Welfare, Workfare, and Labor Supply: A Unified Evaluation*

Francesco Agostinelli Emilio Borghesan

Giuseppe Sorrenti

June 2023

Abstract

We analyze the extent to which labor supply responds to incentives created by social programs in the United States. We find evidence that the incentive and disincentive effects of the EITC and welfare programs on hours worked among single mothers are more extensive than previously found in the literature. We also show that the difference-in-differences design, frequently adopted in the existing literature, fails to identify a meaningful treatment parameter in the context of the welfare-to-workfare transition in the 1990s. Finally, we use our quasi-experimental estimates to identify a structural model of labor supply with multiple tax and transfer programs. Model counterfactuals show that the EITC's effect on labor supply depends on the regime of taxes and welfare system in place.

JEL Classification: H30, I38, J08, J38.

*We thank Domenico Ferraro, Thomas Le Barbanchon, Simon ter Meulen, Aviv Nevo, Evan Rose, Jesse Rothstein, Dan Silverman, Petra Todd, and seminar participants at SOFI (Swedish Institute for Social Research), the Institute for Fiscal Studies, the Melbourne Institute (University of Melbourne), and the University of Bergamo for very useful comments. We are also deeply thankful to Jo Mullins for having shared with us the AFDC/TANF and SNAP benefit formulas.
Agostinelli: Department of Economics, University of Pennsylvania, 133 S 36th St, Philadelphia, PA 19104, USA (email: fagostin@sas.upenn.edu). Borghesan: Department of Economics, University of Pennsylvania, 133 S 36th St, Philadelphia, PA 19104, USA (email: borghesa@sas.upenn.edu). Sorrenti: Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, Amsterdam, 1018 WB, NL (email: g.sorrenti@uva.nl).

1 Introduction

In fiscal year 2022, the United States spent $665 billion, or 11% of the federal budget, on economic security programs (Center on Budget and Policy Priorities 2022). Historically, the largest of these programs have been traditional welfare and the Earned Income Tax Credit (EITC), both of which tie benefit receipt to an individual's income and labor market participation. Understanding the effects of each program in isolation, as well as how they combine to influence labor supply, is challenging as the programs generate conflicting incentives and their benefits vary according to an individual's endogenous labor supply decision and location within the income distribution. The major reforms to the EITC and welfare in the 1990s occurred simultaneously, further complicating analysis. An additional challenge has been the literature's focus on average effects and average elasticities, which, by combining the behavioral responses of individuals who face different incentives owing to their location in the benefit structure of each program, can obscure the effects of these programs on individual labor supply.

This paper aims to address these challenges and evaluate the impact of the EITC and traditional welfare on the labor supply of single mothers. It combines quasi-experimental and structural analyses to provide new evidence of the marginal effects of EITC and welfare benefits on the intensive and extensive margins of labor supply, and it decomposes changes in labor supply during the 1990s into the individual contributions of each program. Reforms to the EITC and traditional welfare during the 1990s transformed the social safety net from a welfare-oriented to a workfare-oriented regime. The EITC, a tax credit for parents who work and have low earnings, was progressively expanded to become the largest income support program in the US. Meanwhile a series of statewide reforms curtailed the generosity of welfare and culminated in the nationwide replacement of Aid for Families with Dependent Children (AFDC) with the stricter Temporary Aid for Needy Families (TANF).[1] By exploiting variation in benefit receipt across space, time, and the income distribution caused by these reforms, we document evidence of extensive and intensive margins labor supply responses that are consistent with standard models of labor supply.

[1] Throughout the paper, we use the term welfare to refer to both AFDC and TANF.

We begin the paper by introducing a simple theoretical model of labor supply, in which single mothers seek to maximize their utility over consumption and hours worked within a framework that incorporates social programs in the form of wage subsidies and unconditional cash transfers. This framework generates two important insights. First, the effectiveness of a program, such as a wage subsidy, in promoting work is contingent upon the income and substitution effects generated by the specific tax regime in place. Second, when estimating the overall labor supply elasticity resulting from a particular reform, it is crucial to differentiate the structural (Frisch) labor supply elasticity from the effects of the tax and transfer environment. This is because aggregate labor supply responses to counterfactual policies are not simple extrapolations of past policies, but are determined jointly by a complex interplay of behavioral parameters and the tax and transfer system.

Our theoretical framework motivates our quasi-experimental evaluation of the EITC and welfare programs. To conduct this analysis, we use data from the March Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS) for the years 1988-2002, a period spanning the major reforms to welfare and the EITC during the 1990s. We examine single mothers as they are particularly exposed to these programs, and we disentangle the separate effects of welfare and the EITC for these individuals. Our analysis reveals that single mothers are highly responsive to both programs. For each $1,000 increase in EITC benefits, we estimate that labor supply increases by about 115 hours worked per year, an eight percent change relative to the pre-reform period. At the same time, $1,000 reduction in the generosity of welfare causes a statistically significant increase of 17 hours per year, a one percent change.[2] These estimates are robust to the inclusion of controls for individual and household characteristics, state fixed effects, state unemployment levels, and the introduction of state specific welfare reforms.

Our empirical analysis contributes to the literature on labor supply responses to social programs in several important dimensions. First, unlike many prior stud-

[2]Rates of EITC take-up in the period we study are high, over 80 percent (Scholz 1994; Blumenthal, Erard, and Ho 2005). Welfare takeup is typically lower, so our empirical estimates should be interpreted as lower bounds on the effects of welfare on employment.

ies of the welfare-to-workfare transition, we use longitudinal data that enable us to eliminate the effect of individual unobserved time-invariant heterogeneity by estimating the model in first differences. Second, we estimate marginal labor supply responses rather than average effects of policy changes. The focus on marginal effects allows us to overcome concerns about averaging heterogeneous effects across individuals who face different effective marginal tax rates across the length of the benefit schedule. By focusing on marginal effects, we are able to find evidence of intensive margin responses to the EITC that are consistent with standard static models of labor supply. A prediction from a standard labor supply model is that the high marginal tax rates created by the phase-out of means-tested programs should have disincentive effects on hours worked. However, this prediction has hitherto found little support in the data (Meyer 2002, Eissa and Hoynes 2006). Our analysis provides direct evidence that mothers reduce hours worked in response to the high marginal tax rates in the phase-out of the EITC schedule.

A third contribution is that we highlight several shortcomings of applying difference-in-differences in this context. When a new program is introduced, DiD identifies a weighted average of traditional treatment parameters corresponding to each level of the treatment, but we show that when a program is merely expanded—as the EITC was in the 1990s—DiD does not generally identify a treatment parameter corresponding to the causal effect of the program, unless strong time-invariant restrictions on the primitives of the behavioral model are imposed. Additionally, the DiD estimand does not provide insights into the individual's marginal behavioral responses to alterations in benefit levels, which are crucial for estimating the elasticity of labor supply. This lack of consideration might explain why the EITC's intensive margin effects have been hitherto overlooked.

In the second part of the paper, we investigate the importance of modeling the interdependence of the tax code and the labor supply incentives created by different social programs when predicting the effects of future fiscal reforms on labor supply. To this end, we specify and estimate a static labor supply model via indirect inference using our empirical quasi-experimental estimates. We use the estimated model to decompose the effects the effect of reforms to the EITC and

welfare in isolation during the 1990s and to analyze the labor supply responses of replacing—in a fiscally neutral way—the current EITC and TANF programs with Universal Basic Income (UBI), a policy that has recently gained currency in the political debate (The Economist 2020).[3]

Similar to early work using structural models of labor supply by Keane (1995) and Keane and Moffitt (1998), we find that employment among single mothers in 1996 would have been 7.5 (5.5) percentage points lower had the EITC (welfare) not been reformed. Second, our model suggests that UBI's effects on labor supply depend on which program is eliminated: Employment and hours worked decrease if UBI replaces the EITC but increase if UBI replaces welfare. These aggregate effects hide vast heterogeneity in individual responses, with hours worked responding nonmonotonically depending on whether the individual was previously benefiting from the now-eliminated social program. Finally, we use the model to show how the aggregate elasticity of labor supply varies with the underlying tax and transfer system. Depending on the parameters of the tax code, the same reform to the EITC may cause anywhere between a 2 percentage point (pp) and 10 pp change in the employment of single mothers.[4] Overall, our results highlight the importance of modeling the entire tax and transfer system when providing quantitative analysis to inform and advise policymakers.

Relationship to Literature. An extensive literature evaluates the effects of tax and transfer regimes on female labor supply, hence we limit our review to those works most closely related to ours. The literature on the effects of the EITC on labor supply is summarized in several excellent reviews by Hotz and Scholz (2003), Nichols and Rothstein (2016), Hoynes and Rothstein (2017), and Albanesi, Olivetti, and Petrongolo (2023). For instance, Hotz and Scholz (2003) discuss the

[3]To the extent that factors like saving, human capital accumulation and fertility induce dynamic effects of the EITC and welfare on labor supply, our model predictions should be interpreted as illustrative of how social program reforms interact with the rest of the tax and transfer system within a context of optimal decision-making.

[4]In addition, the simulated extensive margin elasticity with respect to labor income taxes from the estimated model (0.3) is systematically smaller than the structural labor supply elasticity from which the data is generated (unitary Frisch elasticity). This result reconciles the different conclusions of the nonstructural and structural literatures on the magnitude of the labor supply elasticity (see Keane and Rogerson 2012 and Keane and Rogerson 2015 on the discrepancy between micro and macro elasticities).

importance of complementing reduced-form analyses with structural approaches that parameterize individual preferences and constraints in a theory of optimal decision-making, as we do in this paper. Moffitt (1990), Keane and Moffitt (1998), and Keane (1995) build on this idea to analyze a wide range of policy reforms and expansions of the EITC. According to these studies, the EITC expansions between 1984 and 1996 considerably increased labor force participation, especially for the group of single mothers. Mancino and Mullins (2020) find that EITC expansions generate positive responses for workers transitioning into employment, transitioning to new jobs, and accepting second jobs. Blundell et al. (2016), building on life-cycle models of female labor supply in, e.g., Heckman and Macurdy (1980), Eckstein and Wolpin (1989) and Keane and Wolpin (1997), show that the expansion of tax credits similar to the EITC in the UK increases single mothers' labor supply and marginally reduces educational attainment. We contribute to this strand of literature by providing quasi-experimental evidence that is consistent with the predictions of this class of models and by showing how the same structural elasticity of labor supply can generate markedly different behavioral responses to the EITC depending on the level and progressivity of the tax and transfer system.

Many empirical works evaluating the labor supply effects of the EITC implement various DiD designs that compare a control group of single women—who were ineligible for EITC benefits prior to 1994—to a treatment group of single mothers. For instance, Eissa and Liebman (1996) study the impact of the 1986 EITC reform via a DiD analysis and find that the reform increased labor force participation for single mothers. With a similar empirical strategy, Meyer and Rosenbaum (2001) implement a decomposition analysis and find that a large share of the increase in employment by single mothers between 1984 and 1996 can be attributed to the EITC, with small effects due to welfare reform.[5] Kleven (2020) argues the opposite. He augments an event study of the 1993 EITC reform with controls

[5]These findings align with the results for a sample of California residents in Hotz, Mullin, and Scholz (2006), with studies based on longitudinal data such as Gelber and Mitchell (2012), and with studies exploiting event study setups around the largest EITC reforms such as Hoynes and Patel (2018). Dickert, Houser, and Scholz (1995), Meyer (2002), and Eissa, Kleven, and Kreiner (2008) also find EITC-induced increases in maternal labor force participation and employment with the effect mainly driven by the group of single mothers.

for state-specific welfare reforms and unemployment rates and shows that this shrinks the effect of the EITC on the observed rise in employment in the 1990s to zero.[6] We contribute to this strand of literature in a dual way. First, we show the limitations of DiD and event study designs used in this literature to evaluate multiple programs that are continuously reformed over time, as in the 1990s welfare-to-workfare transition. Second, we provide new evidence of marginal responses of labor supply to both the EITC and welfare using quasi-experimental variation that exploits the reform of these programs across time and space.

Papers that find evidence of intensive margin responses to the EITC are relatively few and use theoretical predictions of bunching at EITC kink points rather than observed labor supply to infer the magnitude of intensive margin responses. Saez (2010) finds bunching only among self-employed taxpayers and only at the first EITC kink, suggesting a manipulation of reported earnings, but no intensive margin labor supply response. Chetty, Friedman, and Saez (2013) exploit variation in the extent of bunching along the kinks in the EITC schedule to estimate intensive margin responses that are similar in magnitude to extensive margin responses. Chetty and Saez (2013) find, however, that a randomized experiment providing personalized information about the EITC had at best marginal effects on the intensive margin of labor supply. Mortenson and Whitten (2020) argue that bunching by wage earners is driven entirely by income misreporting and not by a labor supply response.

A final strand of literature of interest for this study concerns the effects of welfare. Studies on how welfare and welfare reforms affect work incentives, welfare dependency, family structure, and migration are reviewed in Moffitt (1992), Blank (2002), and Grogger and Karoly (2005). The effects of welfare reforms on labor supply are analyzed using a dynamic discrete choice model in Chan (2013) and using a marginal treatment effects framework in Moffitt (2019). Grogger (2003) analyzes how welfare and the EITC jointly affect labor supply in a regression framework. Low et al. (2018) show, using a lifecycle model, that the 1996 welfare reform in the US raised employment and reduced divorce among single mothers who entered the labor force in anticipation of the loss of benefits. Ashenfelter

[6] Schanzenbach and Strain (2021) argue that Kleven's findings are not robust to the choice of labor supply variable or the exclusion of business cycle controls.

(1983), Kline and Tartari (2016), and Bitler, Gelbach, and Hoynes (2006) find intensive and extensive margin responses to welfare that are consistent with standard labor supply models. We document the same for the EITC.

Description of Programs. We now provide a brief description of the two cash transfer programs. We refer the reader to Appendix A for a more detailed description of the changes these programs underwent during the period we study.

To receive the EITC, a recipient must have a dependent child, positive earnings, and adjusted gross income below a threshold that varies with the year and number of dependent children.[7] The schedule of benefits depends on pre-tax income and features three parts: A phase-in where earned income receives a proportional subsidy, a plateau where benefits are neither increased nor reduced, and a phase-out where benefits are withdrawn. The incentives for recipients differ depending on their position in the schedule. Standard models of labor supply predict that the substitution effect created by the phase-in would raise labor supply, while individuals situated on the plateau and phase-out would likely work less. The EITC's schedule of benefits was expanded continually throughout the 1990s, although one of the largest expansions was passed in 1993 and involved yearly increases in the benefit schedule for every year between 1994 and 1996.

In contrast to the EITC, traditional welfare has historically provided benefits to parents who do not work. Concerns that welfare disincentivized labor supply caused states to implement a variety of reforms, primarily between 1994 and 1996. These welfare waiver reforms contained a mix of incentives designed to encourage employment and reduce the number of families receiving benefits. Many characteristics of the waiver reforms were eventually adopted nationwide when TANF replaced AFDC in 1996. Throughout this paper, we refer to the welfare-to-workfare transition as the set of welfare reforms and EITC expansions that were implemented in the United States between 1993 and 1996.

[7] Starting in 1994, individuals without dependent children but satisfying the other criteria were eligible for a small tax credit through the EITC.

2 Identifying Labor Supply Elasticities

To guide our analysis, we start by presenting a simple labor supply model. Single mothers have the following preferences:

$$u_i(c_{i,t}, h_{i,t}) = log(c_{i,t}) - \frac{h_{i,t}^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} \ , \tag{1}$$

where $c_{i,t}$ and $h_{i,t}$ represent private consumption and hours worked and γ represents the Frisch elasticity of labor supply.

In this model, we consider a wage subsidy (ξ), as well as an unconditional cash transfer (T). Agents maximize their utility subject to time and budget constraints:

$$\begin{aligned} &\max_{c_{i,t}, h_{i,t}} \ u_i(c_{i,t}, h_{i,t}) \\ &\quad \text{s.t. } c_{i,t} = h_{i,t} \cdot (\omega_{i,t} + \xi) + T \ , \\ &\qquad h_{i,t} \geq 0 \ , \ c_{i,t} \geq 0 \ . \end{aligned} \tag{2}$$

The wage subsidy may be a fixed fraction of the wage, $\xi_{i,t} = \eta \cdot \omega_{i,t}$ for example, but the important feature for the analysis that follows is that, in contrast to the transfer, the value of the wage subsidy varies with labor supply.

The solution to the problem in (2) is characterized by

$$h^*_{i,t} = \begin{cases} \tilde{h}_{i,t} & \text{if } \log\left(\tilde{h}_{i,t}(\omega + \xi) + T\right) - \frac{\tilde{h}_{i,t}^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} > \log T \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $\tilde{h}_{i,t}$ satisfies the following implicit function:

$$\tilde{h}_{i,t}^{\frac{1}{\gamma}} = \frac{\omega_{i,t} + \xi}{\tilde{h}_{i,t} \cdot (\omega_{i,t} + \xi) + T} \ . \tag{4}$$

An individual's labor supply is a function of the received wage offer ($\omega_{i,t}$), the tax regime in place (ξ, T), and the labor supply elasticity (γ).

In this context, we define an individual's behavioral response to a marginal

change in the wage subsidy (e.g., $\xi' > \xi$), as the counterfactual change in the optimal labor supply induced by the reform: $\Delta h^*_{i,t}(\xi, \xi', T; \omega_{i,t}, \gamma) \equiv h^*_{i,t}(\xi', T; \omega_{i,t}, \gamma) - h^*_{i,t}(\xi, T; \omega_{i,t}, \gamma)$. A large literature has adopted various estimators in an effort to shed light on $\Delta h^*_{i,t}(\xi, \xi', T; \omega_{i,t}, \gamma)$. These estimators are typically functions of labor supply and parameters of the tax regime (or in the case of DiD, the timing of tax reforms), $\beta(T, \xi, \xi', h^*_{i,t}; \gamma)$. Below, we highlight that the extent to which an estimator is informative about the labor supply elasticity of individuals depends on two main factors:

1. The external validity of the same program reform across contexts depends on how alternative tax and transfer programs interact in determining labor supply incentives: $\beta(T, \xi, \xi', h^*_{i,t}; \gamma) \neq \beta(T', \xi, \xi', h^*_{i,t}; \gamma)$ for $T' \neq T$.

2. Identifying parameters (γ, for example) that govern individual behavior within the context of a labor supply model is necessary to forecast the aggregate labor supply elasticities generated by new reforms.

