 18 and becomes ineligible for welfare benefits. On the other hand, if $S_t < H$, then a consumer's problem becomes dynamic. In this case, she has to decide whether or not to use

[12]See Grogger and Michalopoulos (1999), a working paper of GM03, for a full model.

welfare at time t, which affects her future options and expected lifetime utility.

In solving the theory, GM03 predicts that the consumer would become less reluctant to use welfare as the available stock of benefits, S_t, increases relative to the remaining eligibility time horizon, H. Therefore, all else equal, the consumer would be less reluctant to use welfare as the age of the youngest child is older, as her H would be smaller if her youngest child's age is higher.

Findings from the reduced-form studies

Earliest papers that study the effects of the 1996 reform use the state-level variation in each state's socioeconomic characteristics and policy contents and do not report significant effects of time limits on welfare caseloads (Council of Economic Advisers, 1997; Moffitt et al., 1999; Ziliak et al., 2000). However, after the seminal work of Jeffrey Grogger and Charles Michalopoulos, which proposes a DD estimator that additionally exploits the variation in the age of the youngest child together with the state-level variation, significant negative effects were found on welfare use (Grogger, 2002; Grogger and Michalopoulos, 2003; Grogger, 2004; Fang and Keane, 2004; Mazzolari, 2007). The reported sizes of the impacts are substantial. For instance, using Current Population Survey (CPS) data, Grogger (2004) calculates that time limits account for about 12–13 percent of the caseload decline between 1993 and 1999. Fang and Keane (2004) estimate that time limits explain 11 percent of the caseload decline between 1993 and 2002.

Mazzolari (2007) directly look at the role of the available stock of benefits, S_t. GM03 did not directly explore this variable because an individual endogenously determines the current welfare use, which is related to the remaining stock of benefits, making this variable endogenous. To overcome the endogeneity problem, the study proposes a 2SLS estimator. Note that the remaining stock of benefits, S_t, can be expressed as a function of the propensity of welfare use to help understand the IV proposed in the study. In a mathematical formula, $S_t = N - k_i \cdot E_{its}$, where N is the initial stock of benefits, $k_i = \frac{\sum_{j=0}^{t-1} p_j}{E_{its}}$ is the propensity of using

welfare, and E_{its} is the time span passed since the implementation of the time limit policy.[13] Only the k_i is endogenous among the three variables, N, k_i, E_{its}. Mazzolari (2007) use the average rate of welfare use of some socioeconomic group j during a pre-reform period, $\hat{k_j}$ where $i \in j$, as an IV for k_i. Thus, the study constructs the IV for S_t by $Z_t = N - \hat{k_j} \cdot E_{its}$. Using Survey of Income and Program Participation (SIPP) data, the study measures the effects of time limits on welfare use by a 2SLS estimation and finds that time limits explain almost 20 percent of the caseload reduction between 1996 and 2003.

Unlike the ample evidence about the negative effects of time limits on welfare use, findings about the impacts on employment and earnings are fewer. Grogger (2003) report narrowly significant positive impact on labor supply and little and insignificant effect on earnings and income. Wang (2021) reports a positive impact on employment, but the effect is found only in states applying time limits shorter than 60 months.

Findings from the structural studies

Another strand of welfare reform studies adopts the theory-based approach in order to encompass closely interrelated aspects of the 1996 welfare reform. Swann (2005) build a dynamic programming model where a forward-looking consumer makes a decision on work, welfare receipt, and marriage under uncertainty about wage and marriage offer. Fang and Silverman (2009) uses a structural life-cycle model. Chan (2013) move further and incorporate in its model a broader set of policy changes during the welfare reform periods, including time limits, work requirements, Food Stamp participation, Earned Income Tax Credit (EITC), and income tax changes. Unlike the consistent results of reduced-form studies, structural studies report mixed findings on the effects of time limits. This could be because structural models often do not consider economic changes during the reform period.[14] Finally, Chan (2018) try to synthesize the reduced-form approach and the theory-based approach. The

[13]For instance, if 6 months passed from the implementation of the 2-year time limit and if a consumer has received welfare benefits 3 times since then, then $S_6 = 24 - 3/6 \cdot 6 = 21$.

[14]What if a consumer is not forward-looking? A few structural studies allow for behavioral factors in the model, such as present-biased preferences, heterogeneity in discounting, and self-control. The studies indicate a potential for utility gains from imposing the time limit (Fang and Silverman, 2004; Chan, 2017).

paper argues that the DD estimator based on the GM03 strategy could understate the true effects of time limits unless controlling for the remaining stock of available benefits in the regression model.

