Estimating Policy Effects With Staggered Implementation and Multiple Periods: Another Look at Family Caps

Kamila Janmohamed[*]
Advisor: Cormac O'Dea

Yale University, Department of Economics
Senior Essay
2021-2022

Abstract

Over the past three decades, 24 states have denied additional benefits to low-income parents who have children while receiving welfare. Using the Current Population Survey, I exploit variation in the timing and stringency of family cap implementation from 1982 to 2010 to compute the effects of family caps on women at risk of welfare receipt in the United States. I apply the new Sun and Abraham (2021) difference-in-differences estimator that accounts for treatment effect heterogeneity over time and find that on average, family caps do not affect the fertility or labour force participation of women at risk of welfare receipt. Investigating treatment heterogeneity by stringency, I find that the strictest family caps that deny capped children any cash or in-kind benefits lower the probability of employment among at-risk women by 2.1 percentage points. Together, these findings indicate that family caps have failed to achieve their stated aims and may even be counterproductive.

1 Introduction

This paper considers the fertility and labour force participation effects of family caps, a set of welfare reform policies that fix families' welfare benefit levels based on the

number of children in their unit when they first began receiving government assistance. With strong public sentiment against the use of benefits as an income substitute and despite limited evidence to support this claim, family caps were implemented in 24 states over 11 years from 1992 to 2003 with varying degrees of stringency.[1] During this period, the majority of capped states moved to bar capped children from receiving any additional benefits while others continued to afford them a fraction of the standard benefit or in-kind benefits such as food stamps.

Although nine states have since repealed their family caps, the limited empirical evidence on the effects of these policies is conflicting and ambiguous. Two states evaluated the effects of family caps with randomised experiments on samples welfare recipients but several concerns have been raised about the internal validity of their findings.[2] In particular, non-interference violations mean that individuals were aware of their peers' treatment statuses and may have modified their behaviour as a result. Additionally, participants were often confused about their assigned conditions, leading to lower adherence rates and biased estimates (Smith, 2006). On the other hand, quasi-experimental studies have either estimated the effects of family caps on a subset of capped states with few post-implementation observations or indirectly on state-level aggregates of fertility outcomes.[3] The latter approach examines policy effects on the broader population and not on the individuals affected by these policies. The resulting causal estimates are therefore subject to downward bias as well.

Due to limited longitudinal data sources that identify welfare recipients, most research with micro-level data constructs proxies for welfare receipt using race (Jagannathan et al., 2004) or other characteristics associated with welfare receipt (Kaushal and Kaestner, 2001). These samples likely include untreated non-welfare-recipients, resulting in imprecise estimates. The majority of these papers account for the staggered implementation of family caps by estimating their effects with with dynamic two-way fixed effects regressions. However, recent work on difference-in-differences methods with multiple periods indicates that this approach does not control for time-variant treatment effect heterogeneity.[4] At best, these contaminated samples produce lower bound estimates of the effects of family caps. At worst, these results are biased in the opposite direction of the true effect (Goodman-Bacon, 2021).

---

[1] See Smith (2006) for an account.

[2] See Joyce et al. (2004) for a detailed critique.

[3] See Camasso and Jagannathan (2016) for a comprehensive review.

[4] See Goodman-Bacon (2021) and Sun and Abraham (2021).

In this paper, I examine the effect of family caps on the probabilities of birth and employment in a sample of women at risk of welfare receipt from 1982 to 2010. I exploit the natural variation in the timing of family cap implementation across all 24 ever-capped states and use the remaining 27 states as controls. In doing so, I make two major contributions to the literature on the effects of family caps on women at risk of welfare receipt. First, I account for the bias caused by treatment effect heterogeneity over time by estimating family cap effects with the interaction-weighted difference-in-differences estimator proposed by Sun and Abraham (2021). Second, I supplement previous work on the average effect of family caps described above by separately examining the effects of strict family caps that deny capped children incremental benefits of any kind.

Following Kaushal and Kaestner (2001) and Kearney (2004), I define a proxy for women at risk of welfare receipt using marital status and educational attainment from the 1982-2010 March CPS. While this identification strategy is limited in its ability to directly identify welfare recipients, these proxy characteristics are useful for drawing a longitudinal sample of women who were targeted by family caps as a whole. I provide supportive evidence for the parallel trends assumption by demonstrating the exogeneity of treatment status and treatment timing among capped states. Additionally, I show that there is no evidence of treatment anticipation for either outcome by conducting placebo tests on pre-treatment observations. Lastly, I test the robustness of my findings against a more precise but far smaller proxy group for welfare recipients based on household income.

On average, I find no evidence that family caps in general affect the fertility and labour force participation of women at risk of welfare receipt. However, strict family caps that bar capped children from receiving any incremental benefits reduce the probability of employment among women at risk of welfare receipt by 2.1 percentage points. This result is robust to the inclusion of individual controls, macroeconomic state-level controls, measures of state welfare generosity, state fixed effects, and year fixed effects. These findings are difficult to rationalise given the intended effects of family caps and require further investigation. One possible mechanism is that strict family caps reduce the flexibility with which welfare recipients can participate in the labour force by lowering the affordability of its complements, such as childcare.

The remainder of this paper is structured as follows. Section 2 presents the history of family caps and the hypothesised causal pathways behind their effective-

ness. Section 3 discusses previous work on the effects of family caps. Section 4 discusses the data sources and main variables used in my analysis. Section 5 specifies the difference-in-differences model and describes the identification strategy for the interaction-weighted causal estimator. Section 6 presents the results. Section 7 discusses the policy implications of my findings and concludes.

2   Policy Background

Under the Aid to Families with Dependent Children (AFDC) program, low-income families with children were afforded cash payments that increased with family size from 1935 to 1997 (Office of the Assistant Secretary for Planning and Evaluation, 2022). By the early 1990s, proponents of welfare reform had grown increasingly critical of this model because they believed it contributed to welfare dependence. In particular, they claimed that incremental benefits disincentivised marriage and encouraged multiparity as a source of income at the taxpayer's expense (Smith, 2006). As a solution, 20 states applied for waivers to forgo AFDC's incremental benefit model and instead implement child exclusion policies, or family caps. After AFDC was replaced by Temporary Assistance for Needy Families (TANF) in 1996, four additional states implemented family caps at their discretion with no federal oversight. These policies allowed states to freeze families' welfare grants at a level based on the number of children in the family when it began receiving benefits without requiring them to evaluate the cap's efficacy (Smith, 2006).

Figure 1 shows the distribution and timing of the 24 states that implemented family caps.[5] Of these, 17 implemented a full family cap under which children conceived after a family unit begins receiving welfare do not qualify for incremental benefits, three replaced incremental cash benefits for capped children with a voucher of equivalent value[6], two implemented implicit caps under which family units receive the same grant regardless of family size[7], and two implemented partial caps under which capped children receive at most half the standard increment[8].

---

[5]See Table A.1 for the month and year of family cap implementation for these states.

[6]Maryland, Oklahoma, and South Carolina afford capped children vouchers worth the incremental benefit.

[7]Idaho and Wisconsin have implicit family caps.

[8]Capped children in Connecticut receive $50 while capped children in Florida receive half the standard increment.

Figure 1: Family cap implementation by state and year

Notes: Figure shows the timing of family cap implementation by state from 1992 to 2003. Unshaded areas correspond to states that never implemented a family cap. Data from The Urban Institute (2021b).

There is limited evidence to support the claims that motivated family cap implementation. Demographic trends indicate that families on welfare were growing smaller and increasingly comparable in size to the general population (Smith, 2006). Moreover, any positive relationship between incremental benefits and fertility would have been driven by planned births among welfare recipients. This is implausible as the majority of births among welfare recipients in the late 20th Century were unplanned (Fairlie and London, 1997). Furthermore, incremental benefits accrued through welfare payments were significantly smaller than the total costs of raising a child given that most AFDC recipients received benefits for less than two years (Fairlie and London, 1997). Several empirical studies have also tested the assertion that welfare generosity incentivises childbirth by exploiting variation in standard and incremental benefit levels across states over time.[9] These find no evidence of a relationship between welfare generosity and the likelihood of a subsequent pregnancy among female welfare recipients. Combined, these facts suggest that the fertility

---

[9]See Acs (1996), Argys et al. (2000), Fairlie and London (1997), and Grogger and Bronars (1997).

decisions of women on welfare were likely insensitive to incremental benefits and subsequently family caps.

Assuming incremental benefits incentivised additional childbirth, disincentivised marriage, and were used as a source of income, family caps that achieved their stated aim would have two effects. Firstly, nonmarital birth rates among mothers receiving welfare would decrease. Eliminating incremental benefits would increase the marginal cost of raising another child, reducing the incentive for further childbearing (Horvath-Rose et al., 2008; Sabia, 2008). This would trigger increased contraception use or an increased likelihood of abortion to deter childbirth during a period of welfare receipt. Among non-married women, the elimination of additional support could increase their propensity for marriage as a spouse's income might substitute for lost incremental benefits (Sabia, 2008). This would increase the number of nonmarital conceptions that end in marital births, reducing the nonmarital birthrate among women on welfare. Secondly, labour force participation would increase. If leisure is a normal good consumed upon the receipt of incremental benefits, the latter is a substitute for wages. Therefore, their elimination would force welfare participants into employment to recoup the benefits they would have received upon the birth of an additional child (Romero and Agénor, 2009). This constitutes the second goal of the family cap, which is to promote personal responsibility among welfare recipients (Smith, 2006; Romero and Agénor, 2009).

Since 2004 nine states have repealed their family caps, leaving 15 capped states.[10] Assuming family caps operate as intended, it stands to reason that their removal might result in a reversal of the behaviour change they caused. Incremental benefits might reduce the incentive for childbirth aversion through an income effect. Since additional births would result in higher welfare grants, female welfare recipients might reduce their use of contraception or, conditional on pregnancy, choose to deliver a child over termination (Levine, 2002). At the same time, a guaranteed increase in family grant levels might induce individuals to substitute incremental benefits for additional sources of income, such as wages or spousal income, and the behaviour required to attain them (labour force participation and marriage respectively). This study does not evaluate the effects of repeal on individuals in previously capped states due to

---

[10]California (2016), Illinois (2004), Maryland (2004), Massachusetts (2019), Minnesota (2015), Nebraska (2007), New Jersey (2020), Oklahoma (2009), and Wyoming (2008) have repealed their family caps.

data limitations. However, the lack of evidence on family cap effects indicates that the same may be true for their repeal: the removal of family caps cannot reverse trends associated with their implementation if they had no effect to begin with.

3 Literature Review

Few states have individually evaluated the effects of family caps on the behaviour of welfare recipients. Most empirical research has measured the effect of family caps only on fertility outcomes with mixed results. In addition to short post-implementation periods, quasi-experimental research has largely estimated the effects of family caps indirectly using proximal determinants of welfare receipt to define target groups or comparing aggregated statistics between states with and without family caps. As fertility outcomes do not manifest in the short term, these data and empirical limitations may have led to imprecise, lower-bound estimates of family cap effects on a contaminated sample. This section discusses the methods, findings, and limitations of the principal studies on family caps.

3.1 Experimental studies

Two states conducted independent experimental evaluations of the family cap by randomly assigning new female welfare recipients to waiver or control conditions. Under the waiver condition, recipients were informed upon intake that additional children born to them while receiving government assistance would not qualify for incremental benefits. Control recipients were informed that all their children qualified for incremental benefits regardless of welfare receipt status at the time of birth. In Arkansas, Turturro et al. (1997) found no evidence of an effect on fertility from 1994 to 1997. However, the statistical insignificance of these findings could be attributed to the short study period and low statistical power arising from a small sample (N=366). In contrast, a five-year experiment in New Jersey found that family caps lowered birth rates and increased abortion rates, specifically among black welfare recipients (Jagannathan and Camasso, 2003). Later studies corroborate these findings and reveal a positive effect on contraception use, but these are limited to short-term welfare recipients (Camasso, 2004; Jagannathan, 2003).

Though the New Jersey experiment demonstrates effects in the expected direction,

it has several limitations that undermine its credibility. First, over a quarter of caseworkers reportedly assigned individuals to treatment or control conditions at their discretion, violating the randomisation required for an experiment of this design (Joyce et al., 2004). Second, Jagannathan and Camasso (2003) do not control for other waiver policies active during the period and therefore cannot separate the effects of family caps on fertility decisions from those of concurrent reform policies such as time limits. Lastly, neither state experiment enforces non-interference or adherence - welfare recipients were not blinded to the treatment status of other participants and were reportedly confused about their assigned condition (Joyce et al., 2004). If individuals mistakenly believed they were not subject to the family cap, the average treatment effect estimate might be biased by untreated individuals in the treatment group. Similarly, if individuals in the control group believed they were subject to the family cap and adjusted their fertility decisions accordingly, the estimate might be biased by treated control units. These limitations possibly contributed to spillover effects with ambiguous effects on the final estimate.