The first point is particularly apparent when comparing periods with positive and zero transfers. When $T = 0$, ξ affects employment only through the extensive margin, while when $T > 0$, ξ has effects on both the intensive and extensive margins of labor supply. This means that the effectiveness of a program like a wage subsidy in promoting work varies with the income and substitution effects generated by the context-specific tax regime in place. Researchers may therefore estimate seemingly different labor supply responses to the same policy over time, as other aspects of the tax regime may have changed during that period of time. The fact that some estimates of labor supply responses to a particular program have reduced over time do not necessarily suggest that individuals no longer respond to that program anymore, as other features of the tax system and wage offer distribution may have changed. Similarly, Attanasio et al. (2018) find that the empirically estimated aggregate elasticity is not a structural parameter, as it varies over the business cycle because of the heterogeneity of marginal individuals in different aggregate states of the economy.

Our second point follows logically from the first point. If the aggregate labor supply elasticity estimated from a particular reform is an amalgam of the structural

(Frisch) labor supply elasticity and the effects of the tax and transfer environment, then separately identifying the Frisch elasticity and carefully modeling a proposed change to the tax and transfer system is necessary to forecast the counterfactual responses of individuals to various proposed tax programs. We turn to this task in Section 4 after providing descriptive evidence in the effects of ξ and T on the labor supply of single women in the next section.

3 Evaluation of the EITC and Welfare Reforms

There are many possible estimators, $\beta(T, \xi, \xi', h^*_{i,t}; \gamma)$, for capturing behavioral responses to tax credits. In this analysis, we begin with a simple linear estimator. The goal of this exercise is to identify the average marginal behavioral response of labor supply to the benefits provided by each program within our specified sample and period of analysis. Subsequently, we combine these estimates using indirect inference to identify a model to learn about the distribution of behavioral responses and how they change when tax and transfer programs are reformed. Consider the following equation for the labor supply of individual i at time t:

$$Y_{i,t} = \beta_0 + \gamma_0 t + \gamma_1 \xi_{i,t} + \gamma_2 T_{i,t} + \alpha_i + \epsilon_{i,t} \ , \tag{5}$$

where $Y_{i,t}$ is a measure of labor supply, $\xi_{i,t}$ is the value of EITC benefits, $T_{i,t}$ is the value of welfare benefits, α_i represents an unobserved individual-specific and time-invariant preference for work, and $\epsilon_{i,t}$ represents additional unobserved heterogeneity. We allow for a possible time trend, denoted by t. The marginal effect of EITC benefits is given by γ_1, while the marginal effect of welfare benefits is γ_2.

Before estimating (5), we difference the equation to eliminate each individual's unobserved preference component, α_i, yielding

$$\Delta Y_{i,t} = \gamma_0 + \gamma_1 \Delta \xi_{i,t} + \gamma_2 \Delta T_{i,t} + \Delta \epsilon_{i,t} \ , \tag{6}$$

where $\Delta \xi_{i,t}$ and $\Delta T_{i,t}$ are policy-induced longitudinal changes in EITC and welfare benefits. Following this transformation, the average marginal effects of EITC and welfare benefits are identified under the following standard exogeneity assumption.

Assumption 1. *The longitudinal change in the unobserved heterogeneity of labor supply among individuals is mean independent of the policy-induced longitudinal changes in EITC and welfare benefits:* $E[\Delta\epsilon_{i,t} \mid \Delta\xi_{i,t}, \Delta T_{i,t}] = E[\Delta\epsilon_{i,t}] = 0$.

Assumption 1 is similar to the standard parallel trend assumption in DiD, although here the unobserved determinants of labor supply, $\Delta\varepsilon_{i,t}$, must be mean independent of the intensity of the treatment ($\Delta\xi_{i,t}, \Delta T_{i,t}$) rather than whether an individual is in the treated group. Section 3.1 describes the way we construct $\Delta\xi_{i,t}$ and $\Delta T_{i,t}$ so that Assumption 1 is likely to be satisfied.

Under Assumption 1, Ordinary Least Squares (OLS) identifies γ_1 and γ_2. If labor supply is not truly linear in EITC and welfare benefits, then OLS identifies the *average* marginal effect of EITC benefits (welfare benefits) on hours worked, *ceteris paribus* for the level of welfare benefits (EITC benefits):

$$\gamma_1 = \frac{\partial E[\,Y_{i,t} \mid \xi, T]}{\partial \xi} \quad \text{and} \quad \gamma_2 = \frac{\partial E[\,Y_{i,t} \mid \xi, T]}{\partial T}\ .$$

We assess the linearity assumption in Section 3.2 and estimate a model allowing for nonlinear responses in Section 4.

3.1 Data

We use the March Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS). The CPS is a monthly US household survey that asks each member of the household detailed questions related to labor force participation, earnings, and demographic characteristics. Households selected into the CPS are surveyed eight times over a period of 16 months.

We use the ASEC supplement for two reasons. As the survey occurs once per year, we can create a panel with two observations per individual. Moreover, the ASEC measures labor force participation and income in the prior calendar year while the CPS monthly files measure variables in the prior week or prior twelve months. Since the EITC is calculated on the basis of annual earnings and our analysis relies on a precise continuous measure for policy-induced changes in EITC benefits over time, we view the ASEC as more appropriate for our analysis.

Our sample consists of single mothers between the ages of 25 and 50 who are present in two consecutive years of the ASEC supplement. We look at single mothers for three reasons: (i) women with children are a main target of programs such as the EITC and welfare; (ii) single mothers are deemed one of the most responsive groups to tax and welfare reforms; and (iii) it simplifies the analysis by abstracting from potential interaction effects with partners' behavior.

We create longitudinal linkages for single mothers in the ASEC between 1988 and 2002, resulting in a sample of 10,959 unique mothers. Each single mother is observed one time per year for two consecutive years. For each mother, we construct year-on-year changes in total yearly hours worked. We hold family structure constant: Women who have children between the first and second years of the survey are excluded from the analysis.

In the spirit of Dahl and Lochner (2012), we construct exogenous policy-induced changes in EITC and welfare benefits, $\Delta\xi_{i,t}$ and $\Delta T_{i,t}$, for each individual in the sample. This involves predicting the counterfactual level of earnings ($E_{i,t}$) and nonlabor income ($NL_{i,t}$) that would have prevailed in the second period (t) of the sample in the absence of any labor supply response to the EITC and welfare reforms. $\widehat{E}_{i,t}$ (respectively $\widehat{NL}_{i,t}$) is the predicted value from a regression of $E_{i,t}$ ($NL_{i,t}$) on a fifth-order polynomial in its lag and an indicator for positive lagged values. This model provides reasonable predictions of second period income with an R^2 of 0.58. Nevertheless, in the next section we show that our results are robust to a wide range of prediction models.

After predicting earnings and nonlabor income, $\widehat{E}_{i,t}$ and $\widehat{NL}_{i,t}$ then serve as inputs in the computation of second-period welfare and EITC benefits, ($\widehat{\xi}_{i,t}$,$\widehat{T}_{i,t}$), as follows:

$$\widehat{\xi}_{i,t} = EITC_{i,t}(\widehat{E}_{i,t}) \text{ and } \widehat{T}_{i,t} = T_{i,t}^{Welfare}(\widehat{E}_{i,t}, \widehat{NL}_{i,t}),$$

where $EITC_{i,t}(\cdot)$ and $T_{i,t}^{Welfare}(\cdot,\cdot)$ are functions that calculate EITC and AFCD/TANF benefits in year t based on program rules and individual characteristics (number of dependent children and state of residence). Our measures of

exogenous policy-induced changes in benefits are then given by:

$$\Delta\xi_{i,t} = \widehat{\xi}_{i,t} - \xi_{i,t-1} \text{ and } \Delta T_{i,t} = \widehat{T}_{i,t} - T_{i,t-1} \ .$$

EITC benefits are calculated using NBER's TAXSIM and include state as well as federal credits, while welfare benefits are computed from the AFDC/TANF rules in place for each year and state. Appendix E provides further details about the construction of these variables. Descriptive statistics for our sample are presented in Appendix Table D-1.[8]

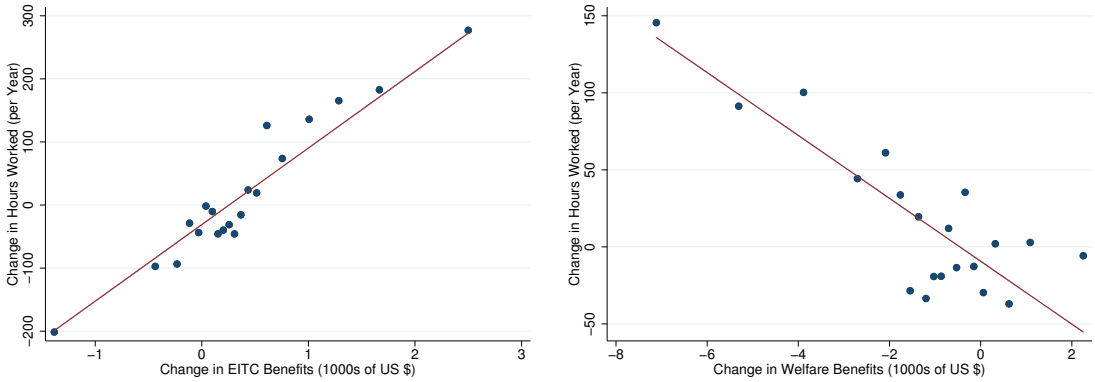
3.2 Empirical Results

In this section we provide estimates of the marginal effects of both the EITC and welfare on single mother's labor supply. We also provide evidence of an approximately linear relationship between labor supply and benefits, present new evidence of intensive margin effects for the EITC that are consistent with labor supply theory, and conduct several robustness tests. We conclude the section by using our estimates of marginal effects to provide an estimate of ATT for each program that is valid under an assumption of linearity.

Figure 1 depicts the binscatter of the nonparametric relationship between the longitudinal change in hours worked by single mothers and the variables $\Delta\xi_{i,t}$ and $\Delta T_{i,t}$. All specifications include controls for state unemployment level, state and year fixed effects, state welfare waivers, and indicator variables for race and the number of dependent children. The analysis of the effect of EITC changes also controls for changes in welfare benefits, and vice versa. The dependent variable, total hours worked per year, is constructed by multiplying the total weeks worked per year by the variable denoting "usual hours worked per week."

[8] Alternatively, Bastian and Lochner (2020) use the variation in the maximum EITC benefits by number of children, states, and years as their measure of EITC expansion during the 2003-2018 period. In our period of analysis, the EITC reforms were mostly at the federal level, with little to no heterogeneity at the state level. For this reason, the variance that is explained by the between-state differences in the maximum EITC benefits—the type of variation that ideally we would want to use—represents only the 3.34 percent of the total variance. This method is therefore not suitable for our analysis as this measure would end up exploiting differences in maximum EITC benefits among mothers with different number of children. Such an approach would confound changes in labor supply due to the EITC with the empirical observation that mothers with more children generally work fewer hours.

Figure 1: Hours Worked, EITC Benefits, and Welfare Transfers



(a) Hours and EITC

(b) Hours and Welfare

The figure shows the relation between policy-induced changes in EITC benefits (Panel (a)) and welfare benefits (Panel (b)) on the change in yearly hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Each panel depicts the binscatter of the nonparametric relation and linear fit line between the y-residuals on x-residuals with specifications containing control variables for mother's race, number of dependent children (indicators), year fixed effects, state fixed effects, state unemployment level and state welfare waivers (indicator).

Two results are worth highlighting. First, Figure 1 suggests a positive (negative) relation between workfare (welfare) policies and single mothers' labor supply. Expansions in EITC benefits induce increases in hours worked by single mothers (Figure 1-a), while reductions in welfare benefits generate increasing levels of labor supply (Figure 1-b). Second, the relations between EITC and welfare benefits and hours worked are well-approximated by a linear specification.

Table 1 presents estimates of equation (6). In column (1), we estimate a model only including longitudinal changes in EITC benefits. In column (2), we focus on changes in welfare benefits in isolation. Column (3) includes both changes in EITC and welfare benefits. Column (4) augments the model in (3) with controls for state and year fixed effects, and includes indicator variables for race and the number of dependent children. Column (5) adds controls for state unemployment rates and an indicator variable for state welfare waivers.

According to the specification in column (1) of Table 1, a $1,000 increase in EITC

Table 1: EITC Benefits, Welfare Benefits, and Hours Worked per Year

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Yearly Hours Worked				
Change in EITC Benefits ($ 1000s)	125.45***		109.96***	115.38***	115.04***
	(7.28)		(8.56)	(8.89)	(8.91)
Change in Welfare Benefits ($ 1000s)		-31.62***	-12.08***	-17.01***	-17.16***
		(2.15)	(2.57)	(2.85)	(2.87)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	No	Yes	Yes
Unemployment and Waiver Controls	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly hours worked by single mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

benefits induces a statistically significant increase of approximately 125 hours worked per year. The effect size is slightly lower, about 115 additional hours per year, and remains positive and highly significant in the specifications with additional controls in columns (2) to (5). Welfare benefits have the opposite effect. A $1,000 decrease in welfare benefits induces a statistically significant increase of about 32 hours worked per year in the specification with welfare benefits in isolation and by about 12-17 hours in more saturated specifications.

The analysis of hours worked highlights two important insights. First, it shows that workfare-oriented policy regimes cause a positive labor supply response among single mothers, while welfare-oriented regimes have the opposite effect. This means that the combination of an increase in the EITC and a reduction in welfare—namely, a welfare-to-workfare transition—causes an unambiguous increase in the aggregate labor supply of single mothers. Second, estimates for the effect of changes in both EITC and welfare benefits are robust to the inclusion

of a wide range of controls for individual characteristics, such as race, number of children, and macroeconomic factors like state unemployment levels and the presence of state welfare waivers.[9]

We can use the point estimates of marginal effects from column (5) of Table 1 to construct a back-of-the-envelope estimate of ATT for each policy. When there is no heterogeneity in labor supply and when marginal effects are constant, ATT is simply the product of the marginal effect and the average EITC benefit rescaled by the proportion of treated individuals. Averaged over all of the years in our sample, this back-of-the-envelope calculation yields an ATT of 159 hours per year for the EITC, while for welfare it is -121 hours per year. These figures are 11 percent and -8.5 percent of total labor supply during the period under consideration. In computing ATT, the much smaller marginal effects for welfare relative to the EITC are offset by the larger average benefit level among single mothers during the sample period. The existence of heterogeneity in labor supply for a fixed benefit profile as well as a nonlinear labor supply function would affect these calculations. We allow for these departures in subsequent analysis.

Threats to Identification. Although our analysis exploits policy-induced shifts in EITC and welfare benefit schedules to provide quasi-experimental variation in benefit levels, there may be lingering concerns about endogeneity caused by the means-tested nature of the two programs. Since income determines eligibility according to program schedules, it is possible that trends in income growth and labor supply between periods $t-1$ and t may be correlated, thereby inducing a mechanical correlation between our imputed changes in benefits and the unobserved term in the regression. To tackle this possible concern, we augment our main specification with a set of controls for the initial position of each mother in the EITC and welfare schedule. The inclusion of these controls allows for cor-

[9] Appendix Tables D-2, D-3, D-4, and D-5 explore the sensitivity of these results to alternative income prediction models. The models are, respectively, (i) a fourth-order polynomial and an indicator for lagged income, (ii) a fifth order polynomial and an indicator for lagged income plus controls for race and number of children, (iii) the same as (ii) with controls included for mother's education, and (iv) the prediction model in the main text estimated on the sample of single women. As single women were minimally affected by the EITC and welfare reforms in the 1990s, their income prediction model may be better able to forecast the income growth that would prevail among single mothers in the absence of any reforms. All results remain similar. Appendix Table D-6 extends the analysis to weeks worked.

Table 2: Robustness Tests: Residual Endogeneity and Mean-Reversion

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: Change in Yearly Hours Worked					
	Controls for Prior (t-1) Income		Earnings (t-1)<$30,000		Earnings (t-1)<$20,000	
Change in EITC Benefits ($ 1000s)	110.24***	109.90***	117.27***	116.88***	117.47***	117.03***
	(9.49)	(9.51)	(9.06)	(9.09)	(9.49)	(9.52)
Change in Welfare Benefits ($ 1000s)	-16.69***	-16.86***	-17.90***	-18.06***	-20.19***	-20.40***
	(3.06)	(3.08)	(2.94)	(2.96)	(3.16)	(3.19)
N	10959	10959	9456	9456	7611	7611
Controls and State F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Unemployment and Waiver Controls	No	Yes	No	Yes	No	Yes

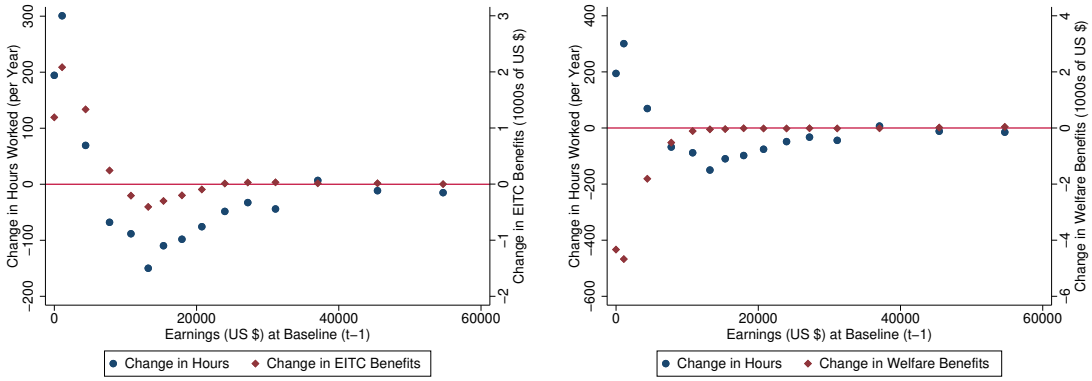
The table replicates baseline estimates by augmenting the set of control variables with lagged earned and unearned income and by restricting the sample to low-income mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. The specification in columns (1) and (2) also includes a set of controls for lagged labor income, lagged business income, lagged farm income, and lagged nonlabor income. The sample in columns (3) and (4) is restricted to single mothers with lagged earnings below $30,000 and $20,000, respectively. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

relation between income growth and unobserved heterogeneity as long as this relationship is constant over time.

Table 2 reports the results for this augmented specification in columns (1) and (2). The augmented specification leaves all the results remarkably similar to baseline estimates. This similarity reassures that our main results are unlikely to be biased by potential correlation between income growth and unobserved determinants of labor supply.

