NBER WORKING PAPER SERIES

THE EITC AND THE EXTENSIVE MARGIN: A REAPPRAISAL

Henrik Kleven

Working Paper 26405
http://www.nber.org/papers/w26405

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2019

I thank Richard Blundell, Raj Chetty, Monica Costa Dias, Arindrajit Dube, Amy Finkelstein, John Friedman, Peter Ganong, Alex Gelber, Jeffrey Grogger, Hilary Hoynes, Theodore Joyce, Larry Katz, Pat Kline, Wojciech Kopczuk, Ilyana Kuziemko, Claus Kreiner, Camille Landais, Adam Looney, Day Manoli, Alex Mas, Larry Mead, Costas Meghir, Bruce Meyer, Pascal Noel, Matthew Notowidigdo, Jesse Rothstein, Dominik Sachs, Emmanuel Saez, Stefanie Stantcheva, Owen Zidar, and Jim Ziliak for comments and discussions. I am especially grateful to Elena Marchetti-Bowick for her outstanding research assistance throughout the project, and I also thank Stephanie Kestelman, Dana Scott, and Shreya Tandon for their excellent research assistance at different stages of the project. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The EITC and the Extensive Margin: A Reappraisal
Henrik Kleven
NBER Working Paper No. 26405
October 2019
JEL No. H20,H24,H31,J20,J21,J22

ABSTRACT

This paper reconsiders the impact of the Earned Income Tax Credit (EITC) on labor supply at the extensive margin. I investigate every EITC reform at the state and federal level since the inception of the policy in 1975. Based on event studies comparing single women with and without children, or comparing single mothers with different numbers of children, I show that the only EITC reform associated with clear employment increases is the expansion enacted in 1993. The employment increases in the mid-late nineties are very large, but they are influenced by the confounding effects of welfare reform and a booming macroeconomy. Based on different approaches that exploit variation in these confounders across household type, space and time, I show that the employment effects align closely with exposure to welfare reform and the business cycle. Single mothers who were unaffected by welfare reform (but eligible for the EITC) did not respond. Overall and contrary to consensus, the case for sizable extensive margin effects of the EITC is fragile. I highlight the presence of informational frictions, widely documented in the literature, as a natural explanation for the absence of extensive margin responses.

Henrik Kleven
Department of Economics
Princeton University
238 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and CEPR
and also NBER
kleven@princeton.edu

An online appendix is available at http://www.nber.org/data-appendix/w26405

1 Introduction

Anti-poverty policy in the United States has changed fundamentally over the last four to five decades. One of the most profound changes has been the expansion of support to the working poor through the Earned Income Tax Credit (EITC) along with the downsizing of traditional cash welfare to the poorest segments of the population. This transition from welfare state to workfare state is illustrated in Figure 1, which plots the number of EITC recipients and the number of welfare recipients over time. The EITC has taken over as the main source of cash support and now dwarfs cash welfare by a factor of almost six. While this shift has distributional consequences, a key motivation has been to encourage labor supply at the extensive margin. Traditional welfare has long been blamed for keeping families out of the workforce (e.g. Murray 1984), while the EITC is supposed to draw them in. This paper revisits the extensive margin effects of the EITC.

In the large literature studying the EITC, there is a strong degree of consensus. Most authors agree that the program has had sizable extensive margin impacts, particularly on single mothers. Starting with Eissa & Liebman (1996) and Meyer & Rosenbaum (2001), the literature has focused mostly on the federal EITC reforms in the 1980s and 1990s and relied on difference-in-differences approaches using variation by the presence and number of children. The most striking evidence brought forward in support of extensive margin responses is the sharp rise in the employment of single mothers, especially single mothers with two or more children, following the large 1993 EITC expansion for these family types (Hotz *et al.* 2006; Meyer 2010).

Given the importance of the historical change in welfare state design and the current proposals to further expand the EITC, it is critical that we have a correct understanding of its effects. The conclusions from the EITC literature are also central to a wider narrative regarding the impact of tax incentives on the extensive margin (see e.g., Blundell & MaCurdy 1999; Chetty *et al.* 2013a). In this paper, I do three things to shed new light on the EITC. First, I take a long-run perspective and consider every EITC reform since its inception in the mid-1970s, including all federal and state reforms. Second, I study all reforms in an event study framework, investigating the dynamics of labor supply changes before and after legislated reforms. Third, I carefully investigate the role of confounding factors, especially for the 1993 reform which underpins much of the consensus. I am far from the first to think about confounders in the 1990s, but I will try to offer some fresh

perspectives.

Like most of the existing literature, I use Current Population Survey (CPS) data and focus on the labor supply of single mothers.[1] For these women, the extensive margin incentives created by the EITC are unambiguously positive. I start from a long-run perspective, documenting the evolution in labor force participation of single women with and without children over the last 50 years. The observed patterns are striking. The participation rates for these two groups have evolved in parallel over this entire period, except for the mid-late 1990s. The participation rate for those with children was about 14 percentage points lower than for those without children in the late 1960s. This gap was about the same 25 years later, in the early 1990s. Then single mothers closed the entire gap in about five years, after which the two groups reverted to parallel trends. The dramatic increase for single mothers followed the 1993 EITC expansion, but the long-run series highlight that this period was an anomaly. If the EITC was important for the extensive margin in the 1990s, then why did it not narrow the gap in other time periods? There were federal EITC reforms in 1975, 1986, 1990 and 2009 along with numerous state EITC reforms between 1984-2018, which significantly increased the incentive to enter the labor market for single mothers.

Turning to event studies of each individual EITC reform further highlights how unusual the 1993 reform looks. Apart from this reform episode, event studies are essentially flat around the various EITC expansions at the federal and state levels. Any compelling narrative regarding the EITC and the extensive margin must reconcile the starkly different patterns around 1993 and elsewhere. One view is that the 1993 reform was different simply by virtue of being bigger and size matters in a world with optimization frictions (Chetty *et al.* 2011; Chetty 2012). A problem with this interpretation is that the 1993 reform was bigger only for those with two or more children, and yet employment also increased substantially for those with one child. Another view is that the anomalous patterns of the 1990s were driven, not by the EITC, but by confounding factors. The importance of such factors has been discussed in the literature (e.g., Ellwood 2000; Meyer & Rosenbaum 2001; Blank 2002; Grogger 2003; Fang & Keane 2004), with authors arguing that the EITC was in fact a key component of what happened in the 1990s even if it wasn't solely responsible.

What were the main confounders? First, there was the collapse of traditional cash welfare as illustrated by the caseload reductions in Figure 1. These caseload reductions were linked to welfare reform, both the federal welfare reform act of 1996 and the numerous state welfare reforms ("wel-

[1] While the existing literature has relied primarily on the March supplement of the CPS, I use linked March and monthly files. The larger dataset is very useful for precision in some of the more demanding event study specifications.

fare waivers") implemented between 1992-96. These policy changes put constraints on welfare receipt through aspects such as time limits, work requirements, training and job search activities. By putting constraints on a benefit that varies by the number of children, welfare reform produced extensive margin incentives correlated with the EITC. Second, the economy was booming at this time and the favorable business cycle may have had heterogeneous effects on single women with and without children. In particular, welfare reform and the business cycle are likely to interact: in a situation where single mothers are pushed off welfare, the labor market dynamics will depend on the tightness of the labor market at the time of the reform. I will argue that the dramatic labor market changes in the mid-late 1990s were driven primarily, if not exclusively, by welfare reform aided by the strong economy.

I conduct a number of exercises that highlight the importance of welfare reform over the EITC. I start by considering event studies around the 1993 reform by family size. The extensive margin effects are monotonically increasing in the number of children: In the years following the 1993 reform, employment increased by about 10pp for single women with one child, 15pp for single women with two children, 20pp for single women with three children, and close to 30pp for those with four or more children. This pattern of heterogeneity is consistent with welfare reform due to the link between AFDC benefits and the number of children. Indeed, AFDC participation levels prior to welfare reform and the subsequent AFDC participation drops were strongly increasing in the number of children. Conversely, the patterns are not consistent with the EITC, which was increased sharply for all families with two or more children and only modestly for families with one child (relative to those without children). This would predict a sharp divergence in the effects between those with one and two children and little divergence elsewhere.

Is it possible that the fanning-out of effects by family size is driven by heterogeneity in extensive margin elasticities? To probe this question, I convert the effects into elasticities assuming that they are driven entirely by the EITC. This analysis produces elasticities that are enormous overall and sharply increasing in family size. The average elasticity across all single mothers equals 1.7, while the elasticity for those with, say, four or more children is about 3. Elasticities of such magnitudes are not credible, implying that the assumption that all or most of the employment effects in the 1990s were driven by the EITC must be far off. To put it differently, if we assume sizable but *potentially* reasonable elasticities, then a very small fraction of what happened in the 1990s was driven by the EITC. This highlights the challenges of using the 1990s to learn about the effects of the EITC. We are trying to pick out a potential EITC effect in a pattern that is mostly not about the

EITC.

Besides the number of children, the age of the youngest child is a strong predictor of AFDC participation (see also Grogger 2002, 2003; Looney & Manoli 2016). The pre-reform level of AFDC participation and its ensuing fall are strongly declining in the age of the youngest child and close to zero for those with older children. As a result, the age of the youngest child can be used as a proxy for welfare treatment, allowing for a comparison of employment effects after the 1993 EITC expansion by intensity of welfare treatment. These employment effects are monotonically declining in the age of the youngest child. Among those with EITC eligible children above 13 years of age, the effects are virtually zero. Building on this idea, I also present an analysis that uses the age of the youngest child together with the number of children and other demographic variables to predict a probability of pre-reform AFDC participation, and consider employment effects by this measure of welfare treatment intensity.[2] There are no employment effects in the lowest deciles of predicted AFDC participation.

The final test focuses on the years 1994-1996, i.e. after EITC reform but before federal welfare reform. During these years, the main confounders are due to welfare waivers and the business cycle, both of which can be potentially controlled for using state-level variables. Augmenting the EITC event study specification with controls for the effect of state welfare waivers (allowed to vary by year and children) and controls for state-level unemployment (allowed to vary by children) renders the employment effects between 1994-1996 small and statistically insignificant. In other words, absorbing the variation that can be explained by waivers and unemployment changes a series that look like an event study of the 1993 EITC expansion into a series that look like an event study of the 1996 welfare reform. This holds across different samples and specifications.

Where does this leave us? The different analyses I present suggest that the dramatic employment increases in the nineties were driven by welfare reform aided by a favorable business cycle. Of course, it would be difficult to prove conclusively that these increases were completely unrelated to the EITC, even if the empirical patterns do not point in that direction. The magnitude of the labor market changes in the nineties may swamp any EITC-related pattern as I demonstrate through simple simulation exercises. This puts the onus on EITC reforms that are less confounded.

[2] Because the analysis is based on EITC eligible children — defined as children living at home who are younger than 18 or between 18-23 and still in education — the number of children and the age of the youngest child are correlated (see also Looney & Manoli 2016). Having more children living at home implies that the last-born tends to be younger. As a result, the univariate analyses of heterogeneity by number of children and age of the youngest child partially overlap, whereas the multivariate analysis of heterogeneity by the probability of AFDC participation exploits the joint predictive power of the two.

When investigating the long-run evolution of labor supply for single mothers and doing event studies of other EITC reforms, it is hard to conclude that the EITC has played any major role for the historic extensive margin changes in the U.S. In fact, when merging all of the EITC reforms at the state and federal level in a stacked event study, I obtain a precisely estimated zero.

How do we reconcile the observational findings presented here with the many randomized evaluations finding extensive margin effects of work incentives?[3] The bulk of these randomized evaluations have considered the impact of welfare reform treatments such as time limits, work requirements and various financial incentives, often bundled together. These are precisely the aspects that featured prominently in both TANF reform and state welfare waivers, and in fact many of the experiments were demonstration projects for statewide waivers. The findings of extensive margin effects in this literature (e.g. Kline & Tartari 2016) are consistent with the arguments in this paper. As for in-work transfers like the EITC, Card & Hyslop (2005) evaluate an earnings subsidy to welfare recipients in Canada — finding effects in the short run, but not in the long run — but this policy was structured quite differently from the EITC. More importantly, the EITC is different from experimental settings in terms of salience, information, administration and claiming. There is a considerable amount of evidence that most potential recipients are either unaware of the EITC or have a limited understanding of the schedule, eligibility and claiming (e.g., Bhargava & Manoli 2015). Given these aspects, the EITC is not an *a priori* likely candidate for finding large effects on labor supply at the intensive or extensive margins.[4] I come back to these issues below.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the policy context and the data. Section 4 provides descriptive evidence on the evolution of the extensive margin for single women over the last fifty years. Section 5 presents a series of exercises trying to understand the historic changes in the nineties. Section 6 presents a combined analysis of all the EITC expansion at the federal and state level. Section 7 concludes.

[3] See Grogger & Karoly (2005) for a comprehensive review of the early literature, and Card & Hyslop (2005), Kline & Tartari (2016) and Miller *et al.* (2018) for more recent experimental evaluations.

[4] While the informational and psychological frictions associated with the EITC are widely acknowledged in the literature, they have been used almost exclusively as an explanation for not observing any *intensive* margin responses. However, theoretical models of intensive and extensive margin responses (see Kleven & Kreiner 2006) imply that such frictions are equally important for the extensive margin.

2 Literature Review

An enormous body of work has examined the labor market impacts of the EITC and welfare reform. This section discusses a number of closely related papers, referring those interested in a more exhaustive summary to one of the many excellent reviews of the literature.[5]

Much of the EITC literature has focused on the extensive margin responses by single mothers, using Current Population Survey (CPS) data and variation from the federal EITC expansions in the 1980s and 1990s. These papers consider difference-in-differences approaches that rely on EITC variation by the presence and/or numbers of children. The early study by Eissa & Liebman (1996) finds sizable employment effects based on comparing single women with and without children before and after the 1986 EITC expansion. Meyer & Rosenbaum (2001) expand the time period to include the 1986, 1990 and 1993 tax reforms, and take a more structural approach based on modeling income taxes and welfare parameters from AFDC, Food Stamps and Medicaid. They estimate that tax changes (including the EITC) account for over 60 percent of the employment increase of single mothers between 1984-96. Hotz *et al.* (2006) use administrative panel data from California, covering families on welfare during the period 1991-2000. Their empirical strategy exploits the differential EITC expansion for families with two or more children relative to families with one child following the 1993 reform. They find sizable employment effects. Gelber & Mitchell (2011) confirm the qualitative findings from these earlier studies using PSID data from 1975-2004. Hoynes & Patel (2018) present event studies of the 1993 reform, comparing single women with zero, one and two or more children. Their focus is one poverty effects, but their online appendix shows employment effects as well. In a recent study of the 1975 introduction of the EITC, Bastian (2018) estimates large effects on maternal employment based on comparing women (single and married together) with and without children.

There are some exceptions to the consensus view described above. Cancian & Levinson (2005) examine Wisconsin's large EITC supplement to families with three or more children, and find no effects on labor supply at the extensive margin. In a paper analyzing the experience effects of employment, Looney & Manoli (2016) emphasize a point closely related to one of the analyses presented in this paper. They highlight that much of the variation in employment by family size (i.e., by number of minor children living at home) is really variation by the age of the youngest child, and that the age of the youngest child is strongly related to AFDC receipt prior to welfare

[5]The EITC literature has been reviewed by Hotz & Scholz (2003), Eissa & Hoynes (2006), Meyer (2010) and Nichols & Rothstein (2015), while the welfare reform literature has been reviewed by Blank (2002) and Grogger & Karoly (2005).

reform and to the drop in welfare receipt through the 1990s.[6]

A very different approach to estimating labor supply responses to the EITC has been proposed by Chetty *et al.* (2013b). Based on the idea that responding to the EITC requires knowledge of the EITC, they estimate behavioral responses by comparing zip-codes that vary by EITC knowledge. Their proxy for knowledge is based on bunching by self-employed individuals around the first kink of the EITC. They back out extensive margin responses by comparing event studies around the birth of the first child in high- and low-bunching areas. Women are more likely to continue working after child birth in high-bunching ("EITC") areas than in low-bunching ("no-EITC") areas. This approach yields a modest extensive margin elasticity of 0.2 using data from after year 2000. Since their proxy for knowledge is increasing over time, the elasticity obtained in this way should be smaller in the 1980s and 1990s.

In the experimental literature, there are relatively few examples of EITC-style policies. Card & Hyslop (2005) study a temporary earnings subsidy to welfare recipients in Canada. Similar to the EITC, the transfer was conditional on working. But the policy was otherwise different because it was time-limited (available for maximum of 3 years) and because eligibility required finding work within 12 months of random assignment, creating a very strong short-term incentive for working in order to obtain the option of any future transfers. Such a program creates intertemporal substitution incentives and is related to the Frisch elasticity (see Chetty *et al.* 2013a). Recently, Miller *et al.* (2018) provides evidence from the Paycheck Plus demonstration in New York City, an EITC-style policy for low-income workers without children. They find relatively modest employment effects.

The paper is also related to the literature on welfare reform, including many observational and experimental studies (see Blank 2002; Grogger & Karoly 2005). Because the federal government required experimental evaluations of the state waiver programs that preceded nationwide TANF reform, there is a substantial body of experimental work assessing programs that involve features like time limits, work requirements, welfare-to-work training, and financial incentives. In general, the literature has found significant extensive margin effects. For example, Grogger & Michalopoulos (2003) provide evidence on the impact of time limits using Florida's Family Transition Program. Leveraging the randomization of a bundle of policies (including time limits) combined with the insight that the impact of time limits should be declining in the age of the youngest child, they find

[6]Outside the EITC context, some papers show zero-responses in labor supply. A recent paper by Martinez *et al.* (2018) use a large and salient two-year income tax holiday in Switzerland to estimate Frisch elasticities along the intensive and extensive margins. They find a precisely estimated zero elasticity at the extensive margin, despite the fact that the Frisch elasticity is an upper bound on the Hicksian elasticity in standard life-cycle models (e.g. Chetty 2012).

large effects of time limits on welfare use. Their findings are broadly consistent with results using the age of the youngest child presented here. More recently, Kline & Tartari (2016) study the extensive margin impacts of Connecticut's Jobs First Program, a randomized waiver demonstration that included a TANF-style bundle of changes such as time limits, work requirements, family caps and earnings disregards (see also Bitler *et al.* 2006). They find substantial effects on women's labor supply at the extensive margin.

There are relatively few studies that directly compare the effects of EITC and welfare reform in the nineties. Ellwood (2000) argued that the independent effects of the EITC, welfare reform, and the strong economy are very difficult to separate due to their overlapping timing. While it is hard to disagree with this general sentiment, I will argue that a number of patterns in the data are strongly suggestive. Grogger (2003) estimates the separate effects of time limits, the remaining waiver/TANF bundle, and the EITC. Using a specification that attributes variation by family size to the EITC rather than to welfare, he argues that the EITC may have been the single most important explanation for the rise in employment among single mothers. As I argue, by putting constraints on a cash benefit that vary by family size, the treatment effect of waiver and TANF initiatives should be allowed to vary by family size, and this changes the evaluation of the comparative effects of EITC and welfare reform. Finally, Fang & Keane (2004) provide a detailed study of all the different policy parameters and macroeconomic variables that may have influenced the behavior of single mothers. Their estimates imply that the rise in employment between 1993-2002 can be explained in roughly equal proportions by the EITC, welfare reform (work requirements and time limits, respectively), and the macro economy.

Of crucial importance to interpretation is the literature on EITC knowledge. In their study of extensive margin responses to the 1986 EITC expansion, Eissa & Liebman (1996) cite evidence from interviews conducted in 1993 showing "virtually no awareness" of the credit among potential recipients (Eissa & Liebman 1993; Olson & Davis 1994). A number of subsequent studies have documented the presence of substantial frictions related to the awareness, understanding, and claiming of the EITC (Romich & Weisner 2000; Smeeding *et al.* 2000; Ross Phillips 2001; Berube *et al.* 2002; Maag 2005; Kopczuk & Pop-Eleches 2007; Jones 2010; Chetty & Saez 2013; Mead 2014; Bhargava & Manoli 2015). For example, Maag (2005) reports that only 58% of low-income families "had heard about the EITC" in a nationally representative sample from 2002. What is more, the understanding of how the schedule is designed is even weaker. Among the families interviewed by Romich & Weisner (2000), most people had heard of the EITC, but virtually no one knew that

they needed to earn a certain amount to maximize the credit. Chetty *et al.* (2013b) show that their proxy for knowledge — the degree of bunching by the self-employed at the first EITC kink — has been increasing over time and was relatively limited in the mid-nineties.[7]

While these informational and psychological frictions are widely acknowledged, they have been used mostly to explain the absence of *intensive* margin responses. The presumption is, it appears, that extensive margin responses can be based solely on knowing about the existence of a tax refund, while intensive margin responses require detailed understanding of the schedule. This view is not consistent with economic theory, nor with basic introspection. Theoretical models predict that the extensive margin decision depends on taxes and transfers at the desired level of earnings, i.e. the intensive and extensive margin decisions are inter-dependent (see e.g., Kleven & Kreiner 2006; Eissa *et al.* 2006, 2008). This inter-dependence is particularly strong for nonlinear incentives like the EITC. In a model with fixed costs of working (due to for example child care), the EITC is an incentive to start working around the refund-maximizing earnings range (for the relevant family size), not in any earnings range. Absent information about where the relevant earnings range is located and the size of the credit in that range, the extensive margin response is not credible.