3.2 Quasi-experimental studies

Several studies have estimated the effects of family caps on women using quasi-experimental methods such as difference-in-differences and triple differences with two-way fixed effects (TWFE) regressions. These papers exploit variation in treatment timing for all or a subset of capped states to compare outcomes between treated and untreated states in their sample. The datasets used include the Panel Study of Income Dynamics (PSID), the Current Population Survey (CPS), the Study of Income and Program Participants (SIPP), the National Survey of Family Growth (NSFG), and the National Longitudinal Survey of Youth (NLSY). Except for the SIPP, no other data sources identify welfare recipients. Thus, most studies indirectly estimate the effects of family caps through proxies for welfare participation such as a history of welfare receipt, education, marital status, and parity.

Five studies directly assess the effect of family caps on female welfare recipients using non-welfare-recipients as a within-state comparison group.[11] These report conflicting estimates of the effect of family caps on fertility and are either statistically underpowered by small sample sizes or biased by their chosen comparison groups.

---

[11]See Hofferth et al. (2002), Mach (2000), Ryan et al. (2006), Wallace (2009), and Wiseman (2000).

Non-welfare-recipients may be an inappropriate within-state comparison group given the concerns that justified family cap implementation in the first place. Single women are over-represented among welfare recipients and nonmarital birth rates were rising during these study periods (Gray et al., 2006). This divergence in pre-implementation trends between welfare recipients and non-recipients violates the parallel trends assumption, contributing to bias.

Ryan et al. (2006) and Wallace (2009) eliminate the bias that could arise from compositional differences between welfare recipients and non-welfare-recipients by restricting their sample to unmarried women. Using the PSID and the SIPP respectively, they find no evidence that family caps affect the subsequent childbearing decisions of unmarried women on welfare. Both studies are likely statistically underpowered because their restrictions result in a small sample of unmarried women. Wallace (2009) in particular may have a disproportionately small "treatment" group as the SIPP is known to underreport welfare receipt. If family caps affect fertility decisions in this sample, welfare recipients erroneously included in the control group likely bias their estimates towards zero.

Using the CPS, Mach (2000) finds that past welfare receipt has a significant negative impact on the likelihood of a subsequent birth for individuals in capped states. Although this is in the expected direction, welfare receipt is likely under-reported in their sample as well: over a 10 year study period, only 25 women in 23 states with family caps had an additional birth (Joyce et al., 2004). Furthermore, Mach (2000) does not control for other policies implemented during the study period that affect the fertility and marriage decisions of welfare recipients. Examples of these changes include conditioning benefits on family planning education and relaxing conditions that previously penalised married couples, such as expanding the earned income disregard to include stepparents (Bitler et al., 2004; Gennetian and Knox, 2004).

Hofferth et al. (2002) is the only study to assess the effect of family caps on the welfare exits of female heads of household. Though they find no evidence of a relationship, their estimates may be biased by their noisily measured outcome. Specifically, they do not distinguish between work exits and other types of exits from public assistance, including marriage (Hofferth et al., 2002). Since family caps were instituted to reduce nonmarital births, it stands to reason that they may disproportionately drive welfare exits by way of marriage. Thus, failing to estimate the effects of family caps on marriage exits alone likely leads to imprecision and downward bias.

More recent studies use state-level data to compare vital statistics such as birth and abortion rates between capped and uncapped states.[12] These further disentangle the effects of family caps from changing welfare policies and public sentiment towards illegitimacy and abortions by including policy, demographic, and time-varying political controls. All of these studies report a significant association between family caps and fertility outcomes. Camasso and Jagannathan (2009) in particular report that family caps are associated with a decrease in nonmarital birth rates and an increase in abortion rates. In a later study with an additional decade of post-reform data, Camasso and Jagannathan (2016) find that the family cap's negative association with nonmarital birth rates is limited to states where Medicaid is used to pay for abortions and black women form a large proportion of the population. Neither study uses measures of welfare participation by race. Instead, they use exogenous measures of states' racial compositions. Consequently, their estimates more accurately describe the fertility behaviour of minority women rather than that of welfare recipients (Camasso and Jagannathan, 2009, 2016). This limits the generalisability of their findings

Horvath-Rose et al. (2008) and Sabia (2008) paint a more comprehensive image of the policy's effects on nonmarital birth rates by considering outcomes at each node on the fertility decision tree: conception, termination, and marriage. They find evidence of a significant association between family caps and a decrease in nonmarital birth rates driven by a reduction in nonmarital pregnancy rates, particularly among black women (Horvath-Rose et al., 2008; Sabia, 2008). Horvath-Rose et al. (2008) also finds evidence of a positive relationship between family caps and nonmarital birth ratios that cannot be entirely explained by family caps inducing unmarried pregnant women to marry, raising questions of policy endogeneity. Interpreting changes in the nonmarital birth ratio is especially challenging because it is unclear whether this is driven by a change in marital or nonmarital births (Horvath-Rose et al., 2008). There is no indication that family caps affect marital fertility rates except by inducing individuals to resolve a potential nonmarital birth through marriage. However, this does not preclude an effect on marital fertility. Horvath-Rose et al. (2008) address this by showing their findings are robust to estimation on nonmarital birth rates.

Sabia (2008) further shows that changes in nonmarital birth rates may be driven by a decline in conception rather than abortion or induced marriages. Like other

---

[12]See Horvath-Rose et al. (2008), Sabia (2008), Camasso and Jagannathan (2009), and Camasso and Jagannathan (2016).

comparisons of state-level aggregates, they assume all individuals in a state are subject to welfare reform policies, including family caps. Since this identification strategy fails to distinguish birthrates among women on welfare from birthrates among non-welfare-recipients, it also provides a lower bound estimate of family cap effects than an estimator that averages differences in outcomes over the target group. Sabia (2008) and Horvath-Rose et al. (2008) justify this approach by claiming it has the advantage of detecting spillover effects associated with the family cap. This might occur if family caps stigmatise nonmarital childbearing among all unmarried women, regardless of parity and welfare status. In this case, estimates would include both the direct effect of the policy on welfare recipients and spillover effects.

Between-state comparisons of family cap effects carry a significant risk of bias as benefit levels, welfare policies, and fertility decisions likely covary across states for unobserved reasons (Kearney, 2004). Such findings may be subject to policy endogeneity, especially if there exist unobserved, time-varying, state-specific factors that are correlated with family cap implementation (Sabia, 2008). This could happen if family cap implementation is driven by anti-illegitimacy sentiments or higher baseline nonmarital fertility rates. Kearney (2004) posit that this endogeneity can be addressed by distinguishing between states that implemented family caps under TANF and those that specifically requested waivers to do so. The latter may be more responsive to family caps if their desire to reduce births among welfare recipients independently of national welfare reform reflects heavier welfare caseloads or higher birth rates among welfare recipients. This approach is infeasible because the majority of states with family caps requested waivers to implement them.[13] Estimates of the differential impact of family caps between these two groups may be underpowered.

An alternative approach is to compare the difference in outcomes between within-state groups at risk of public assistance using proxies for welfare receipt such as age, marital status, education level, and parity. This provides a meaningful counterfactual for policy effects on the target group and accounts for the time-varying contexts of individual states. For example, Levine (2002), Dyer and Fairlie (2004), and Joyce et al. (2004) identify less-educated women with no children as a group at risk of family caps but not affected by them. These individuals serve as a within-state comparison group for less-educated women with at least one child who are both at risk of and

---

[13]20 states requested waivers to implement family caps while the remaining four implemented theirs under TANF.

affected by the family cap. Kaushal and Kaestner (2001) expand on this reasoning by using married, at-risk women with children as an additional within-state comparison group. None of these studies find evidence of a relationship between family cap policies and fertility. However, Dyer and Fairlie (2004) only use five states that implemented family caps over the period 1989 to 1999 while Joyce et al. (2004) use data for 24 states from 1992 to 2000. Both panels end less than five years after the majority of states implemented their caps. These studies may therefore be statistically underpowered by limited observations and post-implementation data.

This paper builds on these approaches by expanding the study period to include 11 years of pre-implementation data and six years of post-implementation data. Quasi-experimental work has accounted for the dynamic implementation of family caps by estimating their causal effect with TWFE regressions. I account for the bias in this method by estimating the effect of family caps with a difference-in-differences estimator that corrects for time-variant treatment effects. Following previous work, I estimate the average effect of family caps on a proxy for welfare recipients. However, I extend my analysis to include the average effects of strict family caps, under which capped children do not receive any cash or in-kind incremental benefits.

4 Data Sources and Main Variables

In my empirical analysis, I combine annual data for all 50 states and DC on (1) individual-level characteristics from the nationally-representative CPS, including gender, employment status, number of hours worked, household income, whether the individual gave birth to a child in the past year, and various socioeconomic characteristics; (2) the dates of family cap implementation and repeal from the Urban Institute, including information on the stringency of the cap such as whether it is implicit, only applies to cash benefits, or includes restrictions on in-kind benefits such as food stamps and Medicaid; (3) measures of welfare generosity from the Urban Institute, namely the maximum monthly welfare benefit for a family of three and the start and end dates of welfare time limit policies; and (4) macroeconomic measures of female labour force participation from the CPS and the Bureau of Labour Statistics, including the median female wage and female unemployment rate. Summary statistics for all variables used in the analysis are presented in Table A.2.

4.1 Data on individual-level characteristics

Data on individual-level characteristics come from the nationally representative March Current Population Survey (CPS) for the years 1982 to 2011 (Flood et al., 2021). This panel of repeated cross-sections provides all outcome variables and individual-level controls correlated with them: age, race, ethnicity, and number of children.

I consider the effects of family caps on fertility and labour force participation through two main outcomes: birth and employment. The March CPS reports the age of the respondent's youngest child at the time of the survey. If this is 12 months or younger, I define the individual as having had a child in the past year. This is equivalent to an indicator for whether a respondent in year t gave birth in year $t-1$. I construct a new dataset with this lagged indicator and the individual's time-invariant characteristics from the 1983-2011 March CPS to estimate the effect of family caps on the probability of birth in the years 1982-2010. Since the CPS reports respondents' employment statuses at the time of the survey, I use the 1982-2010 March CPS to estimate the effect of family caps on employment conditional on giving birth in the previous year. I measure employment with an indicator equal to 1 if the respondent reported having a full or part-time job at the time of the survey and 0 otherwise. As a robustness check, I estimate the effect of family caps on the total number of hours worked at all jobs in the past week, equal to 0 if the individual is unemployed, conditional on giving birth in the previous year.

Fertility and employment data are available for all years between 1982 and 2010, so regressions on these outcomes cover this period and include all 24 capped states. However, data on the number of hours worked at all jobs is only available beginning in 1989, three years before the first treated state (New Jersey) implemented its family cap. To maintain a sample that is balanced in event-time, regressions on this outcome do not include data on New Jersey and span the period 1989 to 2010.[14]

4.1.1 Defining the target group

As the primary outcome of this analysis is the probability of birth, I restrict my sample to adult women between the ages of 18 and 45 years. Individuals below this threshold would be considered minors by welfare eligibility rules and subject to other provisions upon the birth of an otherwise capped child. For example, some states

---

[14]Equivalent to $-7 \leq \tau \leq 6$ in event-time, defined in Section 4.2.

exempt the first child born to a minor if they are part of a family unit that is already receiving welfare (The Urban Institute, 2021b). On the other hand, individuals above the age of 45 are unlikely to conceive and carry a child to term. Including individuals with a significantly lower probability of birth who are unlikely to respond to the policy given biological constraints would bias estimates towards 0.

Although all women in a treated state are exposed to the family cap after implementation, the intended treatment group solely consists of individuals at risk or with a history of welfare receipt. Child exclusion policies should have no impact on the short-term fertility and labour force participation of individuals outside this target group. The treatment effect of family caps on these untreated individuals would be 0 and including them in the sample would bias the final average treatment effect estimate downwards.[15] However, I am unable to directly estimate the effect of family caps on the target group because the March CPS does not include information on welfare receipt. Instead, I construct a proxy for welfare receipt by marital status and education. To ensure my estimates are not limited to my choice of proxy, I define a secondary sample using household size, household income, and need standards for benefit eligibility. Neither definition is precise as they are defined by proximal determinants of welfare receipt. They may include individuals who never receive welfare and are subsequently untreated. Assuming non-welfare-recipient birth rates trend differently than those for welfare recipients, and that welfare recipients respond to family caps, the final estimate may still be biased.