A3. Pre-trend Graphs and Summary Statistics

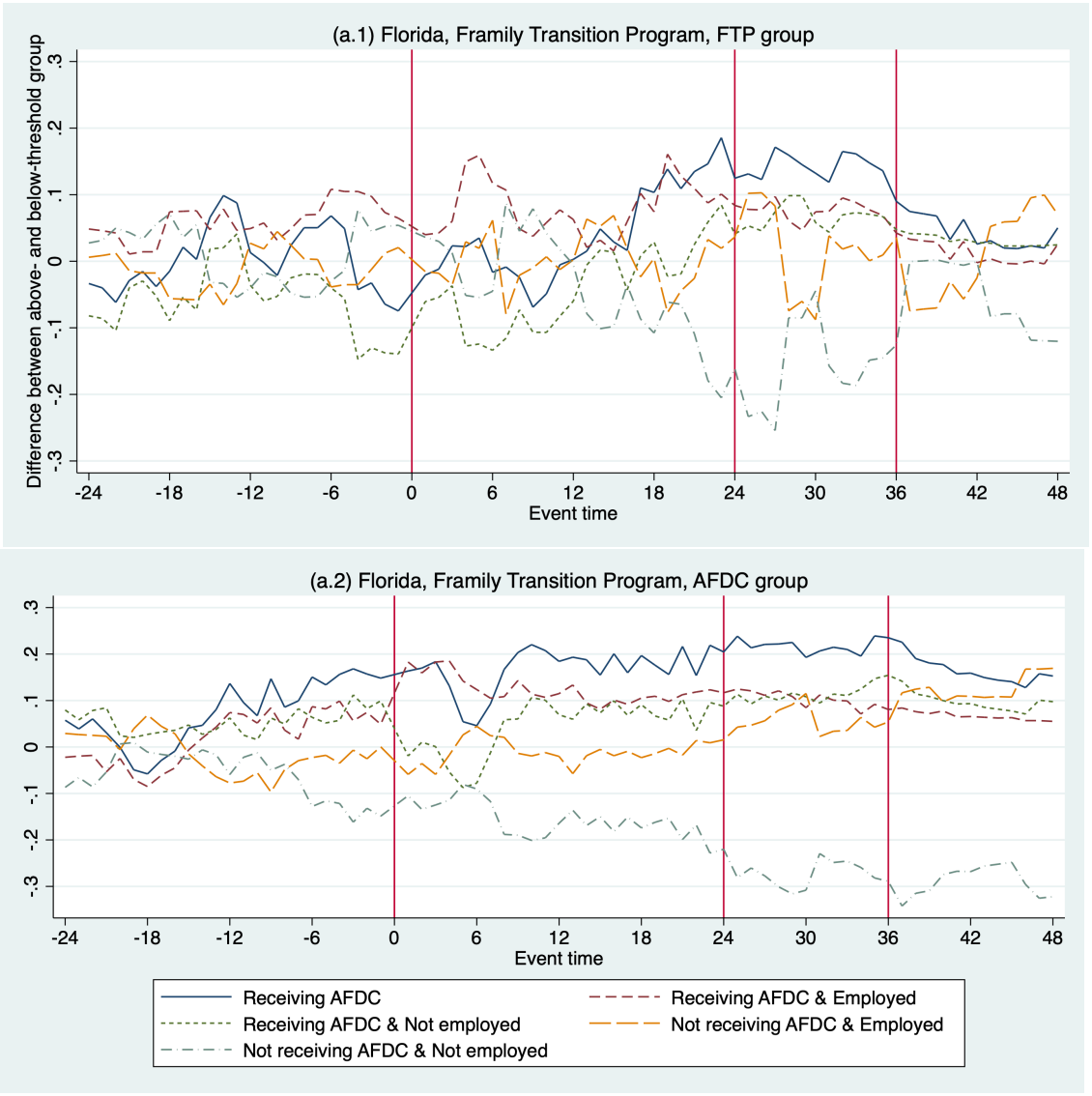

Figure A1-1: Trends of the difference of various outcome variables between the above- and below-threshold age groups, Florida, by treatment status

Notes. The above figures are obtained from the FL's FTP and CT's JFP data provided by MDRC. The first red line in each graph indicates the time of the random assignment. The other red lines represent the time treated people could reach the time limit if they constantly received the benefits since the random assignment.