3 Institutional Background and Data

3.1 The Earned Income Tax Credit

Since its inception in 1975, the Earned Income Tax Credit (EITC) has grown to become the largest cash transfer program in the United States in terms of the number of recipients and revenue costs. The EITC is a refundable tax credit and eligibility depends on having earned income, creating a positive incentive at the extensive margin. As shown in Panel A of Figure 2, the credit amount is a function of household earnings and the number of qualifying dependents (children).[8] For each family size, the credit schedule features a phase-in range, a refund-maximizing plateau range, and a phase-out range. The generosity of the schedule is increasing in family size, with the largest possible credit for families with three or more children. As of 2018, their maximum annual credit

[7] From outside the economics literature, Mead (2014) argues that the EITC was not responsible for moving single mothers into work in the 1990s. Besides the survey evidence on EITC knowledge, his argument is that the academic consensus is inconsistent with the experiences of welfare officials and administrators "on the ground." Based on interviews with welfare officials in Wisconsin and New York, he concludes that those dealing directly with welfare recipients did not think the EITC played any role in the initial decision by welfare mothers to start working.

[8] An EITC qualifying dependent is a relative who is under the age of 19 (24 for full-time students) or permanently disabled, and reside with the tax filer for at least half the year.

equals about $6,400 and is reached at annual earnings in the range of $14,300-18,700.

The parameters of the EITC have been revised substantially over time. Table A.I in the online appendix documents the full set of EITC parameters for each family size and each year between 1975-2018.[9] To simplify, Panel B of Figure 2 focuses on the evolution of just one key parameter — the maximum possible credit — across different family sizes. The credit amount in each year is reported in 2018 US Dollars. As shown in the figure, there have been five federal reforms of the EITC: the introduction in 1975 and the expansions enacted in 1986, 1990, 1993 and 2009.[10] The largest changes are the 1975 EITC introduction for all families with children and the 1993 EITC expansions for families with two or more children. The figure highlights that the most natural research design for studying the impacts of the EITC is to compare different family sizes around the reform episodes. For the 1993 reform in particular, it is natural to compare those with two or more children to those with one child or without any children. This comparison comes with potential placebo checks, because there is relatively little variation between those with one vs zero children and no variation between those with three or more vs two children.

In addition to the federal EITC described above, many states have introduced their own EITC supplements. A total of thirty states instituted EITC supplements between 1984-2018. The details of these state programs are described in Table A.II in the appendix. With the exception of California and Minnesota, state EITCs are specified as a percentage of the federal EITC. In some states the supplement is small, while in others it is quite large. The majority of state EITCs are refundable, just like the federal credit, but some of them are nonrefundable. Overall, when accounting for both the state and federal EITC, there has been a very substantial and ongoing expansion of support to the working poor through the last four decades. I will explore the potential impacts of all of these EITC expansions in this paper.

3.2 Welfare Reform

It would be impossible to assess the EITC without considering the potentially confounding effects from the rest of the welfare system. In particular, the EITC expansion enacted in 1993 and implemented between 1994-96 — the biggest expansion in the history of the program — coincided with dramatic changes to traditional cash welfare. The Personal Responsibility and Work Opportunity

[9] The set of parameters includes the phase-in rate, the maximum credit, the phase-out rate, and the location of the kink points.

[10] These EITC changes were legislated as part of the Tax Reduction Act of 1975, the Tax Reform Act of 1986 (TRA86), the Omnibus Budget Reconciliation Acts of 1990 and 1993 (OBRA90 and OBRA93), and the American Recovery and Reinvestment Act of 2009 (ARRA).

Reconciliation Act of 1996 (PRWORA) implemented welfare reform at the national level, replacing Aid to Families with Dependent Children (AFDC) with the more restrictive Temporary Assistance for Needy Families (TANF). However, this welfare reform did not represent a sudden and unanticipated departure from past policy. It was the culmination of state-led welfare reform efforts starting in the late 1980s and accelerating through the first part of the 1990s. These state reforms were implemented under the heading of *welfare waivers*, permissions from the federal government allowing states to experiment with their welfare programs. While waiver-based reforms had been possible since the early 1960s, they became very popular in the 1990s and set the stage for national reform. I will highlight the key features of state and federal welfare reform below, but refer to Grogger & Karoly (2005) for an exhaustive description of the history and details of these policy changes.

The state waivers involved all the key elements that would later be implemented on a national scale through TANF reform, including time limits, work requirements, and financial incentives to work. Some waivers were pilot programs that affected only a few counties or a subset of welfare recipients, while other waivers were statewide welfare reforms. Between 1987-92, a total of thirty waivers were approved by the federal government, although none of these were statewide reforms. Between 1992-96, under the first term of the Clinton administration, another eighty-three waivers were approved and many of these were statewide reforms. In total, thirty-five states received approval for statewide legislation through waivers. Table A.III in the appendix documents the approval and implementation dates of all statewide waivers.

There were six main types of waiver policies (see e.g., Council of Economic Advisors 1997, 1999; Department of Health and Human Services 1999). *Termination time limits* introduced upper bounds on the length of time that a family or its adult members could receive welfare benefits, even if they were otherwise eligible. These time limits varied from state to state, but a number of places set their limit to 24 months. *Work requirement time limits* imposed mandatory work requirements that kicked in when recipients reached a certain time limit.[11] Furthermore, building on the Job Opportunities and Basic Skills Training Program (JOBS) instituted in 1988, JOBS waivers strengthened the rules regarding participation in education, training, and job search activities. There were two such waivers. *JOBS exemptions* allowed states to eliminate or reduce the exemption of families with young children from JOBS. While previously parents with children under the age of 3

[11] Henceforth, I refer to termination time limits simply an "time limits" and to work requirement time limits as "work requirements".

(or under the age of 6 if the state could not guarantee child care) had been exempt from JOBS, these waivers required parents with very young children (sometimes as young as 12 weeks) to participate in the program. *JOBS sanctions* allowed states to impose harsher sanctions for failure to comply with JOBS requirements. Many states introduced the possibility of suspending the entire family's AFDC grant after a period of non-compliance. Motivated by a perception that welfare benefits were contributing to single motherhood and out-of-wedlock births, *family caps* were used to eliminate or reduce the benefit increase for existing AFDC recipients who had additional children. Finally, *earnings disregards* provided stronger financial incentives to work by disregarding earnings up to a level (such as the federal poverty line) in the calculation of benefit claw back.

The era of waiver-based reform culminated in national reform through PRWORA, signed into law in August 1996. This reform consolidated the AFDC and JOBS programs into the TANF program, which included all the key waiver elements just described. States had considerable latitude in designing their own TANF program under some federal guidelines. For example, it was a federal requirement that states impose a time limit of no more than 60 months, but they were free to choose stricter limits. States that had implemented time limits through waivers could keep the same limit (and many states did) or adjust it up or down. States that did not have time limits had to introduce one. As a result, TANF extended the waiver-type elements that had already been implemented in many states to the remaining states. There was, and still is, considerable variation in the specific program parameters across states.[12]

Many of the policy changes implemented under waivers and PRWORA did not directly reduce the statutory benefit levels, or at least not the maximum benefit under zero earnings, but they imposed much harsher constraints on receiving the benefits. Table A.IV in the appendix shows the maximum monthly benefit by state and number of children in 1993. The benefit is strictly increasing in the number of children in every state. As a result, the predicted treatment effect of welfare reform is increasing in the number of children. For example, the cost of introducing a time limit on benefit receipt will be greater in a large family than in a small family due to their different benefit levels.[13] This is important to keep in mind when interpreting the empirical evidence by family size.

[12] See Grogger & Karoly (2005) for a detailed documentation of TANF parameters in each state.

[13] There is another more fundamental reason why the treatment effect of welfare reform is increasing in the number of children. In a model with fixed work costs per child (due to child care), single mothers with more children are less likely to work are more likely to be on welfare, all else being equal. Hence, their welfare participation rates will be higher at baseline (a prediction borne out in the data), which makes them more treated by welfare reform.

3.3 Data

The analysis is based on the Current Population Survey (CPS). I combine data from the basic monthly files and from the Annual Social and Economic Supplement (ASEC), or "March files". Using the monthly and March files together give much more data than using the March files alone, which is what previous papers in the literature have done.[14] I restrict the dataset to include the monthly files from 1989-2018 and the March files from 1968-2017. Even though the monthly files go back to 1976, they do not allow for accurately identifying the presence or number of children prior to 1989. I focus on the sample of single women aged 20-50.[15] These restrictions result in a sample of 4,674,064 individual-month observations across survey years 1968-2018. Appendix B provides a detailed description of the CPS data, including the link between monthly and March files, variable definitions and measurement.

The CPS allows for different ways of measuring labor supply at the extensive margin. One distinction is between employment and labor force participation, with the latter including unemployed people who report that they are either searching for a job or are temporarily laid off. While employment is the more relevant outcome for welfare calculations, participation has the advantage of not moving with transitions between employment and unemployment and is therefore less sensitive to the business cycle. The other distinction is between annual and weekly measures of the extensive margin. Annual measures are preferred in principle due to fact the EITC is based on annual taxable earnings.[16] On the other hand, focusing on weekly measures offer three key advantages. First, the weekly measures are available in both the monthly files and in the March oversample, whereas the annual measures are available only in the March files. Hence, the weekly measures give much more statistical power. Second, questions about work activities during the previous week likely involve less measurement error that questions about work or earnings during the previous year. Third, the weekly measures largely avoid issues with nonresponses in the CPS. There are very few nonresponses in the weekly labor market and demographic variables, whereas there is significant nonresponse in the annual income variables in the March files. This issue has become more severe over time. As a result, a significant fraction of earnings observations in the CPS are imputed based on a procedure described in appendix section B. I refer to Bollinger

[14] The main advantage of the March supplement is that it contains detailed information on annual income variables during the previous year. I use the March files alone for analyses that requires annual income information.

[15] The definition of "single" includes never married, separated, divorced, and widowed.

[16] On the other hand, the AFDC/TANF program — also very central for the analysis in this paper — is administered at the monthly level and thus depends on monthly earnings.

et al. (2017) for a detailed analysis of this issue.

Figure A.I in the online appendix shows the long-run evolution of all four extensive margin measures. There are level differences between these measures — annual measures are higher than weekly measures and participation is higher than employment — but the time trends are very similar. The employment series fluctuate more than the participation series due to the business cycle effects mentioned above. While this would be an argument for focusing on the participation outcome, and I do start out by considering this outcome in the next section, the main analyses that follow consider employment as the baseline outcome. This is consistent with the focus of the existing literature. I will show results for weekly employment in the baseline specifications, but present results for the three alternative measures in the appendix. None of the main results I present vary fundamentally with the choice of extensive margin measure, but the annual-level analyses are more noisy as they rely exclusively on the March files.

Table 1 provides descriptive statistics in the estimation sample of single women aged 20-50, pooled across all years. Single mothers are somewhat older, have less education, and employment/participation than single women without children. These level differences are unsurprising and, of course, by themselves do not invalidate the difference-in-differences analyses presented here. Figure A.II in the appendix shows the earnings distributions of single mothers, split between low education (high school and below) and high education (some college and above). Earnings are normalized using the first kink point of the EITC in each year (so that the kink is located at zero). While the distribution of the highly-educated is shifted to the right, the graph shows that, in the sample of single mothers, both those with low and high education tend to have earnings within the EITC eligible range. I consider the full sample of single mothers as a baseline, but present results for low-educated single mothers as a robustness check.

4 EITC and the Extensive Margin: The Long View

It is useful to start from a long-run perspective, describing the extensive margin changes for single women over the last fifty years. Figure 3 shows the labor force participation rates of single women with children (blue series) and single women without children (black series) between 1968-2018. Given children is an eligibility requirement for the EITC, we can think of these series as treatment and control groups for assessing the effects of the program.[17] The patterns are very striking. Over

[17] To clarify, childless families could not receive any EITC until 1993 and only a very modest credit thereafter (see Figure 2).

half a century, all of the action in the participation rate of single mothers relative to that of single childless women took place during a single spell in the mid-late 1990s. Outside this spell, the two groups have evolved in parallel. At the beginning of the period, in 1968, the gap in labor force participation between the two groups was equal to 14.3 percentage points. A quarter of a century later, in the early 1990s, the gap was about the same. Then the labor force participation of single mothers rose dramatically in the mid-late 1990s, closing the entire 14 percentage gap in just a few years. After this, the two groups went back to parallel trends — but now at the same levels — and have stayed that way to this day. The graph highlights just how extraordinary the late 1990s were in the history of the U.S. labor market.

What explains this long-run evolution? To think about the role of the EITC and other factors, Figure 4 compares the long-run series to the timing of the five federal EITC reforms (Panel A) and the confounders from welfare reform and the macroeconomy (Panel B). Starting with Panel A, it shows that the EITCs claim to sizable extensive margin effects relies on two concurrent events: the EITC expansion enacted in 1993 (which increased tax refunds between 1994-1996) and the historic increase in the participation rate of single mothers in the mid-late 1990s. At the same time, the figure brings out a major puzzle. If EITC expansion drove the extensive margin increases in the late 1990s, then why do we not see any such effect around the other reforms? The three federal EITC reforms prior to 1993 did not lead to any closing of the gap between single women with and without children.[18] In addition, thirty states introduced EITC supplements between 1984-2018, adding to the puzzle of why there is no action outside the critical period in the 1990s.

Any compelling EITC narrative must reconcile the starkly different patterns in the 1990s and elsewhere. Did the EITC expansion in the mid-1990s have a different impact, because it was bigger and perhaps better advertised than the other reforms? Or were the changes in the 1990s driven by other confounding factors, or by an interaction between the EITC and these other factors? Panel B highlights the two main confounders in the 1990s. First, there was welfare reform: the period of waiver-based reform between 1992-96, culminating in national reform in 1996 (see section 3.2 for details). Second, there was the macroeconomy: the economy was booming during the Clinton era, with the national unemployment rate (dashed series) falling from about eight to four percent between 1992-2000.

Why would the macroeconomy impact single women with and without children differently?

[18]The EITC expansion in 2009 did not lead to any visible effect either, but because this expansion applied only to families with three or more children, its effect may be missed when considering all single mothers. I present long-run series by different family sizes below.

The figure shows that, outside the 1990s, there is no strong correlation between unemployment fluctuations and the gap in labor force participation between the two groups. However, this is not inconsistent with the business cycle having an effect in the 1990s, because of the interaction with the policy shocks happening at that time. Regardless of which policy is primarily responsible for the effects (EITC or welfare reform), the labor market dynamics around the time of the reforms will depend on the tightness of the labor market. For example, when single mothers get pushed off welfare, their ability to find work will depend on the availability of jobs at the time. Hence, the short-term effects of welfare reform will almost surely interact with the business cycle. Moreover, in the presence of employment hysteresis (see e.g., Yagan 2019), the fact that welfare reform happened in a strong economy may play a role even for the longer-term effects.

The long-run series discussed above are based on a specific measure of the extensive margin — weekly labor force participation — and it is worth checking if the empirical patterns are different for alternative measures of the extensive margin. Figure A.III in the online appendix compares all four extensive margin measures available in the CPS data: participation and employment at either the weekly or annual level.[19] The difference between participation and employment is that the former includes unemployed people who are actively searching for work or temporarily laid off. The figure shows that the main insights are robust to the choice of extensive margin measure. If anything, the employment gaps show even less sign of shrinking in the 25 years leading up to the reforms in the 1990s. It is also worth noting that, in general, the employment series are bumpier than the participation series. This is because employment is more sensitive to the business cycle as it is directly affected by movements between employment and unemployment. Figure A.IV shows the evolution of the same four outcomes, but restricted to the sample of those with low education (high school or less). This is motivated by the idea that low-educated single mothers tend to have lower earnings and are therefore more treated by the EITC. The broad patterns are again the same, although there is some shrinking in the participation gaps — but not in the employment gaps — in the decades leading up to the 1990s.

Further insight can be gained by splitting the sample of single mothers into different family sizes. Figure 5 shows the long-run evolution in labor force participation for single women with zero, one, two, and three or more children. The broad pattern is the same as before: the different groups have trended similarly over half a century, except for the mid-late 1990s.[20] However, while

[19] See section 3.3 for details. The annual participation series is shorter than the others, because this measure was not recorded in CPS data until the 1976 March files (i.e., pertaining to calendar year 1975).

[20] These series are somewhat noisier than the aggregate series. This is partly because the long-run graphs — but not

all groups of single mothers increase their labor force participation in the 1990s, the magnitude is strongly increasing in family size. Consider those with three or more children: their participation rate increases by a staggering 23 percentage points over six years, about twice as much as for those with two children. This difference is puzzling under the EITC narrative, because the tax credit expansion was the same for those with two and three children. It is of course possible that larger families have larger extensive margin elasticities and therefore respond more strongly to the same incentive, but this line of reasoning leads to other puzzles: if single women with three or more children are much more elastic, then why do they not increase participation after the 2009 EITC expansion (targeted specifically to them) or after the 1986 and 1990 expansions? Their baseline level was higher in 2009 and there was the Great Recession — both of which might explain the lack of a response — but such arguments do not apply to the 1986 or 1990 reforms. Overall, the long-run evidence suggests that whatever happened to large female-headed families in the late 1990s was something unique to that time period.

Another interesting aspect of the 1990s is that participation rose substantially even for those with just one child. This runs counter to the argument that the EITC had larger effects in the 1990s simply because the 1993 expansion was much larger than the other expansions. While this was true for families with two or more children, it was not true for families with one child. As shown in Figure 2, the increase in the maximum refund for one-child families was not exceptionally large on its own, and it was actually quite modest relative to the concurrent increase for childless families. As a result, the sizable extensive margin increases for single women with one child relative to those without children during the 1990s provide *prima facie* evidence of confounders.

Figures A.V-A.VI in the appendix investigate the robustness of these patterns to the alternative extensive margin measures and to focusing on the low-educated. For each outcome and sample, the stylized pattern is the same: what happened in the nineties was unique to this time period, it happened across all groups of single mothers, and it was monotonically increasing in the number of children. This points to a treatment that was both unique to the 1990s and strictly increasing in family size. Welfare reform satisfies both of these, whereas the EITC satisfies neither. To dig deeper, the next section moves from the long-run descriptive perspective to a detailed econometric analysis of the 1990s.[21]

the analysis of the 1990s in the next section — are based on the March files alone rather than the linked March and monthly files. As mentioned in section 3.3 and described in the data appendix B, the monthly files are used only for post-1989 analyses due to problems with accurately identifying the presence and number children in the CPS monthly files before that time.

[21] Although my main focus is on single mothers, I consider married women in Appendix Figure A.VII. This figure

5 EITC and the Extensive Margin: Making Sense of the 1990s

This section investigates the changes in the 1990s from different angles that will shed light on the role of the EITC vs other factors. The 1990s can be analyzed using the linked March and monthly CPS files, which give more data and precision than when using the March files alone.[22] In particular, the larger dataset allows for flexible event study specifications and granular heterogeneity analyses without loosing too much precision.