1. Proxy by age and education

Kearney (2004) estimates that up to 80% of all welfare recipients are unmarried and have at most a high school education. Following their identification strategy, I construct a proxy for welfare receipt by restricting my sample to unmarried women of childbearing age with at most 12 years of education, equivalent to the completion of high school. The resulting group includes individuals at risk of welfare receipt and whose childbearing behaviour family caps were designed to change. Figure 2 corroborates this statement, showing that birth rates in the proxy group (unmarried, less-educated women) trend upwards in both treated and untreated states towards the

---

[15]I test this assumption by comparing treatment effect estimates on the different samples of women aged 18-45 years to the primary proxy group in Table A.6 and show that this biases the estimate towards 0.

end of the 20th Century, nearly doubling between 1986 and 1990. Conversely, other groups, particularly more-educated women, demonstrate a downward trend. This is consistent with rhetoric surrounding family cap implementation that defined the caps as instrumental to reversing rising nonmarital birth rates.
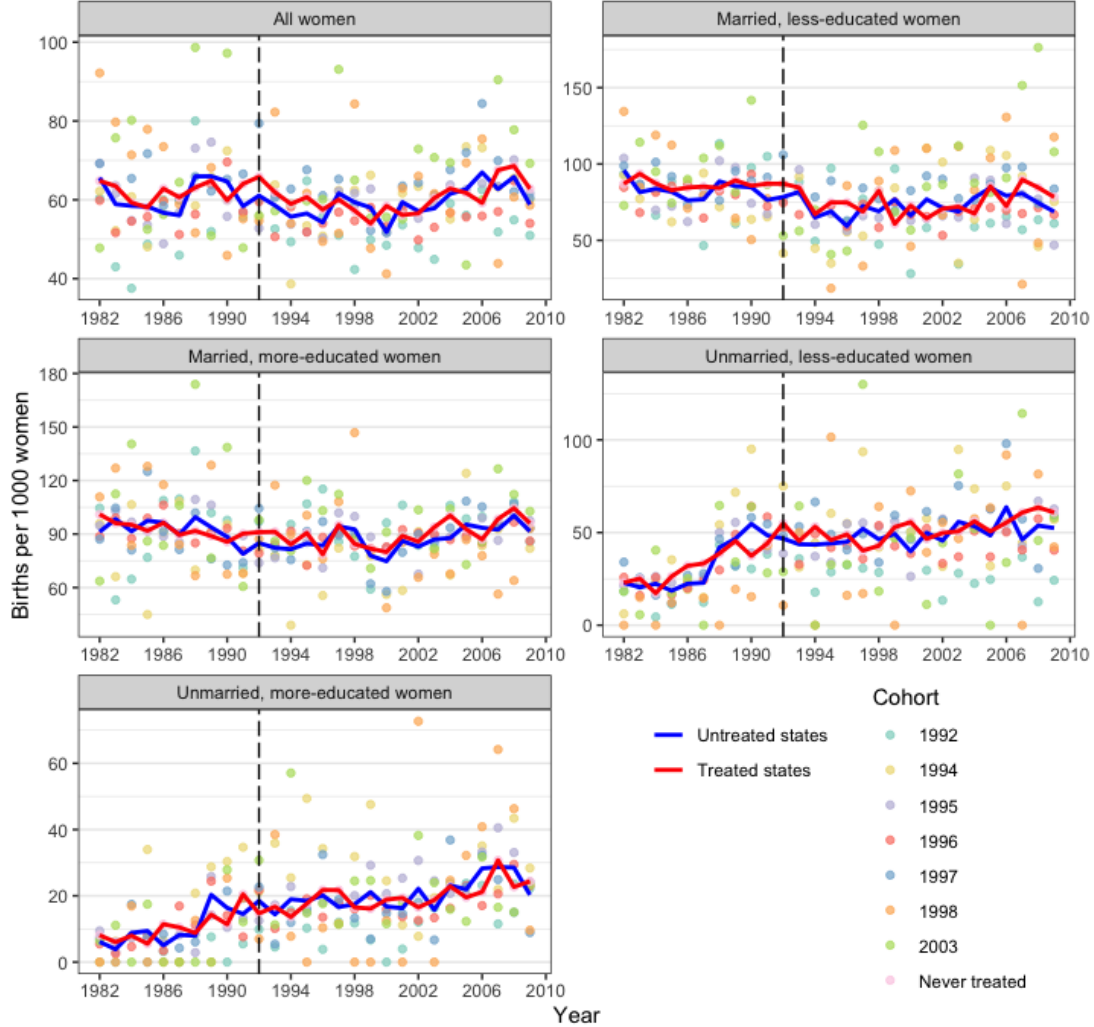

Figure 2: Average annual birth rates by demographic group

Notes: Lines denote the average birth rate for states by family cap implementation status. Points denote average annual birth rates by treatment cohort, defined as states that implemented caps in the same year or states that never implemented caps (never treated). Vertical lines at 1992 indicates the year the first family cap was implemented. Birth rates are defined as the number of births per 1000 women and computed using the March CPS.

Although the y-axes for the plotted subgroups are inconsistent, the driving concern behind family cap implementation was the difference in trends between groups, not levels.

This sample might still include untreated individuals, but Sabia (2008) argues that the inclusion of untreated, unmarried women may capture both the average treatment effect on the treated (ATT) and spillover effects associated with the family cap. If family caps stigmatise nonmarital childbearing among all unmarried women, there may be a treatment effect on unmarried non-welfare-recipients in addition to welfare recipients. This negative spillover effect could reduce nonmarital birth rates for all unmarried women (Sabia, 2008). The final sample includes 147,166 women for regressions on fertility and 149,647 for regressions on employment.

2. Proxy by income

I test the robustness of my estimates on the sample of socioeconomic proxies with an alternative sample that proxies for welfare receipt by household income. Although federal law requires that individuals be "needy" to qualify for TANF benefits, it neither defines this term nor requires states to use federal poverty guidelines to determine eligibility (Center on Budget and Policy Priorities, 2021). Several states place eligibility thresholds below the federal poverty level so this cannot be used to construct the income proxy. Instead, I determine welfare eligibility using the monthly need standard - an income threshold for welfare eligibility based on household size that varies across states and over time.

Data on the need standards for AFDC and TANF benefits come from Version 3 of the Urban Institute's Transfer Income Model (TRIM3) - a simulation model with information on the US government's transfer programs, including AFDC and TANF. The Program Rules function within the simulation tool provides monthly need standards in dollar amounts for one-person to 12-people families in each state from 1982 to 2010. I multiply this by 12 to determine the annual need standard for each family size in each state from 1982 to 2010. For families with more than 12 individuals, I use the need standard for a family of 12. I determine whether individuals meet this requirement using their total combined family income in the past 12 months, reported in the CPS. As family income is reported in bins, I use the upper bound on the income bracket to determine whether individuals meet the maximum need standard.

Although this definition identifies individuals at risk of welfare receipt more closely, I cannot use it as my primary proxy because the resulting samples only include 39,859 women for regressions on fertility and 42,891 women for regressions on employment.

While this might result in low statistical power, it may help rule out effect sizes above some threshold defined by the final estimate and as a falsification test for my estimates with socioeconomic proxies.[16]

4.2 Data on family caps and welfare generosity

Data on family caps, the existence of welfare time limits, and the maximum monthly benefit for a family of three come from two sources: the Urban Institute's Welfare Rules Database and TRIM3. TRIM3 covers the period 1975 to 2001 while the Welfare Rules Database covers the period 1996 to 2019. I use TRIM3 (1982-1995) and the Welfare Rules Database (1996-2010) to identify (1) states that implemented family caps; (2) the year of family cap implementation; (3) cap-specific rules on incremental benefits; (4) the year of welfare time limit implementation (if implemented); and (5) the maximum monthly benefit for a family of three.

I define a measure of time relative to treatment, or event-time, (τ) normalised to 0 in the first full year a treated state has a family cap in place, resulting in a panel of 11 years of pre-treatment observations and six years of post-treatment observations ($-11 \leq \tau \leq 6$). To control for confounding welfare policies, I also use data on the implementation of welfare time limits by constructing an indicator for time limits equal to one if a state has an active welfare time limit in a given year.

4.3 Data on labour market conditions

Data on time-varying state-level macroeconomic factors correlated with labour force participation and the propensity for childbirth include the annual female unemployment rate and the median weekly female wage. Unemployment rates were obtained from the Bureau of Labour Statistics Local Area Unemployment Statistics Office. Weekly wages were computed using the Integrated Public Use Microdata Series (IPUMS) Online Data Analysis System with the CPS Annual Social and Economic supplements from 1982 to 2010. Weekly wages were calculated by dividing the total wage and salary income of women over the age of 16 by the number of weeks they worked in the past year.[17].

---

[16]Note that the correlation coefficient between the membership in either proxy is small but positive (0.17). This is mainly due to missing household income values.

[17]Median weekly wages for each state and year are calculated with person-weights

4.4   Supplementary data sources

To look for evidence of exogeneity in treatment timing and treatment status, I collected data on state-level demographic, economic, and political characteristics in the last year before any treated states implemented family caps. The 1991 State and Metropolitan Area Data Book provides populations, racial compositions, birth rates, marriage rates, abortion rates, and median household incomes from 1979 to 1990. I define dummies for whether the Senate and House of Representatives were majority Democrat, majority Republican, or balanced in 1991 by comparing variables for their composition by party. Similarly, I construct a dummy for the party of the candidate that won the popular vote in each state in the 1988 Presidential election. This is equal to 1 if a Republican candidate won the popular vote and 0 otherwise.

To supplement state-level measures of welfare generosity, I calculate the number of AFDC recipients as a proportion of the total state population in 1990 using populations from the State and Metropolitan Area Data Book and the number of AFDC recipients from the Office of Family Assistance's historical AFDC caseload files.

5   Empirical strategy

The goal of my analysis is to estimate the causal effect of family caps on fertility and labour force participation. In this section, I describe the strategies I use to address the identification challenges inherent in causal inference with dynamic treatment, the risk of treatment endogeneity, and my estimating equation. I also outline the methods I use to identify anticipatory behaviour that might threaten my identifying assumptions and correct for heterogeneity in treatment effects over time.

5.1   A staggered diff-in-diff design

The ideal experiment for identifying the effects of family caps on fertility and employment would be to implement family caps in all states simultaneously but randomly assign welfare recipients to the policy so each state has a within-state control group comprised of uncapped welfare recipients. Such a randomised experiment would estimate the average difference in outcomes between treated and untreated individuals across all states, where untreated welfare recipients provide a strong counterfactual for the untreated potential outcomes of treated welfare recipients. A national ex-

periment of this magnitude would be infeasible to implement without large spillover effects and border on unethical.

Instead, I approximate this ideal experiment with a quasi-experimental research design that exploits the natural variation in the timing of family cap implementation in 24 states over an eleven-year period with repeated cross-sections of individuals who proxy for welfare recipients. I construct counterfactuals for the untreated potential outcomes of these individuals with a similarly defined proxy group in the remaining 27 states that never implement a family cap. The resulting sample consists of a treatment group, composed of unmarried women aged 18-45 years with at most a high school education in states that ever implement a family cap, and a never-treated group, composed of unmarried women aged 18-45 with at most a high school education in states that never implement a family cap. I compute the treatment effect by implementing event studies for these groups and aggregating event-time estimates through the interaction-weighted (IW) difference-in-differences estimator proposed by Sun and Abraham (2021).

The identifying assumption of this difference-in-differences approach is that in the absence of exposure to a family cap, the difference in outcomes between treated individuals in treated and never-treated states will be constant over time. The plausibility of this assumption relies on two degrees of treatment exogeneity. First, the decision to implement a family cap must be random. Second, the timing of family cap implementation among treated states must be random. Together, these assumptions imply that family cap implementation is not endogenous to some state-level characteristic correlated with the outcome of interest. In particular, the parallel trends assumption would be threatened if states with higher non-marital birth rates, larger welfare caseloads, or even less public support for the welfare state, were more likely to implement family caps. Similarly, treatment timing would be endogenous if earlier-treated states implemented family caps sooner than their later-treated counterparts because of different birth trends or some other observable systematic difference.

As states were required to apply for waivers to implement family caps under AFDC, it is possible that early-treated states were systematically different from untreated and later-treated states. The waiver requirement involves the applicant bearing some bureaucratic cost. The decision to apply for a family cap waiver might therefore reflect a willingness to pay driven by some observable state-level characteristic. Since the majority of treated states (83.3%) implemented family caps by

requesting waivers, any tests for systematic differences between AFDC and TANF-era family caps would be statistically underpowered. Instead, I consider whether treatment timing more broadly reflects systematic differences between treated states. Following Hoynes et al. (2016), I show that there is no significant relationship between the year of family cap implementation and state-level characteristics in 1990, shortly before the first treated state implemented a cap through the OLS specification

$$T_s = X_s + \epsilon_s \tag{1}$$

where T_s is the year of cap implementation in treated state s expressed as an index equal to 1 in 1992 and X_s is a vector of demographic, economic and political state-level characteristics from the 1991 State and Metropolitan Area Data Book and the Urban Institute's TRIM3 Database.[18]

I also test for systematic differences between treated and untreated states by computing differences in means for the same state-level characteristics and conducting two-way t-tests for CPS sample characteristics in 1990.[19] I consistently show that there is no evidence of differential trends in birthrates between treated and untreated states and that treatment allocation and timing are not systematically correlated with a broad set of state-level observables. This lends support for the treatment exogeneity assumption and, by extension, the parallel trends assumption.