As a second test, we analyze the possible threat of *mean-reversion* in hours worked discussed in Gruber and Saez (2002). Transitory changes in hours worked for relatively high-income mothers who are unexposed to the reforms can affect our estimates independently on any behavioral responses to the EITC and welfare reforms. We address this potential concern by replicating our baseline analysis on the subsample of low-income mothers who are exposed to the policy reforms. We first analyze the subsample of single mothers who earned at most $30,000 in the previous year, a threshold that (roughly) identifies the EITC earning eligibility criterion in the period of analysis. We then use a more strin-

Figure 2: Distribution of Changes in Hours Worked, EITC and Welfare



(a) Hours Worked and EITC

(b) Hours Worked and Welfare

The figure shows the relation between policy-induced changes in EITC benefits or welfare benefits, change in yearly hours worked by single mothers, and earnings at baseline ($t-1$). Panel (a) illustrates the analysis of policy-induced changes in EITC benefits. Panel (b) illustrates the analysis of policy-induced changes in welfare benefits. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Each panel depicts the binscatter of the nonparametric relation between the year-on-year change in hours worked by single mothers (left y-axis and blue dots) and earnings at baseline and the relation between the policy-induced change in EITC (Panel (a)) or welfare (Panel (b)) benefits (right y-axis and red diamonds) and earnings at baseline.

gent threshold of $20,000 as a robustness check.

Columns (3) through (6) in Table 2 show the results of this analysis limited to low-income mothers. Despite a natural reduction in sample size, the estimated marginal effects remain almost unchanged from the baseline analysis, meaning that mean-reversion in labor supply is unlikely to affect our results.

3.3 Heterogeneity in Policy Exposure and Incentives

In this section we investigate evidence of heterogeneous responses to changes in the EITC and welfare. Figure 2 presents the relationship between family earnings, policy-induced changes in EITC and welfare benefits, and changes in yearly hours worked by single mothers. Figure 2-a shows that policy-induced changes in EITC benefits are large and positive for the poorest single mothers, negative

for mothers earning between \$12,000 and \$22,000, and zero for mothers with still higher earnings. Negative changes in EITC benefits are a consequence of the increase in marginal tax rates associated with the reformed EITC phase-out (higher negative slope) and the imputed (positive) trend in income. Many mothers initially located in the phase-out lose their benefits as a result of income growth. The figure also shows that mothers with the lowest earnings experience the greatest increases in hours worked, up to 300 additional hours per year. Mothers who are located on the plateau or phase-out of the EITC schedule, with earnings between \$10,000 and \$30,000, tend to reduce their hours worked as the EITC is expanded. The number \$10,000 is not coincidental: It is at or beyond the end of the phase-in of the EITC schedule for single mothers with two children in every year that we study.

Figure 2-b depicts the relationship between changes in welfare benefits and changes in labor supply.[10] There is a negative relationship between changes in hours worked and welfare benefits only among working mothers with low income (below \$10,000), suggesting that most of the identifying variation for the effect of welfare on hours comes from the poorest single mothers, an observation consistent with welfare's eligibility rules.

We test for evidence of heterogeneous incentives caused by the EITC program by replicating our analysis on two different subgroups of single mothers. The first subgroup consists of mothers facing positive changes in EITC benefits (see Figure 2-a) because of their location at time $t-1$ in the phase-in of the EITC schedule (i.e. with earnings below \$10,000). The second subgroup of mothers are those located at the plateau or at the phase-out of the EITC program, with earnings in period $t-1$ between \$10,000 and \$30,000. These mothers instead either face a stronger income effect (on the plateau) or a higher effective marginal tax rate (on the phase-out), both of which provide incentives to reduce labor supply. These mothers experience a decrease in the EITC, as shown in Figure 2-a.

This regression analysis reinforces the graphical evidence in Figure 2-a. Columns (1) and (2) of Table 3 display a positive marginal effect for single mothers in the

[10] Fang and Keane (2004), among others, highlight the existence of heterogeneous effects of welfare by showing that about one-quarter of welfare leavers did not start working.

Table 3: Heterogeneous Responses to the EITC Expansion

	(1)	(2)	(3)	(4)
	Outcome: Change in Yearly Hours Worked			
	Phase-In		Plateau or Phase-Out	
	Earnings (t-1)<$10,000		Earnings (t-1)∈($10,000, $30,000)	
Change in EITC Benefits ($ 1000s)	91.68***	90.74***	130.10***	130.36***
	(13.71)	(13.73)	(19.91)	(19.93)
N	4752	4752	4704	4704
Mean Change in Hours Worked	130.25	130.25	-84.05	-84.05
Mean Change in EITC Benefits ($ 1000s)	1.095	1.095	-0.159	-0.159
Controls and State F.E.	Yes	Yes	Yes	Yes
Unemployment and Waiver Controls	No	Yes	No	Yes

The table replicates baseline estimates by splitting the sample between single mothers in the phase-in of the EITC schedule and single mothers in the plateau or phase-out region of the schedule. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC are expressed in thousands of 2015 US dollars. Control variables include mother's race, indicator variables for the number of dependent children, year fixed effects, and policy-induced changes in welfare benefits (in 2015 US dollars). Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. The sample in columns (1) and (2) is restricted to single mothers with lagged earnings below $10,000. The sample in columns (3) and (4) is restricted to single mothers with lagged earnings below $10,000 and $30,000. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

phase-in of the program. Columns (3) and (4) of the table also provide evidence of positive marginal responses for mothers in the plateau or phase-out of the EITC schedule. These mothers, who are exposed to negative changes in their EITC benefits, decrease their labor supply, which explains the positive coefficient on changes in EITC benefits.

Table D-7 in Appendix D provides further evidence that single mothers initially located at both the phase-in and phase-out of the EITC schedule experience statistically significant changes in labor supply because of policy-induced expansions of the program. The finding that mothers facing the high marginal tax rates of the EITC phase-out reduce their labor supply is a standard theoretical prediction that has so far found scant empirical support in the literature (see Eissa and Hoynes 2006). Meyer (2002) considers average aggregate changes in hours worked be-

tween mothers with low and high education during the 1986-2000 period and concludes that the EITC did not generate disincentive effects in hours worked. He interprets his results as evidence of empirical deviations from standard labor theory. Our analysis provides different insights by showing that average effects hide vast heterogeneity in hours responses, with different individuals facing simultaneous incentives and disincentives generated by the EITC schedule. A standard theory of labor supply is consistent with these facts.

3.4 Previously Adopted Estimands in the Literature

Much of the empirical literature analyzing the effects of the EITC on labor market outcomes uses DiD or event study designs. A standard approach partitions women into single women without children, who were ineligible to receive EITC benefits prior to 1994, and single mothers, who *were* eligible, and compares changes in their labor supply around a reform using DiD (see for example Eissa and Liebman 1996; Meyer and Rosenbaum 2001; Kleven 2020; Schanzenbach and Strain 2021).[11] In this section, we highlight some of the challenges of applying DiD in this context. Our focus on the standard DiD design is without loss of generality, as any year-specific coefficient in an event study can be estimated using a DiD design comprising only data from the year in question and the base year.

The standard DiD design estimates the following equation:

$$Y_{i,t} = \beta_0 + \beta_1 Post_t + \beta_2 Treat_i + \beta_3 Treat_i \times Post_t + \varepsilon_{i,t} \ , \tag{7}$$

where $Treat_i$ is a time-invariant binary variable equaling one if the individual receives EITC benefits at any point in the study, and $Post_t$ is a binary variable equaling one in the period after the reform. The parameter of interest is β_3.

Consider a reform of the EITC that occurs between times $t-1$ and t in which the set of possible EITC benefits is the discrete set $\mathcal{J} = \{0, 1, \ldots J\}$. Associated with each benefit level, j, at time t is a potential outcome $Y_{j,t}$. We define separate

[11] Many single mothers are not eligible to receive the EITC because their incomes are too high (noncompliance). In Appendix B we explore the consequences of noncompliance.

indicators, $D_{j,t}$, for each of the $J+1$ EITC benefit levels.[12] $\mathbb{P}(D_{j,t}=1)$ is the probability of receiving benefit level j. At each point in time, the econometrician observes only one of the $J+1$ potential outcomes. We can write this observed outcome, Y_t, as a function of treatment assignments and potential outcomes as follows:

$$Y_t = Y_{0,t} + \sum_{j=1}^{J} D_{j,t}(Y_{j,t} - Y_{0,t}) \ . \tag{8}$$

For reasons that will be clarified below, we introduce an additional indicator, $D^*_{j,t}$, for the benefit level the individual would receive at time t if there had been no EITC reform, and a variable for the associated level of labor supply, $Y^*_{j,t}$. An expansion at time t that causes an individual to receive benefits equal to j instead of h would be characterized by a shift from $D^*_{h,t}=1$ to $D_{j,t}=1$.

We invoke the following assumption that is standard in the program evaluation literature using DiD designs:

Assumption 2 (No Selection on Counterfactual Trends).

$$E[Y_{0,t} - Y_{0,t-1}|Treat] = E[Y_{0,t} - Y_{0,t-1}]$$

Under Assumption 2, the DiD estimand represents the mean difference in longitudinal changes in the outcome Y_t between treatment and control group: $\beta_3^{DiD} = E[Y_t - Y_{t-1}|Treat=1] - E[Y_t - Y_{t-1}|Treat=0]$. Below, we define two main causal parameters of interest.

Definition 1 (The EITC Target Parameters). *Two main parameters of interest with respect to an evaluation of the EITC program are the average effect of treatment on the treated (ATT), and the average treatment effect of the policy reform (ATPR). These pa-*

[12] This analysis does not require that the set $\mathcal{J}$ be ordered. The analysis goes through if $\mathcal{J}$ contains tuples that each represent a benefit level and a marginal tax rate. This is relevant for the EITC as there are two implied marginal tax rates, on the phase-in and phase-out, for each benefit level.

rameters are defined as follows:

$$ATT \equiv \frac{1}{\sum_{j=1}^{J} \mathbb{P}(D_{j,t}=1)} \sum_{j=1}^{J} \mathbb{P}(D_{j,t}=1) E[Y_{j,t} - Y_{0,t} | D_{j,t}=1] \ ,$$

$$ATPR \equiv \frac{1}{\sum_{j=1}^{J} \sum_{h=1}^{J} \phi^*_{j,h}} \sum_{j=1}^{J} \sum_{h=1}^{J} \phi^*_{j,h} E[Y_{j,t} - Y^*_{h,t} | D_{j,t}=1, D^*_{h,t}=1] \ ,$$

where $\phi^*_{j,h} = \mathbb{P}(D_{j,t}=1, D^*_{h,t}=1)$.

ATT is a causal parameter that measures the effectiveness of the program as it is currently implemented. It identifies the program's average effect across all levels of treatment relative to a world with no program *for the people who actually participate in the program.* While ATT evaluates the program itself, ATPR can be used to evaluate a *reform* of the program. It measures the effect of the program's reform for people whose benefit level was altered by the reform.

However, Proposition 1 clarifies that without additional assumptions, the DiD estimand does not identify either of the target parameters.

Proposition 1. *Suppose Assumption 2 holds. Then, the DiD estimand from model (7) is proportional to the difference of weighted sums of treatment on the treated (TT) parameters for each treatment level:*

$$\beta_3^{DiD} = \frac{1}{p_{Treat}} \sum_{j=1}^{J} (p_{j,t} \Delta^{TT}_{j,t} - p_{j,t-1} \Delta^{TT}_{j,t-1}) \ , \tag{9}$$

where $p_{Treat} = \mathbb{P}(Treat = 1)$, *and for each* $j = 1 \ldots, J$, $p_{j,t} = \mathbb{P}(D_{j,t}=1)$, $\Delta^{TT}_{j,t} = E[Y_{j,t} - Y_{0,t} | D_{j,t}=1]$, *and* $p_{j,t-1}$ *and* $\Delta^{TT}_{j,t-1}$ *are defined analogously.*

Proof. See Appendix B. □

Proposition 1 reveals that DiD identifies a difference in weighted ATT parameters, where the weights are given by the proportion of treated individuals in each time period receiving a given level of treatment. The DiD estimand could therefore be zero or negative even if the treatment effect at each treatment level,

$Y_{j,t} - Y_{0,t}$, is positive and hence $\left\{\Delta_{j,t}^{TT} > 0\right\}_{j=1}^{J}$. Moreover, the DiD estimand in equation (9) is silent about the average effect of the policy reform (ATPR).

Additional restrictions are needed in order for the DiD estimand to identify a causal parameter of interest. In Appendix B, we discuss the specific restrictions, which include, among others, the requirement that the potential outcome functions are time invariant. This assumption implies that there are no other programs or policies that exhibit interaction effects with the program of interest. However, Kleven (2020), Grogger (2003), and Meyer and Rosenbaum (2001) discuss how state-level reforms to welfare occurred contemporaneously with the 1993 EITC expansion, and the DiD design confounds the effects of the two policy reforms on employment. They propose overcoming this challenge by exploiting the uneven introduction of welfare waiver reforms across states. In this framework, $W_{i,t} \in \{0, 1\}$ indicates whether individual i resides in a state that had implemented welfare waiver reforms ($W_{i,t} = 1$) or not ($W_{i,t} = 0$) by time t. The DiD model including the effects of waivers is

$$\begin{aligned} Y_{i,t} =& \beta_0 + \beta_1 Post_t + \beta_2 Treat_i + \beta_3 Post_t \times Treat_i + \\ & \beta_4 W_{i,t} + \beta_5 W_{i,t} \times Post_t + \beta_6 W_{i,t} \times Treat_i + \beta_7 W_{i,t} \times Post_t \times Treat_i + \varepsilon_{i,t}\,, \end{aligned} \tag{10}$$

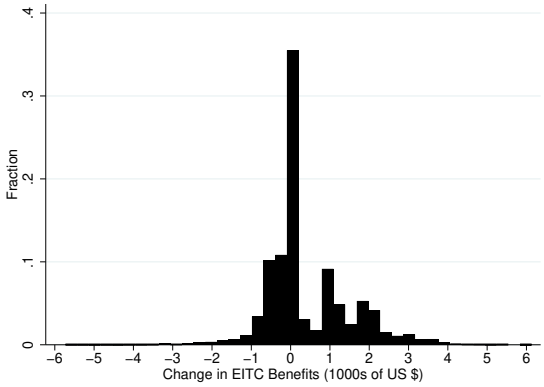
where $Treat_i$ is a time-invariant binary variable equaling one if the individual receives EITC benefits at any point in the study, and $Post_t$ is a binary variable equaling one in the period after the reform.

We show in Appendix B that β_3 in equation (10) still represents a *difference* in weighted ATT parameters, this time for the population with $W_{i,t} = 0$. Moreover, even if we disregard this potential issue, the estimation of (10) through an event study design that substitutes $Post_t$ with year-specific dummy variables still requires conditioning on the subset of states that had not implemented any welfare waivers by the corresponding year to identify each year-specific coefficient. As the set of states that had not implemented welfare waiver reforms by 1994 will differ from the set of states not having implemented them by 1995, the coefficients for 1994 and 1995 in the event study will be identified by labor supply changes in a different set of states. This means that comparisons of the event study co-

efficients across years will confound changes in employment over time resulting from the reform with changes in the average employment level that result from conditioning on a different set of states.[13]

An additional challenge to the DiD design occurs when classifying all single mothers, many of whom had incomes that were too high to be affected by the reforms to the EITC, in the treatment group. Figure 3 plots the distribution of policy-induced changes in EITC benefits due to changes in the program schedule for single mothers. More than one third of single mothers in our sample are unaffected by expansions of the EITC. The practice, common in the literature on the EITC, of assigning unexposed mothers to the treatment group confounds the resulting estimate in much the same way that imperfect compliance in a randomized controlled trial (RCT) affects the interpretation of the RCT evaluation as ITT instead of ATE. Appendix C shows that removing untreated single mothers from the treatment group dramatically increases the estimates obtained via an event study.

Figure 3: Distribution of EITC Benefit Changes for Single Mothers in the ASEC



The figure shows the distribution of policy-induced changes in EITC benefits for single mothers in the CPS-ASEC. Changes in EITC benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC benefits are provided in Section 3.1 and Appendix E.

[13]Appendix B provides a formal derivation of this result, as well as a discussion of the (restrictive) assumptions that are necessary to identify ATT or ATPR. Grogger and Karoly (2005) provide a discussion of the timing of the implementation of the welfare waiver reforms. Athey and Imbens (2006) and De Chaisemartin and D'HaultfŒuille (2017) focus on separate challenges in identifying interpretable treatment parameters using DiD designs.

Altogether, we have documented the existence of (i) substantial population heterogeneity in both the level of EITC and welfare benefits; (ii) heterogeneity in the labor supply responses to policy changes by different subgroups of the population, with new evidence that the income and substitution effects associated with the plateau and phase-out of the EITC schedule cause disincentive effects on hours worked; and (iii), imperfect compliance with the EITC expansion for the group of single mothers. The standard DiD/event study design used in the literature, e.g. Eissa and Liebman (1996) and Meyer and Rosenbaum (2001), reduces all this heterogeneity within the treatment variable to a single margin, that of going from untreated to treated, eliminates the relationship between the intensity of treatment and the intensity of the labor supply response, and aggregates heterogeneous effects into an average effect. These shortcomings may explain some of the more recent conflicting conclusions in the literature that uses DiD designs to evaluate the EITC (Kleven 2020; Schanzenbach and Strain 2021).

4 A Static Model of Labor Supply with Social Programs

The previous section has shown that changes in benefit levels induced by reforms to both the EITC and welfare cause changes in the labor supply of single women. However, regression analysis of equation (6) imposes strong restrictions on labor supply responses that may be heterogeneous in the population. Furthermore, such an equation fails to account for how labor supply, earnings, and thus benefit receipt are determined jointly in response to prevailing wages and the entire set of available social programs. To take these factors into account, we now estimate a structural labor supply model. We use the model to identify the Frisch elasticity of labor supply, which governs behavioral responses to policies, and also to conduct counterfactual analysis of reforms to social programs in new environments characterized by different labor market incentives.

The model we estimate is a generalization of the static model in (2), which allows for greater heterogeneity in preferences and models the tax and transfer system more richly. While there may be dynamic considerations affecting the labor supply of single mothers, we are primarily interested in how the effects of reforms to

the EITC and welfare depend on the tax and transfer regime.[14] We view a static labor supply model that jointly models the tax system and social programs as an appropriate framework to conduct this analysis, especially in light of how our earlier quasi-experimental evaluation supports the predictions of a static labor supply model (Figure 2).