5.1 Event Studies: 1993 Reform vs Other Reforms

I start by putting the data into an event study framework, comparing the 1993 EITC expansion to the other federal EITC reforms. This exercise is not very different from the descriptive evidence presented above — comparing those with and without children over time — but a formal event study approach has several advantages. It allows me to present graphs that highlight the dynamics around each reform more clearly, and it allows me to control for changing demographics and to provide standard errors. Consider the following specification:

$$P_{imt} = \sum_j \alpha_j \cdot Year_{j=t} + \beta \cdot Kids_i + \sum_{j \neq -1} \gamma_j \cdot Year_{j=t} \cdot Kids_i + \boldsymbol{X_i}\boldsymbol{\phi} + \nu_{imt}, \quad (1)$$

where the outcome P_{imt} is an indicator for individual i working or participating in month m of year t. The right-hand side includes dummies for each year, a dummy for having kids, the interaction between year and kids dummies, and a vector of demographic controls $\boldsymbol{X_i}$. The interaction term omits the year just before the reform (denoted by $j = -1$), so that the difference-in-differences coefficient γ_t can be interpreted as the extensive margin effect in year t relative to the pre-reform year. As a baseline, I consider effects on weekly employment without any demographic controls, and then investigate the robustness of the results to alternative extensive margin measures and to controlling for demographics.

The first set of results is presented in Figure 6. This shows event studies around the 1975 EITC introduction (Panel A), the 1986 and 1990 EITC expansions (Panel B), the 1993 EITC expansion

shows labor force participation rates of married women between 1968-2018, considering comparison groups based either on the presence of children or on spousal earnings (using that the EITC is based on family income). The patterns do not point to any clear extensive margin responses for married women (even in the 1990s).

[22] As discussed in section 3.3 and in the data appendix, the recording of children in the monthly files is accurate (and consistent with the March files) only from 1989 onwards. As a result, all analyses using data prior to 1989 are based on the March files alone, while those using data only from after 1989 are based on the linked March-monthly files.

(Panel C), and the 2009 EITC expansion for families with three or more children (Panel D). These are run separately and each graph shows the estimated DiD coefficients $\hat{\gamma}_t$ over time. While Panels A-C compares single mothers with and without children, Panel D compares single mothers with three or more children to those without children.[23]

The figure highlights just how exceptional the patterns around the 1993 reform are. Taken at face value, the 1993 event study looks very compelling: the trends are roughly parallel in the years before the reform, the treatment group starts increasing relative to the control group after the reform, and a large and statistically significant DiD effect builds up over time. No other reform is associated with such patterns. The event studies are essentially flat around the 1975 introduction (which was large) and around the 1986 and 1990 expansions. The event study around the 2009 expansion shows a sizable negative effect, which is presumably due to the confounding effects of the Great Recession in 2008-09. The negative DiD effects obtained during this macroeconomic downturn highlights the problems with interpreting the positive DiD effects during the upturn in the 1990s.

Figure A.VIII in the appendix repeats this analysis with demographic controls to see if the patterns are driven partly by compositional changes in the group of single mothers over time. Specifically, the controls include dummies for the age of the woman (six categories), the age of the youngest child (seven categories), and education level (three categories).[24] The resulting patterns are very similar to those based on raw means. There is one exception worth highlighting, however. Controlling for demographics in the 1993 event study strengthens a feature that was also present (but subtle) in the baseline graph, namely that the inflection point in the DiD series happens a year too early, in 1992. The year 1992 coincides with the turn of the business cycle and the beginning of statewide welfare waivers (as shown in Figure 4), suggesting that the simple DiD estimates around 1993 may substantially overstate the effects of the EITC.

Figure A.IX and A.X in the appendix reproduce the analysis with demographic controls for the alternative extensive margin outcomes: weekly participation and annual employment. The patterns are overall similar when considering these other outcomes, but annual employment is noisier as it is based on the March files alone. The main difference relates to the DiD series of

[23] That is, Panel D is based on an extension of equation (1) that includes separate dummies for 1, 2, and 3+ children, plotting the DiD coefficients on single mothers with 3+ children. Changing the control group to those with one or two children (or their combination) does not affect the results qualitatively.

[24] The binning of these dummy controls is as follows: age of woman (20-24, 25-29, 30-34, 35-39, 40-44, 45-50), age of youngest child (0-1, 2-3, 4-6, 7-9, 10-13, 14-17, 18+), and education (below high school, high school degree, some college and above). Unless otherwise specified, the same binning is used in subsequent analyses.

annual employment around the 1986 and 1990 reforms. This series shows an increase between 1987 and 1989, consistent with finding positive effects when comparing years before the 1986 reform to years in the late 1980s as in Eissa & Liebman (1996).[25] While this could be an effect of the EITC, the evidence is far from conclusive. Not only does the increase start a year too late (which might be explained by frictions), but most of the effect goes away between 1990 and 1991, i.e. after the 1990 reform had enacted further EITC expansions. More broadly, and as discussed in the previous section, there is no narrowing of the annual employment gap between single women with and without children over the 25 years leading up to the 1993 reform, despite the substantial EITC increases over this period.[26] The employment series is quite bumpy from year to year (due to business cycles and other time-varying factors), but there is no longer-run convergence.

Table 2 presents DiD estimates for all five federal EITC reforms across different specifications. The table shows results for employment and labor force participation (weekly measures), for all single mothers and low-educated single mothers (high school and below), and for regressions with and without demographic controls. The estimates represent the average effect over the first three years after each reform. The three-year window is chosen because it avoids overlap between the 1990 and 1993 EITC reforms, and between the 1993 EITC reform and PRWORA. The results are very robust and consistent with the interpretations above. Apart from the 1993 reform, all estimates are either small and statistically insignificant or they are negative. By contrast, the 1993 reform is associated with large and strongly significant DiD estimates that hardly vary across specifications. The estimates are virtually the same for the employment and participation outcomes, and for specifications with and without demographic controls. The estimates are larger in the low-educated sample than in the full sample, but the difference is small: a positive effect of roughly 3.5pp for low-educated single mothers versus roughly 3pp for all single mothers.

To conclude, the empirical patterns that follow the 1993 EITC expansion represent a historic anomaly independently of the outcome, sample, and specification. A natural interpretation, therefore, is that these patterns are driven by confounders that are unique to the 1990s: welfare reforms at the state and federal level, perhaps interacted with the sharp upturn in the US macroeconomy. The following sections investigate these confounders from a variety of angles.

[25] The main specifications in Eissa & Liebman (1996) are based on comparing average annual employment between 1984-86 and 1988-90.

[26] See Panel D of Figure A.III.

5.2 1993 Reform: Heterogeneity by Family Size

This section presents event studies by family size. The long-run descriptive analysis included this dimension of heterogeneity, but using the linked March and monthly files allows me to produce cleaner and more granular evidence for the 1993 reform. The results are based on an extension of equation (1) in which the dummy for having any kids is replaced by separate dummies for having one, two, three, or four or more kids. The graphs presented below plot DiD coefficients γ_t^n for each year t and each number of children n. These coefficients capture the effect in each year relative to the pre-reform year, and for each family size relative to single women without children.

Figure 7 shows results from specifications without demographic controls (Panel A) and with demographic controls (Panel B). In both panels, there is a clear fanning-out of employment effects by number of children. Consider first the raw patterns in Panel A. In the decade following the 1993 reform, the employment rate increased by about 10pp for single women with one child, 15pp for single women with two children, 20pp for single women with three children, and close to 30pp for those with four or more children. Because the baseline level of employment is declining in family size, the fanning-out is even more dramatic in percentage terms. For example, single mothers with four or more children almost doubled their employment rate, from about 30% to about 60%. Panel B of the figure shows that some of the fanning-out can be explained by changes in the demographic composition of single mothers. Including detailed demographic controls, reduces the effect for each group, but by less for small families than for large families. The fanning-out is still strong, however, with increases in the employment rate ranging from about 8pp to 25pp across the four family sizes.[27,28]

These patterns of heterogeneity are puzzling under the notion that they are driven by the EITC. As described in section 3.1, the 1993 EITC expansion was modest for families with one child (relative to those without children), while it was very large for families with two or more children. By itself, this implies that the divergence in the DiD series should occur primarily between one and two children. The DiD series for one child should be relatively flat, while the series for three and four plus children should feature little additional divergence. Conversely, the observed het-

[27] Note that, while demographic controls reduce the fanning-out by number of children, they had virtually no impact on the average employment effects discussed in the previous section. This is because demographic controls matter mostly for single women with three or more children, who represent a small fraction of the sample: about 80% of single mothers have one or two children, about 13% have three children, and the remaining 7% have four or more children. These numbers have remained fairly constant for a long time.

[28] Figure A.XI in the appendix repeats the analysis for labor force participation. If anything, the fanning-out by number of children is even stronger in the participation outcome than in the employment outcome.

erogeneity is natural if the effects are driven by welfare reform. As described in section 3.2, AFDC benefits were increasing in family size — each child triggered additional benefits — and pre-reform caseloads were strongly increasing in the number of children.[29] Hence, the treatment effect of welfare reform — from both waivers and PRWORA — is strictly increasing in family size.

It is useful to zoom in on the years 1994-96, before national welfare reform. It is potentially easier to isolate the impact of the EITC during these years, although welfare waivers and the business cycle still pose a threat to identification. If anything, the evidence is even more puzzling during these years, because single mothers with two children do not begin increasing employment relative to those with one child until later.[30] In other words, in the years prior to PRWORA, there is divergence between each family size except where we should see it. This provides *prima facie* evidence that waivers and the business cycle by themselves were sufficient to swamp any EITC-related pattern.

5.3 1993 Reform: Elasticities and Simulated Responses

To better understand the magnitude of the extensive margin changes in the 1990s, I convert the estimated DiD effects into extensive margin elasticities. The elasticities will be calculated under the (extreme) assumption that the effects are driven solely by the EITC. This exercise clarifies if it is conceivable that the observed effects, including their heterogeneity across family sizes, are explained mainly by the EITC.

The extensive margin elasticity with respect to the net-of-tax rate on labor force participation is defined as follows

$$\varepsilon \equiv \frac{\Delta P / P}{\Delta \left(1-\tau\right) / \left(1-\tau\right)}, \tag{2}$$

where P is the probability of employment or participation and τ is the participation tax rate, i.e. the average tax rate on earnings when accounting for taxes paid and benefits lost upon entry into the labor market. This way of defining the elasticity corresponds to the literature on optimal taxation and welfare measurement (Saez 2002; Kleven & Kreiner 2005; Eissa *et al.* 2006, 2008). Assuming that the 1993 event study effects are driven by the EITC, I relate the total DiD effect ΔP to the change in the participation tax rate $\Delta\tau$ implied by the EITC expansion alone. The baseline net-

[29] Figure A.XII in the appendix shows AFDC/TANF participation rates over the 1990s by number of children. The pre-reform participation rates — and the ensuing drops in those rates — are strongly increasing in family size.

[30] This delay cannot be explained by the gradual phase-in of the EITC expansion for families with two or more children, because even the 1994 expansion on its own was much larger than the total expansion for families with fewer children (see Figure 2).

of-tax rate $1-\tau$, on the other hand, accounts for the combined implications of taxes and welfare benefits. I calculate this tax rate by combining the CPS data with NBER's tax simulation model (TAXSIM) and a welfare benefit calculator that accounts for AFDC and Food Stamp benefits. The tax-benefit simulations are outlined below and the full technical details are provided in Appendix C.[31] While τ and $\Delta\tau$ vary at the individual level, the elasticity ε will be calculated by relating the average treatment effect in the numerator to an average incentive change in the denominator.

Calculating the tax parameters in equation (2) requires information about earnings conditional on working. While the earnings of those who are working are observed in the CPS March files, the latent earnings of those who are not working are not directly observable. Therefore, the first step of the analysis is to predict earnings for non-workers. This is done by running the following regression on the sample of single women with positive (pre-reform) earnings:

$$Y_i = \alpha_a + \beta_n + \gamma_y + \delta_e + \zeta_r + \lambda_s + \eta_{ae} + \theta_{ne} + \vartheta_{ye} + \nu_i, \tag{3}$$

where Y_i is earnings of worker i and the right-hand side includes fixed effects for the age of the woman a (6 categories), the number of children n (7 categories), the age of the youngest child y (7 categories), education e (4 categories), race r (white, non-white), state s, as well as a set of second-level interactions between education, age, number of children, and age of the youngest child. The regression is estimated using pre-reform data (1991-93) for workers, and the estimated coefficients are then used to predict earnings for non-workers. The measurement of tax parameters is based on predicted earnings for non-workers and actual earnings for workers.

Two conceptual issues come up in this type of exercise. First, while the earnings predictions account for selection on observables, they do not account for selection on unobservables. The results are very robust to alternative specifications of the earnings regression, either more parsimonious or richer specifications. The different specifications predict that the majority of single mothers enter the labor market around the peak of the EITC schedule (where the subsidy is largest), which is also what economic theory predicts.[32] Second, I calculate the tax parameters in the denominator of equation (2) using the full sample of workers and non-workers. Both groups are included

[31] The welfare benefit calculator builds on the one developed by Eissa *et al.* (2008).

[32] Panel A of Figure A.XIII shows the predicted earnings distribution in the full sample of single mothers (working and non-working) estimated from equation (3). The kink points of the 1993 EITC schedule are demarcated by the vertical dashed lines. The mode of the earnings distribution is located between the first and second kink points of the EITC (where the refund is maximized) and only a small fraction of workers are predicted to locate above the EITC exhaustion point. Panel B of the figure shows the distribution of participation tax rates implied by the predicted earnings distribution.

because, in general, extensive margin effects may be driven either by higher entry rates among non-workers or by lower exit rates among workers. In theory, the elasticity should be based on the tax parameters of the *marginal* workers, i.e. those close to the indifference margin of entry or exit. By including inframarginal women among both workers and non-workers, there will be offsetting selection effects. These issues related to unobserved earnings and selection are endemic to estimations of extensive margin elasticities.

The results are displayed in Table 3. The first three columns show predicted earnings and tax parameters, the next three columns show employment effects, and the final three columns show participation effects. Each statistic is shown for all single mothers in the first row and separately by number of children in the following rows. The change in the net-of-tax rate ($\Delta(1-\tau)$) and the changes in the employment and participation rates (ΔP) represent difference-in-differences between the treatment and controls groups. For the tax rate changes, the DiDs are calculated at full phase-in of the EITC expansion in 1996. For the extensive margin effects, the DiDs represent 10-year effects as shown in the event studies in Figure 7 (employment) and Figure A.XI (participation).

The following insights emerge from the table. First, the change in the net-of-tax rate is strictly increasing in family size, although the EITC schedule itself does not vary after two children. This is because single mothers with more children are predicted to enter at lower earnings levels where the EITC subsidy is larger. Second, the elasticities are in general very large. Across all single mothers, the elasticity of employment equals 1.67 while the elasticity of participation equals 1.46.[33] Third, the elasticities are strongly heterogeneous across family sizes. The elasticities for large families are huge, with employment elasticities about 2 for those with three children and about 3 for those with four or more children. That is, despite accounting for the larger EITC incentive among larger families (due to variation in the schedule as well as in predicted earnings), the implied elasticities are much larger for these families.[34]

To summarize, assuming that all of the extensive margin increases after the 1993 reform were driven by the EITC implies elasticities that are extremely large. Although extensive margin elasticities are not structural parameters, we do have a sense of a reasonable range from the vast labor

[33] As described in Appendix C, the EITC-induced tax change $\Delta(1-\tau)$ is calculated assuming 100% take-up of the EITC conditional on eligibility. Given incomplete take-up, $\Delta(1-\tau)$ is upward biased and, as a result, the large elasticities presented here tend to be conservative (i.e., conservative given the thought experiment whereby the EITC is assumed to drive the entire extensive margin effect).

[34] The estimates in Table 3 are based on the raw event study specification without demographic controls. Table A.V in the appendix repeats the exercise based on specifications with demographic controls. Adjusting for demographics produces elasticities that are somewhat smaller — especially for the largest family sizes — but the broad findings are the same.

supply literature. The elasticities reported here fall outside any reasonable range. Hence, the view that the EITC was exclusively or mostly responsible for the changes in the 1990s must be very far off.

Instead of calculating elasticities, a related approach is to simulate how much can be explained by the EITC assuming a (potentially) reasonable elasticity. In a static labor supply model where single mothers respond according to the constant extensive margin elasticity ε, we have

$$\Delta P_t = \varepsilon \cdot \frac{\Delta\left(1-\tau_t\right)}{1-\tau_{93}} \cdot P_{93}, \tag{4}$$

where ΔP_t is the extensive margin response in year t relative to the pre-reform year 1993, $\Delta\tau_t$ is the reform-induced tax rate in change in year t relative to 1993 (which features some time variation due to the gradual phase-in of the reform), while τ_{93} and P_{93} are baseline values in 1993.[35] Calculating the net-of-tax rate change $\Delta\left(1-\tau_t\right)$ as the reform-induced change for the treatment group relative to the control group, the implied time path of ΔP_t from (4) can be compared to the observed DiD impacts from the event studies.

This exercise is presented in Figure 8.[36] The figure assumes an extensive margin elasticity of 0.3, corresponding roughly to the preferred estimate in the meta study by Chetty *et al.* (2013a).[37] Even under a sizable elasticity of 0.3, only a small fraction of the observed changes in the nineties would be driven by the EITC: 14% for single women with one child, 23% for single mothers with two children, 16% for single mothers with three children, and 10% for those with four or more children. I am not arguing for using an elasticity of 0.3 — in fact, I will argue that the elasticity with respect to EITC-induced changes is closer to zero — but these simulations give a clear sense of how large the changes in the nineties were. Any statement that the EITC was a major component of what happened in the nineties is implicitly assuming very large elasticities.

5.4 1993 Reform: Heterogeneity by Welfare Treatment Intensity

The next two sections investigate the impact of welfare reform (waivers and PRWORA) on the extensive margin patterns in the 1990s. Such an investigation requires variation in welfare reform

[35] Of course, the simple static model underlying equation (4) may not be realistic. The point of this exercise, however, is to provide a sanity check on the magnitudes in the data, and for such an exercise, the static, iso-elastic model is a natural benchmark to consider.

[36] This figure is based on event study specifications without demographic controls. The corresponding figure with demographic controls is presented in Figure A.XIV in the appendix.

[37] They show that the average extensive margin elasticity across nine studies is equal to 0.25.

treatment across individuals. I start by considering two different proxies for welfare treatment intensity. The first proxy is the age of the youngest child. It turns out that the level of welfare participation prior to waivers and PRWORA and the ensuing drop in welfare participation are strongly related to the age of the youngest child. The second proxy is a predicted probability of pre-reform welfare participation using the age of the youngest child, the number of children, and other demographic variables. Both of these analyses will show that the extensive margin effects in the 1990s are closely aligned with the strength of welfare treatment.

Due to the granularity of the analysis in this section, rather than showing event studies in each cut of the data, I summarize the effects using a standard DiD specification with a post-reform dummy. Specifically, I consider specifications of the following form

$$\begin{aligned} P_{imt} &= \alpha \cdot Post_t + \sum_j \beta_j \cdot Welfare_{j=c} + \sum_j \gamma_j \cdot Post_t \cdot Welfare_{j=c} \\ &\quad + \eta \cdot U_{st} + \theta \cdot U_{st} \cdot Kids_i + \lambda_s + \boldsymbol{X_i}\boldsymbol{\phi} + \nu_{imt}, \end{aligned} \tag{5}$$

where P_{imt} is an indicator equal to one if individual i is employed in month m of year t, $Post_t$ is an indicator equal to one in the years after the 1993 reform, and $Welfare_{j=c}$ is an indicator equal to one if the individual belongs to welfare treatment category c. The welfare treatment categories are based either on the age of the youngest child (7 bins) or on a predicted AFDC probability (10 deciles) as described below. In either case, the omitted category is having no children, so that the welfare category variable subsumes the kids dummy in the previous specifications. The coefficient γ_c represents the average DiD effect (over the post period) for single mothers in welfare category c relative to single women without children. I consider effects over two post-periods: a three-year period (avoiding years after PRWORA) and a ten-year period. I start from the raw DiD effects γ_c obtained from the first line of equation (5), and then consider the implications of the controls in the second line. These controls absorb the effects of state-level business cycles and demographics. Specifically, U_{st} is the aggregate unemployment rate in state s and year t, the effect of which is allowed to vary by children, and λ_s is a state fixed effect.

The results are presented in Figure 9. The left panels show heterogeneity by age of the youngest child, while the right panels show heterogeneity by probability of AFDC participation. Consider first the results for age of the youngest child.[38] The top panel shows that this variable is a very

[38] As elsewhere, I include only EITC-eligible children and consider the following categories for the age of the youngest child: 0-1, 2-3, 4-6, 7-9, 10-13, 14-17, and 18+. Consistent with the EITC rules, the 18+ category includes children aged 18-23, who are living at home and are still in full-time education.

strong predictor of welfare participation before the reform — the younger is the child, the higher is participation — which in turn predicts the drop in welfare participation after the reform. The relationship between pre-reform participation levels and post-reform participation drops — across bins of the age of the youngest child — is almost perfectly linear and has a slope of 0.69. Among single mothers with children aged 0-1, pre-reform AFDC participation was about 50 precent and the subsequent fall was close to 40 percentage points. By contrast, single mothers with older children had very low AFDC participation rates and, as a result, were virtually unaffected by welfare reform.