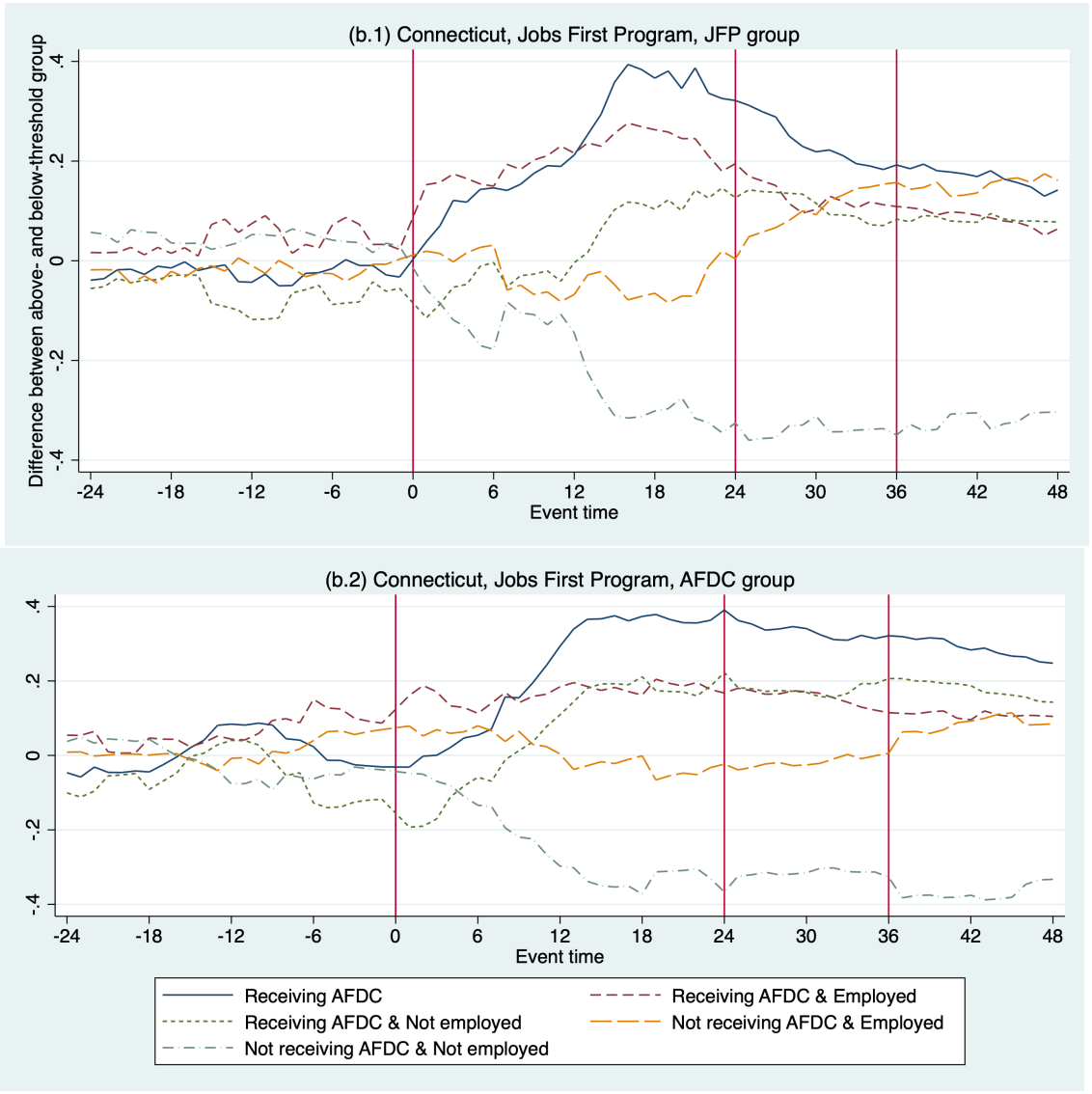

Figure A1-2: Trends of the difference of various outcome variables between the above- and below-threshold age groups, Connecticut, by treatment status

Notes. The above figures are obtained from the FL's FTP and CT's JFP data provided by MDRC. The first red line in each graph indicates the time of the random assignment. The other red lines represent the time treated people could reach the time limit if they constantly received the benefits since the random assignment.

Table A2: Summary statistics of the sample, Florida and Connecticut

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Florida				**Connecticut**			
	All mean/sd	FTP mean/sd	AFDC mean/sd	T-test b/t	All mean/sd	JFP mean/sd	AFDC mean/sd	T-test b/t
Demographic characteristics								
Black	0.50	0.49	0.51	-0.02	0.39	0.39	0.38	0.01
	(0.5)	(0.5)	(0.5)	(-0.6)	(0.5)	(0.5)	(0.5)	(0.7)
Hispanic	0.02	0.02	0.01	0.01	0.22	0.21	0.24	-0.03
	(0.1)	(0.1)	(0.1)	(1.4)	(0.4)	(0.4)	(0.4)	(-1.9)
Years of edu	11.07	11.05	11.10	-0.05	11.15	11.13	11.18	-0.04
	(1.6)	(1.6)	(1.6)	(-0.6)	(1.9)	(1.9)	(2.0)	(-0.5)
Age below 25	0.14	0.14	0.14	0.00	0.07	0.07	0.07	-0.01
	(0.3)	(0.3)	(0.3)	(0.1)	(0.3)	(0.2)	(0.3)	(-0.8)
Num of children	1.95	1.93	1.97	-0.04	1.80	1.81	1.78	0.03
	(1.0)	(0.9)	(1.0)	(-0.8)	(0.8)	(0.8)	(0.8)	(0.8)
Age of the youngest child	7.69	7.54	7.84	-0.30	8.54	8.47	8.62	-0.15
	(4.0)	(3.9)	(4.1)	(-1.4)	(3.8)	(3.8)	(3.9)	(-0.9)
36-month time limit	0.50	0.47	0.52	-0.05*				
(applied only in FL)	(0.5)	(0.5)	(0.5)	(-2.0)				
Welfare use and labor market outcomes during two years prior to the random assignment								
Months of AFDC use	12.38	12.16	12.58	-0.42	15.95	16.58	15.29	1.29**
	(9.7)	(9.7)	(9.7)	(-0.8)	(10.1)	(9.9)	(10.2)	(2.9)
Months of Food Stamp use	15.01	14.73	15.26	-0.53	16.66	17.05	16.25	0.80
	(9.7)	(9.8)	(9.6)	(-1.0)	(9.6)	(9.5)	(9.7)	(1.9)
Quarters of employment	2.38	2.42	2.33	0.09	2.62	2.56	2.69	-0.13
	(2.7)	(2.7)	(2.7)	(0.6)	(3.0)	(2.9)	(3.0)	(-1.0)
Total earnings	4110.12	4125.69	4095.58	30.11	5913.83	5564.65	6278.62	-713.97
	(7340.0)	(6949.2)	(7691.5)	(0.1)	(11173.0)	(10700.9)	(11639.8)	(-1.5)
Welfare use and labor market outcomes after the random assignment, during the first 24 months for FL and 21 months for CT								
Months of AFDC use	11.77	11.65	11.88	-0.23	14.42	15.64	13.14	2.51***
	(8.7)	(8.5)	(9.0)	(-0.5)	(7.7)	(7.1)	(8.0)	(7.6)
Quarters of employment	3.68	3.83	3.55	0.28	3.54	3.83	3.23	0.60***
	(3.0)	(3.0)	(3.1)	(1.7)	(2.9)	(2.8)	(2.9)	(4.8)
Months of getting AFDC	4.40	4.93	3.91	1.02***	6.58	8.48	4.59	3.89***
and work at the same time	(5.6)	(6.1)	(5.1)	(3.4)	(7.2)	(7.8)	(5.8)	(13.0)
Total earnings	7030.71	7399.27	6686.68	712.59	8634.36	8875.63	8382.31	493.31
	(9338.5)	(9684.0)	(8996.9)	(1.4)	(10989.1)	(10469.7)	(11506.4)	(1.0)
Total household income	13608.85	13628.17	13590.83	37.34	18756.86	20035.06	17421.53	2613.53***
	(9219.1)	(9717.4)	(8735.1)	(0.1)	(10562.0)	(10781.5)	(10162.5)	(5.7)
N	1423	687	736	1423	2104	1075	1029	2104

Note. The above table is obtained from Florida's FTP data and Connecticut's JFP data.

Welfare use and labor market variables after the random assignment are defined using the first 24 months after randomization for FL and 21 months for CT. The different time spans are used for FL and CT because the length of the welfare time limit was different between the two states, where FL's time limit is 24 months and CT's time limit is 21 months. Thus, the variables regarding welfare use and labor market outcomes in this table are obtained before nobody reaches the time limit.