In the model, single mothers have preferences over consumption, $c_{i,t}$, and hours worked, $h_{i,t}$, given by

$$u_i(c_{i,t}, h_{i,t}) = \frac{1}{1-\frac{1}{\eta}} c_{i,t}^{1-\frac{1}{\eta}} - \frac{\alpha_i}{1+\frac{1}{\gamma}} h_{i,t}^{1+\frac{1}{\gamma}} \ . \tag{11}$$

The parameter $\eta > 0$ captures the curvature of utility with respect to consumption, where a higher value indicates less concavity, while $\gamma > 0$ is the Frisch elasticity of labor supply. It determines the elasticity of hours worked with respect to changes in wages, holding the marginal utility of consumption fixed. In each period, t, single mothers decide how much to work by solving

$$\begin{aligned} &\max_{c_{i,t}, h_{i,t}} \ u_i(c_{i,t}, h_{i,t}) \\ \text{subject to } c_{i,t} =& \omega_{i,t} \cdot h_{i,t} - Tax_{i,t}(\omega_{i,t} \cdot h_{i,t}) + EITC_{i,t}(\omega_{i,t} \cdot h_{i,t}) + \\ & SNAP_{i,t}(\omega_{i,t} \cdot h_{i,t}) + T_{i,t}^{Welfare}(\omega_{i,t} \cdot h_{i,t}) \ , \quad h_{i,t} \geq 0 \ , \quad c_{i,t} \geq 0 \ . \end{aligned}$$

We allow for heterogeneity in the disutility of working, $\alpha_i > 0$, so that the model can generate self-selection into employment and hours worked on the basis of unobservables. We characterize the various tax and transfer programs as part of the budget constraint. First, we model the three main programs that provide financial assistance to low-income families: (i) the Earned Income Tax Credit (EITC); (ii) Aid to Families with Dependent Children/Temporary Assistance for Needy Families (AFDC/TANF); and (iii) the Supplemental Nutritional Assistance Program (SNAP). We use the benefit formulas for each of the three programs to recover the correct level of benefits claimed by each household. Second, we model the tax on labor income $Tax_{i,t}(\cdot)$ via a parametric function

[14] While we abstract from the impact of these reforms on the next generation of workers, Daruich and Fernández (2020) show that transfers, in the form of Universal Basic Income, have significant intergenerational consequences in a general equilibrium framework.

that maps pre-tax labor income to after-tax labor income. This approach is relatively common in the public finance and labor literature (see for example Benabou 2002; Guner, Kaygusuz, and Ventura 2014; Blundell, Pistaferri, and Saporta-Eksten 2016; Heathcote, Storesletten, and Violante 2017; and Holter, Krueger, and Stepanchuk 2019). It allows us to independently parameterize the level and the progressivity of the tax system as follows:

$$\omega_{i,t} \cdot h_{i,t} - Tax_{i,t}(\omega_{i,t} \cdot h_{i,t}) = \theta_{0,s,t,k} \cdot (\omega_{i,t} \cdot h_{i,t})^{1-\theta_{1,s,t,k}} , \tag{12}$$

where we allow the tax function to vary by state s, year t, and number of children k. The parameters $\theta_{0,s,t,k} \geq 0$ and $\theta_{1,s,t,k} \in [0, 1]$ capture the take-home rate and the progressivity of the tax on labor income, respectively. A higher value of $\theta_{0,s,t,k}$ implies a higher take-home rate (lower level of tax rates), while a higher value of $\theta_{1,s,t,k}$ implies greater progressivity.

In every period t, each individual receives a wage offer, $\omega_{i,t}$. The initial wage offer (ω_0) is distributed according to a conditional log-normal distribution,

$$\ln \omega_0 \sim N(\mu_\omega, \sigma_\omega | k) , \tag{13}$$

which we allow to vary by the number of children. We also model the evolution of the log-wage offer as a random walk:

$$\ln \omega_{i,t} = \ln \omega_{i,t-1} + \nu_{i,t} \quad \text{with} \;\; \nu_{i,t} \sim N(0, \sigma_\nu^2) , \tag{14}$$

where the innovation, ν_t, is assumed to be normally distributed among individuals with mean zero and standard deviation σ_ν.

5 Estimation

We estimate the model via simulated method of moments (SMM). Denote the set of moments we are trying to match by M and the set of model parameters by $\Omega = \{\eta, \gamma, \{\alpha_i\}_i, \mu_\omega, \sigma_\omega^2, \sigma_\nu^2\}$. Given a wage offer and a particular tax and transfer regime, we simulate each individual's optimal labor supply choice. We then use the data created by these simulated choices to construct a set of moments, M_S,

analogous to the moments, M, observed in the data. We estimate the model using the ASEC data from the pre-reform period (through 1993), while the 1995-1996 data from the transition period is used to validate the model. Our SMM estimator is

$$\widehat{\Omega} = \arg\min_{\Omega} \left(M - M_S(\Omega)\right)' W \left(M - M_S(\Omega)\right) \ , \tag{15}$$

where W is a positive semidefinite weighting matrix.[15] In practice, we set the weighting matrix equal to the inverse of the covariance matrix of the moments, $W = \Sigma_M^{-1}$, with Σ_M determined by 100 bootstrap replications of the data set. We target the following 17 moments in the data from 1988-1993 to recover 10 parameters: mean hours and employment at the aggregate level and by number of children, the mean and standard deviation of accepted wages, the autocovariance of log wages at the aggregate level and by number of children, and the causal effects of EITC and welfare on hours worked.[16]

We model the disutility parameter according to the equation

$$\alpha_i = \alpha_k + \alpha \cdot \upsilon_i \ , \tag{16}$$

which allows the disutility of labor to vary with the number of children, k. $\upsilon_i \sim unif\left\{0.1, 2.5\right\}$ is a discrete uniform random variable taking six equally-spaced values between 0.1 and 2.5. This parsimonious approach fits the data well by allowing α_i to take on 18 possible values with only four underlying parameters.

5.1 Parameter Estimates

Preferences. Table 4 shows the estimates for the preference parameters. We estimate a relatively high curvature for the utility over consumption ($\eta = 0.67$) and a Frisch elasticity of one ($\gamma = 1.03$), which is consistent with the previous findings in the literature of a high elasticity of labor supply among single mothers

[15] Two reasons drive the choice of simulated method of moments instead of a likelihood-based method. First, the SMM approach overcomes the additional source of computational burden which arises from the multi-dimensional integration problem associated with the maximum-likelihood estimator of this model. Second, we believe this method highlights more transparently the identifying variation of our model, as it allows us to replicate the causal regression coefficients of the effect of EITC and AFDC/TANF benefits on hours worked.

[16] We target the regression coefficients from the specification in column (1) of Table 2 in a sample comprising only the years prior to the 1993 reform.

Table 4: Estimates for Preferences Parameters

	Preferences
Curvature of Consumption (η)	0.6716 [0.6703,0.6738]
Frisch elasticity (γ)	1.0319 [1.0236,1.0372]
Disutility of Hours Worked (α)	0.0722 [0.0705,0.0774]
Additional Disutility of Hours Worked with One Child ($\%\alpha$)	-0.0048 [-0.0066,-0.0036]
Additional Disutility of Hours Worked with Three Children ($\%\alpha$)	0.0160 [-0.0010,0.0169]

The table shows the estimated preferences parameters; see Equations (11) and (16). The 95 percent confidence intervals in brackets are calculated via 100 bootstrap repetitions. The point estimates are the averages among the bootstrap repetitions.

(Attanasio et al. 2018, Blundell et al. 2016, Keane and Rogerson 2012, and Keane and Rogerson 2015).[17]

The disutility of hours worked is heterogeneous around a point estimate of $\alpha = 0.072$. This translates to a mean value for mothers with two children of 0.09, with the range of values that goes from 0.007 to 0.181.[18] The mean disutility of hours worked is 0.48 percent lower for mothers with one child and 1.6 percent higher for mothers with three children or more.

Wage Process. Table 5 presents estimates for the preference parameters. We find that the mean log-wage offer is $\mu_{\omega,k=1} = 2.56$ for mothers with one child. The average log-wage offer decreases monotonically with the number of children. Mothers with two children receive wage offers that are on average 0.04 log-points lower. The average wage offer is 0.26 log-points lower for mothers with three children or more.

[17]The estimate of η would be equivalent to a coefficient of relative risk aversion of approximately 1.5, although we do not have any source of risk in our framework.

[18]Given our assumption that $\upsilon \sim unif\{0.1, 2.5\}$ and the point estimate of $\alpha = 0.072$, we get $E[\alpha_i] = \alpha \cdot \frac{0.1+2.5}{2} = 0.09$, while the minimum and maximum values are $\alpha \cdot 0.01 = 0.007$ and $\alpha \cdot 2.5 = 0.181$.

Table 5: Estimates for Wage Process

	Wage Process
Mean (One Child, $\mu_{\omega,1}$)	2.5591 [2.5412,2.5874]
Additional Mean with Two Children ($\mu_{\omega,k=2}$)	-0.0381 [-0.0744,-0.0333]
Additional Mean with Three Children ($\mu_{\omega,k=3}$)	-0.2601 [-0.3214,-0.2311]
Initial Standard Deviation (σ_ω)	0.5815 [0.5659,0.5917]
Standard Deviation of Innovation (σ_ν)	0.0183 [0.0123,0.0392]

The table shows the estimated wage process parameters; see Equations (13) and (14). The 95 percent confidence intervals in brackets are calculated via 100 bootstrap repetitions. The point estimates are the averages among the bootstrap repetitions.

Finally, we estimate a fairly large dispersion for the unobserved heterogeneity in the initial wage offer, with a standard deviation of $\sigma_\omega = 0.582$. To put this value in perspective, the estimated standard deviation is more than twice the difference in mean log-wage offers between mothers with one child and mothers with three children or more. Blundell, Pistaferri, and Saporta-Eksten (2018) find a standard deviation of 0.533, although the authors focus on women in intact families in the Panel Study of Income Dynamics (PSID). Moreover, we find a small role of the stochastic innovation in determining the evolution of wage offers, with an estimated standard deviation of $\sigma_\nu = 0.018$.

5.2 In-sample Fit

Tables D-8 and D-9 in Appendix D suggest that the model is successful in replicating the targeted moments. In each table we report both the data moments, M, as well as the simulated moments, M_S, calculated at the model solution.

Table D-8 reports three panels. Table D-8-a shows that the model replicates the mean number of hours worked as well as the negative gradient in hours worked per child during the years 1988-1993. Table D-8-b shows that the model also

replicates the aggregate employment rate (about 0.76) and the negative gradient of employment with respect to the number of children. Mothers with one child are the most likely to work, with an employment rate of 0.84, but employment rates drop to 0.77 and 0.61 for mothers with two and three or more children, respectively. Table D-8-c shows the in-sample fit for accepted wages, an endogenous object in the model. The model successfully replicates the first and second moments of the accepted wage distribution, although the persistence of accepted log-wages is higher in the model than the data.

Table D-9 shows the model fit for the causal regression coefficients of the effects of EITC and welfare on hours worked. The model replicates the positive effect of EITC benefits on hours worked, as well as the negative effect of welfare. During the 1988-1993 period, a $1,000 increase in EITC benefits causes an average increase of about 175 hours worked per year, while the same increase in welfare benefits causes an average reduction of 12 hours per year. Although the model slightly overstates the marginal effect of EITC on hours, the difference is not statistically significant.

5.3 Out-of-sample Predictions

Before examining the model's out-of-sample fit, we must take a stand on the wage offer equation in 1995-1996, a task complicated by the fact that we only observe the distribution of accepted wages in the data. We deal with this challenge by adding two features to the model: a new set of parameters for the wage offer model characterized by (13) and (14) and a utility cost of working, χ_i. The utility cost of working modifies preferences as follows:

$$u_i(c_{i,t}, h_{i,t}) = \frac{1}{1-\frac{1}{\eta}} c_{i,t}^{1-\frac{1}{\eta}} - \frac{\alpha_i}{1+\frac{1}{\gamma}} h_{i,t}^{1+\frac{1}{\gamma}} + \chi_i \mathbb{1}(h_{i,t} > 0) \ .$$

χ_i follows the same distribution of the disutility of working (α_i): $\chi_i = \chi \cdot \upsilon_i$, where χ is a free scale parameter. This specification allows for a "cohort-specific" cost of working and can capture heterogeneity in the unobserved cost of working that is

unrelated to hours and not reflected in wages.[19]

We calibrate χ and the new wage offer model by only matching moments of the accepted wage distribution in 1995-1996.[20] The rest of the preference parameters remain at their estimated values presented in Table 4. This approach lets us test if the preferences we estimate using data prior to 1994 can replicate labor supply statistics that we do not directly target in 1995-1996.

We find that the model predicts increases in both hours worked and employment that are consistent with the data. Figures D-1 and D-2 in Appendix D depict the model's out-of-sample fit. The rise in labor supply appears both at the aggregate level (Panel (a) of Figures D-1 and D-2) and by number of children (Panels (b) through (d)).

6 Counterfactual Analysis

We use the estimated model to analyze three counterfactual reforms. First, we examine the effect of the 1990s reforms to the EITC and welfare in isolation, so that we can disentangle each program's effect on labor supply. Then we consider two new policies: a large expansion of the EITC in 1996 and the replacement of either the EITC or welfare with Universal Basic Income. We show that the response of labor supply to the proposed policies varies considerably with the progressivity of the tax code.

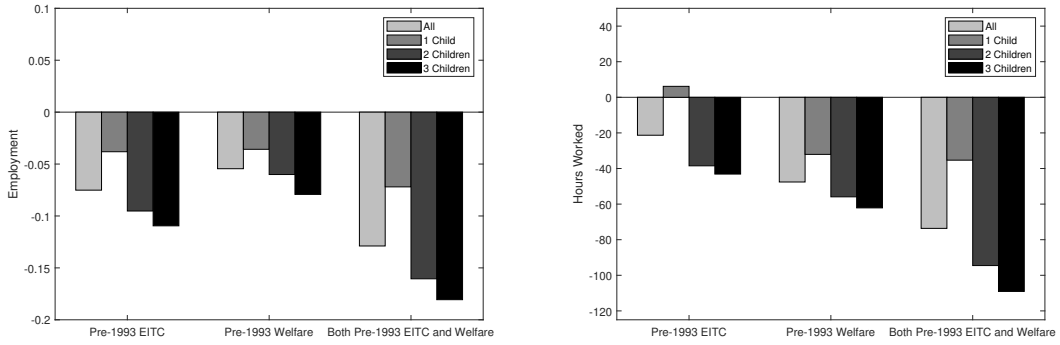
6.1 EITC and Welfare Reform Decomposition

Our first counterfactual exercise consists of several scenarios. The first keeps all tax and transfer programs at their 1996 level with the exception of the EITC,

[19] Because the extensive margin decision is an endogenous choice in the model, we want to avoid estimating the wage offer equation outside the model via some parametric reduced-form models that could be inconsistent with our structural model. An alternative estimation strategy would pool together data from the pre-reform and transition periods and estimate the model by allowing the wage offer parameters and the cost of working to vary by time period. We prefer our estimation strategy as it allows us to use moments from the 1995-1996 data to test the model.

[20] We validate the model using data through 1996, as time limits on the receipt of welfare benefits began taking effect in 1997. We do not model, nor do we observe in the data, an individual's total lifetime benefits, so the model should not be expected to match labor supply once these restrictions take effect. We describe the time limits in detail in Appendix A and list the dates they first take effect in each state in Appendix F.

Figure 4: Policy Contributions in 1996 of Pre-1993 EITC and Welfare Policies



(a) Employment

(b) Hours Worked

The figure shows the counterfactual level of employment (Panel (a)) and hours worked (Panel (b)) for single mothers in 1996 if counterfactually either the EITC or welfare were held at the 1993 regime, while keeping the rest of the other tax and transfer programs at their 1996 levels. The analysis is performed on the whole sample of single mothers, as well as by number of children.

whose benefits are artificially held at their level in 1993. The second scenario repeats the exercise for welfare by combining AFDC's policies from 1993 with the 1996 tax and transfer system. The third scenario keeps both EITC and welfare at their 1993 policy rules, while leaving the rest of tax system at its 1996 level.

Figure 4 presents the results of this exercise. We find that in the absence of reforms to either the EITC or welfare, both employment and hours worked would have been lower. Employment for single mothers would have been 7.5 percentage points lower if the EITC were not expanded as it was, or 5.5 percentage points lower if AFDC were not reformed as it was. Hours worked would have dropped by 20, respectively 48, hours per year if either the EITC were not expanded or welfare not reformed as they were through 1996.

In the absence of both reforms, the 1996 employment rate for single mothers would have been 12 percentage points lower, and the average hours worked would have been reduced by 73 hours. This is an interesting result as the model suggests that the counterfactual level of employment and hours would have decreased after 1993 due to rises in SNAP benefits and changes in wage offers and preferences for working. This finding is consistent with the fact that, through

1993, both employment and hours were trending downward for single mothers, and the model suggests that these would have kept falling without the reforms.

6.2 The Interaction of Social Programs and the Tax Code

Our remaining counterfactual exercises demonstrate how the exact same reform can generate different effects on labor supply depending on the tax regime and method of financing the reform. Denote a specific social program by a vector of parameters, Υ_j, that fully characterize the program, and denote the entire tax and transfer system by a set of social programs: $\Upsilon = \{\Upsilon_1, \ldots, \Upsilon_J\}$. We show that the labor supply response to a single program, Υ_j, can depend on the other taxes and transfers already in place, Υ_{-j}.

We evaluate a reform to program j $\left(\Upsilon_j^{'} \neq \Upsilon_j\right)$ using a standard potential outcomes framework, where for each individual i we define the pair of potential outcomes, $\left(\mathrm{Y}_i(\Upsilon_j^{'}, \Upsilon_{-j}), \mathrm{Y}_i(\Upsilon_j, \Upsilon_{-j})\right)$. We then examine the reform's average treatment effect (ATE):

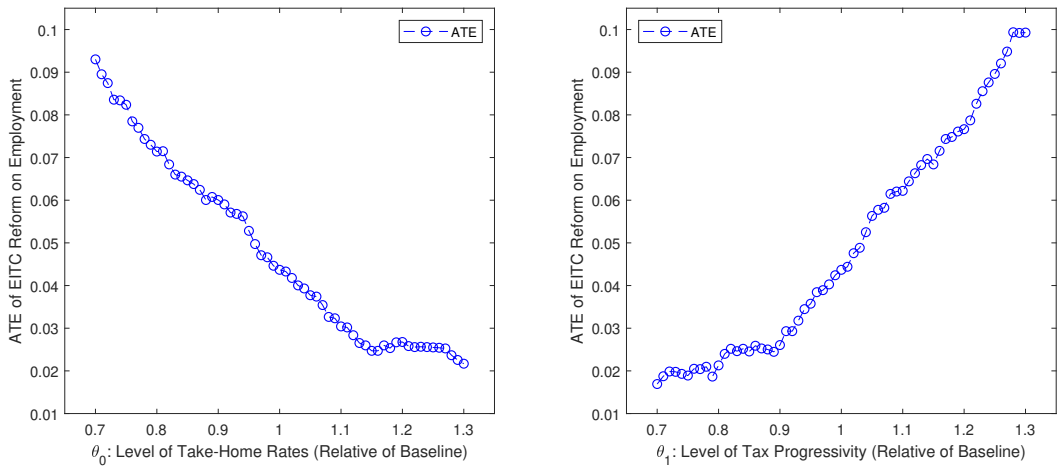
$$ATE_j(\Upsilon_{-j}) = E\left[\mathrm{Y}_i(\Upsilon_j^{'}, \Upsilon_{-j}) - \mathrm{Y}_i(\Upsilon_j, \Upsilon_{-j})\right] . \tag{17}$$

We use our estimated model to simulate changes in the policy regime Υ_j and observe the entire distribution of potential outcomes, $\left\{\mathrm{Y}_i(\Upsilon_j^{'}, \Upsilon_{-j}), \mathrm{Y}_i(\Upsilon_j, \Upsilon_{-j})\right\}_i$, for a fixed parameterization of the tax code ($\theta_{0,s,t,k}$ and $\theta_{1,s,t,k}$ in Equation (12)).