The middle panel shows raw DiD estimates by age of the youngest child over three years (solid black) and over ten years (dashed black) together with pre-reform AFDC participation rates (solid red). The extensive margin effects are strongly and monotonically declining in the age of the youngest child. The three-year effects fall to about zero (and become statistically insignificant) for those whose youngest child is older than 13 years, while the ten-year effects fall to zero when the youngest child is older than 17. These estimates include any confounding effects of the business cycle and changing demographics. As shown in the bottom panel, controlling for these factors reduces the treatment effect at each age and therefore strengthens the results. The three-year effects are non-positive above age 9 and the ten-year effects are non-positive above age 13.

Table 4 shows estimated impacts in two age groups (0-13 years and 14+ years) across different specifications and outcomes. The main insight is very robust: while the extensive margin effects are very large for single women with younger children, there are no effects for those with older children. Among those with children aged 14 and over, the three-year effects are either small and statistically insignificant or they are negative. The ten-year effects are significant only when no controls are included. Once unemployment controls are added, there is no longer any effect.

Because the analysis is based on EITC eligible children — that is, children living at home who are below age 18 or below age 24 if still in education — the age of the youngest child is correlated with the number of children. Having more children implies that the last-born tends to be younger. As a result, the analysis using the age of the youngest child overlaps with the previous analysis using the number of children. I now move to a multivariate analysis using the joint predictive power of the age of the youngest child, the number of children, and other demographic variables. Specifically, I estimate the probability of pre-reform welfare participation based on the following specification

$$AFDC_i = \alpha_a + \beta_n + \gamma_y + \zeta_r + \lambda_s + \nu_i, \tag{6}$$

where an indicator for receiving welfare benefits, $AFDC_i$, is regressed on fixed effects for the age of the mother a, the number of children n, the age of the youngest child y, race r, and state s.[39] The binning of these demographic variables is the same as before. The regression is run on the CPS March files (which contains information on welfare receipt) in the pre-reform year, 1993. From this regression, I predict the probability of AFDC participation for each single mother in the estimation sample, and define an indicator for being in different deciles of the distribution of these AFDC probabilities.

The results are shown in the right panels of Figure 9. The top panel shows that pre-reform welfare participation varies widely across the distribution, from 8% to more than 60%, and is strongly related to the post-reform welfare participation drop. The relationship between pre-reform levels and post-reform drops is stronger in this specification than when using only the age of the youngest child. The next panels show DiD impacts on employment over three years and ten years by deciles of the predicted AFDC probability. The employment impacts are strongly increasing in the AFDC probability. In the raw data, the ten-year effect is about 20pp in the top decile, but only 2pp in the bottom decile. Adding controls for business cycle effects and demographic changes makes the estimates smaller: the ten-year effect is now 13pp in the top decile and virtually zero in the bottom decile. In fact, there are no significant ten-year effects in the bottom two deciles and no significant three-year effects in the bottom four deciles. Table 5 shows average effects in the bottom two deciles and in the top eight deciles of welfare treatment intensity across different specifications. In the bottom deciles, there are no significant three-year effects in any specification and no significant ten-year effects once controls are added.

To summarize, the large extensive margin increases in nineties were driven by single mothers who had high AFDC participation prior to welfare reform and were therefore strongly treated by waivers and PRWORA. Single mothers with low pre-reform AFDC participation did not respond. These results are consistent with welfare reform, but harder to reconcile with EITC reform.

Building on these insights, it is useful to consider a different outcome variable: the fraction of single mothers who are either employed or on AFDC/TANF (henceforth referred to as the "employment-welfare rate"). This is useful because the EITC and welfare reform are expected to affect this outcome differently. Cutting welfare pushes people from welfare into work or into

[39] Equation (6) does not include education even though this variable would help with predicting AFDC participation. Education is excluded because it is a relatively direct proxy for earnings conditional on working and therefore for EITC eligibility. As a result, while using education information would strengthen my results, this is misleading as highly-educated single mothers who did not participate in welfare (thus being untreated by welfare reform) were also less likely to be eligible for the EITC (thus being untreated by EITC reform as well).

searching for work. If everyone finds jobs, the employment-welfare rate will be unaffected. Otherwise, it will decrease. By contrast, expanding the EITC provides work incentives to people from all non-working states, including those in the AFDC/TANF program, those in other social assistance programs, and those receiving no social assistance. If people respond to the EITC, the employment-welfare rate will increase.

The results are presented in the online appendix. Figure A.XV illustrates the approach in the full sample of single women. Panel A shows the basic DiD comparing the employment series for single women with and without children around the 1993 reform, while Panel B adds AFDC/TANF participants to the employment series for single mothers.[40] Strikingly, once welfare recipients are included, there is no visible treatment effect in the 1990s. That is, the employment effects are fully accounted for by movements from welfare to employment, with no additional effects coming from any other non-employed state. Figure A.XVI-A.XVIII shows that this is true in subsamples that vary by number of children, age of the youngest child, and states with high and low welfare caseload drops. While the employment increase varies greatly across subsamples, the employment-welfare rate series is always flat (or slightly falling) across time. In each subsample and year, the extensive margin increase corresponds essentially to a movement between AFDC/TANF and employment. These patterns are again consistent with welfare reform, but puzzling under the EITC narrative. This narrative requires that, while the EITC had sizable effects on people inside the AFDC program, people outside the program were unresponsive to the EITC. This would be an unusual pattern of heterogeneity, even if not inconceivable.[41]

5.5 1993 Reform: Controlling for Waivers and the Business Cycle

The confounding factors that make it difficult to identify the impact of the 1993 EITC expansion operate at both the state and national levels. This section focuses specifically on state-level factors — welfare waivers and state business cycles — as these can be estimated. Waivers were introduced in some states but not others, and the timing of their approval and implementation varied from place to place (see section 3.2 for details). Similarly, while the economy was strong across the country, the labor market implications of the booming economy varied spatially. Exploiting the variation in these factors across time and space, I investigate how much of the 1990s employment increase for

[40] To avoid double counting, I add only AFDC/TANF participants who are not also employed.

[41] The reason why the pattern is not inconceivable is selection. It is possible that single women in other social assistance programs — in particular, Disability Insurance (DI) and Supplemental Security Income (SSI) — and those receiving no social assistance are a selected sample of people who are unresponsive to taxes.

single mothers can be explained by waiver-based reform along with falling unemployment. I focus primarily on the years 1994-96, i.e. after the EITC reform but before nationwide TANF reform took full effect in 1997. After the implementation of TANF, welfare confounders are operating at a national scale and there is no quasi-experimental approach to separate welfare from the EITC. We still have the cross-sectional variation in welfare treatment intensity analyzed above, but there is no separate event study approach for the EITC and TANF reforms.

The results in this section are based on variants of the following specification:

$$\begin{aligned} P_{imt} &= \sum_j \alpha_j \cdot Year_{j=t} + \beta \cdot Kids_i + \sum_{j \neq -1} \gamma_j \cdot Year_{j=t} \cdot Kids_i \\ &\quad + \sum_j \delta_j \cdot Year_{j=t} \cdot Waiver_{sj} + \sum_j \zeta_j \cdot Year_{j=t} \cdot Kids_i \cdot Waiver_{sj} \\ &\quad + \eta \cdot U_{st} + \theta \cdot U_{st} \cdot Kids_i + \lambda_s + \boldsymbol{X_i \phi} + \nu_{imt}, \end{aligned} \tag{7}$$

where the first line is the basic DiD event study of the 1993 reform, while the second and third lines control for welfare waivers, the business cycle, and demographics. Specifically, $Waiver_{st}$ is an indicator equal to one if state s has approved a statewide waiver in year t. The waiver indicator is interacted with a kids dummy and a full set of year dummies. That is, the impact of waivers is allowed to vary between those with and without children (as only the former are eligible for welfare) and from year to year. To the extent that waiver states saw larger employment increases for single mothers after the introduction of waivers than non-waiver states, this effect gets absorbed in the waiver controls and reduces the DiD coefficients γ_t. Similarly, if states with larger drops in (aggregate) unemployment saw larger employment increases for single mothers, this gets absorbed in the unemployment controls and further reduces the DiD coefficients.

It is useful to highlight a few conceptual points about the specification. First, since it controls only for state variation, any national confounders are still absorbed in the DiD coefficients. The strong macroeconomy of the mid-late 1990s was a national phenomenon, and PRWORA affected all states after its enactment in August 1996. Moreover, most states without any major waiver reform introduced smaller pilot programs, and it is therefore conceivable that both treatment and control states were affected by welfare reform prior to PRWORA. For these reasons, even if the state-level controls in (7) had no impact on the estimates, this would not necessarily imply that confounders are not affecting the estimates. Second, while there were six main types of waiver policies, these are lumped together in a single any-waiver dummy. Below I consider an extension that allows for separate coefficients on different types of waivers. Third, the waiver indicator may

switch on either at the time of waiver approval or at the time of waiver implementation (see Table A.III for for approval and implementation dates of each state waiver). The baseline specification is based on approval dates, but I consider a robustness check based on implementation dates. Finally, the impact of waivers is allowed to vary non-parametrically over time through the interaction with year dummies. This specification choice reflects that waivers are unlikely to have the same impact after 1996 where TANF introduced waiver-style initiatives across the country. As a result, a specification with simple before-after waiver controls will have much less impact.

The first set of results is presented in Figure 10. Panel A compares the raw DiD series to a specification that includes unemployment controls, but no other controls.[42] Panel B compares the raw DiD series to a specification with both unemployment and waiver controls. The figure shows that both of these controls reduce the extensive margin effects after 1993. While the unemployment controls reduce the estimates throughout the mid-late 1990s, waivers reduce the estimates just before TANF reform in 1997. This is natural given that most statewide waivers were approved in 1995 and 1996, just before TANF reform. The striking feature of Panel B is that the state-level controls change the DiD series from what looks like a compelling event study of the 1993 EITC reform to what looks like an event study of the 1996 TANF reform. This finding is consistent with welfare reform (in combination with the macroeconomy) being the main driver of the extensive margin effects.

The appendix provides a number of robustness checks. Figure A.XIX repeats the analysis for different outcomes (employment vs participation) and samples (all women vs low-educated women). None of these aspects affect the results in any significant way. Figure A.XX investigates heterogeneity by number of children. The results are particularly strong for families with two children, where the event study adjusted for business cycle and waiver effects is completely flat until 1996, the year before PRWORA takes full effect. The results are a little less strong for families with three or more children, but qualitatively similar. Figures A.XXI and A.XXII vary the specification of waivers. The two figures consider the same waiver specifications, but the first includes both waiver and unemployment controls while the other includes only waiver controls. In each figure, the baseline specification of waivers (in Panel A) is compared to specifications where the any-waiver indicator is replaced by indicators for each waiver type (Panel B), where the date of implementation is used instead of the date of approval (Panel C), and where the time-varying

[42] To clarify, I always introduce unemployment controls together with state fixed effects. This ensures that the business cycle controls capture only within-state variation in unemployment over time as opposed to also cross-state variation in structural unemployment levels.

waiver controls are replaced by basic before-after waiver controls (Panel D). The graphs produce two main insights. The first is that waivers on their own reduce the estimates mainly in 1996; the other is that the estimates are very robust to specification, except for the before-after waiver dummy. The second insight is related to the first: because the impact of waivers has a very specific timing, introducing a basic before-after dummy has little impact.[43]

Table 6 summarizes the estimated effects of the 1993 reform after three years (before PRWORA) and after ten years (after PRWORA) with waiver and business cycle controls under different specifications. The table consider different extensive margin outcomes, different education levels, and different sets of controls. The table demonstrates the robustness of the main message: while the raw estimates are very large, the three-year effects become small and statistically insignificant as soon as unemployment and waivers controls are added. Of course, the ten-year effects remain significant as these include the impact of national welfare reform.

Like the results in the previous section, the analysis presented here points to the crucial role of welfare reform and the macroeconomy in the 1990s. The results show that the extensive margin effects between 1994-96 — the years that offer the best chance of separately identifying any EITC effects — can be explained by a combination of waiver-based reform and the business cycle. These results highlight the fragility of EITC estimates that rely on variation in the 1990s.

6 EITC and the Extensive Margin: Putting Everything Together

Having analyzed the 1993 reform in detail, I consider all EITC reforms at the state and federal level together in a stacked event study. Combining all the reforms increases precision and is arguably better identified than an event study of any individual reform. In particular, because the time series of employment is bumpy from year to year (due to business cycles and other time-varying factors), event studies of individual reforms may easily create the illusion of either positive or negative effects depending on the timing of the reform. Combining reforms with different timing alleviates this problem.

6.1 State EITC Supplements

This section presents evidence from state introductions of EITC supplements. The roll-out of these supplements is illustrated in Figure 11 and the details of each program are provided Table A.II.

[43]The small impact of a before-after waiver dummy is consistent with results presented in an online appendix to Hoynes & Patel (2018). They include a post-waiver indicator and find that it does not significantly impact their results.

A total of 30 states instituted EITC supplements between 1986-2018, and I restrict attention to the 27 states that have maintained their supplement for at least 3 years.[44] Except for California and Minnesota, the state EITCs are specified as percentages of the federal EITC. They are refundable credits in most states, but non-refundable in some. An empirical advantage of these policies is that they generate variation across space, in addition to the variation across family size considered above.

The analysis is implemented by creating a synthetic control state for each state with an EITC supplement, and then running a stacked event study comparing treatment and synthetic control states around the time of EITC introductions. The synthetic control states are constructed from the states that never had a supplement, matching on the level of the employment rate in each of the five pre-reform years. Table A.VI shows the composition of each synthetic state, and the details of the approach and implementation are provided in section D of the online appendix.

In the first specification, the estimation is implemented in the sample of single women with children. Here the empirical strategy is a difference-in-differences estimation that compares different states over time, conditioning on children. The results are shown in Figure 12. Panel A of the figure includes all 27 reforms and shows no effect of the EITC. The treatment and control states track each other in the pre-reform years (by construction) and continue to do so in the post-reform years. A concern with the analysis in Panel A is that many of the state supplements are quite small, making it difficult to detect any effect. Panel B therefore restricts attention to 13 "large reforms", specifically reforms that introduced *refundable* supplements equal to *at least 10%* of the federal EITC.[45] Focusing on large EITC reforms does not alter the main result: states that introduce EITC supplements do not see employment increases relative to states without such supplements.

As a robustness check, Figure A.XXIII in the appendix shows results from a specification that exploits the variation both across states and across single women with and without children within states. Here the empirical strategy is a triple-differences estimation that compares treatment and control states conditional on children relative to those same states conditional on no children. The figure shows that this refinement has no substantive impact on the empirical patterns. There is still no effect of state EITC supplements on employment.

[44] I drop Hawaii and South Carolina as they introduced their EITC supplements too recently (in 2018). I drop Washington as they enacted, but de facto never implemented their EITC supplement.

[45] To obtain a set of relatively sharp policy experiments, I also require that the state EITC reached 10% within three years of enactment. This excludes states that have sizable and refundable credits, but where this level was reached too incrementally for a compelling event study design. The 13 states with large EITC reforms are California (2015), Colorado (1999), Connecticut (2011), Dist. of Columbia (2000), Kansas (1998), Massachusetts (1997), Michigan (2008), Minnesota (1991), New Jersey (2000), New Mexico (2007), New York (1994), Vermont (1988), and Wisconsin (1989).

6.2 Stacked Event Study: All State and Federal Reforms

The final piece of the investigation combines the state and federal reforms in a stacked event study framework. I include 31 reforms in the analysis: the 27 state reforms analyzed in the previous section along with the 4 federal reforms enacted in 1975, 1986, 1993, and 2009. The 1990 reform is not included as a separate event due to its close proximity in time to the (larger) 1986 and 1993 reforms. All federal reforms are weighted equally, while state reforms are weighted according to the state's share of the national population in 2016. Therefore, although there are many more state reforms than federal reforms, all of the state reforms together weigh less than a single federal reform. The qualitative results are robust to alternative weighting schemes (e.g, equal weights for all reforms).

Normalizing the first year after each reform to zero, I consider an event time window that runs from -5 to +5. There are two exceptions to this: California's 2015 reform allows for an event time window running only to +2, while the federal 1993 reform is truncated at event time +2 for identification reasons. The latter specification choice is done to avoid overlap between the 1993 EITC reform and the 1996 TANF reform, which would otherwise confound the estimates. I include controls for welfare waivers and the business cycle (specified as in equation 7) to deal with the confounders already at play during event times 0 to +2 for the 1993 reform.

Two additional specification points are worth noting. First, the estimates are based on comparing single women with and without children, except for the 2009 federal reform targeted to families with three or more children. For this reform, the treatment indicator is based on having at least three children (relative to zero children). Second, in order to combine the state and federal reforms within one event model, the effect of the state reforms is based on comparing single women with and without children within each state. That is, in contrast to the synthetic control approach considered above, the spatial variation in EITC supplements is not used. The full details regarding the specification are provided in section E of the appendix.

The results are presented in Figure 13. Panel A shows the federal reforms alone, Panel B shows the state reforms alone, while Panel C combines the federal and state reforms. The federal reforms — and by implication the federal and state reforms combined — feature small positive post-reform coefficients, but at the same time there is a pre-trend. There is no inflection point in the series after the reforms and therefore no visible EITC effect. To make this clearer, Panel D adjusts for linear, group-specific trends estimated on the pre-reform data. This pre-trend adjustment makes the DiD series completely flat. Each panel provides an estimate of the average EITC effect over the post-

reform period (event times 0 to 5) along with the confidence band in parenthesis. The estimates are close to zero and precisely estimated: in Panel D, we can rule out effects on employment larger than 0.8 percentage points across all 31 reforms. Finally, Figure A.XXIV in the appendix repeats the analysis in the sample of single women with low education (high school degree or less). The main finding is the same: the EITC effect is a tightly estimated zero.

7 Conclusions

In this paper I have analyzed every EITC reform at the state and federal level within one modern empirical framework. From this comprehensive and long-run perspective, the EITC has not had any clear effects on labor supply at the extensive margin. Apart from the expansion enacted in 1993, EITC reforms are not associated with increases in the employment of single mothers relative to single women without children. The 1993 reform, on the other hand, is associated with very large employment increases, but these increases align closely with the confounding effects of welfare reform and a booming macroeconomy. Exploiting variation in these confounders across household type, space and time, I have shown that the effects are driven exactly by those affected most strongly by welfare reform and the business cycle. A stacked event study of all EITC reforms, controlling for welfare reform and the business cycle, produces a precisely estimated zero.

There are two interpretations of this null result. One is that the extensive margin elasticity with respect to taxes is small, and perhaps especially the elasticity with respect to a highly nonlinear tax refund like the EITC. The other is that, even if the elasticity is not zero in general, the informational and psychological frictions specific to the EITC have reduced its impact. These frictions may be related to the complex nonlinear schedule or to the procedures for filing taxes and claiming the credit. As reviewed above, a number of studies have documented that most potential recipients are either unaware of the EITC or have a limited understanding of the schedule and eligibility requirements.[46] Given these frictions, the EITC is not an *a priori* likely candidate for finding large labor supply impacts.

While these frictions are widely acknowledged in the literature, they have been used mainly to explain the absence of *intensive* margin responses. The reasoning seems to be that extensive margin responses can be based solely on knowing about the existence of a tax refund without understanding the specifics of the schedule and eligibility rules. Even ignoring the fact that many

[46] See e.g., Romich & Weisner (2000), Smeeding *et al.* (2000), Ross Phillips (2001), Maag (2005), Chetty & Saez (2013), and Bhargava & Manoli (2015).

potential recipients are unaware of the EITC, this argument is surprising. Economic theory predicts that intensive and extensive margin decisions are interdependent, and this interdependence is particularly strong for nonlinear incentives (see Kleven & Kreiner 2006; Eissa *et al.* 2006, 2008). The EITC is an incentive to enter the labor market around the refund-maximizing earnings range, which is relatively narrow and vary by family size. Without precise information about the location of the refund-maximizing earnings range and the size of the credit in that range, the extensive margin response is not very credible.