A4. Estimation Results for the Full Set of Variables Used in Table 3

Below three Appendix Tables A3-A, A3-B, and A3-C contain the estimated coefficients of all variables used in estimating Equation (1). Each table corresponds to Panel A, B, and C of Table 3. The estimates of coefficients for the control variables show reasonable signs and magnitudes.

Table A3-A: Effects of welfare time limit on welfare use and employment

Panel A. Florida						
	(1) AFDC	(2) Empl	(3) AFDC & Empl	(4) AFDC & No empl	(5) No AFDC & Empl	(6) No AFDC & No empl
Interaction (β_{TL})	-2.477	-0.422	-1.290	-1.187	0.0238	2.453
	(1.61)	(0.63)	(0.88)	(1.56)	(1.70)	(1.69)
E (β_E)	2.350	0.638	2.248***	0.102	-0.335	-2.015
	(1.56)	(0.61)	(0.83)	(1.52)	(1.66)	(1.65)
Below-threshold (β_A)	3.221***	0.516	1.521***	1.699	0.0275	-3.248***
	(1.19)	(0.43)	(0.54)	(1.20)	(1.24)	(1.17)
Age youngest child	-0.0272	-0.0283	-0.116***	0.0883	0.0306	-0.00342
	(0.06)	(0.02)	(0.04)	(0.06)	(0.06)	(0.06)
36-month time limit	1.561***	-0.177	0.427	1.135**	-0.956**	-0.605
	(0.49)	(0.17)	(0.32)	(0.45)	(0.45)	(0.45)
Black	1.466***	0.508***	1.841***	-0.376	-0.317	-1.149***
	(0.44)	(0.16)	(0.29)	(0.41)	(0.41)	(0.40)
Hispanic	0.858	-0.754	0.505	0.354	-2.765***	1.907
	(1.52)	(0.50)	(1.13)	(1.28)	(1.04)	(1.47)
Years of schooling	-0.187	0.255***	0.207**	-0.394***	0.557***	-0.369***
	(0.13)	(0.05)	(0.08)	(0.12)	(0.12)	(0.13)
Age below 25	-0.513	-0.0878	-0.336	-0.178	0.0724	0.441
	(0.69)	(0.23)	(0.45)	(0.61)	(0.59)	(0.59)
Num of children	0.0845	0.0912	0.304*	-0.220	-0.0308	-0.0537
	(0.23)	(0.08)	(0.16)	(0.22)	(0.20)	(0.20)
Prior welfare use	0.371***	0.0122	0.158***	0.212***	-0.121***	-0.249***
	(0.04)	(0.01)	(0.03)	(0.04)	(0.04)	(0.04)
Prior FS use	0.0367	-0.0192	-0.0309	0.0676*	-0.0267	-0.0101
	(0.04)	(0.01)	(0.02)	(0.04)	(0.04)	(0.04)
Prior employment	-0.0357	0.464***	0.736***	-0.772***	0.655***	-0.619***
	(0.12)	(0.04)	(0.10)	(0.10)	(0.12)	(0.11)
Prior earning	-0.0000829**	0.00000857	-0.000125***	0.0000420	0.000151***	-0.0000677
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cons	4.688**	-0.839	-3.351***	8.039***	0.834	18.48***
	(2.10)	(0.77)	(1.26)	(2.05)	(2.05)	(2.07)

Note. This table contains the estimation results of all variables from Equation (1). Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3-B: Effects of welfare time limit on welfare use and employment

Panel B. Connecticut (16+ as a control group)						
	(1) AFDC	(2) Empl	(3) AFDC & Empl	(4) AFDC & No empl	(5) No AFDC & Empl	(6) No AFDC & No empl
Interaction (β_{TL})	0.373	0.194	1.507*	-1.134	-0.924	0.551
	(1.20)	(0.45)	(0.78)	(1.19)	(1.04)	(1.16)
E (β_E)	1.696	0.446	2.364***	-0.668	-1.026	-0.670
	(1.16)	(0.43)	(0.73)	(1.15)	(1.00)	(1.14)
Below-threshold (β_A)	3.766***	0.573*	1.706***	2.060**	0.0125	-3.778***
	(0.92)	(0.33)	(0.54)	(0.95)	(0.80)	(0.90)
Age youngest child	-0.0680	-0.0424**	-0.119***	0.0512	-0.00799	0.0760**
	(0.05)	(0.02)	(0.05)	(0.05)	(0.04)	(0.04)
Black	0.953***	-0.285**	-0.232	1.186***	-0.623**	-0.331
	(0.34)	(0.12)	(0.32)	(0.34)	(0.29)	(0.23)
Hispanic	-0.597	-0.112	-0.315	-0.283	-0.0202	0.617*
	(0.42)	(0.15)	(0.38)	(0.43)	(0.36)	(0.32)
Years of schooling	-0.388***	0.146***	0.0856	-0.474***	0.353***	0.0349
	(0.08)	(0.03)	(0.08)	(0.09)	(0.06)	(0.06)
Age below 25	0.348	0.0240	0.564	-0.215	-0.492	0.143
	(0.59)	(0.22)	(0.58)	(0.58)	(0.53)	(0.39)
Num of children	0.0638	0.0402	0.0861	-0.0223	0.0346	-0.0984
	(0.20)	(0.08)	(0.19)	(0.21)	(0.18)	(0.14)
Prior welfare use	0.233***	0.0218**	0.0997***	0.133***	-0.0342	-0.199***
	(0.03)	(0.01)	(0.03)	(0.03)	(0.02)	(0.03)
Prior FS use	0.00704	-0.0207**	0.00208	0.00496	-0.0641**	0.0571**
	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
Prior employment	-0.155**	0.497***	1.009***	-1.163***	0.483***	-0.328***
	(0.07)	(0.02)	(0.08)	(0.07)	(0.07)	(0.06)
Prior earning	-0.0000522**	-0.0000239***	-0.000117***	0.0000653***	0.0000457*	0.00000651
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cons	11.28***	0.308	-0.491	11.77***	1.416	8.302***
	(1.54)	(0.57)	(1.34)	(1.63)	(1.29)	(1.29)