5.2 Empirical specification

Difference-in-differences analyses with dynamic treatment have historically estimated treatment effects with the TWFE difference-in-differences (TWFEDD) estimator. An event study of the effects of family caps would use the following dynamic TWFE specification:

$$Y_{ist} = \alpha_{ist} + \alpha_{st} + \theta_s + \theta_t + \sum_{\tau=-11}^{6} \delta_\tau \mathbf{1}\{t - D_s = \tau\} + \epsilon_{ist} \tag{2}$$

where Y_{ist} is the outcome of interest for individual i in state s in year t, α_{ist} is a vector of individual characteristics for a respondent i surveyed in year t, α_{st} is a vector of time-varying state controls, θ_s are state fixed effects, and θ_t are year fixed effects. D_s

---

[18]Results shown in Table A.3.

[19]Results shown in Table 1.

is the first full year state s had a family cap. $\mathbf{1}\{t - D_s = \tau\}$ is an event-time dummy equal to 1 if state s is τ years away from D_s and 0 otherwise. The causal parameter δ_τ corresponds to the average treatment effect upon exposure to treatment for τ years.

The average treatment effect estimate δ^{DD} would, in turn, be estimated by

$$Y_{ist} = \alpha_{ist} + \alpha_{st} + \theta_s + \theta_t + \delta^{DD} D_{st} + \epsilon_{ist} \tag{3}$$

where D_{st} is a dummy equal to 1 if state s has an active family cap in year t and 0 otherwise.

This TWFEDD estimate is a weighted average of all possible 2×2 DD estimators, each of which is computed with a treatment group, whose treatment status changes, and a control group, whose treatment status does not change, for some window in the whole study period (Goodman-Bacon, 2021). In this model, each state, including those that are eventually treated, is used as a control group in at least one 2×2 DD estimator. Although some 2×2 DD estimators compare never-treated units to treated units, others use later-treated units as controls for earlier-treated units. Recent work has shown that this can lead to bias in cases of dynamic treatment and heterogeneous treatment effects over time. The existence of such time-varying treatment effects violates the parallel trends assumptions for these comparisons, subsequently biasing the treatment effect estimate away from the sign of the true effect (Goodman-Bacon, 2021).[20] This is a pressing concern in the case of family caps as any policy effects[21] are unlikely to manifest in the short-term and may increase over time if the caps contribute to gradual changes in norms surrounding the fertility and labour force participation of women at risk of welfare receipt.

In the absence of treatment effect heterogeneity, the TWFEDD estimator is equivalent to the ATT. However, the weights assigned to group-specific ATTs are a function of sample shares and treatment variance (Goodman-Bacon, 2021). Since the latter is highest for units in the middle of the panel, the estimator places less weight on units treated earlier or later in the study period (Goodman-Bacon, 2021). As a result, the TWFEDD estimator directly depends on when treatment is rolled out relative to the

---

[20]Appendix C implements the Goodman-Bacon decomposition for the TWFEDD estimator of family cap effects with state aggregates and compares it to the IW estimator. While neither estimate is statistically significant, the TWFEDD estimate demonstrates a significant downward bias on the IW estimate.

[21]At least on fertility.

study period; that is, the number of periods in the panel before and after treatment. This dependence renders the TWFEDD estimator an unreliable estimate of the true ATT even if there is no reason to suspect treatment effect heterogeneity.

5.2.1 Applying recent advances in the literature on difference-in-differences

Sun and Abraham (2021) propose an extension of the TWFEDD estimator that accounts for heterogeneous treatment effects over time by considering the dynamic treatment effect for each cohort, defined as a group of units with the same treatment time. This interaction-weighted (IW) estimator is a weighted average of the cohort-specific average treatment effect on the treated (CATT) for each post-treatment period (Sun and Abraham, 2021). Deriving the CATT by interacting event-time dummies with cohort membership limits the control group to never-treated units. This eliminates the risk of bias from using already-treated units as controls (Sun and Abraham, 2021). The result is a CATT for each treatment cohort which can then be weighted and aggregated to compute a time-varying average treatment effect.

Previous attempts to measure the causal effects of family caps have used TWFE estimators.[22] This approach may be inappropriate if family caps result in time-variant treatment effects. The main contribution of this paper is that it measures the causal impact of family caps without the bias associated with treatment effect heterogeneity. I estimate the effect of child exclusion policies using three steps outlined in Sun and Abraham (2021) and the linear regression:

$$Y_{ist} = \alpha_{ist} + \alpha_{st} + \alpha_{st-1} + \theta_s + \theta_t + \sum_{e \neq N} \sum_{\tau=-11, \neq -1}^{6} \delta_{e\tau}(\mathbf{1}\{E_s = e\} \cdot D_{st}^{\tau}) + \epsilon_{ist} \quad (4)$$

where Y_{ist} is the outcome of interest for individual i in state s in year t. For binary birth and employment outcomes, this regression is a linear probability model. τ denotes the year relative to family cap implementation. For each state, $\tau = 0$ corresponds to the first full year of an active family cap. Since this model assumes treatment is absorbing, units must remain treated after their exposure to treatment. Therefore, I drop all observations after treated states repealed their caps, resulting in a panel of repeated cross sections over the period 1982 to 2010. This corresponds to $-11 \leq \tau \leq 6$ in event-time. Equation (4) includes a vector of individual characteris-

---

[22]See Horvath-Rose et al. (2008), Joyce et al. (2004), Kaushal and Kaestner (2001), Kearney (2004), and Sabia (2008).

tics (α_{ist}), a vector of time-varying state controls (α_{st}), a vector of lagged time-varying state controls (α_{st-1}), state fixed effects (θ_s), and year fixed effects (θ_t). Since I derive a lagged birth dummy from the CPS, regressions on labour force participation outcomes include α_{ist-1}, a dummy for whether an individual surveyed in year t gave birth in the previous year.

Individual observations in treated states are grouped into one of seven cohorts, denoted by e. This is the year of family cap implementation in their state. $\delta_{e\tau}$ corresponds to the conditional average treatment for cohort e, τ periods from treatment ($CATT_{e\tau}$). It is computed by interacting $\mathbf{1}\{E_s = e\}$ with event-time dummies equivalent to $\mathbf{1}\{t - D_s = \tau\}$ from Equation (2). $\mathbf{1}\{E_s = e\}$ is a set of dummies for all cohorts except the never-treated group (N) equal to 1 if state s first has a full year with a family cap in year e. The final IW estimator is the weighted average of all $CATT_{e\tau}$ estimates, where weights are computed as the sample of shares for each cohort in event-time τ (Sun and Abraham, 2021).

5.2.2 Control variables

Following Kaushal and Kaestner (2001), I control for individual characteristics that are proximal determinants of childbearing decisions and labour force participation. These include the mother's age, race, and total number of children. Individual-level controls and outcome variables are obtained from the nationally representative March Current Population Survey (CPS) for the years 1982 to 2011 (Flood et al., 2021). Although current employment status might be correlated with having a child under the age of 1, including it as a control would result in reverse causality. Lagged employment, which could influence the probability of birth through the opportunity cost or income effect arguments would arguably be a better control but is unavailable as the panel is composed of repeated cross-sections.[23] Instead, its effect on the probability of birth is captured by the residual ϵ_{ist}.

A body of literature suggests that birth rates in the United States are pro-cyclical. On average, unemployment risks during recessions disincentivise childbirth in the short term, affecting the timing of childbearing but not overall fertility rates (Coskun

---

[23]The opportunity cost argument posits that unemployment reduces the opportunity cost of bearing a child in that there are no foregone wages involved, incentivising childbirth. The income effect argues that individuals might delay the decision to have a child until they have access to the resources to raise them. The ultimate effect of unemployment is subsequently indeterminate. See Coskun and Dalgic (2020) and Sobotka et al. (2011) for a review.

and Dalgic, 2020; Sobotka et al., 2011). To capture the effect of labour market conditions on the probabilities of birth and employment, I include time-varying state-level controls for median weekly female wages from the Annual Social and Economic Supplement (ASEC) of the CPS and female unemployment rate from the Bureau of Labour Statistic's Local Area Unemployment Statistics Office. I also control for confounding welfare policies with potential effects on fertility and labour force participation through two state-level variables. These are the maximum monthly benefit for a family of three and a dummy variable for whether a state has an active welfare time limit policy in a given year. Both are obtained from the Urban Institute's TRIM3 and Welfare Rules databases (The Urban Institute, 2021a,b).

As conception occurs up to nine months before birth, the outcome of interest is likely correlated with state welfare generosity and macroeconomic characteristics in the year before birth. Therefore, I include lags for female unemployment rates, median female wages and the maximum monthly benefit for a family of three at $t-1$.

5.2.3 Identifying assumptions of the IW estimator

In addition to the parallel trends assumption discussed in the previous section, the IW estimator relies on treatment effect homogeneity between cohorts and the absence of anticipatory behaviour before treatment. This section describes how violations of these assumptions might threaten the causal interpretation of the final estimate and the methods used to test for them.

The treatment homogeneity assumption stipulates that for each relative time period τ, the $CATT_{e\tau}$ is independent of cohort e (Sun and Abraham, 2021). This does not preclude the possibility of treatment effect heterogeneity in τ, where treatment effects change over the course of treatment. Instead, it assumes that all cohorts share the same treatment profile. Table A.3 examines pre-treatment state-level characteristics and does not indicate systematic pre-treatment differences between cohorts. Therefore, it is reasonable to assume this will hold for post-treatment trends.

Since family caps were announced before their implementation, it is not possible to rule out treatment anticipation. However, it is reasonable to assume individuals did not change their behaviour in anticipation of treatment. Anticipatory behaviour would require an expectation of future welfare receipt. In particular, non-welfare-recipients would have to expect welfare receipt at least nine months after conception and hinge their decision to have a child on the prospect of future benefits. Simi-

larly, non-recipients would need to conceive a child at least nine months before cap implementation to qualify for incremental benefits. In both cases, anticipatory behaviour change involves having a child that would not be conceived otherwise. This would indicate that incremental benefits are correlated with fertility, a hypothesis that has been discredited empirically by Acs (1996) and Fairlie and London (1997). Collectively, this suggests there is little evidence in favour of anticipatory behaviour.

Nevertheless, I empirically test this assumption with a placebo test for anticipatory behaviour by restricting the sample to pre-treatment observations ($\tau < 0$). Using a similar specification to Equation (4), I define a new event-time variable τ^* normalised to 0 when $\tau = -6$ and estimate

$$Y_{ist} = \alpha_{ist} + \alpha_{st} + \alpha_{st-1} + \theta_s + \theta_t + \sum_{e \neq N} \sum_{\tau^* = -5, \neq -1}^{4} \delta_{e\tau^*}(1\{E_s = e\} \cdot D_{st}^{\tau^*}) + \epsilon_{ist} \quad (5)$$

for anticipatory changes in fertility and labour force participation in the window $(-5 \leq \tau^* \leq 4)$. $D_{st}^{\tau^*}$ is a set of dummies equal to 1 when individual i in state s in year t is $\tau*$ years away from treatment. Equivalently,

$$D_{s,t}^{\tau^*} = \mathbf{1}\{t - (D_s - 6) = \tau + 6 = \tau^*\} \quad (6)$$

A treatment effect of 0 in the pre-treatment period implies that the no-anticipatory assumption holds, further supporting the parallel trends assumption.

6 Results

In the first part of my analysis, I provide suggestive evidence in favour of the parallel trends assumption and show that there is no evidence of anticipatory behaviour. I then show that family caps have no significant effect on the fertility or labour force participation of the target group and that these findings are robust to the inclusion of individual controls, state controls, and fixed effects, but not to the exclusion of states with lenient family caps.

6.1 Descriptive trends and summary statistics

The IW estimates of the effect of family caps on the probabilities of birth and employment are based on the assumption that counterfactual trends for each outcome in treated states are the same as observed trends in untreated states. As this assumption cannot be tested, I determine its plausibility by examining pre-treatment trends in both outcomes and comparing average treatment group characteristics before the first treated state implemented a family cap. If there is no evidence of diverging trends before treatment, this is supportive of the parallel trends assumption after treatment.