By contrast, welfare reform is not subject to these concerns about salience and information. The debate about welfare reform and welfare culture was extremely prominent in the nineties, as exemplified by Bill Clinton's famous campaign pledge to "end welfare as we know it." State waivers and national TANF reform imposed drastic changes on *existing* program participants, so information was essentially automatic. What is more, the changes to welfare were mechanical and brute-force. They included lifetime limits on welfare receipt (down to two years in a number of states), work requirements, community service and training. Those who did not satisfy these requirements were kicked off the welfare rolls. These initiatives stand in sharp contrast to the complex tax refund incentives introduced by the relatively unknown EITC program.

These findings do not necessarily imply that the EITC is a bad policy. Absent labor supply responses at either the extensive or intensive margins, the EITC is a non-distortionary transfer to the working poor.[47] The optimality of such a transfer depends on the social welfare function and on who pays for it. For example, if the EITC is financed by a lump-sum tax on all individuals and if the government puts a larger weight on the working poor than on the average individual, then the policy is socially optimal. This is the reasoning underlying the results in Saez (2002).[48] Conversely, if transfers to the working poor are financed by reducing welfare benefits to the poor, then the equity effect is negative under standard social preferences. In this case, the desirability of the EITC requires social preferences that put more weight on the working poor than on the non-working poor, even if the latter are worse off. Leaving aside these optimal tax considerations, the political argument for the EITC has relied heavily on its supposed employment effects. The empirical findings presented here do not lend support to this argument.

[47] Of course, this paper has not shown the absence of intensive margin responses to the EITC. The literature has struggled to find any clear evidence of intensive margin responses to the policy, perhaps because of the frictions just discussed. Chetty *et al.* (2013b) estimate significant intensive margin responses to the EITC *conditional on knowledge*, which is still consistent with small average responses across informed and uninformed taxpayers.

[48] In Saez (2002), the optimal EITC (defined as a negative tax rate on participation) reflects a trade-off between the equity gain just described and the efficiency *loss* from extensive margin responses to an EITC that make people work too much.

References

BASTIAN, JACOB E. 2018. The Rise of Working Mothers and the 1975 Earned Income Tax Credit. Working Paper. 6

BERUBE, ALAN, KIM, ANNE, FORMAN, BENJAMIN, & BURNS, MEGAN. 2002. The Price of Paying Taxes: How Tax Preparation and Refund Loans Erode the Benefits of the EITC. 8

BHARGAVA, SAURABH, & MANOLI, DAYANAND. 2015. Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment. *American Economic Review*, **105**(11), 1–42. 5, 8, 35

BITLER, MARIANNE, GELBACH, JONAH, & HOYNES, HILARY. 2006. What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. *American Economic Review*, **96**(4), 988–1012. 8

BLANK, REBECCA M. 2002. Evaluating Welfare Reform in the United States. *Journal of Economic Literature*, **40**(4), 1105–1166. 2, 6, 7

BLUNDELL, RICHARD, & MACURDY, THOMAS. 1999. Labor Supply: A Review of Alternative Approaches. *In:* ASHENJELTER, O., & CARD, D. (eds), *Handbook of Labor Economics*, vol. 3A. Elsevier. 1

BOLLINGER, CHRISTOPHER, HIRSCH, BARRY, HOKAYEM, CHARLES, & ZILIAK, JAMES. 2017. Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch. Working Paper. 13, 101

CANCIAN, MARIA, & LEVINSON, ARIK. 2005. Labor Supply Effects of the Earned Income Tax Credit: Evidence from Wisconsin Supplemental Benefit for Families with Three Children. NBER Working Paper No. 11454. 6

CARD, DAVID, & HYSLOP, DEAN R. 2005. Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers. *Econometrica*, **73**(6), 1723–1770. 5, 7

CHETTY, RAJ. 2012. Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, **80**(3), 969–1018. 2, 7

CHETTY, RAJ, & SAEZ, EMMANUEL. 2013. Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients. *American Economic Journal: Applied Economics*, **5**(1), 1–31. 8, 35

CHETTY, RAJ, FRIEDMAN, JOHN N., OLSEN, TORE, & PISTAFERRI, LUIGI. 2011. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, **126**(2), 749–804. 2

CHETTY, RAJ, GUREN, ADAM, MANOLI, DAY, & WEBER, ANDREA. 2013a. Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities. *NBER Macroeconomics Annual*, **27**(1), 1–56. 1, 7, 25

CHETTY, RAJ, FRIEDMAN, JOHN N., & SAEZ, EMMANUEL. 2013b. Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, **103**(7), 2683–2721. 7, 9, 36

COHANY, SHARON, POLIVKA, ANNE, & ROTHGEB, JENNEFER. 1994. Revisions in the Current Population Survey Effective January 1994. *Employment & Earnings*, **41**. 100

COUNCIL OF ECONOMIC ADVISORS. 1997. Explaining the Decline in Welfare Receipt, 1993-1996. Washington: Executive Office of the President. Retrived at https://clintonwhitehouse4.archives.gov/WH/EOP/CEA/Welfare/Report.html. 11

COUNCIL OF ECONOMIC ADVISORS. 1999. The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update. Washington: Executive Office of the President. Retrived at https://clintonwhitehouse4.archives.gov/WH/EOP/CEA/html/welfare/. 11

DEPARTMENT OF HEALTH AND HUMAN SERVICES. 1999. State Implementation of Major Changes to Welfare Policies, 1992-1998. Washington: US Department of Health and Human Services. Assistant Secretary for Planning and Evaluation. Retrieved at https://aspe.hhs.gov/report/state-implementation-major-changes-welfare-policies-1992-1998. 11

EISSA, NADA, & HOYNES, HILARY. 2006. Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply. *Tax Policy and the Economy*, **20**, 73–110. 6

EISSA, NADA, & LIEBMAN, JEFFREY. 1993. The End of Welfare as We Know It? Behavioral Responses to the Earned Income Tax Credit. *Department of Economics, Harvard University, mimeo*. 8

EISSA, NADA, & LIEBMAN, JEFFREY B. 1996. Labor Supply Response to the Earned Income Tax Credit. *Quarterly Journal of Economics*, **111**(2), 605–637. 1, 6, 8, 20

EISSA, NADA, KLEVEN, HENRIK, & KREINER, CLAUS. 2006. Welfare Effects of Tax Reform and Labor Supply at the Intensive and Extensive Margins. *In:* AGELL, JONAS, & SORENSEN, PETER (eds), *Tax Policy and Labor Market Performance*. MIT Press: Cambridge, MA. 9, 22, 36

EISSA, NADA, KLEVEN, HENRIK, & KREINER, CLAUS. 2008. Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers. *Journal of Public Economics*, **92**(3–4), 795–816. 9, 22, 23, 36, 109

ELLWOOD, DAVID T. 2000. The Impact of the Earned Income Tax Credit and Social Policy Reforms on Work, Marriage, and Living Arrangements. *National Tax Journal*, 1063–1105. 2, 8

FANG, HANMING, & KEANE, MICHAEL P. 2004. Assessing the Impact of Welfare Reform on Single Mothers. *Brookings Papers on Economic Activity*, **2004**(1), 1–116. 2, 8

FLOOD, SARAH, & PACAS, JOSÉ. 2016. Using the Annual Social and Economic Supplement with Current Population Survey Panels. University of Minnesota Working Paper No. 2016–4. 98

GELBER, ALEXANDER, & MITCHELL, JOSHUA. 2011. Taxes and Time Allocation: Evidence from Single Women and Men. *The Review of Economic Studies*, **79**(3), 863–897. 6

GROGGER, JEFF, & KAROLY, LYNN A. 2005. *Welfare Reform: Effects of a Decade of Change*. Harvard University Press: Cambridge, MA. 5, 6, 7, 11, 12

GROGGER, JEFFREY. 2002. The Behavioral Effects of Welfare Time Limits. *American Economic Review Papers and Proceedings*, **92**(2), 385–389. 4

GROGGER, JEFFREY. 2003. The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income Among Female-Headed Families. *Review of Economics and Statistics*, **85**(2), 394–408. 2, 4, 8

GROGGER, JEFFREY, & MICHALOPOULOS, CHARLES. 2003. Welfare Dynamics Under Time Limits. *Journal of Political Economy*, **111**(3), 530–554. 7

HOTZ, V. JOSEPH, & SCHOLZ, JOHN KARL. 2003. The Earned Income Tax Credit. *In:* MOFFITT, ROBERT A. (ed), *Means-Tested Transfer Programs in the United States*. Chicago: University of Chicago Press. 6

HOTZ, V. JOSEPH, MULLIN, CHARLES H., & SCHOLZ, JOHN KARL. 2006. Examining the Effect of the Earned Income Tax Credit on the Labor Market Participation of Families on Welfare. NBER Working Paper No. 11968. 1, 6

HOYNES, HILARY, & PATEL, ANKUR. 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. 6, 32

JONES, DAMON. 2010. Information, Preferences and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401(k) Savings. *American Economic Journal: Applied Economics*, **2**(2), 147–63. 8

JONES, MAGGIE, & ZILIAK, JAMES. 2019. The Antipoverty Impact of the EITC: New Estimates from Survey and Administrative Tax Records. Working Paper. 101

KLEVEN, HENRIK J., & KREINER, CLAUS T. 2005. Labor Supply Behavior and the Design of Tax and Transfer Policy. *Danish Journal of Economics*, **143**(2005), 321–358. 22

KLEVEN, HENRIK J., & KREINER, CLAUS T. 2006. The Marginal Cost of Public Funds: Hours of Work Versus Labor Force Participation. *Journal of Public Economics*, **90**(10–11), 1955–1973. 5, 9, 36

KLINE, PATRICK, & TARTARI, MELISSA. 2016. Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach. *American Economic Review*, **106**(4), 972–1014. 5, 8

KOPCZUK, WOJCIECH, & POP-ELECHES, CRISTIAN. 2007. Electronic Filing, Tax Preparers and Participation in the Earned Income Tax Credit. *Journal of Public Economics*, **91**(7–8), 1351–1367. 8

LOONEY, ADAM, & MANOLI, DAY. 2016. Are There Returns to Experience at Low-Skill Jobs? Evidence from Single Mothers in the United States over the 1990s. Upjohn Institute Working Paper No. 16–255. 4, 6

MAAG, ELAINE. 2005. Paying the Price? Low-Income Parents and the Use of Paid Tax Preparers. New Federalism: National Survey of America's Families, Series B, No. 64. 8, 35

MARTINEZ, ISABEL, SAEZ, EMMANUEL, & SIEGENTHALER, MICHAEL. 2018. Intertemporal Labor Supply Substitution? Evidence from the Swiss Income Tax Holidays. NBER Working Paper No. 24634. 7

MEAD, LAWRENCE. 2014. Overselling the Earned Income Tax Credit. *National Affairs*, **21**, 20–33. 8, 9

MEYER, BRUCE. 2010. The Effects of the Earned Income Tax Credit and Recent Reforms. *In:* BROWN, JEFFREY R. (ed), *Tax Policy and the Economy, Volume 24*. Chicago: University of Chicago Press. 1, 6

MEYER, BRUCE, & MITTAG, NIKOLAS. 2019. Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net. *American Economic Journal: Applied Economics*, **11**(2), 176–204. 102

MEYER, BRUCE, & ROSENBAUM, DAN. 2001. Welfare, The Earned Income Tax Credit, and the Labor Supply of Single Mothers. *Quarterly Journal of Economics*, **116**(3), 1063–1114. 1, 2, 6

MEYER, BRUCE, MOK, WALLACE, & SULLIVAN, JAMES. 2015. Household Surveys in Crisis. *Journal of Economic Perspectives*, **29**(4), 199–226. 101, 102

MILLER, CYNTHIA, KATZ, LAWRENCE F., AZURDIA, GILDA, ISEN, ADAM, SCHULTZ, CAROLINE B., & ALOISI, KALI. 2018. Boosting the Earned Income Tax Credit for Singles: Final Impact Findings From the Paycheck Plus Demonstration in New York City. *New York: MDRC*. 5, 7

MURRAY, CHARLES. 1984. *Losing Ground : American Social Policy, 1950-1980*. Basic Books: New York, NY. 1

NICHOLS, AUSTIN, & ROTHSTEIN, JESSE. 2015. The Earned Income Tax Credit. *In:* MOFFITT, ROBERT A. (ed), *The Economics of Means-Tested Transfer Programs in the United States, Volume 1*. Chicago: University of Chicago Press. 6

OLSON, LYNN M., & DAVIS, AUDREY. 1994. The Earned Income Tax Credit: Views from the Street Level. *Center for Urban Affairs and Policy Research*. 8

ROMICH, JENNIFER L, & WEISNER, THOMAS. 2000. How Families View and Use the EITC: Advance Payment versus Lump Sum Delivery. *National Tax Journal*, 1245–1265. 8, 35

ROSS PHILLIPS, KATHERIN. 2001. Who Knows About the Earned Income Tax Credit? New Federalism: National Survey of America's Families, Series B, No. 64. 8, 35

SAEZ, EMMANUEL. 2002. Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses. *Quarterly Journal of Economics*, **117**(3), 1039–1073. 22, 36

SMEEDING, TIMOTHY, ROSS PHILLIPS, KATHERIN, & O'CONNOR, MICHAEL. 2000. The EITC: Expectation, Knowledge, Use, and Economic and Social Mobility. *National Tax Journal*, **53**(4), 1187–1209. 8, 35

YAGAN, DANNY. 2019. Employment Hysteresis from the Great Recession. *Journal of Political Economy*, **127**(5), forthcoming. 16

TABLE 1: DESCRIPTIVE STATISTICS

	Single Women		
	All	With Children	Without Children
	(1)	(2)	(3)
Age	32.05	34.34	30.68
Age of Youngest Child	7.51	7.51	.
Number of Children	0.67	1.79	0.00
High School & Below	0.41	0.55	0.33
Some College	0.34	0.32	0.36
College Degree	0.24	0.13	0.31
Employment Rate	0.73	0.68	0.75
Labor Force Participation Rate	0.78	0.76	0.80
Welfare Participation Rate	0.20	0.39	0.03
Earnings	23,519	21,365	24,984
Observations	4,674,064	1,746,591	2,927,473

Notes: This table shows means of demographics, labor market variables, and earnings (in 2018 USD) for single women. Column (1) shows all single women, column (2) shows single women with EITC-eligible children, while column (3) shows single women without EITC-eligible children. EITC-eligible children are defined as children under 19 (24 if in full-time education) and living at home for at least half of the year. Employment and labor force participation rates are based on respondents' activities last week. The sample includes single women aged 20-50 using the March and monthly CPS files combined, pooling all years from 1968 to 2018.

TABLE 2: THREE-YEAR EFFECTS OF FEDERAL EITC REFORMS ON SINGLE MOTHERS

	Employment		Participation	
	(1)	(2)	(3)	(4)
Panel A: 1975 Reform				
All Educations	-0.713	-0.678	0.708	0.706
	(1.377)	(1.338)	(1.330)	(1.297)
Low-Education	-0.118	-0.248	1.087	0.991
	(1.687)	(1.654)	(1.617)	(1.582)
Panel B: 1986 Reform				
All Educations	-1.028	-0.178	-1.571	-0.877
	(1.084)	(1.036)	(1.002)	(0.969)
Low-Education	-0.411	0.293	-0.879	-0.339
	(1.439)	(1.405)	(1.334)	(1.309)
Panel C: 1990 Reform				
All Educations	-0.347	-1.608	-1.115	-2.330**
	(1.068)	(1.005)	(0.998)	(0.947)
Low-Education	-0.170	0.925	-0.804	0.075
	(1.479)	(1.430)	(1.391)	(1.347)
Panel D: 1993 Reform				
All Educations	3.060***	2.949***	3.055***	2.965***
	(0.508)	(0.475)	(0.467)	(0.440)
Low-Education	3.538***	3.627***	3.604***	3.592***
	(0.758)	(0.716)	(0.711)	(0.674)
Panel E: 2009 Reform				
All Educations	-3.554***	-3.920***	-2.162**	-2.490***
	(1.074)	(1.030)	(0.966)	(0.932)
Low-Education	-2.170	-2.786*	-2.001	-2.476*
	(1.486)	(1.443)	(1.374)	(1.337)
Demographic Controls:		X		X

Notes: This table shows DiD estimates of the effects of the five federal EITC reforms after three years. The estimates are obtained from a modified version of equation (1) in which the year dummies are collapsed into a post dummy, equal to one for the three years after each reform. Panels A-D are based on comparing single women with children to those without children, while Panel E is based on comparing single women with 3+ children to those without children. The columns show results for different outcomes (weekly employment and weekly participation) and different controls, while the rows show results for the full sample and for those with low education (defined as having a high school degree or less). The sample includes single women aged 20-50. Panels A-C use the March CPS files alone, while Panels D-E use the March and monthly files combined. Robust standard errors are clustered at the individual level.

TABLE 3: EXTENSIVE MARGIN ELASTICITIES WHEN IGNORING CONFOUNDERS

	Earnings and Tax Parameters			Employment Effects			Participation Effects		
	Earnings	τ	$\Delta(1-\tau)$	P	ΔP	ε	P	ΔP	ε
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Any Children:	14,685	0.399	0.077	0.610	0.130	**1.67**	0.691	0.129	**1.46**
1 Child:	16,197	0.387	0.034	0.681	0.085	**2.21**	0.756	0.083	**1.96**
2 Children	14,703	0.404	0.105	0.613	0.138	**1.29**	0.695	0.137	**1.12**
3 Children:	12,110	0.414	0.124	0.477	0.196	**1.93**	0.570	0.202	**1.67**
4+ Children:	8,327	0.428	0.159	0.318	0.262	**2.96**	0.413	0.252	**2.19**

Notes: This table shows estimates of the extensive margin elasticities based on the 1993 reform, assuming that the entire DiD effect between 1993-2003 can be attributed to the EITC. Columns (1)-(3) show predicted earnings and tax parameters, columns (4)-(6) show employment effects, and columns (7)-(9) show participation effects. Each statistic is shown for all single mothers in the first row and separately by number of children in the following rows. The earnings measure in column (1) is based on predicted earnings for non-workers (estimated using equation 3) and actual earnings for workers. The changes in employment and participation rates (ΔP) as well as the EITC-induced change in the net-of-tax rate ($\Delta(1-\tau)$) represent difference-in-differences comparing single women with and without children. The elasticities in columns (6) and (9) are calculated using equation (2). See section C in the appendix for additional details.

TABLE 4: EFFECTS OF THE 1993 EITC REFORM BY AGE OF YOUNGEST CHILD

	Employment			Participation		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average 3-Year Effects (Before PRWORA)						
Youngest Aged 0-13	4.516***	3.771***	2.161***	4.335***	3.723***	2.259***
	(0.396)	(0.403)	(0.383)	(0.365)	(0.371)	(0.355)
Youngest Aged 14+	0.557	-0.032	-1.580**	0.176	-0.327	-1.718***
	(0.662)	(0.662)	(0.638)	(0.599)	(0.599)	(0.579)
Panel B: Average 10-Year Effects (After PRWORA)						
Youngest Aged 0-13	11.595***	9.523***	7.611***	11.072***	9.210***	7.492***
	(0.322)	(0.346)	(0.330)	(0.296)	(0.318)	(0.307)
Youngest Aged 14+	2.533***	0.560	-1.151**	2.316***	0.514	-1.017**
	(0.531)	(0.542)	(0.527)	(0.479)	(0.490)	(0.478)
Controls:						
State Fixed Effects		X	X		X	X
Unemployment × Kids		X	X		X	X
Demographics			X			X

Notes: This table shows effects of the 1993 reform by the age of youngest child across different specifications and outcomes. The estimates are obtained from versions of equation (5) where the welfare category variable corresponds to bins of the age of the youngest child. Panel A shows 3-year effects (before PRWORA) while Panel B shows 10-year effects (after PRWORA). Each panel shows the effects among mothers with youngest children aged 0-13 and youngest children aged 14+. The different columns show different outcomes (weekly employment and weekly participation) and specifications with different controls: no controls in columns (1) and (4), adding unemployment-by-kids and state fixed effects in columns (2) and (5), and finally adding demographics in columns (3) and (6). The sample includes single women aged 20-50 using the March and monthly CPS files combined. Robust standard errors are clustered at the individual level.