Note. This table contains the estimation results of all variables from Equation (1). Standard errors are in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3-C: Effects of welfare time limit on welfare use and employment

Panel C. Connecticut (17+ as a control group)						
	(1)	(2)	(3)	(4)	(5)	(6)
	AFDC	Empl	AFDC & Empl	AFDC & No empl	No AFDC & Empl	No AFDC & No empl
Interaction (β_{TL})	2.565*	0.183	2.924***	-0.359	-2.375	-0.190
	(1.41)	(0.72)	(0.72)	(1.58)	(1.82)	(1.76)
E (β_E)	-0.502	0.457	0.942	-1.444	0.430	0.0719
	(1.38)	(0.72)	(0.66)	(1.55)	(1.80)	(1.74)
Below-threshold (β_A)	6.106***	0.490	2.147***	3.959***	-0.678	-5.428***
	(0.92)	(0.40)	(0.58)	(0.95)	(0.98)	(1.14)
Age youngest child	-0.0577	-0.0420**	-0.115**	0.0571	-0.0113	0.0689*
	(0.05)	(0.02)	(0.05)	(0.05)	(0.04)	(0.04)
Black	1.031***	-0.296**	-0.249	1.280***	-0.639**	-0.392*
	(0.34)	(0.12)	(0.33)	(0.35)	(0.30)	(0.23)
Hispanic	-0.615	-0.118	-0.347	-0.268	-0.00602	0.621*
	(0.43)	(0.16)	(0.39)	(0.44)	(0.37)	(0.32)
Years of schooling	-0.406***	0.152***	0.0802	-0.487***	0.376***	0.0303
	(0.08)	(0.03)	(0.08)	(0.09)	(0.06)	(0.06)
Age below 25	0.392	0.0334	0.594	-0.202	-0.493	0.102
	(0.59)	(0.22)	(0.58)	(0.58)	(0.53)	(0.39)
Num of children	0.0691	0.0492	0.106	-0.0371	0.0413	-0.110
	(0.20)	(0.08)	(0.19)	(0.21)	(0.18)	(0.14)
Prior welfare use	0.245***	0.0222**	0.108***	0.137***	-0.0414	-0.203***
	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
Prior FS use	-0.00509	-0.0209**	-0.00357	-0.00152	-0.0591**	0.0642**
	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
Prior employment	-0.178**	0.495***	1.005***	-1.183***	0.479***	-0.301***
	(0.07)	(0.03)	(0.08)	(0.07)	(0.07)	(0.06)
Prior earning	-0.0000465**	-0.0000233***	-0.000114***	0.0000673***	0.0000440*	0.00000250
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cons	9.073***	0.312	-0.981	10.05***	1.918	10.01***
	(1.55)	(0.62)	(1.38)	(1.65)	(1.42)	(1.46)

Note. This table contains the estimation results of all variables from Equation (1). Standard errors are in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A5. Robustness Checks

A5.1. Effects of the Time Limit on a Quarterly Basis

In this Appendix Section A5, I explore the quarterly effects of the time limit by estimating the equation using outcome variables split by each quarter during the 24 months for Florida and 21 months for Connecticut. I plot the quarterly impacts of the time limit on AFDC use in Figure A2 and the quarterly effects on combining AFDC benefits and employment in A3. For instance, Panel A of Figure A2 plots the coefficients of the interaction term, β_{TL}, from Equation (1) using outcome variables of AFDC use during the first three quarters after randomization (Q1), AFDC use during the month 4-6 (Q2), AFDC use during the month 7-9 (Q3), ..., and AFDC use during the month 21-24 (Q8). Panel B of Figure A2 contains the coefficients of the interaction term using outcome variables from Q1 to Q7 because the outcome variables for Connecticut in Table 3 are defined as AFDC use and employment during the first 21 months after the random assignment, a time span where anybody reaches the time limit.

This quarterly analysis has two purposes. First, the analysis is to check whether aggregating outcome variables for the first 24 months for Florida and 21 months for Connecticut mask some critical trends during the period. Figure A2 and A3 show that the overall trends in time limit's effects on using AFDC benefits and combining welfare and work were not very fluctuated during the period of interest, and the results presented in Table 3 are robust. In Figure A2, we can observe that the effects of the time limit on AFDC use were generally negative for the Florida sample. However, none of the estimates are statistically significant. The effects for the Connecticut sample were positive (but not statistically significant) during the first three quarters, but the estimates became negative and closer to zero during the later quarters. Figure A3 contains the effects of the time limit on months of combining welfare and work. Panel A indicates that the time limit policy did not make the Florida sample who faced the time limit combine welfare and work more than those who were not bound by the

time limit. Except for the first quarter, the effects are statistically insignificant and close to zero. On the other hand, Panel B shows that the time limit's impacts on combining welfare and work were positive throughout all quarters, which yields the negative estimate of the time limit's effects during the first 21 quarters in Panel B of Table 3.

Second, the analysis also addresses the concern that families with the youngest kid older than 16 (or 17) years old will mechanically lose eligibility within 24 months (or 21 months). For instance, if a kid was 17 years old at the time of the randomization, the kid's family will lose AFDC/TANF eligibility one year after due to the age eligibility criterion. Theoretically, this mechanical drop does not directly bias the estimates since the automatic drop influences the outcome variables of the above-threshold age groups of both waiver and AFDC groups, thus being canceled out in the DD setting. The first estimate of the coefficient of the interaction term in each Panel, which is the effects of the time limit on outcomes during the first three months (Q1), generally shows similar estimates compared to the results in Table 3. This result implies that the mechanical drop has little impact on the estimation results of Table 3.