The parallel trends assumption might be violated if treated and untreated states differ significantly in their observable characteristics during the pre-treatment period. More importantly, treatment would be endogenous if states elected to have family caps due to differential trends in welfare caseloads or birth rates among welfare recipients. Figure 3 displays trends and differences in means for birth and employment rates by state treatment status.[24] Column A suggests that birth rates in both groups trended upwards before family caps were implemented. The differences in means suggest that treated states had lower average birth rates than untreated states in 1988, but there are no other indications of diverging trends in the pre-treatment period. Column B indicates that employment rates among treated states were systematically higher than those in untreated states until 2004. Such a pattern would not bias a difference-in-differences estimator because both groups ultimately follow the same trend at different levels. If family caps had a significant effect on either outcome, one would expect the differences in means to be negative for birth rates, positive for employment rates, and statistically significant, especially after 2003 (by which point all treated states had implemented caps). As there are no consistent, significant changes in differences in means in the post-treatment period, Figure 3 suggests that family caps have no discernible effects on fertility or labour force participation in this sample.

Table 1 further shows that treated and untreated state characteristics are balanced at baseline.[25] On average, treated states have lower median female wages and larger proportions of black residents but also lower poverty rates, fewer AFDC recipients and lower welfare generosity. These differences do not systematically indicate selection

---

[24]See Figure B.2 for weighted group means for all outcomes, i.e. the probability of birth, probability of employment and the number of hours worked.

[25]I use 1990 as a pre-treatment baseline year as it falls shortly before any treated states implemented family caps.

into treatment. Furthermore, both groups have similar birth and unemployment rates. This suggests that family cap implementation is not driven by disparities in welfare expenditure, fertility, or employment. Table A.2 shows that at $\tau = -1$, a more immediate and uniform measure of pre-treatment trends among the treated, averages for the main outcome and control variables used in the regressions among treated states are similar to column (2) in Table 1.

Figure 3: Differences in means for unweighted birth rates (Column A) and employment rates (Column B) among women at risk of welfare receipt.

Notes: The top row shows average birth and employment rates for all states from 1982 to 2010, as derived from the March CPS. Values calculated for a sample of unmarried women with at most a high school education (between ages 18 and 45). Red lines correspond to average outcomes for states that ever implemented a family cap. Blue lines correspond to average outcomes for states that never implemented a family cap. The bottom row shows differences in means for each outcome. Vertical bars denote 95% confidence intervals for t-tests on the differences in means. Grey vertical lines denote the year 1992, when the first treated state implemented a family cap.

Family caps were implemented over 11 years so differences in means from 1992 to 2003 include treated states that are as yet untreated. Two treated states repealed their caps in 2004, rendering differences in means after that point "contaminated" with treated states that are no longer treated. Although this reduces the reliability of these estimates as a measure of policy effects, that average birth rates in the treated group consistently mirror the positive trend in the untreated group is consistent with a null effect.

Table 1: Differences in pre-treatment means by individual and state-level characteristics

	(1)	(2)	(3)	(4)
	Untreated states	Treated states	Diff.	P-val on Diff.
	A. Individual outcome measures			
Gave birth in the past year (1990)	0.05	0.04	0.00	0.87
Employed (1990)	0.62	0.64	0.02	0.20
Hours worked in the past week (1990)	34.64	35.04	0.41	0.56
	B. Individual characteristics			
Age (1990)	28.08	28.02	-0.06	0.87
Number of children (1990)	0.79	0.78	-0.01	0.87
Race				
Black (1990)	0.20	0.24	0.04	0.00
White (1990)	0.75	0.72	-0.03	0.01
Hispanic (1990)	0.00	0.00	0.00	0.56
Other (1990)	0.04	0.03	-0.01	0.31
N	3852	3979		
	C. State level characteristics			
Log population (1990)	14.71	15.12	0.40	See notes.
Percent aged ≤ 5 (1990)	7.63	7.47	-0.16	-
Percent aged 65+ (1990)	12.21	12.62	0.40	-
Births per 1000 (1988)	16.33	15.55	-0.78	-
Percent urban-dwelling	67.44	67.75	0.31	-
Percent Black (1990)	8.98	12.50	3.53	-
Percent White (1990)	83.16	82.38	-0.78	-
Percent Hispanic (1990)	5.67	5.03	-0.64	-
Median female wage (1990)	306.33	311.04	4.72	-
Female unemployment rate (1990)	5.49	5.27	-0.22	-
Percent below poverty (1979)	12.64	12.58	-0.06	-
Percent AFDC recipients (1990)	4.30	3.84	-0.46	-
Percent food stamp recipients (1988)	7.68	6.68	-0.99	-
Percent Medicaid recipients (1988)	8.40	7.29	-1.11	-
Maximum monthly benefit for a family of 3 (1990)	$408.74	$359.08	-$49.66	-
Majority popular vote (1988)				
Republican	0.70	0.88	0.17	-
Democrat	0.30	0.12	-0.17	-
Senate majority (1991)				
Republican	0.26	0.17	-0.09	-
Democrat	0.26	0.42	0.16	-
Balanced	0.48	0.42	-0.06	-
House of Representatives majority (1991)				
Republican	0.19	0.21	0.02	-
Democrat	0.48	0.75	0.27	-
Balanced	0.33	0.04	-0.29	-
N	27	24		

Notes: Summary statistics for states that implemented a family cap between 1992 and 2003 (treated) and states that never implemented a family cap (untreated). Columns (1) and (2) show averages for untreated and treated states, column (3) shows the difference in means between groups, and column (4) shows the P-value on the difference in means for sample characteristics. The sample consists of unweighted observations of unmarried women with at most a high school education from the 1990 March CPS.

Sources: State level characteristics from the 1991 State and Metropolitan Area Data Book, the Urban Institute's TRIM3 Database, the Office of Family Assistance's historical AFDC caseload files, the Office of Family Assistance's historical AFDC caseload files, and Bureau of Labour Statistics.

*P-values on differences are not reported for state-level characteristics as thse are population measures.

As family caps were implemented dynamically, the parallel trends assumption might also be threatened if systematic differences between treated states drove the timing of family cap implementation. To test this, I regress an index of the year of family cap implementation on several observable state characteristics from 1990 as specified in Equation (1). Table A.3 shows that there is no evidence of an association between the year of family cap implementation and population, demographic trends, income, welfare generosity, welfare caseloads or the compositions of the Senate and House of Representatives immediately before family cap implementation. The negative and statistically significant correlation between implementation year and whether a Republican candidate won the popular vote in the 1988 Presidential election may reflect public alignment with the Republican stance on welfare reform.[26] As this relationship is not driven by factors that could cause differential trends in nonmarital birth rates among welfare recipients, the year of family cap implementation may not be endogenous. This is supported by an F-test, which indicates that these variables are not jointly significant predictors of cap timing.[27]

I test the final threat to the parallel trends assumption by conducting a placebo test for treatment anticipation in the window $-11 \leq \tau \leq -1$, assuming treatment occurred in the period $\tau = -6$. Figure 4 shows placebo event study estimates for both binary outcomes. There is no evidence of anticipatory behaviour on the probabilities of birth or employment for at-risk women as event study estimates in the five years before treated states first implemented family caps are not statistically significant. This is consistent with the differences in means shown in Figure 3 and the statistically insignificant IW estimates for overall anticipation effects. These are reported in Table A.4.

---

[26]Both the incumbent and elected administrations were Republican and advocated for welfare reform that enforced personal responsibility among recipients (Stoesz and Karger, 1993).

[27]Results reported in Table A.3.
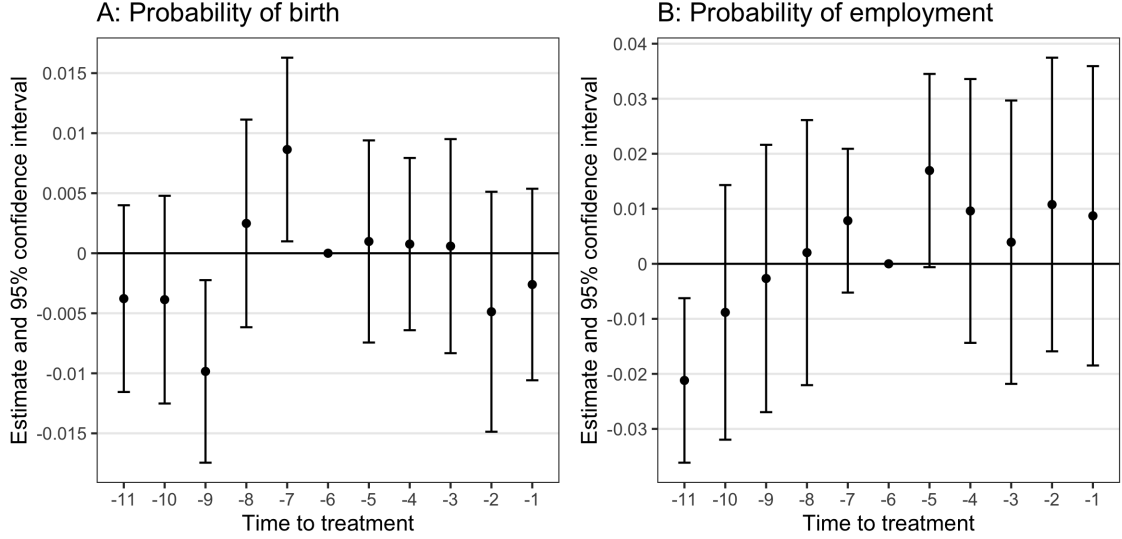
Figure 4: Placebo tests for treatment anticipation effects on the probabilities of birth and employment

Notes: Figure shows event study estimates for placebo tests of the effects of family caps on the probability of birth (Panel A) and the probability of employment (Panel B) for unmarried women with at most a high school education (between ages 18 and 45). Placebo treatment assumed to occur at $\tau = -6$ and all estimates are expressed relative to this period.

6.2 Diff-in-Diff Results

Family caps were intended to eliminate the financial incentive for additional childbearing and encourage personal responsibility through labour force participation among welfare recipients. Table 2 shows that there is no evidence of a statistically significant effect on the probability of birth among individuals at risk of welfare receipt as estimated by IW (Panel A) and TWFE (Panel B) difference-in-differences estimators. After controlling for individual characteristics, state welfare generosity, macroeconomic labour market indicators, and state and year fixed effects, both estimators for the effect of family caps on the probability of birth are positive but small in magnitude and not statistically significant. While this does not imply an effect size of 0, I can rule out effect sizes outside the range (-0.007, 0.009) for the IW estimate and (-0.008, 0.016) for the TWFEDD estimate. Consistent with Goodman-Bacon (2021), the TWFEDD estimates are quantitatively different and less precise than the IW

estimates. Panel A in Figure 5 plots the event study estimates for both estimators relative to $\tau = -1$. Notably, there is no evidence of a significant treatment effect for either estimator on the probability of birth at any post-treatment period. Together these findings suggest that family caps did not affect the fertility of women in the target group.[28]

Table 2: Effect of family caps on birth and employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: IW Estimator			
Probability of birth	0.003		0.003	-0.004	-0.004	-0.001	0.001
	(0.003)		(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
R^2	0.002		0.083	0.084	0.084	0.085	0.085
Within R^2						0.084	0.082
Probability of employment	0.037**	0.038**	0.041***	0.015	0.015	0.005	-0.011
	(0.015)	(0.015)	(0.013)	(0.010)	(0.010)	(0.011)	(0.010)
Birth in past year dummy		-0.272***	-0.155***	-0.157***	-0.157***	-0.157***	-0.156***
		(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
R^2	0.004	0.016	0.059	0.064	0.064	0.070	0.07051
Within R^2						0.057	0.055
				Panel B: TWFEDD Estimator			
Probability of birth	0.000		-0.001	-0.001	-0.006	-0.007	0.004
	(0.003)		(0.002)	(0.002)	(0.002)	(0.005)	(0.006)
R^2	0.000		0.082	0.083	0.083	0.084	0.084
Probability of employment	0.024	0.024	0.029	0.013	0.012	-0.016	-0.021
	(0.019)	(0.019)	(0.019)	(0.014)	(0.013)	(0.014)	(0.015)
Birth in the past year dummy		-0.273***	-0.156***	-0.157***	-0.157***	-0.157***	-0.156***
		(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
R^2	0.001	0.013	0.056	0.063	0.062	0.069	0.070
Individual controls			X	X	X	X	X
State controls				X	X	X	X
Lagged state controls					X	X	X
State fixed effects						X	X
Year fixed effects							X