TABLE 5: EFFECTS OF THE 1993 EITC REFORM BY PROBABILITY OF AFDC PARTICIPATION

	Employment			Participation		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average 3-Year Effects (Before PRWORA)						
Deciles 3-10 of AFDC Prob.	4.375***	3.826***	2.137***	4.188***	3.728***	2.191***
	(0.398)	(0.405)	(0.379)	(0.366)	(0.373)	(0.352)
Deciles 1-2 of AFDC Prob.	0.505	-0.114	-1.074*	0.339	-0.170	-1.078*
	(0.645)	(0.647)	(0.624)	(0.586)	(0.589)	(0.569)
Panel B: Average 10-Year Effects (After PRWORA)						
Deciles 3-10 of AFDC Prob.	11.484***	9.598***	7.445***	10.972***	9.264***	7.325***
	(0.323)	(0.348)	(0.329)	(0.298)	(0.321)	(0.305)
Deciles 1-2 of AFDC Prob.	2.340***	0.790	-0.012	2.325***	0.827*	0.088
	(0.519)	(0.531)	(0.514)	(0.470)	(0.481)	(0.468)
Controls:						
State Fixed Effects		X	X		X	X
Unemployment × Kids		X	X		X	X
Demographics			X			X

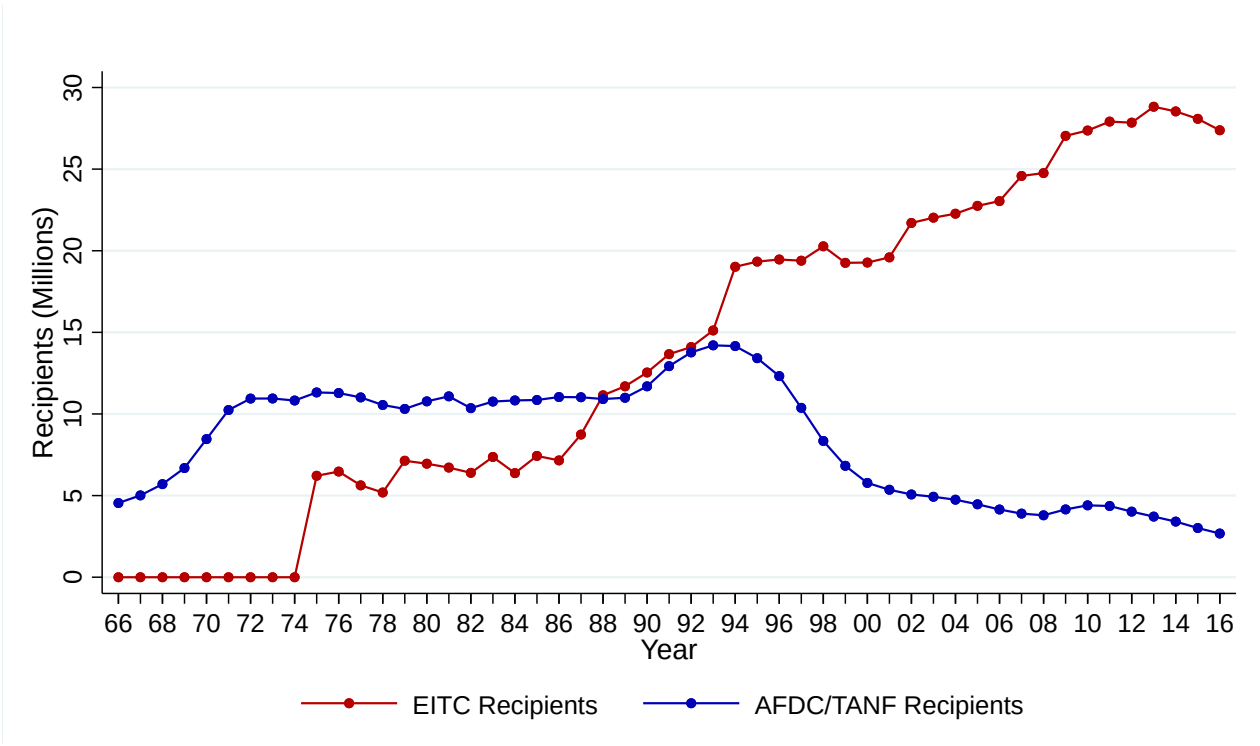
Notes: This table shows effects of the 1993 reform by probability of AFDC participation across different specifications and outcomes. The estimates are obtained from versions of equation (5) where the welfare category variable is estimated from (6). Panel A shows 3-year effects (before PRWORA) while Panel B shows 10-year effects (after PRWORA). Each panel shows the effects in the bottom two deciles and in the top eight deciles of the predicted probability of AFDC participation. The different columns show different outcomes (weekly employment and weekly participation) and specifications with different controls: no controls in columns (1) and (4), adding unemployment-by-kids and state fixed effects in columns (2) and (5), and finally adding demographics in columns (3) and (6). The sample includes single women aged 20-50 using the March and monthly CPS files combined. Robust standard errors are clustered at the individual level.

TABLE 6: EFFECTS OF THE 1993 EITC REFORM WITH WAIVER AND BUSINESS CYCLE CONTROLS

	Employment			Participation		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Year-3 Effects (Before PRWORA)						
All Educations	4.684***	-0.053	0.836	5.074***	-0.027	0.795
	(0.671)	(1.076)	(0.996)	(0.614)	(0.996)	(0.930)
Low-Education	5.494***	-0.105	1.060	5.923***	-0.205	0.893
	(1.017)	(1.614)	(1.511)	(0.950)	(1.527)	(1.431)
Panel B: Year-10 Effects (After PRWORA)						
All Educations	13.009***	9.753***	9.017***	12.877***	9.550***	9.005***
	(0.648)	(1.036)	(0.982)	(0.590)	(0.945)	(0.904)
Low-Education	16.050***	12.976***	12.021***	16.120***	13.501***	12.837***
	(1.003)	(1.633)	(1.563)	(0.935)	(1.527)	(1.468)
Controls:						
State Fixed Effects		X	X		X	X
Unemployment × Kids		X	X		X	X
Welfare Waivers × Kids		X	X		X	X
Demographics			X			X

Notes: This table shows effects of the 1993 reform across different samples, specifications, and outcomes. The estimates are obtained from versions of equation (7). Panel A shows 3-year effects (before PRWORA) while Panel B shows 10-year effects (after PRWORA). Each panel shows the effect among single women of all education levels and single women with a high school education or below. The different columns show different outcomes (weekly employment and weekly participation) and specifications with different controls: no controls in columns (1) and (4), adding unemployment-by-kids and state fixed effects in columns (2) and (5), and finally adding demographics in columns (3) and (6). The sample includes single women aged 20-50 using the March and monthly CPS files combined. Robust standard errors are clustered at the individual level.

FIGURE 1: LONG-RUN EVOLUTION OF EITC AND CASH WELFARE

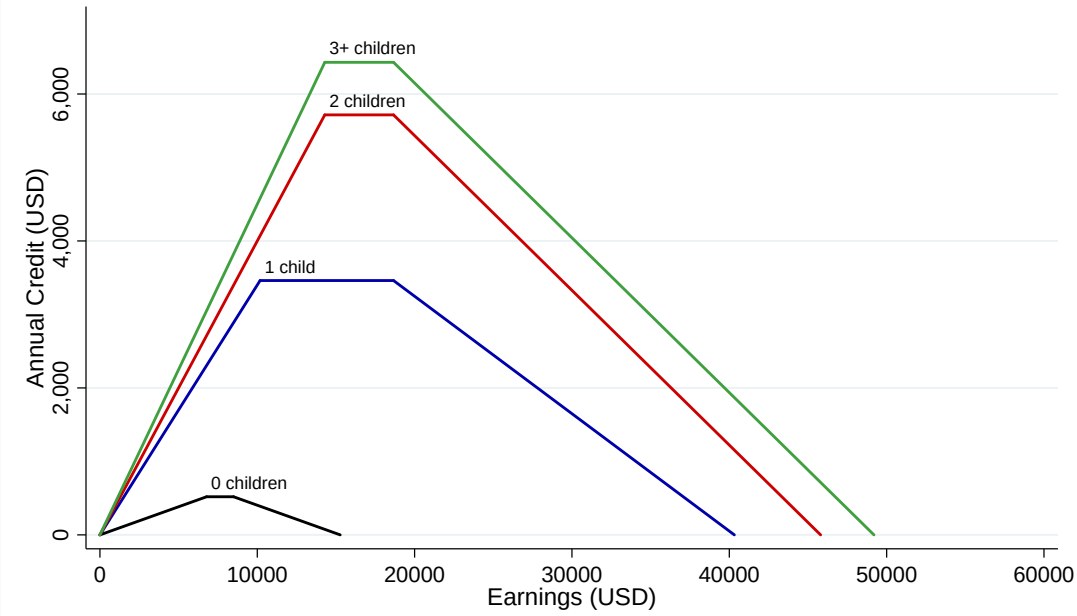


Source: Internal Revenue Service (EITC) and Department of Health and Human Services (AFDC/TANF).

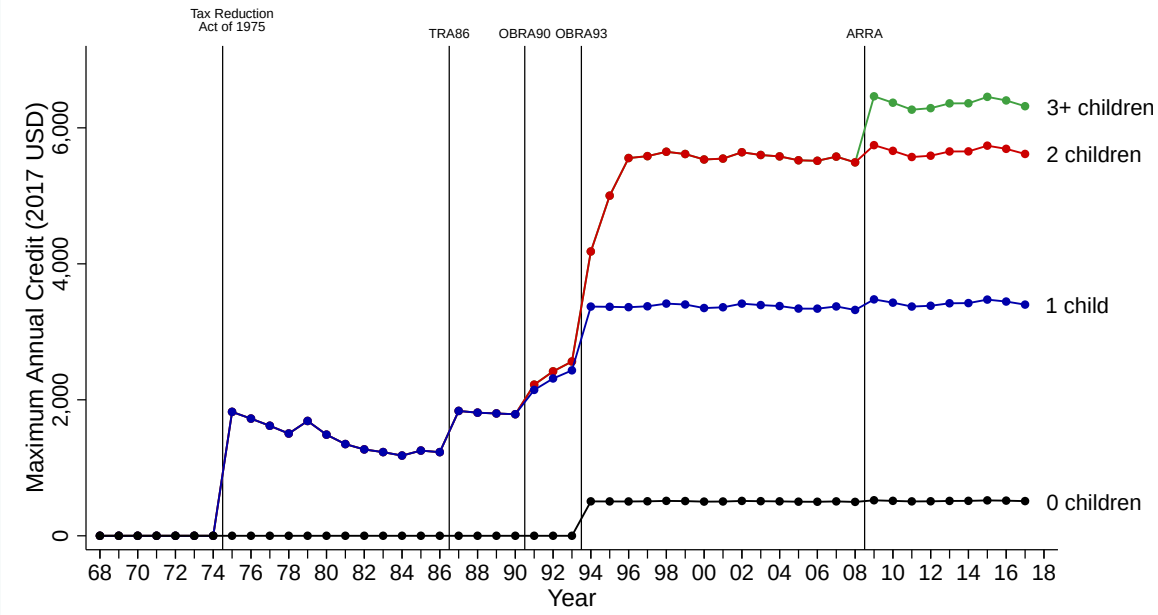
Notes: The red series show the annual number of federal EITC recipients between 1966-2016. The blue series show the average monthly number of Aid to Families with Dependent Children (AFDC) recipients between 1966-1996, and the average monthly number of Temporary Assistance for Needy Families (TANF) recipients between 1997-2016.

FIGURE 2: EITC PARAMETERS

A: EITC Schedule in 2018

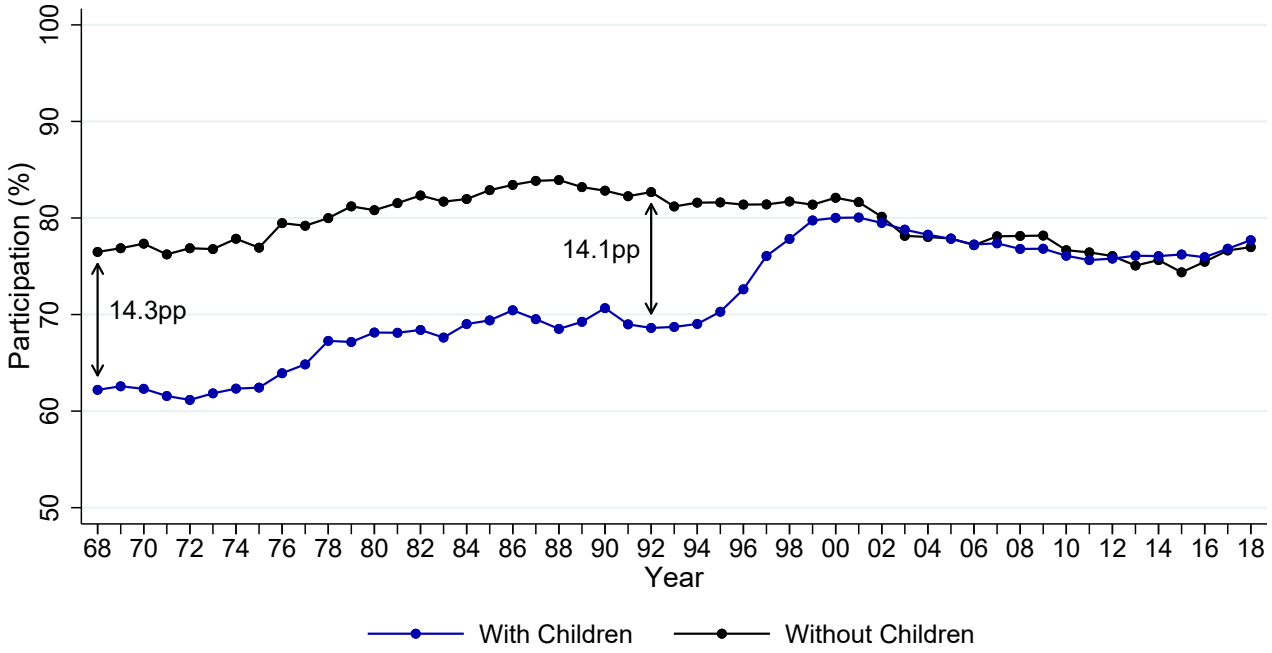


B: EITC Maximum Credit Over Time



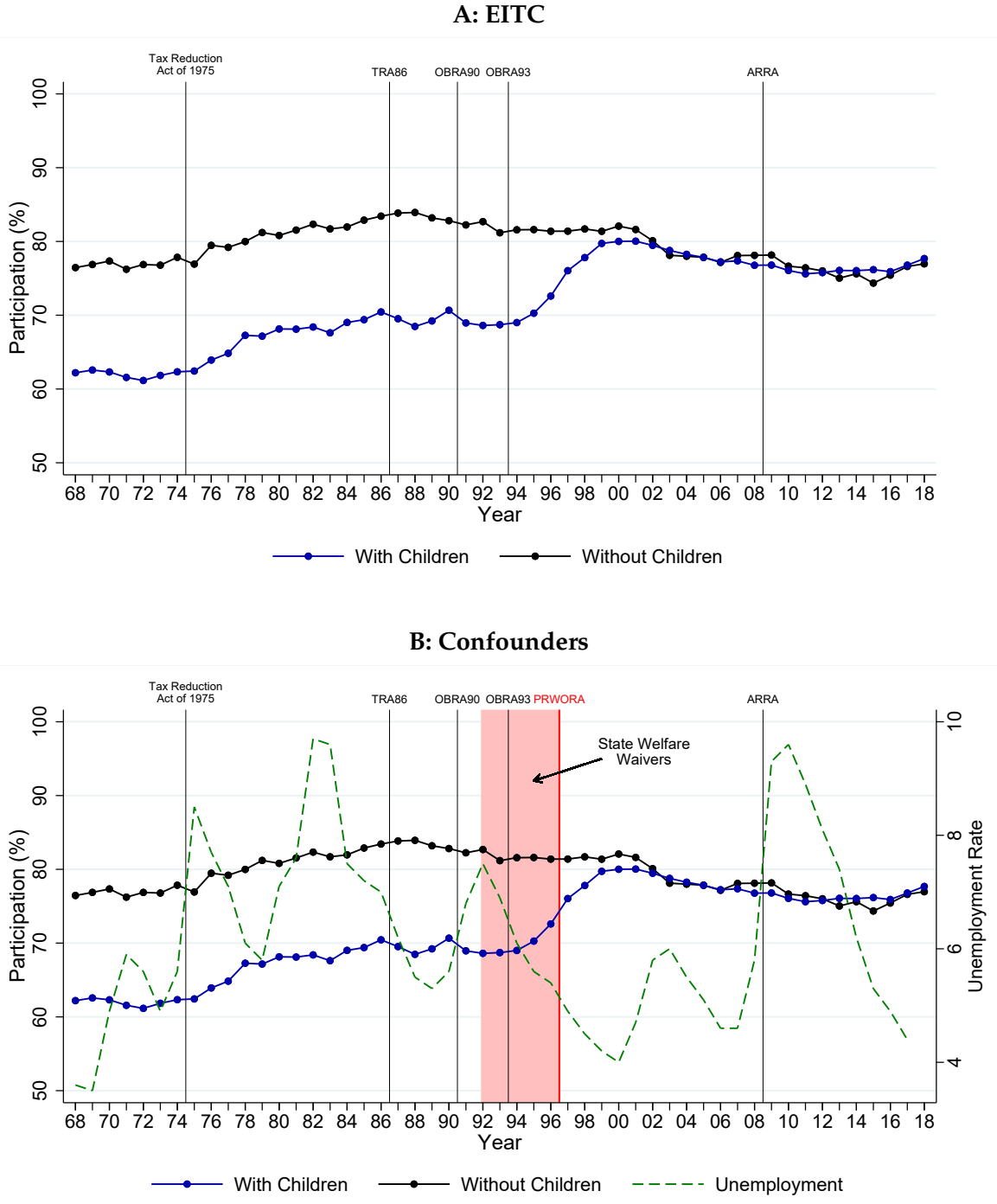
Notes: This figure shows federal EITC parameters for different family sizes. Panel A shows the 2018 EITC schedule as a function of total family earnings for families with 0, 1, 2, and 3+ EITC-eligible children. Panel B shows the maximum annual credit for families with 0, 1, 2, and 3+ EITC-eligible children between 1968 and 2018, in 2018 USD.

FIGURE 3: FIFTY YEARS OF LABOR FORCE PARTICIPATION FOR SINGLE WOMEN



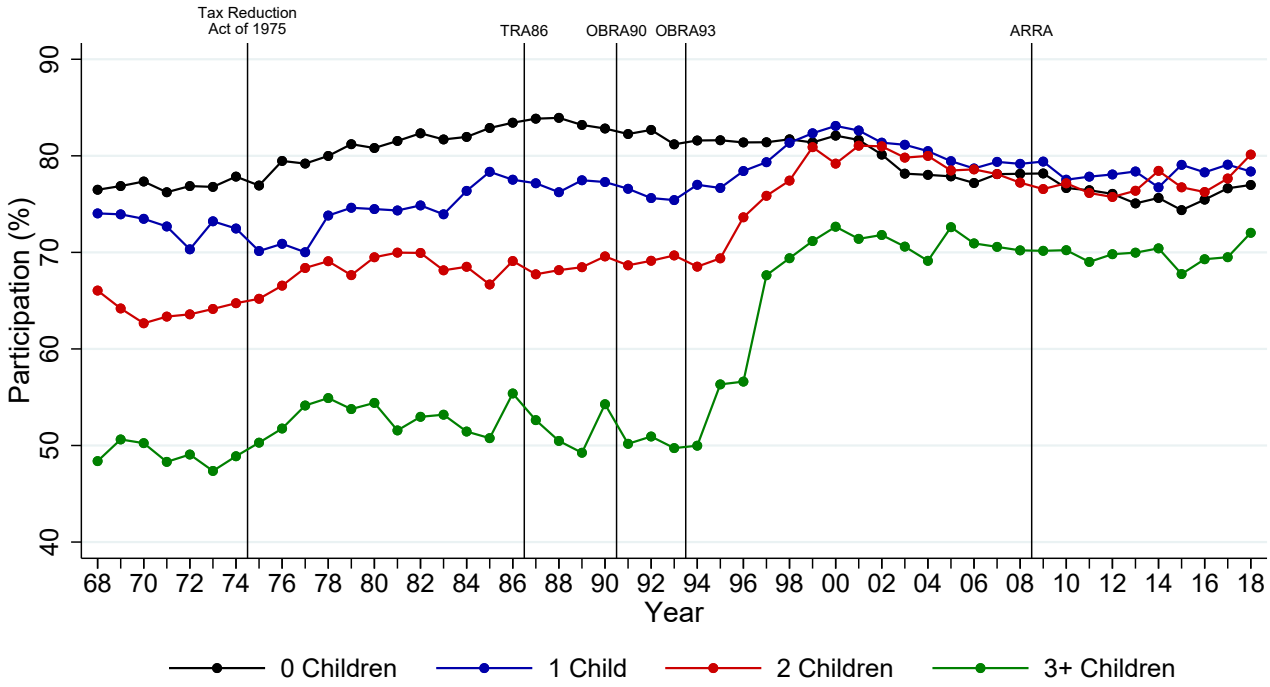
Notes: This figure shows the weekly labor force participation rate of single women with and without children between 1968 and 2018. The sample includes single women aged 20-50 using the March CPS files.