We first examine the response of the extensive margin to an expansion of the EITC that increases maximum benefits by \$2,000 and the federal income limit by \$5,000 in 1996, while simultaneously varying the tax on labor income. This expansion resembles the 1993 EITC reform in terms of changes in benefits and income limits. Figure 5-a shows how the ATE of the reform depends on the take-home rate, θ_0. The *x-axis* is defined relative to the original 1996 take-home rate, which means that a level of 1.1 is a take-home rate that is ten percent higher than baseline.[21] ATEs are higher when take-home rates are lower (tax rates are higher). For example, the ATE on employment goes from 7 to 2.5 percentage points if we

[21] Tax rates are heterogeneous by state and number of children. We proportionally change the various tax rates by the same factor (x-axis).

Figure 5: Simulated ATE by Different Tax Regimes



(a) ATE and Tax Rates

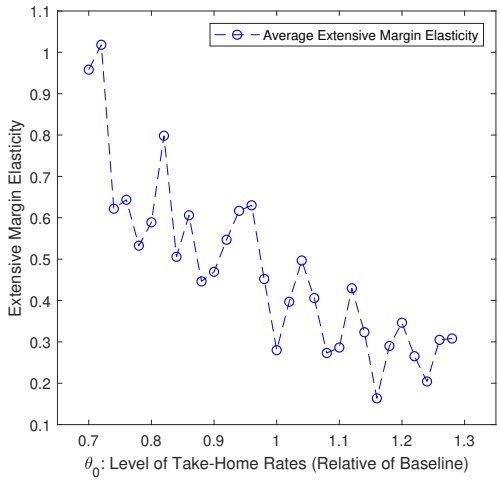
(b) ATE and Tax Progressivity

The figure shows the ATE (blue dots) of an EITC reform on the probability of employment as a function of the level (Panel (a)) and the progressivity (Panel (b)) of the tax regime. The simulated EITC reform includes an expansion of the federal exhaustion point of the EITC schedule of $5,000, as well as an increase of the maximum federal credits of $2,000.

vary the take-home rate from 0.8 to 1.2 times the baseline 1996 level. This result is sensible as the marginal benefit of a tax credit like the EITC depends on the tax rate, with workers facing an extensive margin decision benefiting more from the credits when income taxes are higher.

Figure 5-b shows a similar pattern for the progressivity of the tax system, determined by θ_1: The ATE is higher when individuals face steeper marginal tax rates as their income rises (θ_1 increases). These results suggest that the estimated ATE of the EITC on employment is not invariant to changes in taxes. Despite this fact, a large body of the empirical public economics literature (e.g. Saez 2002 and Eissa, Kleven, and Kreiner 2008) aims to recover the average extensive margin elasticity for the population of interest. Figure 6 demonstrates that this elasticity varies widely according to the tax system, for the same set of structural parameters. Each dot represents the average percentage change in the probability of being employed induced by a one percent change in the take-home rate. First, the graph shows that the simulated aggregate elasticities from the estimated model

Figure 6: Simulated Extensive Margin Elasticity by Different Tax Regimes



The figure shows how the estimated aggregate extensive margin elasticity to taxes varies by the level of the tax rates. Each elasticity (blue dot) is the percentage change in the aggregate employment rates caused by a small ($\frac{0.01}{\theta_0}$) change in the tax rates.

are smaller than the estimated structural Frisch elasticity. This result is similar in spirit to the result in Rogerson and Wallenius (2009), where the authors show that in a life cycle model of labor supply with intensive and extensive margins, micro and macro elasticities are effectively unrelated. Second, the wide variation in estimated elasticities is caused by changes in the composition of individuals at the margin of employment as the tax code changes.[22]

Our final counterfactual exercise examines the effects of Universal Basic Income on the labor supply of single mothers. UBI has recently gained traction in policy circles as a result of the dislocations created by rapid technological development and COVID-19. We replace either the EITC or welfare with a UBI program that targets the population of single mothers in a budget-neutral way.[23] Table 6 shows

[22]Similarly, Moffitt (2019) finds that the marginal treatment effects of welfare reforms on labor supply change over time because of the preference heterogeneity of the marginal individual.

[23]Kearney and Mogstad (2019) provide a review of the UBI proposals and highlight that they risk increasing inequality and being expensive and inefficient. We adopt a concept of UBI that resembles the one in Hoynes and Rothstein (2019), namely that the cash transfer is universally provided to the whole population of single mothers but, due to the budget neutral criterion to finance it, it is not sufficiently generous to allow recipients to live on it without additional earnings. On the latter aspect, our definition is similar to the one in Banerjee, Niehaus, and Suri (2019).

Table 6: Simulated Effects of UBI by Different Ways of Financing the Program

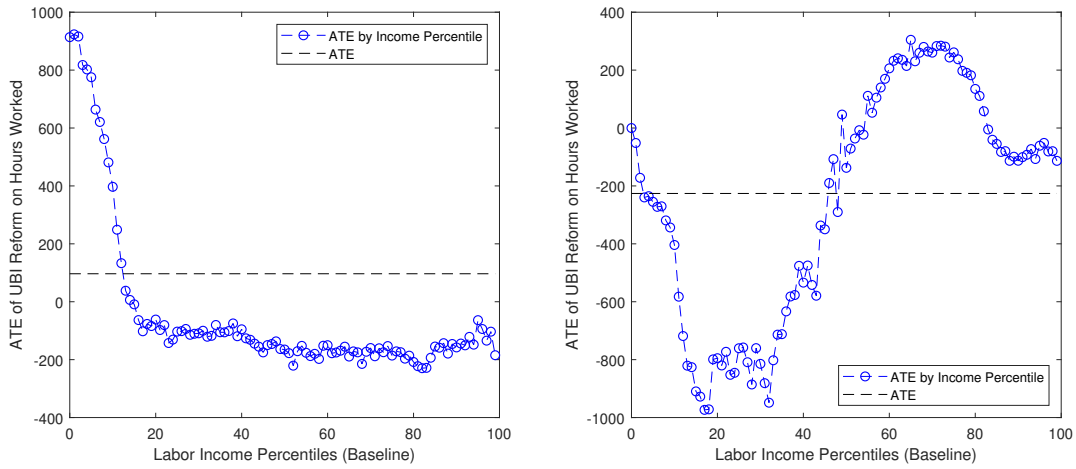
	(1)	(2)
	Different Ways of Financing UBI:	
	UBI replaces AFDC/TANF	UBI replaces EITC
ATE on Employment	0.13	-0.24
ATE on Hours	96.65	-226.03

The table shows the average response of employment and hours worked in the population (ATE) to a reform that substitutes UBI for welfare (Column 1) or EITC (Column 2).

the effect of UBI on employment and hours worked. Contrary to common wisdom, UBI can generate positive or negative effects on labor supply depending on the program it replaces. If UBI completely replaces the EITC, employment and hours worked would fall by 24 percentage points and 226 hours per year, respectively. While these results are large, the considered reform is also massive, pooling all of the EITC money received by single mothers in 1996 and redistributing it in (approximately) $3,000 checks to each individual. If UBI instead replaces welfare, the effect on hours and employment is positive. Given our empirical analysis, this finding is hardly surprising: The reform removes a disincentive to enter the labor force by eliminating benefits targeting people who do not work and equalizes the cash transfer to everybody regardless of employment status.

Figure 7 shows that the responses of hours worked to the UBI reform are heterogeneous and nonmonotone. When UBI replaces welfare, shown in Figure 7-a, there are strong positive effects on labor supply for individuals at the lowest quintile of the income distribution, those who directly lose access to welfare benefits. The rest of the population reduces labor supply as a consequence of the unconditional transfer. On the other hand, when UBI replaces the EITC, we find opposing intensive margin effects depending on whether the individual was either at the phase-in or at the phase-out of the EITC schedule. Individuals at the phase-in lose the negative marginal tax rates on earnings, causing a reduction in hours worked, while individuals at the phase-out experience the opposite. These two forces explain the nonmonotonic effects in Figure 7-b.

Figure 7: Heterogeneous Effects of UBI on Hours Worked by Financing Options



(a) UBI replaces TANF

(b) UBI replaces EITC

The figure shows the response of hours worked in the population to a welfare reform that substitutes UBI for TANF (Panel (a)) or EITC (Panel (b)). Each dot represents the ATE by labor income percentile at baseline. The dashed line represents the ATE in the population.

7 Conclusion

An important goal of evidence-based policymaking is to learn from prior policies to predict the effects of future policy regime changes. This paper has exploited variation in social program benefit receipt induced by policy changes to the EITC and welfare throughout the 1990s to show how behavioral labor supply responses are consistent with the predictions of standard theoretical labor supply models. We explain how several features of the commonly-used DiD design – such as simultaneously including both individuals who have pre-existing benefits and those who are untreated in the treatment group, as well as estimating an event study while controlling for welfare waivers – have concealed the marginal labor supply responses of individuals to the benefits provided by economic security programs.

Guided by our quasi-experimental findings, we estimate a structural model of labor supply with heterogeneous exposure to multiple tax and transfer programs. The estimated model shows that the effect of an EITC expansion on employment

depends heavily on the level and progressivity of labor income taxes. This result suggests that the evaluation of past EITC reforms does not speak for itself about the effects of future EITC expansions. The evolution of the tax code over time affects the choices made by individuals, which in turn determine the aggregate labor supply response. For this reason, policymakers considering new reforms are advised to carefully consider how the reforms would interact with existing programs and generate new incentives when forecasting their effects on labor supply.

References

Albanesi, Stefania, Claudia Olivetti, and Barbara Petrongolo. 2023. "Families, Markets, and Policy." In *Handbook of the Economics of the Family*, edited by Shelly Lundberg and Alessandra Voena, Volume 1, 255–326. Amsterdam: Elsevier.

Angrist, Joshua D., and Guido Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90:431–442.

Ashenfelter, Orley. 1983. "Determining Participation in Income-Tested Social Programs." *Journal of the American Statistical Association* 78 (383): 517–525.

Athey, Susan, and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2): 431–497.

Attanasio, Orazio, Peter Levell, Hamish Low, and Virginia Sánchez-Marcos. 2018. "Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply." *Econometrica* 86 (6): 2049–2082.

Banerjee, Abhijit, Paul Niehaus, and Tavneet Suri. 2019. "Universal Basic Income in the Developing World." *Annual Review of Economics* 11 (1): 959–983.

Bastian, Jacob, and Lance Lochner. 2020. "The EITC and Maternal Time Use: More Time Working and Less Time with Kids?" NBER Working Paper 27717.

Benabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2): 481–517.

Bitler, Marianne P, Jonah B Gelbach, and Hilary W Hoynes. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review* 96 (4): 988–1012.

Blank, Rebecca M. 2002. "Evaluating Welfare Reform in the United States." *Journal of Economic Literature* 40 (4): 1105–1166.

Blumenthal, Marsha, Brian Erard, and Chih-Chin Ho. 2005. "Participation and Compliance with the Earned Income Tax Credit." *National Tax Journal*, pp. 189–213.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica* 84 (5): 1705–1753.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. "Consumption Inequality and Family Labor Supply." *The American Economic Review* 106 (2): 387–435.

———. 2018. "Children, Time Allocation, and Consumption Insurance." *Journal of Political Economy* 126 (S1): S73–S115.

Center on Budget and Policy Priorities. 2022. "Where Do Our Federal Tax Dollars Go?" Technical Report, Washington, DC: Center on Budget and Policy Priorities.

Chan, Marc K. 2013. "A Dynamic Model of Welfare Reform." *Econometrica* 81 (3): 941–1001.

Chetty, Raj, John N Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.

Chetty, Raj, and Emmanuel Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics* 5 (1): 1–31.

Dahl, Gordon B., and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *The American Economic Review* 102 (5): 1927–1956.

Daruich, Diego, and Raquel Fernández. 2020. "Universal Basic Income: A Dynamic Assessment." Technical Report, National Bureau of Economic Research.

De Chaisemartin, Clément, and Xavier D'HaultfŒuille. 2017. "Fuzzy Differences-in-Differences." *The Review of Economic Studies* 85 (2): 999–1028 (08).

Dickert, Stacy, Scott Houser, and John Karl Scholz. 1995. "The Earned Income Tax Credit and Transfer Programs: A Study of Labor Market and Program Participation." *Tax Policy and the Economy* 9:1–50.

Eckstein, Zvi, and Kenneth I. Wolpin. 1989. "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience." *Review of Economic Studies* 56 (3): 375–390.

Eissa, Nada, and Hilary W. Hoynes. 2006. "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply." *Tax Policy and the Economy* 20:73–110.

Eissa, Nada, Henrik Kleven, and Claus T. Kreiner. 2008. "Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers." *Journal of Public Economics* 92 (3): 795–816.

Eissa, Nada, and Jeffrey B. Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *The Quarterly Journal of Economics* 111 (2): 605–637.

Fang, Hanming, and Michael P. Keane. 2004. "Assessing the Impact of Welfare Reform on Single Mothers." *Brookings Papers on Economic Activity* 2004 (1): 1–95.

Gelber, Alexander M., and Joshua W. Mitchell. 2012. "Taxes and Time Allocation: Evidence from Single Women and Men." *The Review of Economic Studies* 79 (3): 863–897.

Grogger, Jeffrey. 2003. "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income among Female-Headed Families." *The Review of Economics and Statistics* 85 (2): 394–408.

Grogger, Jeffrey, and Lynn A. Karoly. 2005. *Welfare Reform: Effects of a Decade of Change*. Harvard University Press.

Gruber, Jon, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1): 1–32.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2014. "Income Taxation of U.S. Households: Facts and Parametric Estimates." *Review of Economic Dynamics* 17 (4): 559–581.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Heckman, James J., and Thomas E. Macurdy. 1980. "A Life Cycle Model of Female Labour Supply." *The Review of Economic Studies* 47 (1): 47–74.

Holter, Hans A., Dirk Krueger, and Serhiy Stepanchuk. 2019. "How do Tax Progressivity and Household Heterogeneity Affect Laffer Curves?" *Quantitative Economics* 10 (4): 1317–1356.

Hotz, V. Joseph, Charles H. Mullin, and John K. Scholz. 2006. "Examining the Effect of the Earned Income Tax Credit on the Labor Market Participation of Families on Welfare." NBER Working Paper 11968.

Hotz, V. Joseph, and John K. Scholz. 2003. "The Earned Income Tax Credit." In *Means-Tested Transfer Programs in the United States*, edited by Robert A. Moffitt, 141–197. Chicago: University of Chicago Press.

Hoynes, Hilary W., and Ankur J. Patel. 2018. "Effective Policy for Reducing Poverty and Inequality? The Earned Income Tax Credit and the Distribution of Income." *Journal of Human Resources* 53 (4): 859–890.

Hoynes, Hilary W., and Jesse Rothstein. 2017. "Tax Policy Toward Low-Income Families." In *The Economics of Tax Policy*, edited by Alan J. Auerbach and Kent Smetters. Oxford: Oxford University Press.

———. 2019. "Universal Basic Income in the United States and Advanced Countries." *Annual Review of Economics* 11 (1): 929–958.

Keane, Michael P. 1995. "A New Idea for Welfare Reform." *Federal Reserve Bank of Minneapolis Quarterly Review* 19 (2): 2–28.

Keane, Michael P., and Robert A. Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review* 39 (3): 553–589.

Keane, Michael P., and Richard Rogerson. 2012. "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom." *Journal of Economic Literature* 50 (2): 464–476.

———. 2015. "Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective." *Annual Review of Economics* 7 (1): 89–117.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3): 473–522.

Kearney, Melissa S., and Magne Mogstad. 2019. "Universal Basic Income (UBI) as a Policy Response to Current Challenges." *Report, Aspen Institute*.

Kleven, Henrik. 2020. "The EITC and the Extensive Margin: A Reappraisal." Unpublished Manuscript, Princeton University.

Kline, Patrick, and Melissa Tartari. 2016. "Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach." *The American Economic Review* 106 (4): 972–1014.

Low, Hamish, Costas Meghir, Luigi Pistaferri, and Alessandra Voena. 2018. "Marriage, Labor Supply and the Dynamics of the Social Safety Net." Technical Report, National Bureau of Economic Research.

Mancino, Antonella, and Joseph Mullins. 2020. "Frictional Adjustment to Income Tax Incentives: An Application to the Earned Income Tax Credit." Unpublished Manuscript.

Meyer, Bruce D. 2002. "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked." *The American Economic Review* 92 (2): 373–379.

Meyer, Bruce D., and Dan T. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *The Quarterly Journal of Economics* 116 (3): 1063–1114.

Moffitt, Robert A. 1990. "The Econometrics of Kinked Budget Constraints." *The Journal of Economic Perspectives* 4 (2): 119–139.

———. 1992. "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature* 30 (1): 1–61.

———. 2019. "The Marginal Labor Supply Disincentives of Welfare Reforms." NBER Working Paper 26028.

Mortenson, Jacob A, and Andrew Whitten. 2020. "Bunching to Maximize Tax Credits: Evidence from Kinks in the US Tax Schedule." *American Economic Journal: Economic Policy* 12 (3): 402–32.

Nichols, Austin, and Jesse Rothstein. 2016. "The Earned Income Tax Credit." In *Economics of Means-Tested Transfer Programs in the United States, Volume I*, edited by Robert A. Moffit. University of Chicago Press.

Rogerson, Richard, and Johanna Wallenius. 2009. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." *Journal of Economic Theory* 144 (6): 2277–2292.

Saez, Emmanuel. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *The Quarterly Journal of Economics* 117 (3): 1039–1073.

———. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

Schanzenbach, Diane, and Michael R Strain. 2021. "Employment Effects of the Earned Income Tax Credit: Taking the Long View." *Tax Policy and the Economy* 35 (1): 87–129.

Scholz, John Karl. 1994. "The Earned Income Tax Credit: Participation, Compliance, and Antipoverty Effectiveness." *National Tax Journal*, pp. 63–87.