Notes: Standard errors clustered at the state level in parentheses. Column (1) estimates a restricted model of the effect of family caps on each outcome with no controls. Columns (2)-(7) in Panel B estimate the effect of family caps on the probability of employment with a control for whether the individual had a child in the past year. Columns (3)-(7) incorporate individual controls including age, race and number of children. Columns (4)-(7) incorporate state controls including annual female unemployment rates, median weekly wages for women, maximum monthly benefits for a family of three, and a dummy for welfare time limits. Columns (5-7) includes lagged state level controls for annual female unemployment rates, median weekly wages for women, and maximum monthly benefits for a family of three. Columns (6)-(7) include state fixed effects. Column (7) includes year fixed effects. Regressions on the probability of birth include 147,166 observations. Regressions on the probability of employment include 149,647 observations. Both samples consist of unmarried women aged 18-45 years with at most a high school education. *p<0.1; **p<0.05; ***p<0.01
–

Childbirth and labour force participation are negatively correlated, particularly for women who serve as primary childcare providers in the absence of feasible sub-

[28]Table A.5 shows coefficients for the IW estimate and covariates used in columns (2)-(7) of Table 2.
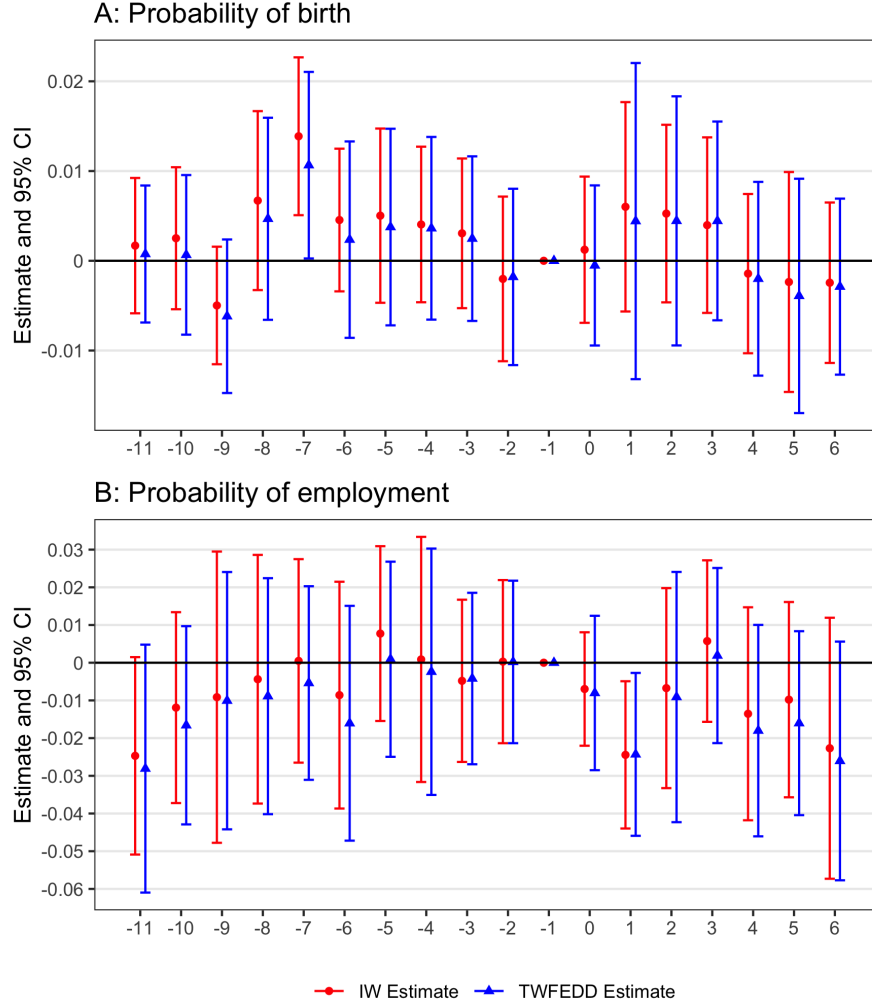
Figure 5: TWFEDD and IW event study estimates for the effects of family caps

Notes: Figure shows TWFEDD (red) and IW (blue) estimates for the effect of family caps on the probability of birth (Panel A) and the probability of being employed (Panel B) for unmarried women with at most a high school education (between ages 18 and 45). Points correspond to event study estimates and bars denote 95% confidence intervals. All estimates are relative to the period $\tau = -1$, the last full year before family cap implementation.

stitutes. Family caps could affect labour force participation measures among welfare recipients directly by encouraging personal responsibility for their livelihood and, assuming family caps affect the probability of birth, endogenously through their fertility decisions. To test the effect of family caps on labour force participation and demon-

strate the robustness of my estimate on the probability of birth, I compute the effect of family caps on the probability of employment with the same specification as Equation (4), with an additional dummy for whether the individual had a child in the past 12 months.

Consistent with a null effect on the probability of birth, Table 2 also indicates that including a dummy for birth does not affect the treatment effect estimate on the probability of employment. Column (3) shows that conditional on time-invariant individual controls, family caps are associated with a 4.1 percentage point increase in the probability of employment but this effect is not robust to the inclusion of state controls and fixed effects. The final IW estimate on the unrestricted model (column (7)) is negative, small, and not statistically significant but rules out effects outside the range (-0.031, 0.009). This implies that the true treatment effect could be negative, zero, or positive, but still small in absolute value. The TWFEDD estimate for the same model is also negative, but larger in absolute value and statistically significant at the 5% level. Given that the TWFEDD estimates on the probability of employment are biased in columns (1) through (7), and that the IW estimates are consistently measured with more precision, I disregard this result.

I also estimate the effect of family caps on a secondary measure of labour force participation, namely the number of hours usually worked per week. However, the CPS began reporting this variable in 1989, three years before the first treated state (New Jersey) implemented a family cap. I subsequently drop observations from New Jersey from regressions on this outcome to maintain a panel that is balanced in event-time. Consistent with the effect of family caps on the probability of employment, Table A.8 reports small, positive relationship between family caps and the number of hours worked but this is not statistically significant.

6.3   The role of strictness

Family caps were enforced with varying levels of stringency. Idaho and Wisconsin have implicit family caps, where all units receive the same benefits regardless of family size. Rather than bar capped children from qualifying for any additional payments, other states afford them a nominal benefit less than the standard increment. For example, capped children in Connecticut received a $50 benefit increment before repeal in 2016 while Florida affords them half the standard benefit. It could be argued that such

measures are less effective at modifying behaviour because they do not completely preclude the substitution of wages with incremental benefits. As a result, target individuals in these "lenient" states do not fully receive treatment and including them in the treatment group biases treatment effect estimates downwards.

To test whether these "lenient" states mask the effects of more stringent family caps, I compute the unrestricted IW estimator for the effect of family caps on the probabilities of birth and employment with a sample of observations from untreated and strict treated states. I define "strict" states as those with explicit family caps and that do not afford capped children any cash benefits. This sample excludes observations from Connecticut, Florida, Idaho, and Wisconsin. I also test a more stringent definition that further excludes treated states that afford capped children in-kind benefits (vouchers) in lieu of cash. This sample excludes observations from Connecticut, Florida, Idaho, Maryland, Oklahoma, South Carolina, and Wisconsin. Table 3 provides suggestive evidence of this mechanism: IW estimates for both outcomes are marginally larger in absolute value than those with the full sample of treated states. Furthermore, strict, explicit family caps that do not afford capped children any incremental benefits lower the probability of employment among at-risk women by 2.1 percentage points. This contradicts the notion that family caps might induce individuals to work more as they substitute wages for incremental benefits.

The negative relationship between family caps and employment does not reflect differential trends in macroeconomic conditions for the female labour market as the model includes controls for the female unemployment rate and median female wages. Given the intended effects of family caps, this result is difficult to rationalise. One possible explanation might be an income effect associated with family cap implementation. The null effect on the probability of birth for women in the target group implies that children are born to at-risk women in capped states regardless of the policy. These children are denied benefits that their counterparts in untreated states receive, meaning that childbirth among at-risk women in treated states leads to lower material benefit levels per person in treated units. This is effectively a negative income shock that could affect expenditure on services directly related to labour force participation. For example, reduced marginal benefits might lower the affordability of child-care, making employment infeasible if the parent is unable to find an outside substitute. By this logic, family caps would lower labour force participation for welfare recipients with capped children.

Table 3: Effect of strict family caps on fertility and labour force participation

	(1)	(2)	(3)
	Probability of birth in the past year	Probability of being employed	
Panel A: Excluding implicit caps and marginal benefit increases (cash)			
IW estimate	0.001	-0.017**	-0.016*
	(0.004)	(0.008)	(0.008)
Observations	137,883	140,481	140,481
R^2	0.086	0.067	0.070
Within R^2	0.082	0.052	0.055
Panel B: Excluding implicit caps and marginal benefit increases (cash and in-kind)			
IW estimate	0.000	-0.022***	-0.021***
	(0.003)	(0.008)	(0.008)
Observations	133,584	135,878	135,878
R^2	0.085	0.067	0.071
Within R^2	0.082	0.051	0.055
Birth in the past year dummy			X
Individual controls	X	X	X
State controls	X	X	X
Lagged state controls	X	X	X
State fixed effects	X	X	X
Year fixed effects	X	X	X

Notes: Standard errors clustered at the state level shown in parentheses. Panel A excludes capped states that provide a flat benefit regardless of cap status or afford capped children incremental benefits below the standard amount (Connecticut, Florida, Idaho, and Wisconsin). Panel B further excludes states that provide units with vouchers in the amount of the incremental benefit (Maryland, Oklahoma, and South Carolina). Column (1) shows the IW treatment effect estimate on the probability of birth. Columns (2)-(3) show the IW treatment effect estimate on the probability of employment. Columns (1)-(3) include individual controls, state controls, state fixed effects, and year fixed effects. Column (3) includes a dummy for birth in the past year. All regressions run on samples of unmarried women with at most a high school education (between ages 18 and 45).
*p<0.1; **p<0.05; ***p<0.01

6.4 An alternative proxy

The modest evidence for family cap effects may be due to a highly contaminated sample. All estimates thus far are based on a sample that proxies for welfare receipt by education and marital status. In reality, not all of these individuals are at risk of welfare receipt so subsequent estimates may be biased downwards. To test this, I construct a closer proxy for welfare receipt using the needs standards and household income data for each state and year reported in the CPS. The resulting sample only includes individuals whose household incomes fall within a bracket less than or equal to the state and year-specific need standard for the respondent's family size. Due to a large number of missing household income values, the resulting sample is over three times smaller than the samples that proxy by education and marital status. Table A.9 reports the effects of family caps on the probabilities of birth and employment

in this sample by cap strictness. These estimates are generally smaller in magnitude than those on the sample of demographic proxies and are not statistically significant. This is likely due to small sample sizes. However, these sample estimates rule out average family cap effects outside the range (-0.028, 0.004) for the probability of birth and (-0.021, 0.029) and for the probability of employment. Similarly, they rule out effects of strict family caps outside the range (-0.031, 0.001) for the probability of birth and (-0.026, 0.025) for the probability of employment.

7   Policy implications and conclusion

Past estimates of the effects of family caps on women at risk of welfare-receipt are conflicting, ambiguous and carry a significant risk of bias. Often, these have been limited by a short post-implementation period and low power arising from the use of a subset of capped states. In particular, quasi-experimental approaches that have computed TWFEDD estimates do not account for treatment effect heterogeneity over time and may subsequently be biased away from the true effect. Furthermore, previous work does not distinguish between strict family caps that deny capped children any incremental benefits, and implicit or lenient caps that afford capped children in-kind or reduced benefits. Assuming welfare generosity affects the fertility and labour force participation decisions of welfare recipients, lenient family caps may not provide strong incentives for behaviour change. Their inclusion may therefore bias the causal estimate towards zero.

In this paper, I employ frontier methods in difference-in-differences estimation to compute the causal effect of family caps on the fertility and labour force participation of women at risk of welfare receipt. This approach exploits variation in the timing and stringency of family cap implementation and considers observations 11 years before and six years after family cap implementation. Using a panel of repeated cross-sections from 1982 to 2010 allows for the use of controls for individual characteristics not reflected in state-level aggregates and addresses problems associated with limited post-treatment observations.

I find that on average, family caps have no effect on the fertility or labour force participation of women at risk of welfare receipt. This likely reflects the weakness of incremental benefits as an incentive for additional childbearing among welfare recipients. The costs of raising a child far exceed the cumulative incremental benefits

associated with their birth. Assuming individuals conceived and delivered a child while receiving welfare benefits, they would receive the corresponding incremental benefits for less than five years, after which they would bear the full, unsubsidised cost of childrearing.[29] Combined, these limitations render incremental benefits an implausible wage substitute.