FIGURE 4: WHAT EXPLAINS THE LONG-RUN EVOLUTION?



Notes: This figure shows the weekly labor force participation rate of single women with and without children between 1968 and 2018. Panel A marks the timing of the five federal EITC reforms by vertical black lines. Panel B adds the timing of statewide welfare waivers (red shaded area), federal welfare reform (red line), and the national unemployment rate (green dashed line). The sample includes single women aged 20-50 using the March CPS files.

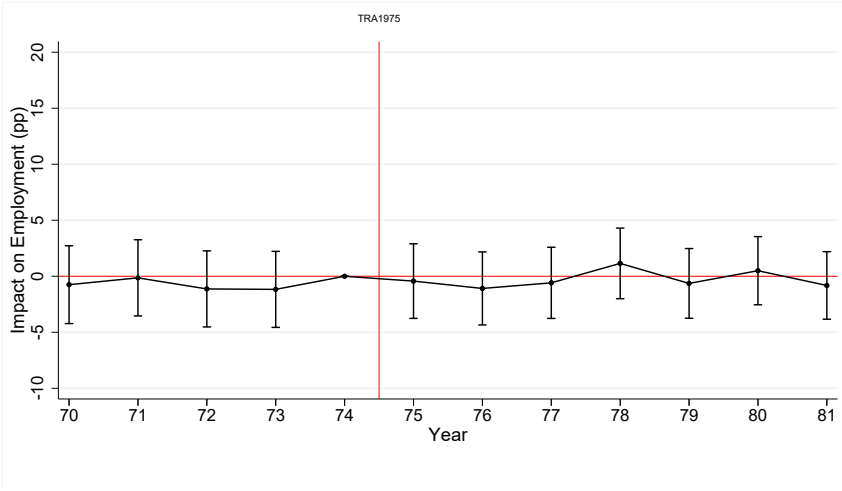
FIGURE 5: FIFTY YEARS OF PARTICIPATION FOR SINGLE WOMEN, BY FAMILY SIZE



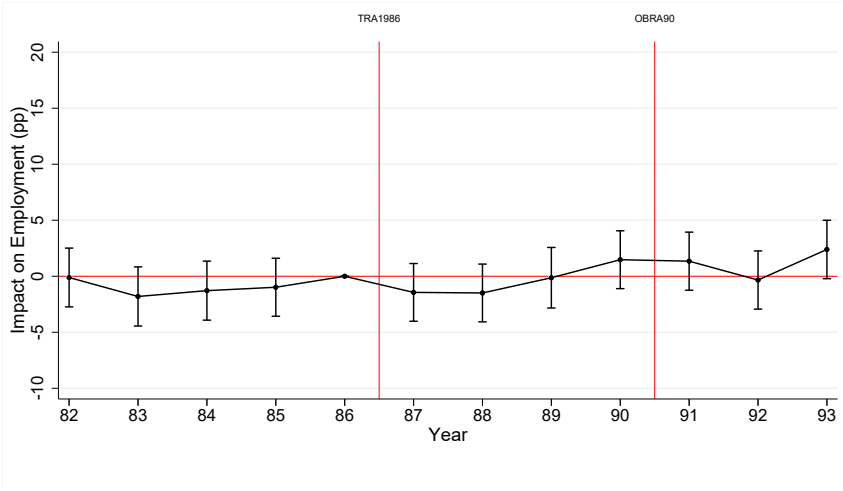
Notes: This figure shows the weekly labor force participation rate of single women with 0, 1, 2, and 3+ children between 1968 and 2018. The timing of the five federal EITC reforms are marked by vertical black lines. The sample includes single women aged 20-50 using the March CPS files.

FIGURE 6: DID EVENT STUDIES OF ALL FEDERAL EITC REFORMS

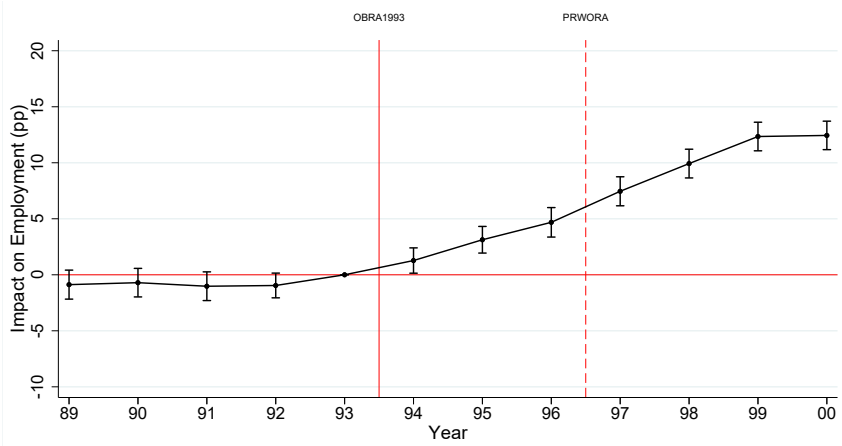
A: 1975 Reform, With vs Without Children



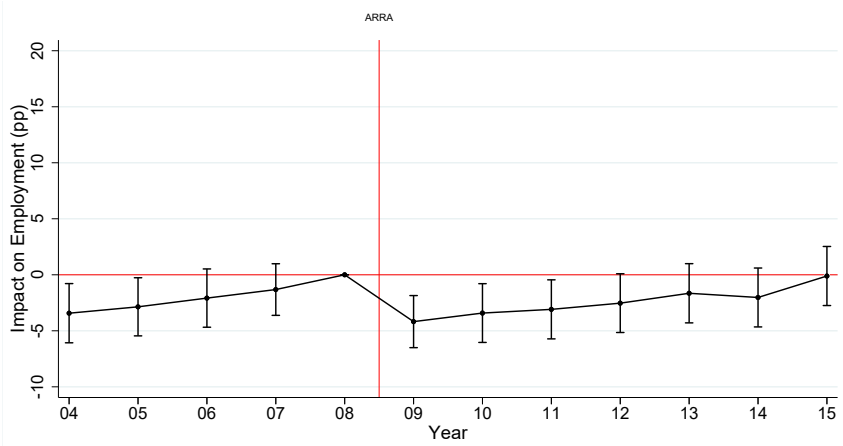
B: 1986 and 1990 Reforms, With vs Without Children



C: 1993 Reform, With vs Without Children

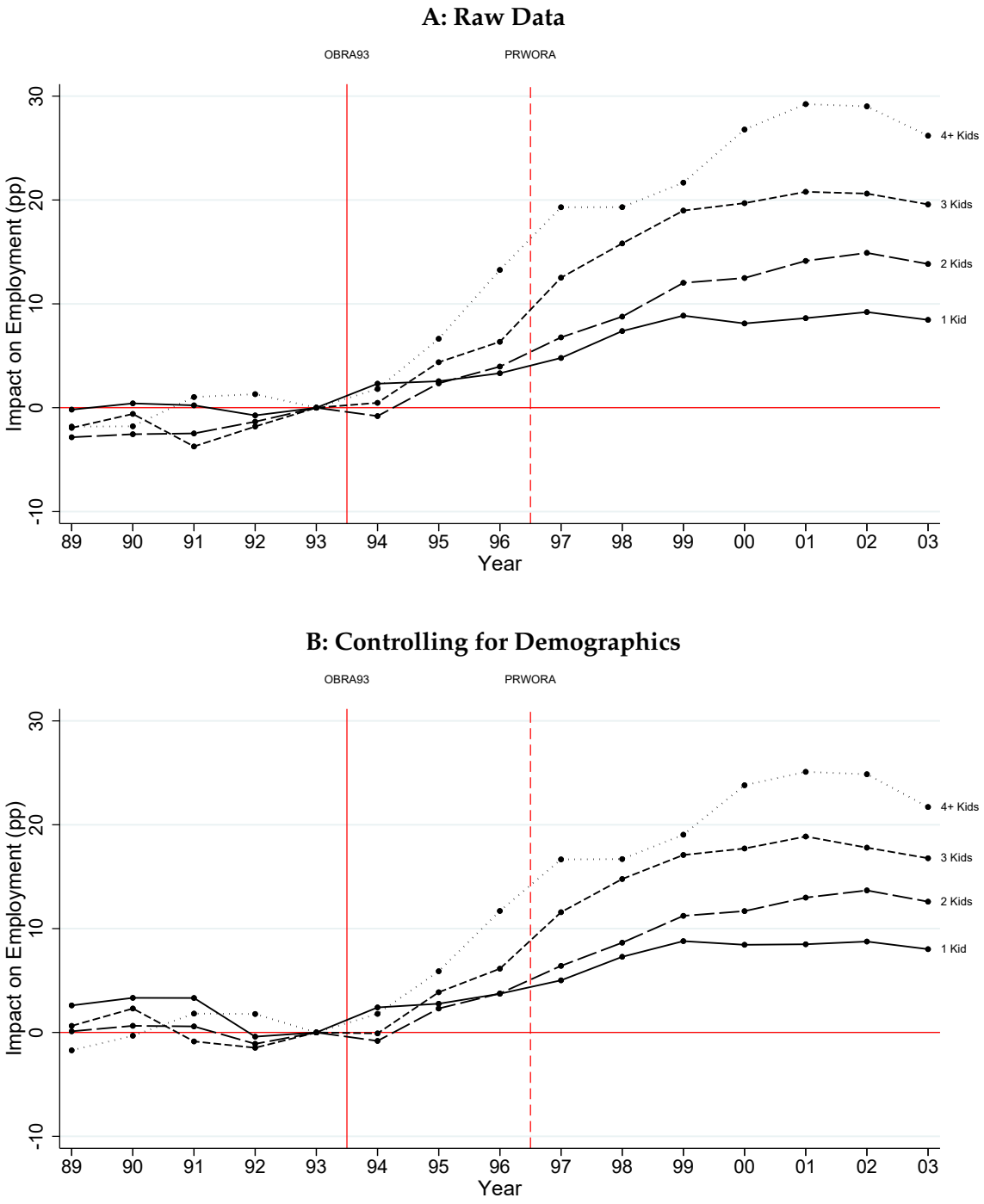


D: 2009 Reform, 3+ vs Without Children



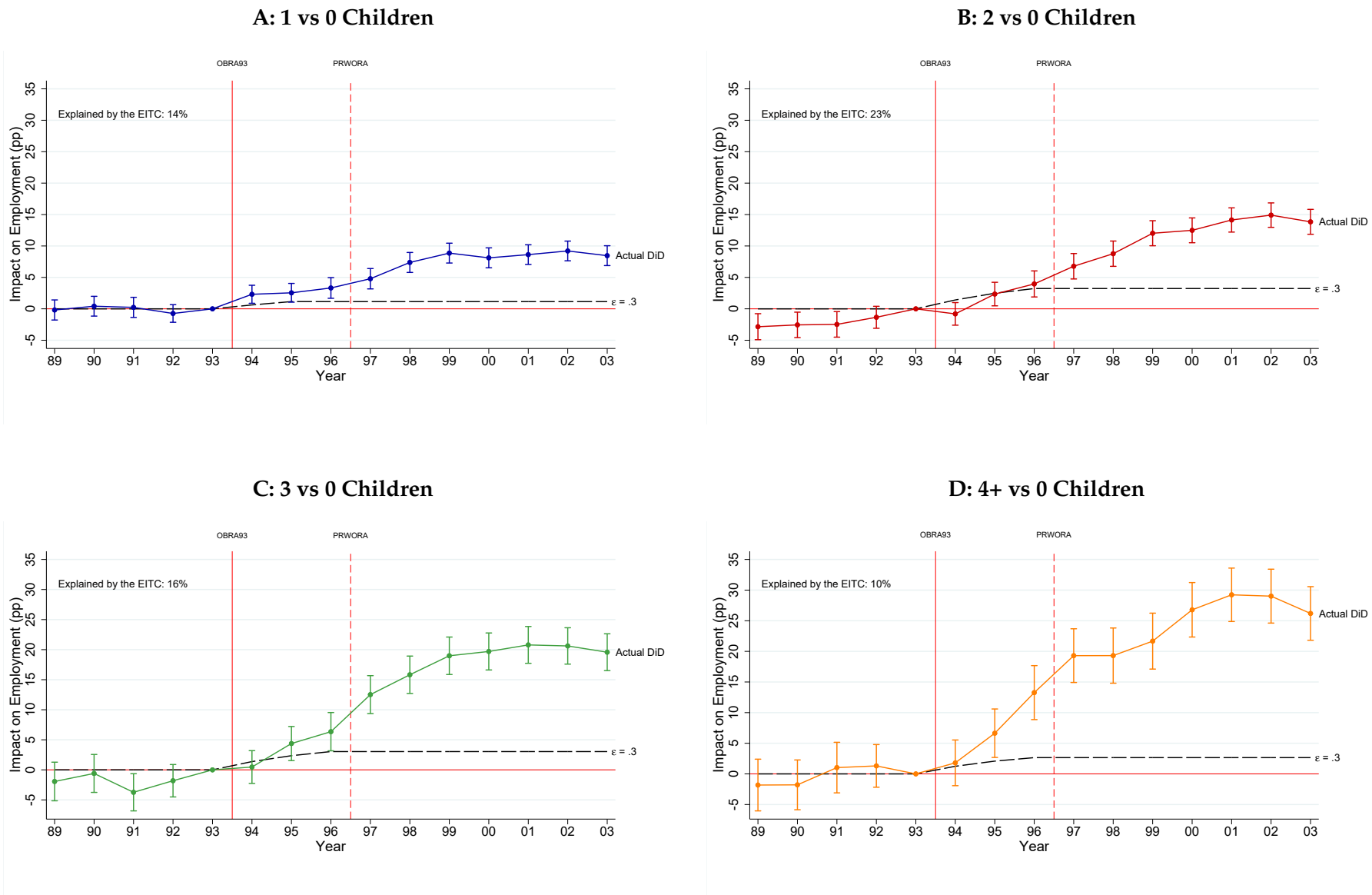
Notes: This figure shows DiD event studies for the five federal EITC reforms. The graphs plot estimates of γ_t based on specification (1) without demographic controls. Panels A-C are based on comparing single women with and without children, while Panel D is based on comparing single women with 3+ children to those without children. In each panel, the difference in the pre-reform year is normalized to zero. The dependent variable is weekly employment. The sample includes single women aged 20-50. Panels A-B use the March CPS files alone, while Panels C-D use the March and monthly files combined. The 95% confidence intervals are based on robust standard errors clustered at the individual level.

FIGURE 7: A FANNING-OUT BY NUMBER OF CHILDREN



Notes: This figure shows DiD event studies for the 1993 reform by number of EITC-eligible children (1, 2, 3, 4+). The graphs plot DiD coefficients γ_t based on an extension of specification (1) that includes dummies for each family size. Hence, each series shows the difference between single mothers with a given number of children and single women without children, normalized to zero in 1993. Panel A shows raw estimates, while panel B controls for demographic composition: dummies for the age of the woman (six categories), dummies for the age of the youngest child (seven categories), and dummies for education (three categories). The dependent variable is weekly employment. The sample includes single women aged 20-50 using the March and monthly CPS files combined.

FIGURE 8: ACTUAL VS SIMULATED RESPONSES TO THE 1993 REFORM

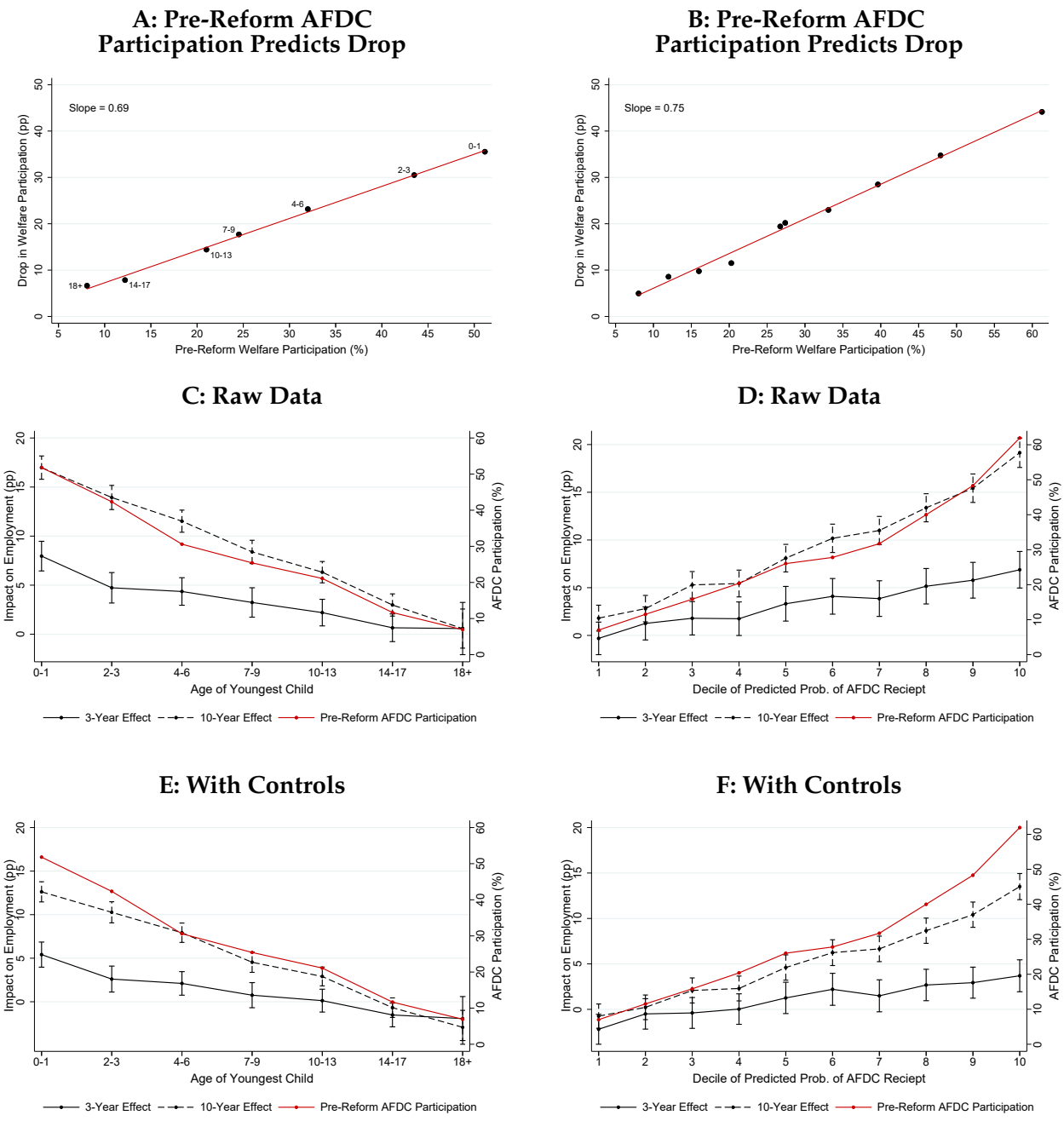


Notes: This figure shows actual and simulated DiD event studies for the 1993 reform, by number of EITC-eligible children. The actual DiD series plot DiD coefficients γ_t based on an extension of specification (1) with separate dummies for each family size. The specification does not include demographic controls and the dependent variable is weekly employment. The simulated DiD series (black dashed lines) plot ΔP_t calculated from equation (4), assuming an elasticity of 0.3. See section C in the appendix for additional details. The fraction explained by the EITC equals the simulated DiD estimate in 2003 divided by the actual DiD estimate in 2003. The sample includes single women aged 20-50 using the March and monthly CPS files combined. The 95% confidence intervals are based on robust standard errors clustered at the individual level.