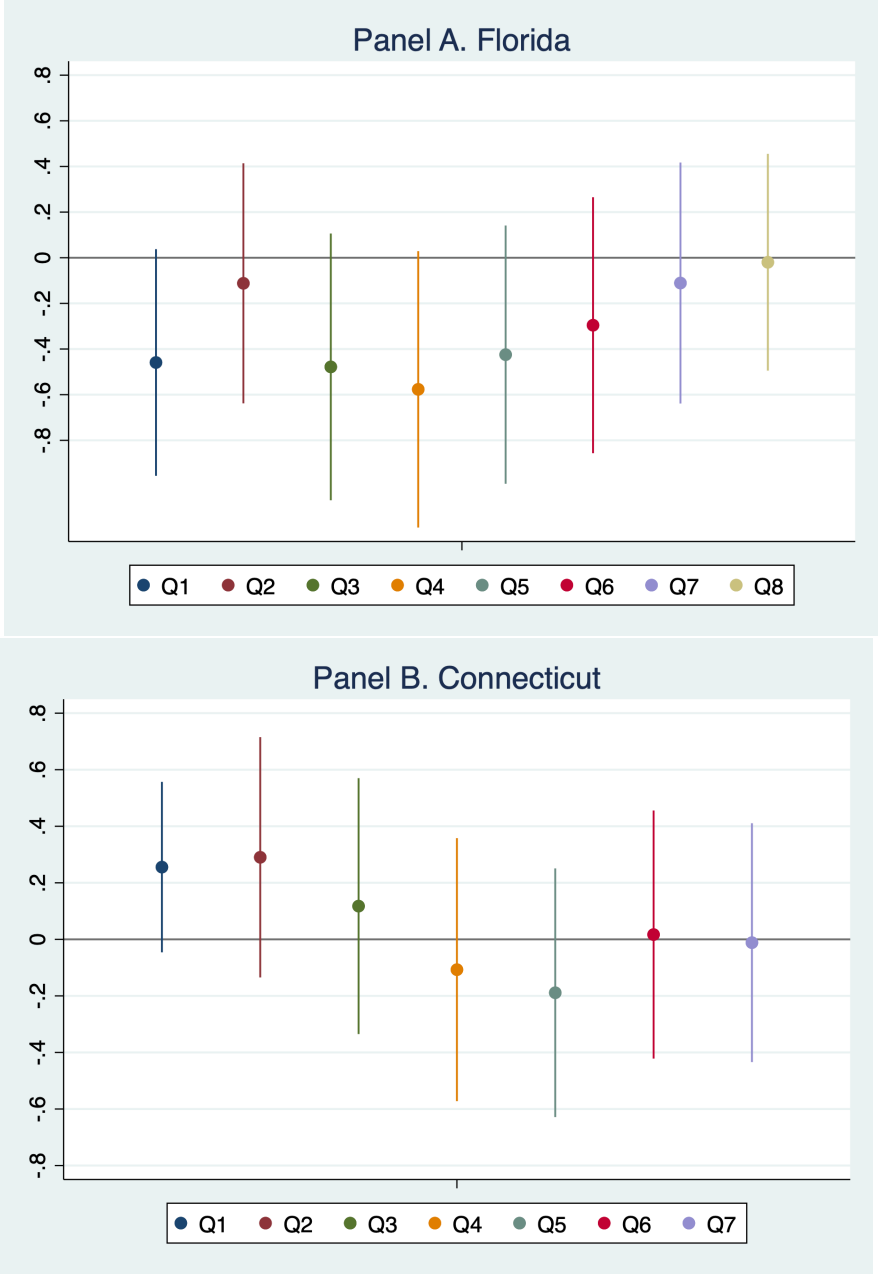

Figure A2: Effects of the time limit on AFDC use on a quarterly basis

Notes. Above two Panels plot the coefficients of the interaction term, β_{TL}, from Equation (1) using quarterly AFDC use. Panel A contains the effects of the time limit during the first three quarters after randomization (Q1), during the month 4-6 (Q2), ..., and during the month 21-24 (Q8). Panel B contains the same estimates except for AFDC use during the month 21-24 (Q8) because the outcome variables for Connecticut in Table 3 are defined as the first 21 months after the random assignment, a time span where nobody reaches the time limit.