Conversely, strict family caps have a modest but negative effect on their labour force participation. Over the entire sample, strict family caps lower the probability of employment among women at risk of welfare by 2.1 percentage points. This estimate is robust to numerous specifications, including those that control for individual characteristics, state-level female labour market characteristics, welfare generosity, state fixed effects, and year fixed effects. These findings are more difficult to rationalise given the intended effects of family caps. One possible argument is that under family caps, childbirth among welfare recipients lowers marginal benefits. The resulting negative income shock may lower the affordability of complements to labour force participation, including outside childcare options, lowering the probability of employment. However, this warrants further investigation.

Family caps were intended to encourage employment and discourage benefit-driven childbirth. This paper has shown that they were not successful in achieving these goals. However, it does not address whether the threat of welfare dependence outweighs the social gains from incremental welfare benefits. Given the positive effect of extra family income on child development for impoverished children (see for example Dahl and Lochner (2012)), it stands to reason that family caps may have significant negative and lasting effects on capped children and their families. Future research would do well to characterise the effects of family cap policies on the prosperity of affected individuals.

References

Gregory Acs. The impact of welfare on young mothers' subsequent childbearing decisions. *Journal of Human Resources*, 31, 1996. doi: 10.2307/146151.

Laura M Argys, Susan L Averett, and Daniel I Rees. Welfare generosity, pregnancies

---

[29] Per the federal lifetime limit on cash assistance.

and abortions among unmarried AFDC recipients. *Journal of Population Economics*, 13:569–594, 2000. URL `https://about.jstor.org/terms`.

Marianne P. Bitler, Jonah B. Gelbach, Hilary W. Hoynes, and Madeline Zavodny. The impact of welfare reform on marriage and divorce. *Demography*, 41(2):213–236, 2004. ISSN 00703370, 15337790. URL `http://www.jstor.org/stable/1515164`.

Michael J Camasso. Isolating the family cap effect on fertility behavior: Evidence from New Jersey's family development program experiment. *Contemporary Economic Policy*, 22:453, 2004. URL `https://ideas.repec.org/a/bla/coecpo/v22y2004i4p453-467.html`.

Michael J. Camasso and Radha Jagannathan. How family caps work: Evidence from a national study. *Social Service Review*, 83:389–428, 9 2009. ISSN 00377961. doi: 10.1086/648190.

Michael J Camasso and Radha Jagannathan. The future of the family cap: Fertility effects 18 years post-implementation. *Social Service Review*, 2016. doi: 10.1086/687368.

Center on Budget and Policy Priorities. Policy basics: Temporary assistance for needy families, 2021. URL `https://www.cbpp.org/research/family-income-support/temporary-assistance-for-needy-families`. Accessed: 2022-02-25.

Sena Coskun and Husnu Dalgic. The emergence of procyclical fertility: The role of gender differences in employment risk. Technical report, University of Mannheim, 2020.

Gordon B. Dahl and Lance Lochner. The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5):1927–56, May 2012. doi: 10.1257/aer.102.5.1927. URL `https://www.aeaweb.org/articles?id=10.1257/aer.102.5.1927`.

Wendy Tanisha Dyer and Robert W Fairlie. Do family caps reduce out-of-wedlock births? Evidence from Arkansas, Georgia, Indiana, New Jersey and Virginia. Discussion Paper, Economic Growth Centre, 2004 [Online], 2004.

Robert W Fairlie and Rebecca A London. The effect of incremental benefit levels on births to AFDC recipients. *Journal of Policy Analysis and Management*, 16: 575–597, 1997. URL `https://about.jstor.org/terms`.

Evan Flack and Edward Jee. *bacondecomp: Goodman-Bacon Decomposition*, 2020. URL `https://CRAN.R-project.org/package=bacondecomp`. R package version 0.1.1.

Sarah Flood, Miriam King, Ranae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. *Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]*. Minneapolis, MN: IPUMS, 2021. doi: 10.18128/ D030.V9.0.

Lisa A. Gennetian and Virginia Knox. Staying single: The effects of welfare reform policies on marriage and cohabitation. Working paper, MDRC, 2003 [Online], 2004. URL `https://www.mdrc.org/sites/default/files/full_513.pdf`.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021. doi: 10.1016/ j.jeconom.2021.03.014.

Jo Anna Gray, Jean Stockard, and Joe Stone. The rising share of nonmarital births: Fertility choice or marriage behavior? *Demography*, 43(2):241–253, 2006.

Jeff Grogger and Stephen G. Bronars. The effect of welfare payment on the marriage and fertility behavior of unwed mothers: Results from a twins experiment. *NBER Working Paper Series*, 1997.

Sandra L Hofferth, Stephen Stanhope, and Kathleen Mullan Harris. Exiting welfare in the 1990s: Did public policy influence recipients' behavior? *Population Research and Policy Review*, 21:433–472, 2002.

Ann E. Horvath-Rose, H. Elizabeth Peters, and Joseph J. Sabia. Capping kids: The family cap and nonmarital childbearing. *Population Research and Policy Review*, 27:119–138, 2008. ISSN 15737829. doi: 10.1007/s11113-008-9076-7.

Hilary Hoynes, Diane Whitmore Schanzenbach, and Douglas Almond. Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):

903–34, April 2016. doi: 10.1257/aer.20130375. URL https://www.aeaweb.org/articles?id=10.1257/aer.20130375.

Radha Jagannathan. New Jersey's family cap and welfare births: An examination of racial differences in fertility within the framework of proximate determinants. *Journal of Urban Affairs*, pages 357–375, 2003. doi: 10.1111/1467-9906.00166.

Radha Jagannathan and Michael J Camasso. Family cap and nonmarital fertility: The racial conditioning of policy effects. *Journal of Marriage and Family*, 65:52–71, 2003. URL https://www.jstor.org/stable/3600050.

Radha Jagannathan, Michael J Camasso, Mark Killingsworth, Sara McLanahan, James Trussell, Anne Case, Jennifer Hochschild, Sara Curran, Michael Greenberg, and Charles Murray. New Jersey's family cap experiment: Do fertility impacts differ by racial density? *Journal of Labor Economics*, 22, 2004. doi: 10.1086/381256.

Ted Joyce, Robert Kaestner, Sanders Korenman, and Stanley Henshaw. Family cap provisions and changes in births and abortions. *Population Research and Policy Review*, pages 475–511, 2004. URL http://www.jstor.org/stable/40230874.

Neeraj Kaushal and Robert Kaestner. From welfare to work: Has welfare reform worked? *Journal of Policy Analysis and Management*, 20:699–719, 2001. URL https://www.jstor.org/stable/3325779.

Melissa Schettini Kearney. Is there an effect of incremental welfare benefits on fertility behavior? a look at the family cap. *The Journal of Human Resources*, 39:295–325, 2004. URL https://about.jstor.org/terms.

Phillip B Levine. The impact of social policy and economic activity throughout the fertility decision tree. Working Paper 9021, National Bureau of Economic Research, June 2002. URL http://www.nber.org/papers/w9021.

Traci L. Mach. Three essays on welfare reform, 2000. URL https://www.proquest.com/dissertations-theses/three-essays-on-welfare-reform/docview/304634389/se-2?accountid=15172.

Office of the Assistant Secretary for Planning and Evaluation. Aid to Families with Dependent Children (AFDC) and Temporary Assistance for Needy Families (TANF) - Overview, 2022. URL https://aspe.hhs.gov/aid-families-

`dependent-children-afdc-temporary-assistance-needy-families-tanf-overview`. Accessed: 2022-03-19.

Diana Romero and Madina Agénor. Us fertility prevention as poverty prevention. an empirical question and social justice issue. *Women's Health Issues*, 19:355–364, 11 2009. ISSN 10493867. doi: 10.1016/j.whi.2009.08.004.

Suzanne Ryan, Jennifer Manlove, and Sandra L. Hofferth. State-level welfare policies and nonmarital subsequent childbearing. *Population Research and Policy Review*, 25:103–126, 2 2006. ISSN 01675923. doi: 10.1007/s11113-006-0004-4.

Joseph J. Sabia. Blacks and the family cap: Pregnancy, abortion, and spillovers. *Journal of Population Economics*, 21:111–134, 1 2008. ISSN 09331433. doi: 10.1007/s00148-005-0049-4.

Rebekah J Smith. Family caps in welfare reform: Their coercive effects and damaging consequences. *Harvard Journal of Law & Gender*, 29:151–200, 2006. URL `https://heinonline.org/HOL/License`.

Tomáš Sobotka, Vegard Skirbekk, and Dimiter Philipov. Economic recession and fertility in the developed world. *Population and Development Review*, 37(2):267–306, 2011. doi: https://doi.org/10.1111/j.1728-4457.2011.00411.x.

David Stoesz and Howard Jacob Karger. Deconstructing welfare: The reagan legacy and the welfare state. *Social Work*, 38(5):619–628, 1993. ISSN 00378046, 15456846. URL `http://www.jstor.org/stable/23717160`.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2): 175–199, 2021. doi: 10.1016/j.jeconom.2020.09.

The Urban Institute. Trim3 project website, 2021a. URL `trim3.urban.org`. Accessed: 2022-01-08.

The Urban Institute. Welfare rules database project website, 2021b. URL `https://wrd.urban.org/wrd/query/query.cfm`. Accessed: 2022-01-08.

C Turturro, B Benda, and H Turney. *Arkansas Welfare Waiver Demonstration Project: Final Report*. 1997.

United States. Bureau of the Census. State and metropolitan area data book [united states]: 1991. inter-university consortium for political and social research [distributor], 2008. URL `https://doi.org/10.3886/ICPSR06398.v1`. Accessed: 2022-02-05.

Geoffrey L. Wallace. The effects of family caps on the subsequent fertility decisions of never-married mothers. *Journal of Population Research*, 26:73–101, 2009. ISSN 14432447. doi: 10.1007/s12546-008-9006-x.

Michael Wiseman. Welfare's children. Discussion Paper, Wisconsin Univ., Madison. Inst. for Research on Poverty, 2000, [Online], 2000. URL `https://eric.ed.gov/?id=ED446183`.

A   Supplementary tables

Table A.1: Family Cap Implementation and Repeal By State

State	Implemented	Repealed
Arizona	November 1995	
Arkansas	July 1994	
California	August 1997	June 2016
Connecticut[a]	January 1996	
Delaware	October 1995	
Florida[b]	October 1996	
Georgia	January 1994	
Idaho[c]	July 1997	
Illinois	January 1996	January 2004
Indiana	May 1995	
Maryland[d][e]	April 1996	September 2004
Massachusetts	October 1995	April 2019
Minnesota	July 2003	January 2015
Mississippi	November 1995	
Nebraska	November 1995	October 2007
New Jersey[f]	October 1992	October 2020
North Carolina	July 1996	
North Dakota	July 1998	
Oklahoma[e]	November 1997	November 2009
South Carolina[e]	January 1997	
Tennessee	September 1996	
Virginia	July 1995	
Wisconsin[c]	January 1996	
Wyoming	February 1997	October 2008

Notes: Data from The Urban Institute (2021b)
[a] Capped children receive $50
[b] Capped children receive half the standard increment
[c] State has an implicit caps: all units receive the same benefits regardless of unit size
[d] Incremental benefit goes to a third party for use on the child's behalf
[e] Units receive vouchers in the amount of the incremental benefit
[f] Units in which at least one adult member of the unit is working are not subject to the family cap.

Table A.2: Summary statics for the target group in capped states when $\tau = -1$

Variable	N	Mean	Std. Dev.
	A. Individual outcome measures		
Birth in the past year	2861		
Did not give birth in the past year	2742	95.8%	
Gave birth in the past year	119	4.2%	
Employment status	2861		
Employed	1773	62%	
Unemployed	1088	38%	
Hours spent at work last week	1773	34.14	14.01
	B. Individual characteristics		
Age	2861	28.90	8.49
Race	2861		
Black	669	23.4%	
Hispanic	17	0.6%	
Other	110	3.8%	
White	2065	72.2%	
Number of children	2861	0.91	1.25
Age of youngest child	1311	7.23	6.35
Age of eldest child	1311	10.54	6.92
	C. State controls		
Median female wage	24	388.23	69.13
Female unemployment rate	24	5.10	1.08
Maximum welfare benefit for a family of 3	24	369.04	148.79
Welfare time limits	24		
In place	17	70.8%	
Not in place	7	29.2%	

Note: Basic Monthly CPS samples from 1976-1988 only recognised "White", "Black" and "Other" as racial groups. More detailed categories from subsequent samples for non-Hispanic individuals have been merged into "Other" for comparability. The target group is defined as unmarried women with at most a high school education (between ages 18 and 45).