The Economist. 2020. "From Unthinkable to Universal: Universal Basic Income Gains Momentum in America." *The Economist*, Aug.

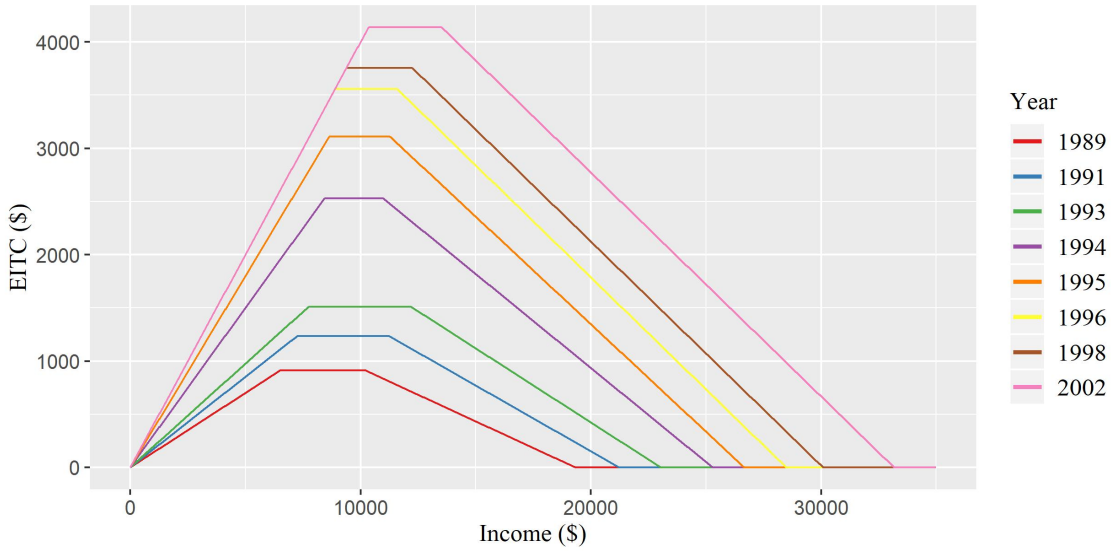
US Department of Health and Human Services. 1997. "Setting the Baseline: A Report on State Welfare Waivers." *Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation, June*.

Appendices

A Institutional Background: The Welfare-to-Workfare Transition

The Earned Income Tax Credit (EITC) is one of the largest income support programs in the United States. Enacted in 1975 to provide a modest supplement to the income of working families, it has been expanded significantly by the federal government in several rounds, most notably in 1986, 1993, and 2009. While the EITC expanded in each year during the period we study, 1988-2002, Figure A-1 shows that the largest year-to-year expansions of the program occurred in precisely those years, 1993-1996, that we define to be the welfare-to-workfare transition. In the 2000s, many states also implemented and expanded their own EITC programs. Throughout the paper, we calculate an individual's EITC benefits to be the sum of federal and state EITC credits.

Figure A-1: EITC Schedule: Selected Years



The figure shows the relationship between earned income and the federal Earned Income Tax Credit in selected years. Amounts are expressed in nominal US dollars.

Unlike the EITC, welfare has historically provided benefits to mothers who do not work. In response to growing concerns that welfare's incentives were contributing to high unemployment and out-of-wedlock births, many states implemented reforms between 1992 and 1996. These so-called welfare waiver reforms contained a mix of punishments and incentives to get mothers off of the welfare rolls and into employment. They had five main characteristics: time limits, changes in exemptions from program requirements, sanctions for recipients who violated program requirements, family caps, and earnings disregards.

While Aid For Families with Dependent Children (AFDC) did not impose limits on how long beneficiaries could receive welfare, many of the welfare waivers restricted the receipt of benefits to specific periods of time, such as 24 out of every 48 months. There were three types of time limits. "Termination" time limits resulted in the loss of benefits after the limit had been reached, while "reduction" time limits caused a reduction in benefits, and "work requirement" limits did not cut off aid so long as the beneficiaries complied with state-stipulated work requirements. These time limits were not retroactive. As a result, very few people were kicked off the welfare rolls before 1997.

AFDC required states to run education and jobs training programs (JOBS) for welfare recipients, and participation in JOBS (or similar activities such as secondary education and job search) was mandatory for nonexempt individuals. Federal policy exempted recipients if their youngest child was under the age of 3, but many state waivers lowered the age exemption and imposed sanctions on individuals who violated this requirement. In the most severe cases, repeated violations could result in the lifetime termination of benefits.

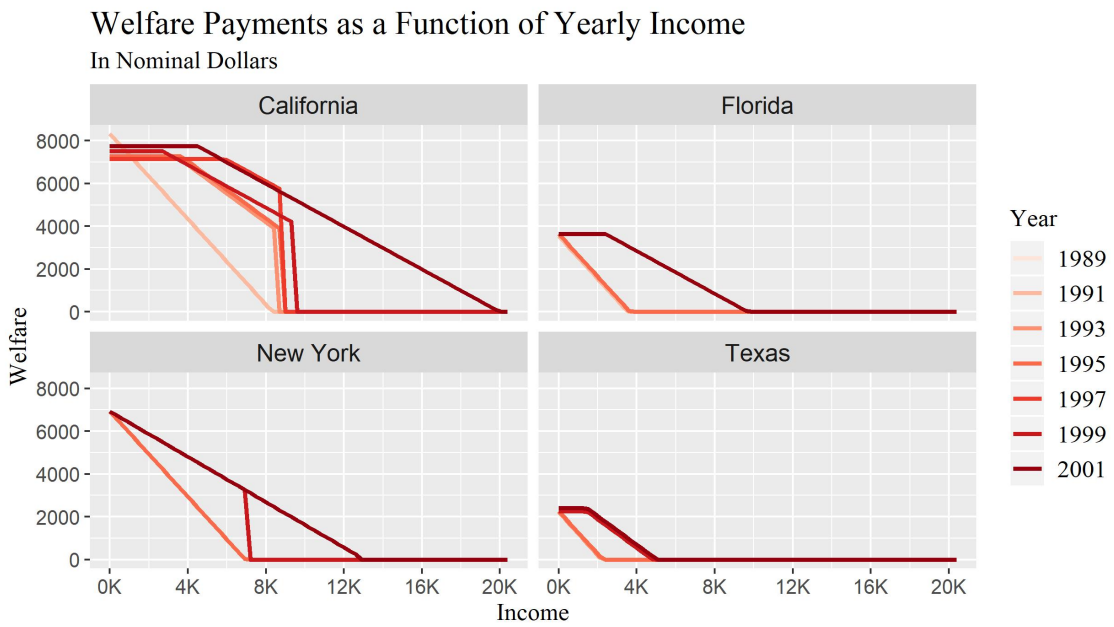
While AFDC stipulated that benefits increase with the number of children, several states instituted family caps that froze benefit levels if a recipient had a child while currently receiving welfare.

Lastly, under AFDC, welfare recipients faced a 100 percent marginal tax rate: Benefits were reduced one-to-one with each dollar earned through employment. Many state waivers countered this disincentive to work in two ways. Earnings *dollar* disregards allowed recipients to earn a fixed amount of money before the benefits were reduced, while earnings *rate* disregards reduced the marginal tax

rate on remaining earnings to below 100 percent. Many states implemented both dollar and rate disregards simultaneously. Michigan, for example, disregarded the first $200 of monthly income and lowered the marginal tax rate to 80 percent on the remaining income. Unlike the previous four characteristics of the waivers, disregards represent the use of carrots, rather than sticks, to provide incentives for welfare recipients to participate in the labor market.

Figure A-2 displays welfare's schedule of benefits for a mother with two children and no nonlabor income in four large states: California, Florida, New York, and Texas. Earnings dollar disregards introduce kinks in the schedule, while earnings rate disregards reduce the magnitude of the slope.

Figure A-2: Welfare Functions by State and Year



The first states to implement the welfare waivers—New Jersey, California, and Michigan—did so in 1992. However, the majority—23 of 30 statewide reforms—were implemented between 1994 and 1996, precisely when the EITC experienced the most dramatic expansions.

The welfare waivers culminated in the passage by Congress of the Personal Work Opportunity and Reconciliation Act of 1996 (PRWORA). PRWORA replaced AFDC with TANF. Under TANF, each state received a block grant and

was given substantial leeway in designing its TANF program. All five of the main characteristics of the statewide waivers found their way into PRWORA. States were no longer required to provide more benefits to larger families, and family caps were implemented in many states. A federal lifetime limit of five years on the receipt of benefits was instituted, although many states imposed more stringent limits. Earned income disregards became the rule rather than the exception, and many states adopted graduated sanctions, some of which could ultimately result in the lifetime loss of benefits. Many of the states that implemented waiver reforms retained these policies as part of their TANF programs. Others made modifications.

Parceling out the effects of each of the five types of welfare waiver reforms on labor supply is a difficult task and one that we do not pursue in this paper. However, it at least seems likely that the effects of the welfare waiver reforms on labor supply between 1994 and 1996 stressed in Kleven (2020) were not caused by lifetime limits. The welfare waivers imposed lifetime limits on benefits that were not retroactive, meaning that few people were kicked off the welfare rolls prior to 1997 (see Appendix F and US Department of Health and Human Services (1997)). The disregards took effect earlier and could account for increased labor supply during this period. However, in reducing the effective tax on earned income, they made welfare operate more like the EITC in its use of financial incentives designed to draw beneficiaries into the labor force.

B Formal Analysis of DiD Estimand

Proof of Proposition 1.

Proof. We consider a two-period model and define an indicator variable $D_{0,0}$ that equals 1 if an individual receives no subsidy in either period and 0 otherwise. The variable denoting treatment is then given by $Treat = 1 - D_{0,0}$. We define the following terms: for $j = 1, \ldots, J$, $p_{j,t} \equiv \mathbb{P}(D_{j,t} = 1)$, $p_{j,t-1} \equiv \mathbb{P}(D_{j,t-1} = 1)$, and $p_{0,0} \equiv \mathbb{P}(D_{0,0} = 1)$. Note that $1 - p_{0,0} = p_{Treat}$.

The DiD estimand is

$$\beta_3^{DiD} = \frac{cov(Y_t - Y_{t-1}, Treat)}{var(Treat)} \ . \tag{B-1}$$

The denominator in B-1 is equal to $p_{0,0}(1 - p_{0,0})$. The numerator can be written as

$$\begin{aligned} cov(Y_t - Y_{t-1}, Treat) &= cov(Y_t - Y_{t-1}, 1 - D_{0,0}) \\ &= cov(Y_{0,t} - Y_{0,t-1}, 1 - D_{0,0}) \\ &+ cov(\sum_{j=1}^{J} D_{j,t}(Y_{j,t} - Y_{0,t}), 1 - D_{0,0}) \\ &- cov(\sum_{j=1}^{J} D_{j,t-1}(Y_{j,t-1} - Y_{0,t-1}), 1 - D_{0,0}) \ , \end{aligned} \tag{B-2}$$

The first expression in B-2 is zero because of the parallel trend assumption. The second term can be simplified as follows:

$$cov(\sum_{j=1}^{J} D_{j,t}(Y_{j,t} - Y_{0,t}), 1 - D_{0,0}) = -\sum_{j=1}^{J} cov(D_{j,t}(Y_{j,t} - Y_{0,0}), D_{0,0}) \ ,$$

where

$$\begin{aligned} cov(D_{j,t}(Y_{j,t} - Y_{0,t}), D_{0,0}) &= \mathbb{E}[D_{j,t}(Y_{j,t} - Y_{0,t})D_{0,0}] - \mathbb{E}[D_{j,t}(Y_{j,t} - Y_{0,t})]\mathbb{E}[D_{0,0}] \\ &= -\mathbb{E}[Y_{j,t} - Y_{0,t} | D_{j,t} = 1] p_{j,t} p_{0,0} \ . \end{aligned}$$

The second equality follows from the first because $\mathbb{E}[D_{j,t}(Y_{j,t} - Y_{0,t})D_{0,0}] = 0$ for

all $j \geq 1$.

The third term in B-2 similarly simplifies to

$$\sum_{j=1}^{J} cov(D_{j,t-1}(Y_{j,t-1} - Y_{0,t-1}), D_{0,0}) = -\sum_{j=1}^{J} \mathbb{E}[Y_{j,t-1} - Y_{0,t-1}|D_{j,t-1} = 1]p_{j,t-1}p_{0,0} \ .$$

Combining all the terms in the numerator with the denominator yields

$$\beta_3^{DiD} = \frac{1}{1 - p_{0,0}} \sum_{j=1}^{J} (p_{j,t}\Delta_{j,t}^{TT} - p_{j,t-1}\Delta_{j,t-1}^{TT}) \ . \tag{B-3}$$

□

No Pre-Existing Policy Regime. If a program is introduced at time t for the first time, so that no individuals in the treatment group were exposed to the program prior to the reform ($p_{j,t-1} = 0$ for all $j \neq 0$), DiD identifies ATT: $\beta_3^{DiD} = \sum_{j=1}^{J} \omega_{j,t}\Delta_{j,t}^{TT}$. The weights, $\omega_{j,t} = \frac{p_{j,t}}{\sum_{j=1}^{J} p_{j,t}}$, are given by the fraction of the treated population receiving each level of treatment, and they sum to one.

DiD Interpretation with Time-Invariant Potential Outcomes. Suppose that the potential outcomes are time-invariant, $Y_{j,t} = Y_{j,t-1} = Y_j \ \forall j \in \mathcal{J}$. In this case, the DiD estimand is a function of the marginal effects on the treated generated by the reform: $\beta_3^{DiD} = \frac{1}{p_{Treat}} \sum_{j=1}^{J} \sum_{h \neq j} \phi_{j,h} E[Y_j - Y_h | D_{j,t} = 1, D_{h,t-1} = 1]$, where $p_{Treat} = \mathbb{P}(Treat = 1)$, and $\phi_{j,h} = \mathbb{P}(D_{j,t} = 1, D_{h,t-1} = 1)$.

With time-invariant potential outcomes, DiD identifies ATPR when both $D_{h,t}^* = D_{h,t-1}$ for $h = 1, \ldots, J$ and for $\sum_{j=1}^{J} \sum_{h \neq j} \phi_{j,h} = p_{Treat}$. The first assumption means that individuals must have the same treatment level in time $t-1$ as they would have in time t absent the reform, while the second means that every individual must change the amount of EITC benefits they receive between periods $t-1$ and t because of the reform. These assumptions are unlikely to hold true in many empirical applications in the EITC literature: In Section 3.3 we show that the fraction of single mothers with no change in EITC benefits over time is sizable. Additionally, restricting potential outcomes over time may be inconsistent with changes in the rest of the tax and transfer programs, where an individual's

labor supply for a given level of EITC benefits can change because of incentives from other social program reforms.[24]

Similar restrictions are made in De Chaisemartin and D'HaultfŒuille (2017), who analyze the identification of treatment effects in a fuzzy DiD design. Although we acknowledge the relevance of their results, we also believe that their proposed estimators are unlikely to be applicable in the case of social welfare reforms. First, both single mothers and single women who receive no EITC benefits before the reform are likely to differ in their subsequent labor supply after the reform, violating their conditional common trend assumption. This can be driven by differences in preferences, as well as by other changes in the tax and transfer system, such as the TANF reform, that affect single mothers but not single women. Moreover, De Chaisemartin and D'HaultfŒuille (2017) consider the average causal response (ACR) as the targeted causal parameter of the proposed estimands. This causal parameter is not well defined in our framework, as social welfare reforms like the EITC reform typically cause more than a one-unit increment in the treatment intensity, generating potential overlapping of complier groups among different margins of the treatment (see Angrist and Imbens 1995).

If a fraction of individuals in the treatment group experience no change in EITC benefits, then DiD in the presence of time-invariant potential outcomes only identifies a quantity that is proportional to the ATPR. This occurs because $\sum_{j=0}^{J}\sum_{h\neq j}\phi_{j,h} < p_{Treat}$. In this case, the researcher may rescale the estimated DiD parameter to identify ATPR= $\frac{p_{Treat}}{\sum_{j=0}^{J}\sum_{h\neq j}\phi_{j,h}}\widehat{\beta}_3^{DiD}$. This result is similar to the intention-to-treat analysis in RCTs with imperfect compliance, and the estimated DiD cannot be directly interpreted as the ATPR without this transformation.

Imperfect Compliance. Suppose that only a subsample of individuals in the treatment group receive any treatment. This is the setting where single motherhood—rather than actual benefit receipt—is defined as the treatment indicator. We define treatment indicators D_j^{τ}, for every EITC benefit $j \in \{0, 1, \ldots J\}$ of treated individuals and retain $D_{0,t} \in \{0, 1\}$ to denote membership in the control

[24] An alternative restriction on the structure of the model would be to assume that trends are independent of the level of the benefits. Similarly, this restriction is not suitable when the labor supply incentives generated by the EITC for the treatment group interact with simultaneous changes in other tax and transfer programs.

group. The treatment indicators are mutually exclusive and collectively exhaustive for each period t so that $D_{0,t} + \sum_{j=0}^{J} D_{j,t}^{\tau} = 1$. We can write the observed outcome as a function of treatment assignments and potential outcomes:

$$Y_t = Y_{0,t} + \sum_{j=0}^{J} D_{j,t}^{\tau}(Y_{j,t}^{\tau} - Y_{0,t}). \tag{4}$$

Proposition 2. *Suppose Assumption 2 holds. Then, the DiD estimand in this case is equal to:*

$$\beta_3^{DiD} = \frac{1}{p_{Treat}} \sum_{j=0}^{J} (p_{j,t}^{\tau} \Delta_{j,t}^{TT,\tau} - p_{j,t-1}^{\tau} \Delta_{j,t-1}^{TT,\tau}) \ , \tag{5}$$

where $p_{Treat} = \mathbb{P}(Treat = 1)$, $p_{j,t}^{\tau} = \mathbb{P}(D_{j,t}^{\tau} = 1)$, *and* $\Delta_{j,t}^{TT,\tau} = E[Y_{j,t}^{\tau} - Y_{0,t} | D_{j,t}^{\tau} = 1]$ *for* $j \in \mathcal{J}$.

Proof. Analogous to proof of Proposition 1. □

Proposition 2 is similar to Proposition 1 in that it highlights that the DiD estimand fails to identify any of the causal parameters of interest without additional restrictions. However, the implications differ if the policy regime did not exist prior to the reform. With imperfect compliance and no pre-existing policy regime, the DiD estimand still does not identify ATT unless one makes the additional assumption that the behavior of the treatment group and the control group are identical in the case of no EITC benefits, $Y_{0,t} = Y_{0,t}^{\tau}$. When analyzing the EITC, this means that single mothers and single women without children are assumed to have the same counterfactual outcomes without any tax credits. This restriction is much stronger than Assumption 2 and unlikely to be satisfied, as the literature on female labor supply has documented large differences in labor supply between single women and mothers.