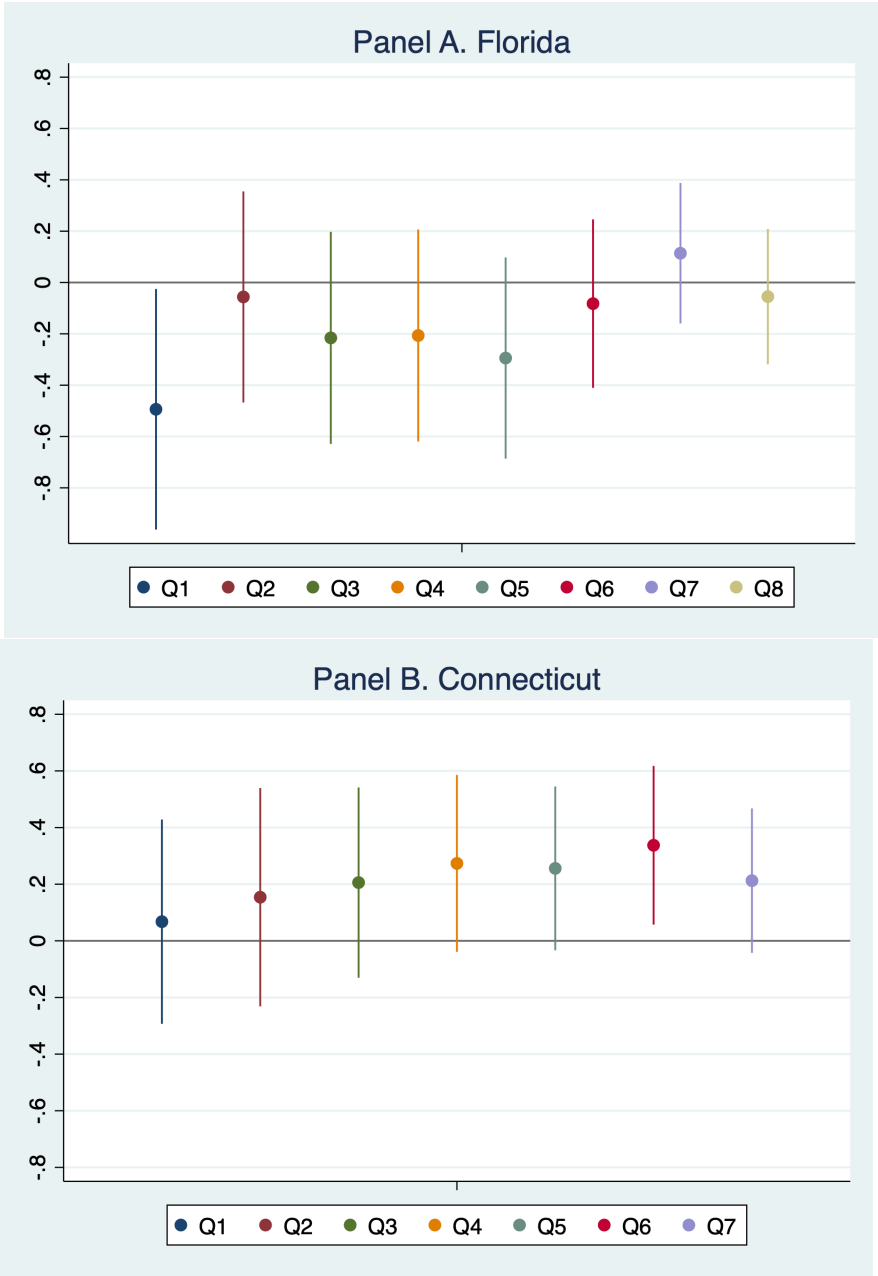

Figure A3: Effects of the time limit on combining AFDC and employment on a quarterly basis

Notes. Above two Panels plot the coefficients of the interaction term, β_{TL}, from Equation (1) using months of combining AFDC and employment at the quarterly level. Panel A contains the effects of the time limit during the first three quarters after randomization (Q1), during the month 4-6 (Q2), ..., and during the month 21-24 (Q8). Panel B contains the same estimates except for AFDC use during the month 21-24 (Q8) because the outcome variables for Connecticut in Table 3 are defined as the first 21 months after the random assignment, a time span where nobody reaches the time limit.

A5.2. Alternative Age Heterogeneity Analyses Assuming Quadratic Age Effects

In this Section, I present alternative estimation results of equation (2), assuming heterogeneous age effects. Specifically, I estimate the following equation:

$$\begin{aligned} y_i =& \delta_0 + \delta_{1,TL}\, E_i \cdot 1\{A_{i0} < \bar{A}\} \cdot (A_{i0} - \bar{A}) + \delta_{2,TL}\, E_i \cdot 1\{A_{i0} < \bar{A}\} \cdot (A_{i0} - \bar{A})^2 + \delta_E\, E_i \\ &+ \delta_A\, 1\{A_{i0} < \bar{A}\} + X_i\Delta + \epsilon_i. \end{aligned} \tag{1}$$

The estimation results of the coefficients for age effects ($\delta_{1,TL}, \delta_{2,TL}$) and the coefficient for the effect of waiver policies other than the time limit (δ_E) are presented in Appendix Table A4 and A5. The results in Appendix Table A4 show that the coefficients for the quadratic term, $\delta_{2,TL}$, are close to zero and statistically insignificant in Panel A and B. In Column (1), (3), and (6) of Panel C, however, the estimates of the coefficient for the quadratic term are statistically significant from zero, indicating the violation of the assumption of the linear age effect.

The results in Appendix Table A5 contain the effects of the time limit policy on other outcomes assuming quadratic age effects. Same as Table 5, it includes results on monthly Food Stamp use, total earnings, household income, and overall self-sufficiency during the first 24 months for the Florida sample and during the first 21 months for the Connecticut sample. The coefficients for the quadratic age effect term, $\delta_{2,TL}$, are negative for the Florida sample in all four outcomes. On the other hand, for the Connecticut sample, the age effects on Food Stamp use seem to follow a negative parabola. The positive (but insignificant) effect of TANF on monthly Food stamp use found in Table 5 seems to increase up to age 8.7 and then decrease after that.

Table A4: Age heterogeneity of the effects of welfare time limit on welfare use and employment (quadratic age effects)

	(1) AFDC	(2) Empl	(3) AFDC & Empl	(4) AFDC & No empl	(5) No AFDC & Empl	(6) No AFDC & No empl
Panel A. Florida						
Age effects ($\delta_{1,TL}$)	0.301	0.0265	0.0218	0.279	0.0579	-0.359
	(0.33)	(0.12)	(0.21)	(0.30)	(0.32)	(0.31)
Age effects squared ($\delta_{2,TL}$)	0.0141	0.00169	0.000117	0.0140	0.00497	-0.0191
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
E (δ_E)	1.300	0.318	1.193*	0.107	-0.238	-1.062
	(1.07)	(0.40)	(0.63)	(1.03)	(1.10)	(0.99)
Panel B. Connecticut (16+ as a control group)						
Age effects ($\delta_{1,TL}$)	-0.399	-0.0555	-0.380	-0.0189	0.214	0.185
	(0.25)	(0.09)	(0.24)	(0.26)	(0.20)	(0.21)
Age effects squared ($\delta_{2,TL}$)	-0.0251	-0.00288	-0.0202	-0.00485	0.0116	0.0135
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
E (δ_E)	0.809	0.414	2.336***	-1.527*	-1.096	0.286
	(0.83)	(0.30)	(0.67)	(0.83)	(0.69)	(0.75)
Panel C. Connecticut (17+ as a control group)						
Age effects ($\delta_{1,TL}$)	-0.640***	-0.127	-0.590**	-0.0506	0.210	0.430**
	(0.23)	(0.10)	(0.23)	(0.25)	(0.20)	(0.20)
Age effects squared ($\delta_{2,TL}$)	-0.0343**	-0.00735	-0.0305*	-0.00376	0.00846	0.0258**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
E (δ_E)	-0.516	0.185	1.369*	-1.886**	-0.814	1.330
	(0.88)	(0.36)	(0.74)	(0.91)	(0.81)	(0.84)

This table contains the estimation results of selective variables from Equation (3).