FIGURE 9: EFFECTS OF THE 1993 EITC REFORM BY WELFARE TREATMENT INTENSITY

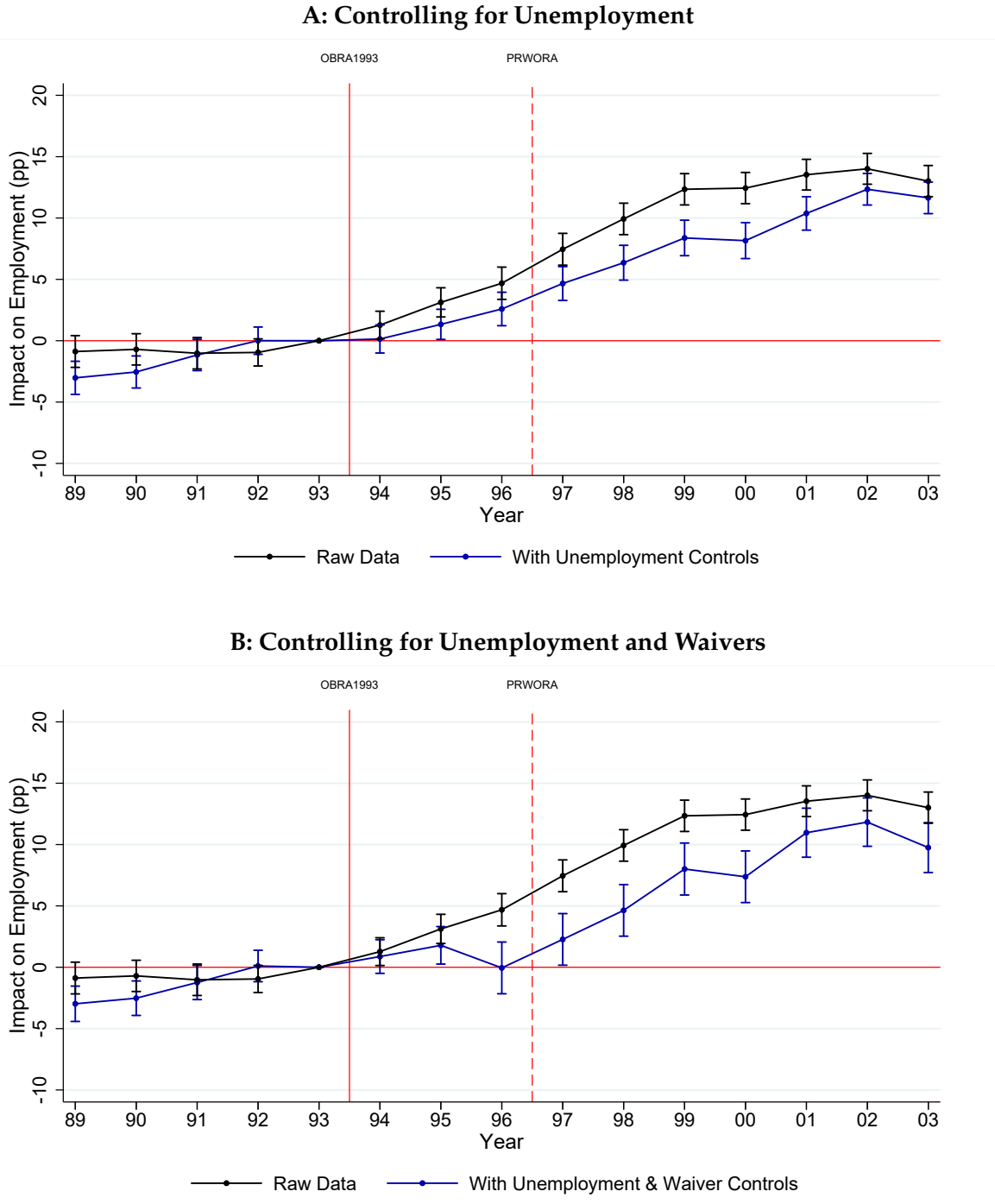
BY AGE OF YOUNGEST CHILD

BY PROBABILITY OF AFDC PARTICIPATION



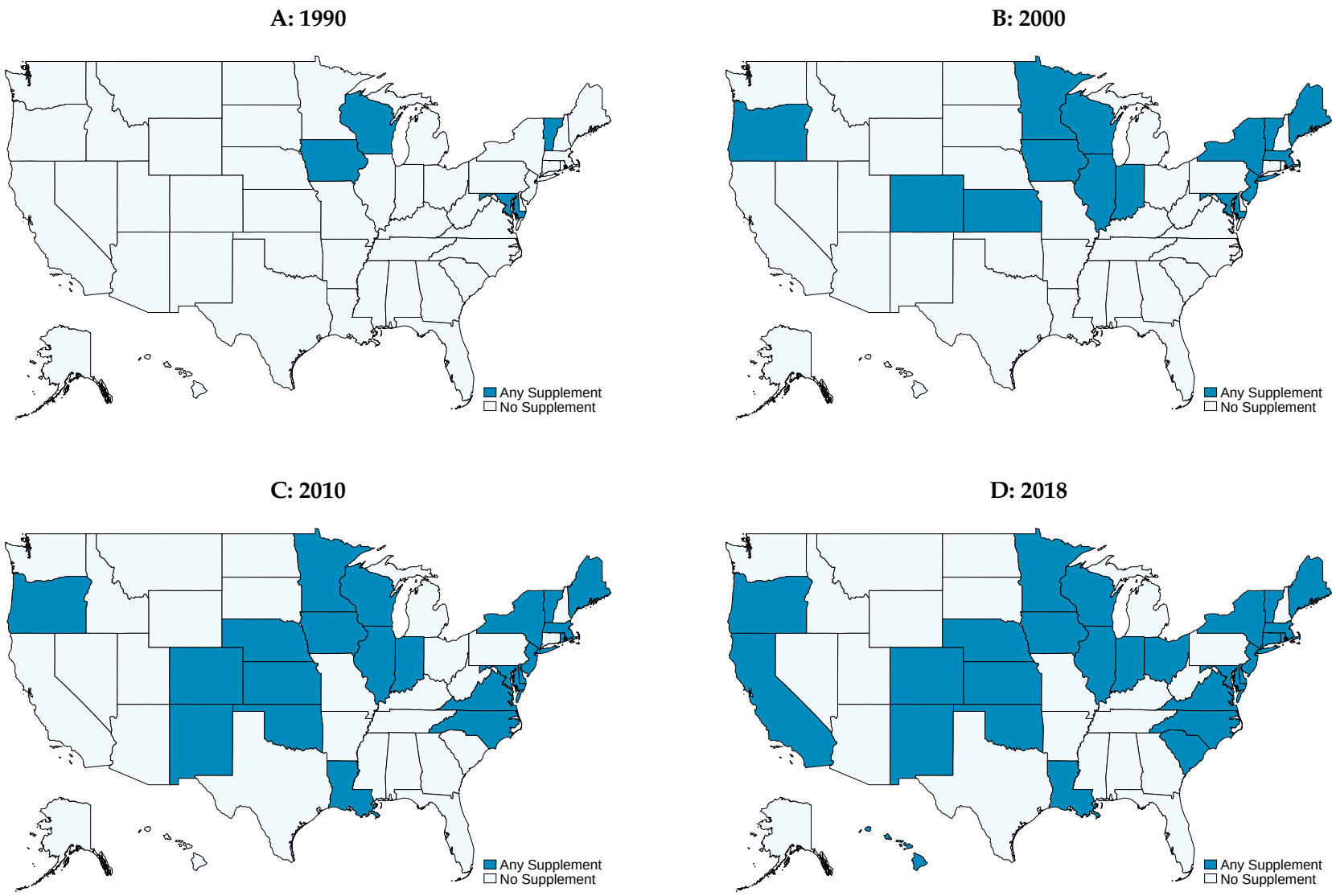
Notes: This figure shows DiD effects of the 1993 reform by welfare treatment intensity, estimated from equation (5). The left panels use the age of youngest child as a proxy for welfare treatment, while the right panels use deciles of a predicted probability of AFDC receipt (estimated from eq. 6) as the proxy for welfare treatment. The top row plots the level of AFDC participation before the reform (1993) against the drop in welfare participation after the reform (1993-2003) across bins of the welfare treatment proxy. The next rows plot the 3-year effects (solid black) and 10-year effects (dashed black) by welfare treatment intensity in the raw data (middle row) and after controlling for business cycle effects and demographics (bottom row). The 3-year and 10-year effects correspond to the coefficient on a 1994-1996 dummy and a 1994-20003 dummy, respectively, The dependent variable is weekly employment. The sample includes single women aged 20-50 using the March and monthly CPS files combined. The 95% confidence intervals are based on robust standard errors clustered at the individual level.

FIGURE 10: HOW MUCH CAN BE EXPLAINED BY THE BUSINESS CYCLE AND WAIVERS?



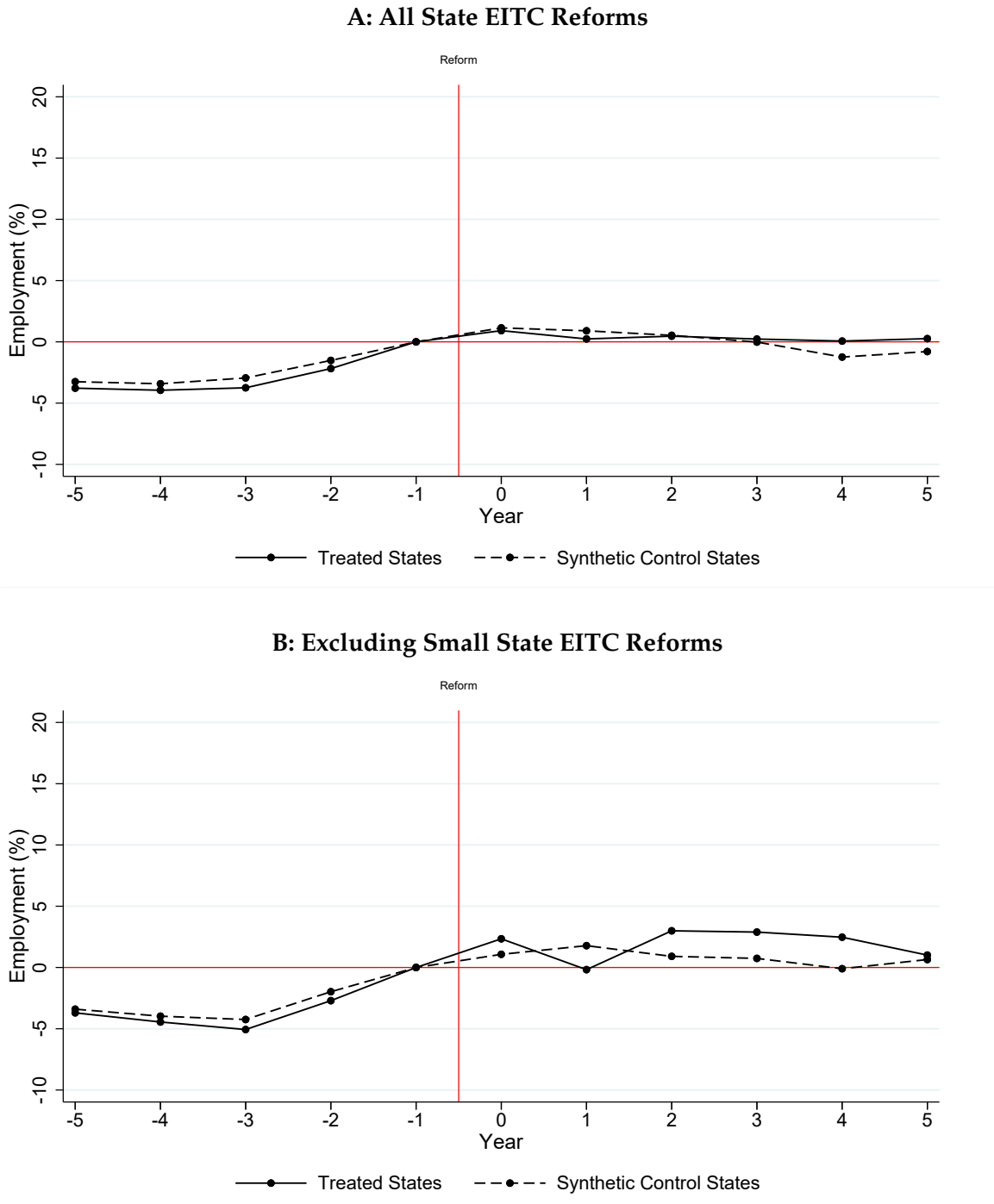
Notes: This figure shows DiD event studies for the 1993 reform with controls for unemployment and waivers. The graphs plot DiD coefficients γ_t based on equation (7). In both panels, the black series show the raw DiD without controls. In panel A, the blue series includes controls for state unemployment by kids and state fixed effects. In panel B the blue series adds controls for waivers by kids. The dependent variable is weekly employment. The sample includes single women aged 20-50 using the March and monthly CPS files combined. The 95% confidence intervals are based on robust standard errors clustered at the individual level.

FIGURE 11: STATES WITH EITC SUPPLEMENTS



Notes: This figure shows which states had an EITC supplement (dark blue) and which states did not have a supplement (light blue) in 1990, 2000, 2010, and 2018.

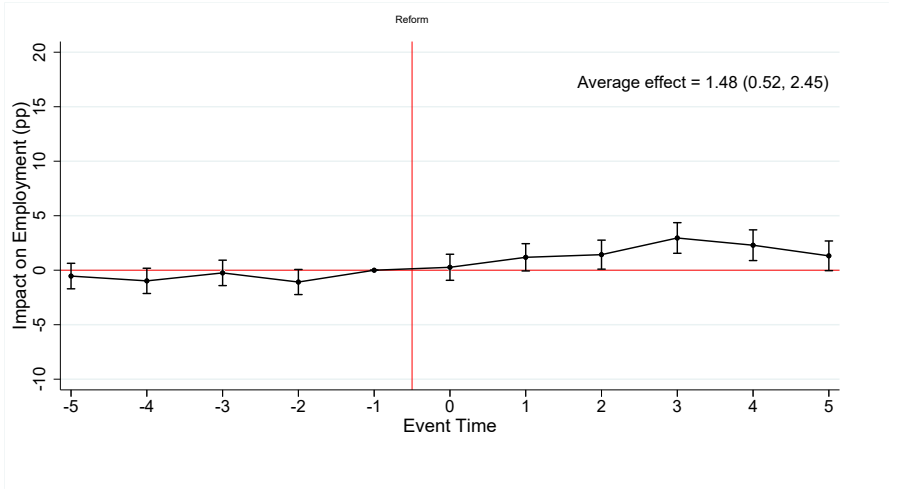
FIGURE 12: SYNTHETIC CONTROL ANALYSIS OF STATE EITC REFORMS



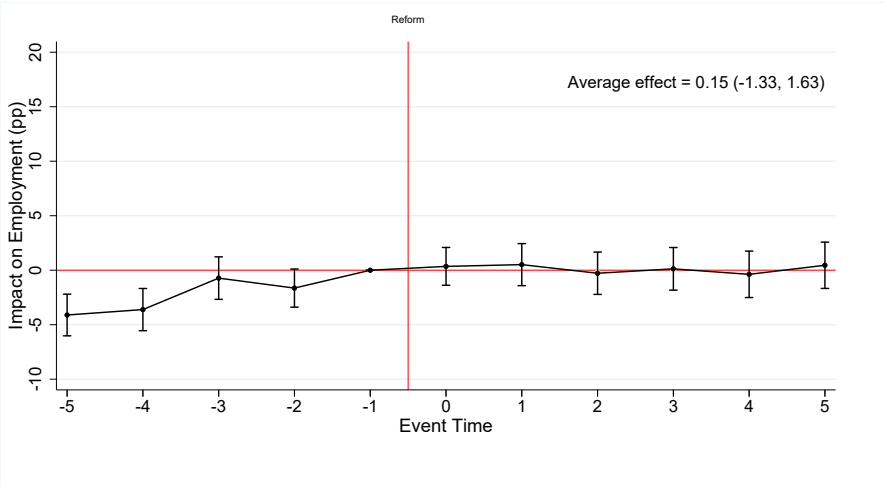
Notes: This figure shows stacked event studies of state EITC reforms using a synthetic control approach. The graphs plot employment rates for single mothers in treatment and synthetic control states across event time, normalized to zero in the last pre-reform year (event time -1). Panel A includes all states that instituted an EITC supplement before 2015, while panel B includes only states that instituted a "large" EITC supplement. Large supplements are defined as refundable credits equal to at least 10% of the federal credit. In both panels, the synthetic control states are constructed from states that never instituted an EITC supplement. For each treatment state, a synthetic control state is constructed by matching on the employment rate in the five pre-reform years. Table A.VI shows the make-up of each synthetic state. For states with supplements enacted before 1993 the sample is based on March CPS files alone, while for states with supplements enacted after 1993 the sample is based on and March and monthly CPS files combined. The outcome is weekly employment and the sample includes single mothers aged 20-50. See section D in the appendix for additional details.

FIGURE 13: STACKED EVENT STUDIES OF ALL STATE AND FEDERAL EITC REFORMS

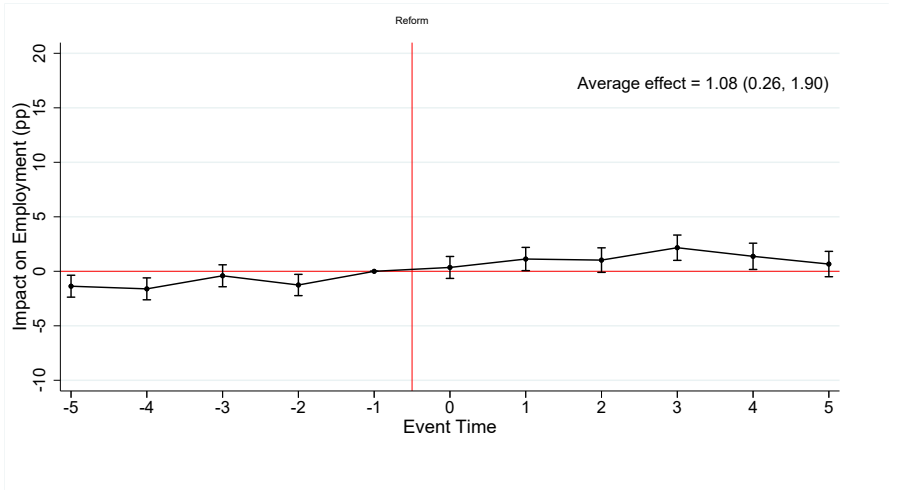
A: Federal Reforms



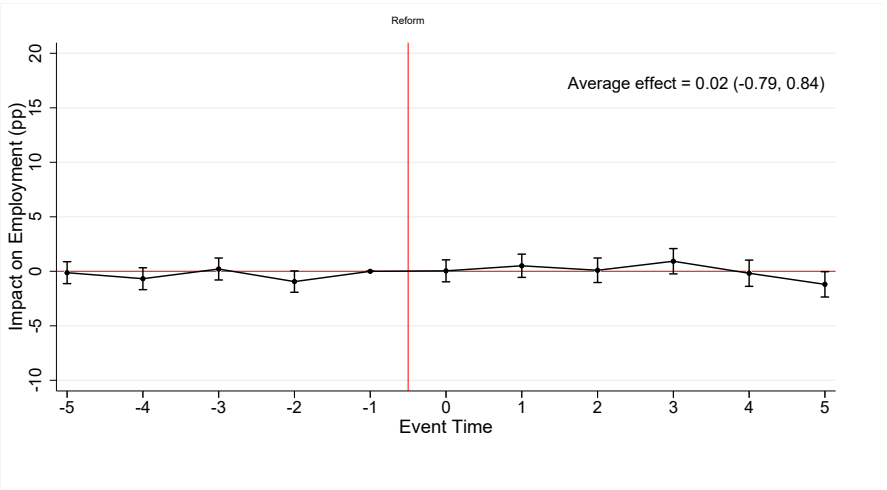
B: State Reforms



C: State and Federal Reforms



D: State and Federal Reforms, Adjusted for Pre-Trends



Notes: This figure shows stacked event studies of all state and federal EITC reforms. The graphs plot DiD coefficients based on comparing single women with and without children across event time, normalized to zero in the pre-reform year (event time -1). The specifications control for demographics, and for the impact of waivers and unemployment around the 1993 federal reform (i.e., the controls in equation (7), interacted with an OBRA93 indicator). Panel A includes all federal reforms, Panel B includes all state reforms. while Panel C includes all state and federal reforms together. Panel D is similar to panel C, but adjusts for group-specific linear pre-trends. Each panel reports the average effect across the post-reform years, with 95% confidence intervals in parentheses. For reforms enacted before 1993 the sample is based on the March CPS files alone, while for reforms enacted after 1993 the sample is based on the March and monthly CPS files combined. The outcome variable is weekly employment and the sample consists of single women aged 20-50. See section E in the appendix for additional details.