Exploiting Variation in Welfare Waivers. In the main text we discuss how Kleven (2020), Grogger (2003), and Meyer and Rosenbaum (2001) exploit the uneven introduction of welfare waiver reforms across states in an effort to identify the effect of the EITC on labor supply. In this framework, let $W_{i,t} \in \{0, 1\}$ denote

whether individual i resides in a state that had implemented welfare waiver reforms ($W_{i,t} = 1$) or not ($W_{i,t} = 0$) by time t. The DiD model including the effects of waivers is

$$\begin{aligned} Y_{i,t} =& \beta_0 + \beta_1 Post_t + \beta_2 Treat_i + \beta_3 Post_t \times Treat_i + \\ & \beta_4 W_{i,t} + \beta_5 W_{i,t} \times Post_t + \beta_6 W_{i,t} \times Treat_i + \beta_7 W_{i,t} \times Post_t \times Treat_i + \varepsilon_{i,t} \,, \end{aligned} \tag{6}$$

where $Treat_i$ is a time-invariant binary variable equaling one if the individual receives EITC benefits at any point in the study, and $Post_t$ is a binary variable equaling one in the period after the reform.

In this setting, Proposition 3 reveals that the DiD estimand when controlling for waivers is still a difference in weighted treatment on the treated parameters for each margin of the EITC program.

Proposition 3. *Suppose assumption 2 holds. Then, the DiD estimand from model (10) is:*

$$\beta_3^{DiD} = \frac{1}{p_{Treat}} \sum_{j=1}^{J} p_{j,t}^{W=0} \Delta_{j,t}^{TT,W=0} - p_{j,t-1}^{W=0} \Delta_{j,t-1}^{TT,W=0} \,,$$

where $p_{Treat} = \mathbb{P}(Treat = 1 | W_{i,t} = 0)$*, and for each* $j = 1 \ldots, J$*,* $p_{j,t}^{W=0} = \mathbb{P}(D_{j,t} = 1 | W_{i,t} = 0)$*,* $\Delta_{j,t}^{W=0} = \mathbb{E}(Y_{j,t} - Y_{0,t} | W_{i,t} = 0)$*,* $p_{j,t-1}^{W=0} = \mathbb{P}(D_{j,t-1} = 1 | W_{i,t} = 0)$*, and* $\Delta_{j,t-1}^{W=0} = \mathbb{E}(Y_{j,t-1} - Y_{0,t-1} | W_{i,t} = 0)$*.*

Proof. β_3^{DiD} can be consistently estimated by a DiD design that conditions on the subsample with $W_{i,t} = 0$. The rest of the proof is then analogous to the proof of proposition 1 □

As explained in the main text, implementing (10) via an event study design that replaces $Post_t$ with dummy variables for each year confounds changes in employment over time resulting from the reform with changes in the average employment level that result from conditioning on a different set of states. This occurs because each coefficient in the event study is identified by conditioning

on the subset of states that had not implemented any welfare waivers by the year in question. Implementing DiD when controlling for waivers therefore fails to identify either ATT or ATPR and may actually be more biased than a simple DiD that does not control for waivers.

C Reconciling Previous Literature with Our Results

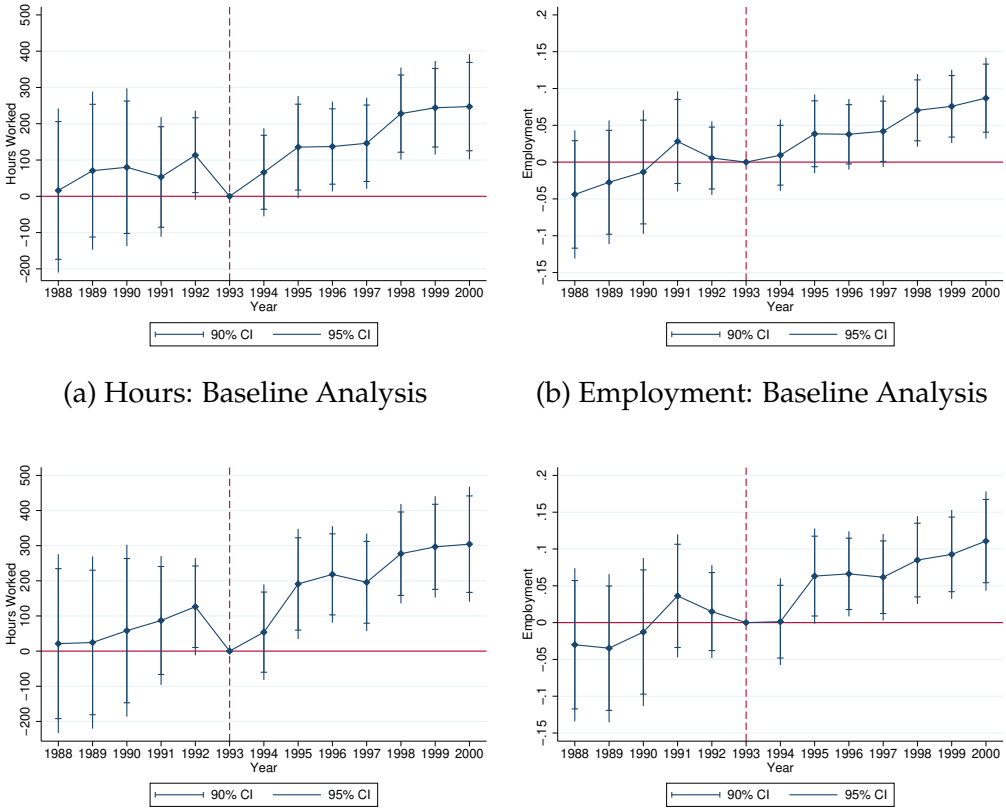
In this section, we analyze how estimates of labor suppply responses are affected by a DiD/event study design that neglects heterogeneity in social programs benefits and considers all single mothers as treated and all single women as controls. To this aim, we replicate the research design in Kleven (2020) and test whether it is robust to the exclusion of single mothers who received no benefits from the treatment group. If the labor supply of single mothers were unaffected by the EITC expansion, the event study would be robust to this exclusion. To perform the event study, we have augmented the sample by including single women without children, a group that was not part of our empirical analysis in Section 3.2. Single women without children form the control group since, prior to 1993, the presence of at least one dependent child was a requirement for EITC eligibility.

Figure C-1 displays event studies of the 1993 EITC expansion. The results are obtained by regressing each outcome on the interactions between the indicator variable for treatment and indicator variables for each year, after controlling for state waivers and unemployment levels. The only difference between the analysis here and that in Kleven (2020) is our use of the ASEC instead of the full CPS.

Figure C-1-a shows that in the post-reform period, the treatment effect of being a single mother is positive (135 additional hours per year) and statistically significant at the ten percent level in 1995 and at the five percent level in 1996. Figure C-1-b replicates the analysis on the extensive margin done in Kleven's work and produces point estimates similar to those in his study, but with larger standard errors due to the smaller sample size. The figure provides the same qualitative conclusions as the analysis of hours worked, but the point estimates are noisy, and we cannot reject the hypothesis that the 1993 EITC reform had no impact on the extensive margin.

Figures C-1-c and C-1-d replicate the above analysis on a sample that excludes single mothers with no change in EITC benefits according to the variable plotted in Figure 3. Already in 1995, single mothers display a large and statistically significant increase of about 200 yearly hours worked relative to the control group.

Figure C-1: Event Study Analysis of the 1993 EITC Reform



(a) Hours: Baseline Analysis

(b) Employment: Baseline Analysis

(c) Hours: "As Treated" Analysis

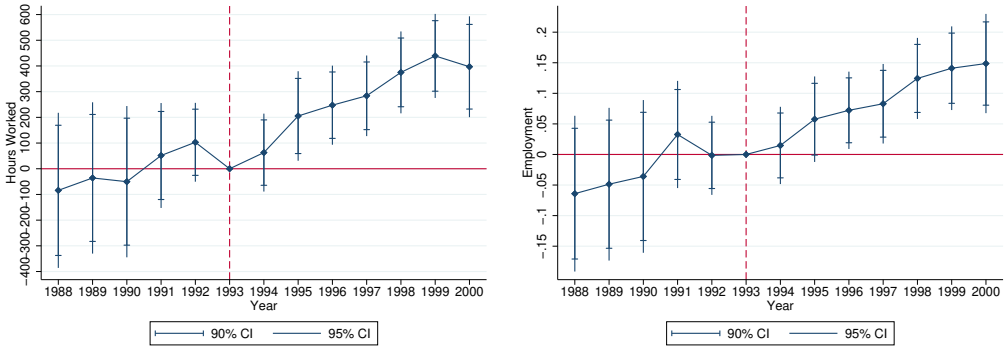
(d) Employment: "As Treated" Analysis

The figure compares two event study analyses of the effect of the 1993 EITC reform on yearly hours worked (Panels (a) and (c)) and employment (Panels (b) and (d)). In Panels (a) and (b), the treatment group is comprised of single mothers and the control group is comprised of single women without children. The specifications in Panels (c) and (d) exclude single mothers who were unexposed to changes in EITC benefits caused by the 1993 EITC reform. Details on the definition of the group of single mothers unexposed to policy-induced changes in EITC benefits are provided in Section 3.3. Yearly hours worked (in Panels (a) and (c)) and an indicator variable for employment status (in Panels (b) and (d)) are regressed on a set of interaction terms of the indicator variable for the treatment group and indicator variables for each year in the period 1988-2000. The event study specification in each panel also contains control variables for the number of dependent children (indicators), state fixed effects, state unemployment level and state welfare waivers (indicator). The year of the reform's passage, 1993, is the reference year for the analysis. The red vertical line separates the pre-reform (1993 and earlier) period from the post-reform period. Each panel shows the point estimates for the treatment effect of the reform together with the 90 and 95 percent confidence intervals based on standard errors clustered at the individual level.

Results are similar for employment, with a statistically significant six percentage point increase in the probability of being employed for exposed single mothers relative to single women, representing an eight percent change in employment probability relative to the pre-reform mean. The pre-trends have not changed qualitatively relative to our baseline analysis in Figures C-1-a and C-1-b. The analysis displays similar results if single mothers unaffected by the policy reform, e.g. with earnings above the EITC threshold, were assigned to the control group. As a further test, Figure C-2 replicates the analysis on the subsample of single mothers and single women without children reporting labor income below the EITC eligibility threshold of $30,000. The analysis displays the absence of differential trends through 1993 by treatment status and sizable positive treatment effects in the post-reform period for both hours worked (Figure C-2-a) and employment status (Figure C-2-b).

Altogether, this section shows that the finding of an insignificant effect of the EITC on labor supply depends on the inclusion of a large number of untreated individuals in the treatment group. Despite the strong positive marginal effects of the EITC on labor supply that we find throughout this paper, DiD estimates zero empirical effect of the policy because the design averages the effects of people who are expected to respond to it with those for whom no response is expected.

Figure C-2: Event Study Analysis of the 1993 EITC Reform

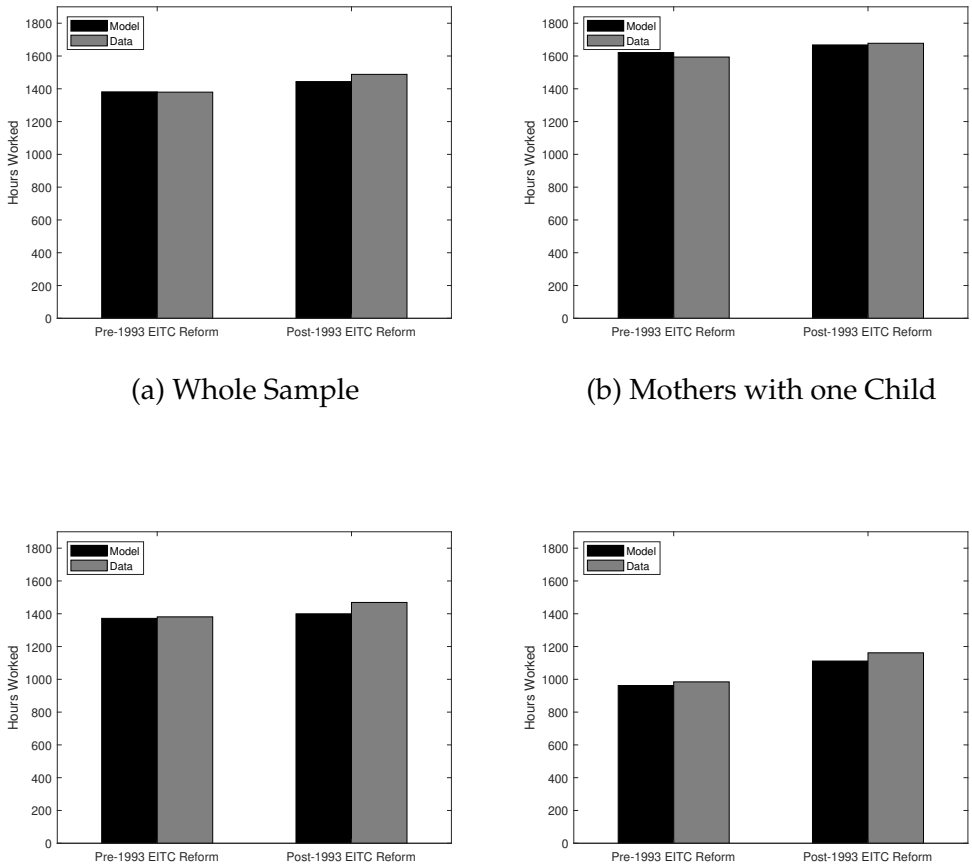


(a) Hours: Low-Income Subgroup (b) Employment: Low-Income Subgroup

The figure shows the event study analysis of the effect of the 1993 EITC reform on yearly hours worked (Panel (a)) and employment (Panel (b)) of single mothers. The analysis is run on the subsample of single mothers and single women without children with labor income below $30,000 (EITC eligibility threshold). The treatment group is made of single mothers and the control group is made of single women without children. Yearly hours worked (Panel (a)) and an indicator variable for employment status (Panel (b)) are regressed on a set of interaction terms between the indicator variable for the treatment group (single mothers) and indicator variables for each year in the period 1988-2000. The event study specification in each panel also contains control variables for the number of dependent children (indicators), state fixed effects, state unemployment level and state welfare waivers (indicator). The year of the reform, 1993, is the reference year for the analysis. The red horizontal line separates the pre-reform (pre-1993) period from the post-reform period. Each panel shows the point estimates for the treatment effect of the reform together with the 90 and 95 percent confidence intervals based on standard errors clustered at the individual level.

D Additional Tables and Figures

Figure D-1: Validation: Predicted Hours Worked Pre- and Post-1993 EITC Reform



(a) Whole Sample

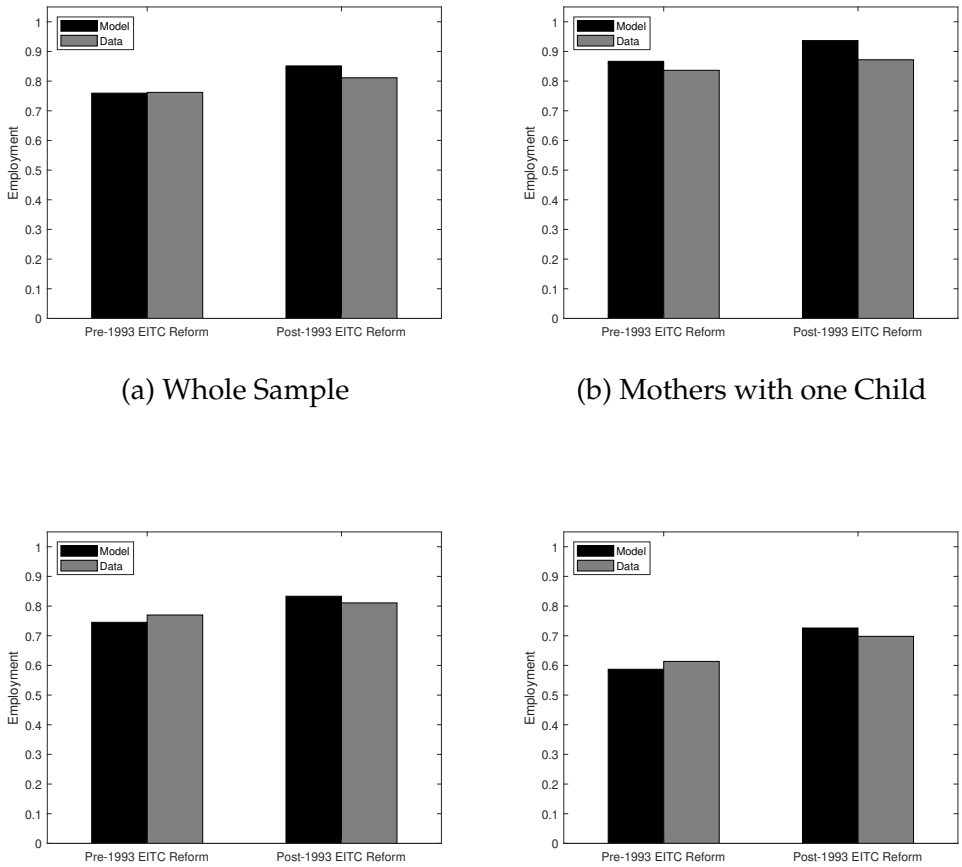
(b) Mothers with one Child

(c) Mothers with two Children

(d) Mothers with three Children

The figure shows the model's predictions for hours worked. In each graph, the first set of bars shows the fit of the model for the years 1988-1993, prior to the implementation of the 1993 EITC reform, while the second set of bars shows the performance of the model for the untargeted moments of hours worked in 1995-1996. The analysis is performed on the whole sample of single mothers (Panel (a)) and the sample of single mothers with one child (Panel (b)), two children (Panel (c)), and three or more children (Panel (d)). The figure displays yearly hours worked by single mothers as predicted by the model (black bars) and as observed in the data (gray bars).

Figure D-2: Validation: Predicted Employment Pre- and Post-1993 EITC Reform



(a) Whole Sample

(b) Mothers with one Child

(c) Mothers with two Children

(d) Mothers with three Children

The figure shows the model's predictions for employment. In each graph, the first set of bars shows the fit of the model for the years 1988-1993, prior to the implementation of the 1993 EITC reform, while the second set of bars shows the performance of the model for the untargeted moments of employment in 1995-1996. The analysis is performed on the whole sample of single mothers (Panel (a)) and the sample of single mothers with one child (Panel (b)), two children (Panel (c)), and three or more children (Panel (d)). The figure displays the employment rate of single mothers as predicted by the model (black bars) and as observed in the data (gray bars).