Standard errors are in parentheses.

A6. Methods for Calculating "Simulated" Monthly Benefit Receipt of the Connecticut Sample

The goal of this exercise is to construct the "simulated" monthly benefit use variable, which gives 1 if an individual among the Connecticut sample is likely to have ever received welfare during each quarter under the assumption that (1) she faced the Florida's EID rule and (2) no behavioral response regarding labor supply decision due to changes of the EID rule.

1. Calculating the possible number of monthly welfare use within each quarter

As information about employment and earned income is available at the quarterly level, I make assumptions to convert the quarterly earnings into monthly. Note that I cannot identify the number of children of families more than three for CT (four for FL). For families of 3+, I calculate the number of monthly benefit based on the rule that families of three faced.

1. Assume that each monthly earning is quarterly earning /3

First, I assume that each monthly earning is quarterly earning/3 and use that to determine monthly welfare eligibility.

Let's say e is the amount of monthly earning at and above which the welfare benefit becomes zero. Let's assume that a family's quarterly income is x.

(i) If $x < 3e$, then I consider the family received welfare all 3 months within the quarter.

(ii) If $x >= 3e$, then we can infer that the the family did not receive welfare at all.

2. Calculate the upper bound and lower bound of the number of the monthly welfare use

(i) If $x < e$, then we can infer that the family received welfare all 3 months within the quarter.

(ii) If $e <= x < 2e$, then the upper bound of the number of monthly welfare use is 3 (e.g., monthly incomes are $x/3, x/3, x/3$), and the lower bound is 2 (e.g., monthly incomes are $e, x-e, 0$).

(iii) If $2e <= x < 3e$, then the upper bound is 3 (e.g., $x/3, x/3, x/3$), and the lower bound

Table A5: Effects of the time limit policy on monthly Food Stamp use, total earnings, total family income, and overall self-sufficiency (quadratic age effects)

	(1) Food Stamp use	(2) Total earnings	(3) Total family income	(4) Self-sufficiency index
Panel A. Florida				
Age effects ($\delta_{1,TL}$)	0.287	-354.7	-240.6	-0.0356
	(0.33)	(437.60)	(434.12)	(0.04)
Age effects squared ($\delta_{2,TL}$)	0.00686	-23.18	-18.79	-0.00179
	(0.02)	(30.62)	(30.72)	(0.00)
E (δ_E)	1.366	-322.0	-245.3	-0.0804
	(1.04)	(1455.54)	(1405.01)	(0.14)
Panel B. Connecticut (16+ as a control group)				
Age effects ($\delta_{1,TL}$)	-0.570**	-390.5	-667.1**	-0.000793
	(0.24)	(334.41)	(336.78)	(0.03)
Age effects squared ($\delta_{2,TL}$)	-0.0393**	-26.78	-39.08	0.000156
	(0.02)	(25.16)	(25.20)	(0.00)
E (δ_E)	-0.636	-270.1	320.9	0.0375
	(0.80)	(1041.42)	(1044.19)	(0.10)

This table contains the estimation results of selective variables variables from Equation (3).
Standard errors are in parentheses.

Table A6: Benefit amount of a family of three under different earned income disregard rules

Monthly income	Benefit amount under the FL's rule	Benefit amount under the CT's rule	CT's "simulated" benefit amount under the FL's rule	CT's "simulated" benefit use under the FL's rule	FL's "simulated" benefit amount under the CT's rule	FL's "simulated" benefit use under the CT's rule
\$0	303	543	543	1	303	1
\$200	303	543	543	1	303	1
\$400	203	543	443	1	303	1
\$600	103	543	343	1	303	1
\$800	3	543	243	1	303	1
\$1,000	0	543	143	1	303	1
\$1,200*	0	0	43	1	0	0
\$1,400	0	0	0	0	0	0

Note. This table illustrates a amount of welfare benefit and welfare use under different earned income disregard rules. The fourth column contains the simulated benefit amount a family of three in Connecticut had received if they would have been faced the less generous earned income disregard rule of Florida. The sixth column contains the simulated benefit amount a family of three in Florida had received if they would have been faced the less generous earned income disregard rule of Connecticut.
*\$1,200 is an income that exceeds the federal poverty line in 1998, which is \$1,138.

is 1 $(e, e, x - 2e)$.

(iv) If $x >= 3e$, then the upper bound is 2 (e.g., $3e, 0, 0$), and the lower bound is 0 (e.g., $e, e, x - 2e$).

Using the average, upper bound, and lower bound counts of simulated welfare use at each quarter, I calculating the simulated number of monthly welfare use during first 21 months after the random assignment.

2. Making indicators for quarterly welfare use

I make two additional indicators for quarterly welfare use. The first indicator gives 1 if it is considered that a family received welfare for all three months based on the average, upper bound, and lower bound counts. The second indicator gives 1 if it is considered that a family received welfare at least one month based on the lower bound counts.[15]

[15]I did not use the average and upper bound counts of the number of monthly welfare use within each quarter to make the second indicator. The reason is that the average counts always give me 0 or 3 months of welfare use for each quarter, which makes the second indicator exactly same as the first indicator. Similarly, the upper bound counts always give me 2 or 3 months of welfare use for each quarter, identifying every quarter as using welfare at least once. I also did not use average counts because average counts always gives me 0 or 3 months of welfare use for each quarter, which makes indicator (1) and (2) exactly same.