Table D-1: Summary Statistics

Sample of Single Mothers

	Mean	Standard Deviation
Employment Rate	0.792	0.406
Yearly Hours Worked	1440.430	940.395
Earnings ($ 1000s)	24.552	22.924
EITC Benefits ($ 1000s)	1.008	1.458
Welfare Benefits ($ 1000s)	1.974	3.909
One Child	0.426	0.494
Two Children	0.348	0.476
Three or More Children	0.226	0.418
White	0.661	0.473
Black	0.300	0.458
Other Races	0.039	0.194

This table presents descriptive statistics for the sample of single mothers used in estimation. All monetary values are expressed in thousands of 2015 US dollars.

Table D-2: EITC Benefits, Welfare Benefits and Hours Worked per Year, Alternative Prediction Model I

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Yearly Hours Worked				
Change in EITC Benefits ($ 1000s)	126.20***	110.30***	115.84***	115.07***	110.76***
	(7.42)	(8.69)	(9.01)	(9.02)	(9.62)
Change in Welfare Benefits ($ 1000s)		-12.45***	-17.57***	-17.62***	-17.00***
		(2.55)	(2.84)	(2.84)	(3.06)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	Yes	Yes	Yes
Unemployment and Waiver Controls	No	No	No	Yes	Yes
Prior (t-1) Income	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly hours worked by single mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. The table differs from Table 1 in the main text in that it uses a fourth order polynomial and an indicator for lagged income to predict second period income for the purposes of computing policy-induced changes in EITC benefits and welfare benefits. See Section 3.1 and Appendix E for details. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-3: EITC Benefits, Welfare Benefits and Hours Worked per Year, Alternative Prediction Model II

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Yearly Hours Worked				
Change in EITC Benefits ($ 1000s)	115.19***	98.48***	102.24***	101.48***	92.57***
	(7.17)	(8.35)	(8.65)	(8.66)	(9.13)
Change in Welfare Benefits ($ 1000s)		-14.04***	-16.65***	-16.92***	-15.06***
		(2.72)	(2.92)	(2.94)	(3.17)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	Yes	Yes	Yes
Unemployment and Waiver Controls	No	No	No	Yes	Yes
Prior (t-1) Income	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly hours worked by single mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. The table differs from Table 1 in the main text in that it includes additional controls for race and the number of children in the income prediction equation. See Section 3.1 and Appendix E for details. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-4: EITC Benefits, Welfare Benefits and Hours Worked per Year, Alternative Prediction Model III

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Yearly Hours Worked				
Change in EITC Benefits ($ 1000s)	101.15***	81.91***	84.58***	83.99***	73.72***
	(7.15)	(8.15)	(8.32)	(8.32)	(8.58)
Change in Welfare Benefits ($ 1000s)		-20.70***	-22.92***	-23.05***	-19.49***
		(3.47)	(3.66)	(3.67)	(3.91)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	Yes	Yes	Yes
Unemployment and Waiver Controls	No	No	No	Yes	Yes
Prior (t-1) Income	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly hours worked by single mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. The table differs from Table 1 in the main text in that it includes additional controls for race, number of children, and educational attainment (less than high school, high school, some college, college) in the income prediction equation. See Section 3.1 and Appendix E for details. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-5: EITC Benefits, Welfare Benefits and Hours Worked per Year, Alternative Prediction Model IV

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Yearly Hours Worked				
Change in EITC Benefits ($ 1000s)	94.78***	89.29***	92.06***	91.74***	88.33***
	(4.71)	(5.93)	(6.09)	(6.10)	(6.41)
Change in Welfare Benefits ($ 1000s)		-3.98**	-8.34***	-8.19***	-7.99***
		(1.98)	(2.24)	(2.25)	(2.43)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	Yes	Yes	Yes
Unemployment and Waiver Controls	No	No	No	Yes	Yes
Prior (t-1) Income	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly hours worked by single mothers. The dependent variable is the year-on-year change in hours worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. The table differs from Table 1 in the main text in that it predicts second period income for single mothers using the estimates from an income prediction model estimated on the sample of single women. The income prediction model is a fifth-order polynomial in lagged income together with a dummy variable for positive lagged values. See Section 3.1 and Appendix E for details. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-6: EITC Benefits,Welfare Benefits and Weeks Worked per Year

	(1)	(2)	(3)	(4)	(5)
	Outcome: Change in Weeks Worked per Year				
Change in EITC Benefits ($ 1000s)	3.016***		2.578***	2.688***	2.680***
	(0.180)		(0.212)	(0.218)	(0.219)
Change in Welfare Benefits ($ 1000s)		-0.800***	-0.342***	-0.467***	-0.471***
		(0.056)	(0.066)	(0.071)	(0.071)
N	10959	10959	10959	10959	10959
Controls and State F.E.	No	No	No	Yes	Yes
Unemployment and Waiver Controls	No	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits and welfare benefits on the change in yearly weeks worked by single mothers. The dependent variable is the year-on-year change in weeks worked by single mothers. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Control variables include mother's race, indicator variables for the number of dependent children, and year fixed effects. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-7: EITC Benefits and Hours Worked per Year by EITC Schedule Position

	(1)	(2)	(3)	(4)
	Outcome: Change in Yearly Hours Worked			
EITC Phase-In × Change in EITC Benefits ($ 1000s)	74.82***	71.48***	70.96***	62.83***
	(14.51)	(16.02)	(16.03)	(15.86)
EITC Plateau × Change in EITC Benefits ($ 1000s)	32.09	45.49	43.37	41.63
	(34.59)	(34.96)	(35.06)	(35.04)
EITC Phase-Out × Change in EITC Benefits ($ 1000s)	140.99***	155.01***	154.56***	175.51***
	(23.75)	(23.92)	(23.91)	(24.51)
N	7721	7721	7721	7721
Mean Change in Hours Worked (Phase-In)	172.74	172.74	172.74	172.74
Mean Change in EITC Benefits ($ 1000s, Phase-In)	1.353	1.353	1.353	1.353
Mean Change in Hours Worked (Plateau)	-108.37	-108.37	-108.37	-108.37
Mean Change in EITC Benefits ($ 1000s, Plateau)	-0.069	-0.069	-0.069	-0.069
Mean Change in Hours Worked (Phase-Out)	-104.06	-104.06	-104.06	-104.06
Mean Change in EITC Benefits ($ 1000s, Phase-Out)	-0.247	-0.247	-0.247	-0.247
Controls and State F.E.	No	Yes	Yes	Yes
Unemployment and Waiver Controls	No	No	Yes	Yes
Prior (t-1) Income	No	No	No	Yes

The table shows the causal effect of changes in policy-induced EITC benefits on the change in yearly hours worked by single mothers, controlling for the initial period position $(t-1)$ in the EITC benefit schedule. The dependent variable is the year-on-year change in hours worked by single mothers. Control variables include mother's race, indicator variables for the number of dependent children, year fixed effects, and policy-induced changes in welfare benefits (in 2015 US dollars). Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. Unemployment and waiver controls include controls for state unemployment level and indicator variables for state welfare waivers. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D-8: In-sample Fit for Hours, Employment and Wages

	(1) Model	(2) Data
	Panel A: Hours Worked	
Mean Hours Worked	1381.2	1379.6
Mean Hours Worked (One Child)	1621.0	1593.7
Mean Hours Worked (Two Children)	1371.9	1381.0
Mean Hours Worked (Three Children)	962.3	984.4
	Panel B: Employment Rate	
Employment Rate	0.759	0.762
Employment Rate (One Child)	0.867	0.837
Employment Rate (Two Children)	0.745	0.770
Employment Rate (Three Children)	0.587	0.614
	Panel C: Accepted Wages	
Mean Accepted Wage	16.09	16.11
SD Accepted Wage	9.62	9.60
Mean Accepted Wage (One Child)	16.46	17.02
Mean Accepted Wage (Two Children)	16.79	16.33
Mean Accepted Wage (Three Children)	13.77	13.41
Autocovariance Accepted Log-Wages	0.315	0.222
SD Accepted Log-Wage	0.56	0.64

The table shows the in-sample fit for hours worked by single mothers (Panel (a)), employment rate (Panel (b)), and accepted wage (Panel (c)). The table displays outcomes as predicted by the model (column 1) and as observed in the data (column 2). All monetary values are expressed in 2015 US dollars.

Table D-9: In-sample Fit for Regression of Hours on EITC and Welfare

	(1)	(2)
	Outcome: Yearly Hours Worked	
	Model	Data
Change in EITC Benefits ($ 1000s)	214.02	175.29
Change in Welfare Benefits ($ 1000s)	-19.44	-11.74

The table shows the in-sample fit for the regression of year-on-year changes in hours worked by single mothers on policy-induced changes in EITC and welfare benefits. Policy-induced changes in EITC and welfare benefits are expressed in thousands of 2015 US dollars. Details on the construction of the variable for policy-induced changes in EITC and welfare benefits are provided in Section 3.1 and Appendix E. The table displays outcomes as predicted by the model (column 1) and as observed in the data (column 2).

E Tax and Transfer Rules and Construction of Variables

Construction of independent variables

The challenge in estimating the linear labor supply model in Equation (6),

$$\Delta h_{i,t} = \gamma_0 + \gamma_1 \Delta \xi_{i,t} + \gamma_2 \Delta T_{i,t} + \Delta \epsilon_{i,t} \,,$$

using observed changes in EITC and welfare benefits is that they depend on labor supply. The EITC and welfare influence maternal income in two ways: (i) directly through the transfer, and (ii) indirectly through the labor supply response. This second channel is the source of the endogeneity. Consider an individual with pre-tax earnings, $I_{i,t}^{pre-tax} \equiv \omega_{i,t} \cdot h_{i,t}$, and nonlabor income, $NL_{i,t}$. The benefit formulas for both subsidies depend on labor supply through her pre-tax earnings:

$$\begin{aligned} \xi_{i,t} &= \xi_{i,t}(I_{i,t}^{pre-tax}) = \xi_{i,t}(\omega_{i,t} \cdot h_{i,t}) \,, \\ T_{i,t} &= T_{i,t}(I_{i,t}^{pre-tax}, NL_{i,t}) = T(\omega_{i,t} \cdot h_{i,t}, NL_{i,t}) \,. \end{aligned}$$

To eliminate this source of endogeneity, we calculate policy-induced changes in benefits for each individual caused by variation in the EITC and welfare schedules over time. These policy-induced changes are calculated on the basis of *predicted* earnings and nonlabor income:

$$\Delta \xi_{i,t} \left(I_{i,t-1}^{pre-tax} \right) = \xi_{i,t} \left(\widehat{E} \left[I_{i,t}^{pre-tax} | I_{i,t-1}^{pre-tax} \right] \right) - \xi_{i,t-1} \left(I_{i,t-1}^{pre-tax} \right) \,, \tag{E-1}$$

$$\Delta T_{i,t} \left(I_{i,t-1}^{pre-tax} \right) = T_{i,t} \left(\widehat{E} \left[I_{i,t}^{pre-tax} | I_{i,t-1}^{pre-tax} \right], \widehat{E} \left[NL_{i,t} | NL_{i,t-1} \right] \right) - T_{i,t-1} \left(I_{i,t-1}^{pre-tax}, NL_{i,t-1} \right) \,. \tag{E-2}$$

We follow Dahl and Lochner (2012) and use a fifth-order polynomial in the lagged variable as well as an indicator for a positive lagged value to construct the conditional expectation.

$\Delta \xi_{i,t}$ and $\Delta T_{i,t}$ represent policy-induced changes in the benefits a mother would expect to receive based on first-period income. To the extent that these differ from zero, it is due to factors—such as shifts in policy—that are exogenous with respect to the mother's labor supply decision.

EITC, Welfare, and Food Stamp Formulas

Given $I_{i,t-1}^{pre-tax}$, $NL_{i,t-1}$, and estimates of $\widehat{E}[I_{i,t}^{pre-tax}|I_{i,t-1}^{pre-tax}]$ and $\widehat{E}[NL_{i,t}|NL_{i,t-1}]$, we calculate EITC benefits using NBER's TAXSIM and welfare benefits using the AFDC/TANF rules in effect for each year and state in which we observe the mother. Table A.1 in Kleven (2020) provides a detailed reference for the federal EITC parameters during the period we study.

The computation of welfare benefits depends on earnings, nonlabor income, number of children, and the individual's state of residence. Each state bases eligibility on whether both gross and net income fall below a threshold specific to the number of children in the family. If a family is eligible, they receive a benefit that depends on several parameters set by the state: the maximum allowable benefit (MB), the dollar and rate disregards to earnings (EDD and ERD) described in Appendix A, and the payment standard (PS). Given earnings ($I_{i,t}^{pre-tax}$) and nonlabor income ($NL_{i,t}$), net income is given by $Net_{i,t} = (I_{i,t}^{pre-tax} - EDD)(1-ERD) + NL_{i,t}$, and benefits are

$$Benefit_{i,t} = \max\{\min\{MB, PS - Net_{i,t}\}, 0\} \ . \tag{E-3}$$

All parameters vary substantially across states and years. Figure A-2 plots the welfare benefit function for four states in selected years between 1988 and 2002 for a mother with two children and no nonlabor income. MB determines the y-intercept, while ERD influences the slope of the function and a positive EDD induces benefits to be constant in earnings at low levels.

SNAP benefits (food stamps) enter into the individual's budget constraint in the model in Section 4. SNAP benefits depend on an individual's earnings, nonlabor income, and welfare benefits received. Provided that gross earnings are below 130 percent of the federal poverty line and earnings and nonlabor income net of welfare benefits ($NE_{i,t}$) are below 100 percent of the poverty line, an individual with k children receives SNAP benefits according to the formula:

$$SNAP_{i,t}(k) = \max\{MB_{i,t}(k) - 0.3 * NE_{i,t}(k), 0\} \ , \tag{E-4}$$

where $MB_{i,t}(k)$ is the maximum benefit for a family with k children in year t.

Parameters of Estimated Tax Function

We approximate mother's after-tax income by the parametric function in (12). For each year, state, and number of children, we use NBER's TAXSIM program to simulate the aftertax earnings of mothers with incomes at intervals of \$1,000 between \$0 and \$100,000. Then we estimate $\theta_{0,s,t,k}$ and $\theta_{1,s,t,k}$ $\forall s, t, k$ by minimizing the sum of squared residuals between actual after-tax income and the after-tax income predicted by the right-hand-side of Equation (12). Estimation is done by Nonlinear Least Squares.

F Welfare Time Limits

The table in this section documents the earliest possible date on which time limits might result in a welfare recipient being kicked off the welfare rolls (US Department of Health and Human Services 1997).

State	Extent	First Cases Reach Limit	Consequence
Arizona	Statewide	November-97	Adult portion of grant is terminated.
California	Statewide	August-97	Adults must participate in CWEP for 100 hours per month.
Colorado	Five counties	May-96	Nonexempt adults must be working at least 30 hours per week or actively participating in a JOBS training program.
Connecticut	Two cities: New Haven and Manchester	June-96	End of cash assistance.
Connecticut	Statewide	September-97	End of cash assistance.
Delaware	Statewide	November-97	Adult must enter pay-after-performance work experience program.
Delaware	Statewide	November-99	End of cash assistance.

Florida	Escambia & Alachua counties. Later expanded to six more counties.	February-96	End of cash assistance. Transitional employment will be offered to for those who have diligently completed plans, are unable to find employment and have not voluntarily quit or been discharged for misconduct.
Georgia	Ten counties	Between December-98 and December-99	Recipients must work 20 hours per month in a work experience program for a state, local government, federal agency or nonprofit organization, subject to availability of work slots.
Hawaii	Statewide	Between November-01 and November-02	End of cash assistance.
Illinois	Statewide	November-96	Recipients whose youngest child is 13 or older must accept up to 60 hours per month of work subsidized by AFDC grant.
Illinois	Statewide	November-97	End of cash assistance; family ineligible to reapply for aid for two years.

Indiana	Statewide (initially limited to 12,000 adult recipients)	May-97	Adult portion of grant is terminated.
Iowa	Statewide	Unknown	Benefits will be phased out for failure to make satisfactory progress towards self-sufficiency.
Louisiana	Statewide	February-99	End of cash assistance.
Massachusetts	Statewide	January-96	Recipients who can not find work will be placed in a community service position for 20 hours per week.
Michigan	Statewide	November-95	After one year of noncompliance with work requirements, penalty increases to loss of all AFDC benefits.
Missouri	Statewide	June-97	At the time limit, recipients will be assigned to job search and work experience.

Missouri	Statewide	May-98	The state will deny AFDC to an individual who received benefits for at least 36 months and who reapplies after completing a self-sufficiency agreement entered into after July 1, 1997, if the individual was responsible for becoming unemployed. Other eligible members of the family will receive benefits.
Montana	Statewide	February-98 for single parents	Individuals who reach time limit but have not achieved self-sufficiency will be required to participate in Community Services Program for 20 hours per week in order to receive benefits.
Nebraska	Two counties in 1995. Expanded statewide in 1996.	November-97	End of cash assistance.
New Hampshire	Statewide	Between February-97 and February-98	Requires job search for up to 26 weeks followed by work-related activities for 26 weeks. These cycles will repeat until the recipient is off AFDC.

North Carolina	Statewide	Between March-98 and March-99	End of cash assistance. Family becomes ineligible for 36 months.
North Dakota	Ten counties	Unknown	Placement in a work experience program or extension of benefits, based on an evaluation of the recipient's circumstances.
Ohio	Statewide	Between July-99 and January-00.	End of cash assistance.
Oklahoma	Six counties	May-99	Mandatory workfare participation of at least 24 hours a week.
Oregon	Statewide	August-97	End of cash assistance.
South Carolina	Statewide	Between July-98 and July-99	End of cash assistance.
South Dakota	Statewide	May-96 or May-99	If adult is not employed at least 30 hours per weeks, must perform 30 hours of approved volunteer service each week (fewer if good cause shown).
Tennessee	Statewide	Between April-98 and April-99 for continuous recipients	End of cash assistance. After receiving AFDC for 18 months, a household must wait at least three months before re-applying.

Texas	Statewide	May-97	Adults who reach the time limits may not receive cash assistance for a five-year period. The children will continue to be eligible for benefits.
Vermont	Statewide	November-95 for 15-month group. March-97 for 30-month group.	Requires participation in subsidized employment.
Virginia	Statewide. Phased in over four years.	August-97	End of cash assistance.
Washington	Statewide	February-00	Imposes a 10 percent grant reduction for families who have received assistance for 48 out of 60 months, and imposes an additional 10 percent grant reduction for every 12 months thereafter.
Wisconsin	Two counties: Fond du Lac and Pierce	February-97	End of cash assistance.