Table A.3: Determinants of State Family Cap Implementation Year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log population (1990)	-0.249 (0.426)											0.053 (0.691)
Percent of population black 1990		-0.061 (0.041)										0.039 (0.097)
Marriages per 1000 people (1988)			-0.118 (0.179)									-0.061 (0.266)
Births per 1000 people (1988)				0.166 (0.352)								0.650 (0.533)
Legal abortions per 1000 women (15-44 years) (1988)					-0.031 (0.041)							-0.190 (0.108)
Log median household income (1979)						1.283 (2.862)						1.822 (4.975)
Maximum monthly benefit for a family of 3 (1990)							0.005 (0.003)					0.009 (0.009)
Percent of population AFDC recipients (1990)								-0.048 (0.340)				-0.192 (0.619)
Senate majority - Dem. (1991)									-1.300 (0.888)			-1.721 (1.069)
Senate majority - Rep. (1991)									-0.750 (1.175)			-2.354 (1.953)
House of Representatives majority - Dem. (1991)										-0.167 (2.142)		-1.125 (2.588)
House of Representatives majority - Rep. (1991)										-0.200 (2.284)		-1.882 (2.854)
Presidential popular vote - Rep. (1988)											-2.476** (1.141)	-0.537 (2.015)
Constant	8.597 (6.454)	5.602*** (0.643)	5.997*** (1.812)	2.250 (5.487)	5.523*** (1.002)	-7.624 (27.785)	3.167** (1.130)	5.016*** (1.369)	5.500*** (0.628)	5.000** (2.085)	7.000*** (1.068)	-19.109 (51.556)
R^2	0.015	0.094	0.019	0.010	0.025	0.009	0.101	0.001	0.093	0.0004	0.176	0.520
F Statistic	0.341	2.295	0.435	0.223	0.569	0.201	2.476	0.020	1.075	0.004	4.706**	0.834
df	1	1	1	1	1	1	1	1	2	2	1	13
Res. df	22	22	22	22	22	22	22	22	21	21	22	10

Notes: All data are at the state level. The dependent variable is equal to the year the state first implemented a family cap (normalised to 1 in 1992). Control variables come from the 1991 State and Metropolitan Area Data Book, the Urban Institute's TRIM3 Database and the Office of Family Assistance's historical AFDC caseload files. Population weighted versions of these regressions are available upon request.
*p<0.1; **p<0.05; ***p<0.01

Table A.4: Anticipation effects of family caps, full placebo regression results

	Probability of birth	Probability of employment
	(1)	(2)
IW Estimate	-0.001	0.010
	(0.003)	(0.010)
Age	-0.004***	0.010***
	(0.0001)	(0.0004)
Race		
Hispanic	-0.005	0.042
	(0.011)	(0.037)
Other	0.0004	0.006
	(0.005)	(0.022)
White	-0.008***	0.162***
	(0.002)	(0.013)
Number of children	0.046***	-0.060***
	(0.001)	(0.005)
Gave birth in the past year		-0.174***
		(0.008)
Welfare time limit	0.0001	-0.006
	(0.004)	(0.015)
Female unemployment rate, t	0.0009	-0.008***
	(0.0007)	(0.003)
Median female weekly wage, t	-1.89×10^{-5}	-0.0003**
	(3.62×10^{-5})	(0.0001)
Maximum monthly benefit for a family of 3, t	2.57×10^{-5}	0.0001**
	(4.66×10^{-5})	(5.74×10^{-5})
Female unemployment rate, t-1	0.0002	-0.007***
	(0.0009)	(0.002)
Median female weekly wage, t-1	$-7.62 \times 10^{-5*}$	2.86×10^{-5}
	(4.02×10^{-5})	(0.0001)
Maximum monthly benefit for a family of 3, t-1	-2.58×10^{-5}	-0.0001**
	(3.11×10^{-5})	(5.81×10^{-5})
State fixed effects	X	X
Year fixed effects	X	X
Observations	110,115	112,074
R^2	0.08644	0.08541
Within R^2	0.08250	0.06912

Notes: Standard errors clustered at the state level shown in parentheses. Anticipation effects of family caps on unmarried women aged 18-45 with at most a highschool education in all states from $-11 \leq \tau \leq -1$. Column (1) estimates anticipation effects on the probability of birth. Column (2) estimates anticipation effects on the probability of employment. All columns include state and year fixed effects.
*p<0.1; **p<0.05; ***p<0.01

Table A.5: Effect of family caps on the probability of birth, full regression results

	Probability of birth					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.042***	0.131***	0.123***	0.124***		
	(0.002)	(0.004)	(0.005)	(0.005)		
IW Estimate	0.003	0.003	-0.004	-0.004	-0.001	0.001
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
Age		-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Race						
Hispanic		0.006	0.002	0.002	0.001	0.0004
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
White		-0.004***	-0.004***	-0.004***	-0.005***	-0.005***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Other		0.006	0.005	0.005	0.002	0.002
		(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Number of children		0.048***	0.048***	0.048***	0.047***	0.047***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female unemployment rate, t			-0.001***	-0.0006	-0.001*	3.18×10^{-5}
			(0.0004)	(0.0006)	(0.0006)	(0.0007)
Median female weekly wage, t			4.76×10^{-5}***	3.35×10^{-5}	4.28×10^{-5}*	-8.15×10^{-6}
			(9.75×10^{-6})	(2.22×10^{-5})	(2.26×10^{-5})	(2.33×10^{-5})
Maximum monthly benefit, t for a family of 3		-6.9×10^{-6}	-1.3×10^{-5}	6.64×10^{-6}	-4.36×10^{-6}	
			(8.28×10^{-6})	(1.79×10^{-5})	(2.84×10^{-5})	(2.83×10^{-5})
Welfare time limit			0.002	0.002	-0.002	0.0008
			(0.003)	(0.003)	(0.003)	(0.003)
Female unemployment rate, t-1				-0.0007	-0.0009	0.0003
				(0.0008)	(0.0008)	(0.0008)
Median female weekly wage, t-1				1.3×10^{-5}	7.33×10^{-6}	-6.54×10^{-5}**
				(2.25×10^{-5})	(2.31×10^{-5})	(2.95×10^{-5})
Maximum monthly benefit, for a family of 3, t-1				6.17×10^{-6}	1.61×10^{-5}	-7.89×10^{-6}
				(1.87×10^{-5})	(2.16×10^{-5})	(2.07×10^{-5})
State fixed effects					X	X
Year fixed effects						X
Observations	147,166	147,166	147,166	147,166	147,166	147,166
R^2	0.002	0.083	0.084	0.084	0.085	0.08536
Within R^2					0.084	0.082

Notes: Standard errors clustered at the state level shown in parentheses. Effect of family caps on the probability of birth of unmarried women aged 18-45 with at most a high school education in all states from $-11 \leq \tau \leq 6$. Columns (5)-(6) include state fixed effects. Column (6) includes year fixed effects.
*p<0.1; **p<0.05; ***p<0.01

Table A.6: Policy effect estimates with different samples of women

	Probability of birth			
	(1)	(2)	(3)	(4)
Sample	Unmarried, less-educated women	Less-educated women	Unmarried Women	All women
IW estimate	0.001	0.000	0.002	0.0006
	(0.004)	(0.003)	(0.002)	(0.002)
Age	-0.004***	-0.007***	-0.003***	-0.006***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Race				
Hispanic	0.0004	0.014**	0.0007	0.008*
	(0.009)	(0.006)	(0.006)	(0.004)
White	-0.005***	0.008***	-0.008***	0.011***
	(0.001)	(0.0010)	(0.001)	(0.0009)
Other	0.002	0.007**	-0.004	0.011***
	(0.005)	(0.003)	(0.003)	(0.002)
Number of children	0.047***	0.043***	0.043***	0.045***
	(0.001)	(0.0006)	(0.0007)	(0.0004)
Female unemployment rate, t	3.18×10^{-5}	-0.0004	-0.0003	-0.0003
	(0.0007)	(0.0007)	(0.0005)	(0.0004)
Median weekly female wage, t	-8.15×10^{-6}	-2.71×10^{-5}	8.27×10^{-6}	6.49×10^{-6}
	(2.33×10^{-5})	(2.76×10^{-5})	(1.51×10^{-5})	(1.81×10^{-5})
Maximum monthly benefit for a family of 3, t	-4.36×10^{-6}	-1.67×10^{-5}	1.52×10^{-5}	5.69×10^{-7}
	(2.83×10^{-5})	(1.83×10^{-5})	(1.48×10^{-5})	(1.34×10^{-5})
Welfare time limit	0.0008	-0.001	0.001	0.0004
	(0.003)	(0.003)	(0.002)	(0.002)
Female unemployment rate, t-1	0.0003	0.0002	0.0004	-0.0005
	(0.0008)	(0.0006)	(0.0004)	(0.0004)
Median weekly female wage, t-1	$-6.54 \times 10^{-5**}$	3.59×10^{-5}	$-4.67 \times 10^{-5**}$	1.83×10^{-5}
	(2.95×10^{-5})	(3.1×10^{-5})	(2×10^{-5})	(1.94×10^{-5})
Maximum monthly benefit for a family of 3, t	-7.89×10^{-6}	-1.14×10^{-5}	$-2.3 \times 10^{-5*}$	-1.17×10^{-5}
	(2.07×10^{-5})	(1.78×10^{-5})	(1.34×10^{-5})	(1.29×10^{-5})
State fixed effects	X	X	X	X
Year fixed effects	X	X	X	X
Observations	147,166	326,304	301,484	677,715
R^2	0.085	0.074	0.076	0.065
Within R^2	0.082	0.073	0.074	0.064

Notes: Standard errors clustered at the state level shown in parentheses. Effect of family caps on the probability of birth for different samples of women in all states from $-11 \leq \tau \leq 6$. Column (1) estimates the probability of birth on a sample of unmarried women aged 18-45 with at most a high school education. Column (2) uses a sample of women aged 18-45 with at most a high school education. Column (3) uses a sample of unmarried women aged 18-45 . Column (4) uses all women aged 18-45. All columns include state and year fixed effects.
*p<0.1; **p<0.05; ***p<0.01

Table A.7: Effect of family caps on the probability of employment, full regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of employment						
Constant	0.603***	0.614***	0.284***	0.496***	0.496***		
	(0.013)	(0.013)	(0.015)	(0.024)	(0.024)		
IW Estimate	0.037**	0.038**	0.041***	0.015	0.015	0.005	-0.011
	(0.015)	(0.015)	(0.013)	(0.010)	(0.010)	(0.011)	(0.010)
Birth		-0.272***	-0.155***	-0.157***	-0.157***	-0.157***	-0.156***
		(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Age			0.009***	0.009***	0.009***	0.009***	0.009***
			(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Race							
Hispanic			0.087***	0.087***	0.087***	0.080***	0.082***
			(0.026)	(0.024)	(0.024)	(0.024)	(0.024)
White			0.151***	0.144***	0.144***	0.144***	0.144***
			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Other			0.026	0.017	0.017	-0.002	-0.002
			(0.025)	(0.025)	(0.025)	(0.021)	(0.021)
Number of children			-0.048***	-0.048***	-0.048***	-0.048***	-0.048***
			(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female unemployment rate, t				-0.010***	-0.010***	-0.009***	-0.009***
				(0.002)	(0.002)	(0.002)	(0.002)
Median female weekly wage, t				-0.0002**	-0.0002**	-0.0002***	-0.0002***
				(9.27×10^{-5})	(9.27×10^{-5})	(6.83×10^{-5})	(7.31×10^{-5})
Maximum monthly benefit for a family of 3, t				1.93×10^{-5}	1.93×10^{-5}	0.0001	0.0002**
				(5.24×10^{-5})	(5.24×10^{-5})	(8.47×10^{-5})	(8.92×10^{-5})
Welfare time limit				0.023*	0.023*	0.034***	0.002
				(0.012)	(0.012)	(0.010)	(0.010)
Female unemployment rate, t-1				-0.010***	-0.010***	-0.007***	-0.006***
				(0.003)	(0.003)	(0.002)	(0.002)
Median female weekly wage, t-1				5.39×10^{-5}	5.39×10^{-5}	-1.76×10^{-5}	7.23×10^{-5}
				(9.23×10^{-5})	(9.23×10^{-5})	(7.84×10^{-5})	(9.82×10^{-5})
Maximum monthly benefit for a family of 3, t-1				-8.02×10^{-5}	-8.02×10^{-5}	-0.0001***	-0.0001**
				(6.43×10^{-5})	(6.43×10^{-5})	(5.01×10^{-5})	(4.72×10^{-5})
State fixed effects						X	X
Year fixed effects							X
Observations	149,647	149,647	149,647	149,647	149,647	149,647	149,647
R^2	0.004	0.016	0.059	0.064	0.064	0.070	0.072
Within R^2						0.057	0.055

Notes: Standard errors clustered at the state level shown in parentheses. Effect of family caps on the probability of unemployment for unmarried women aged 18-45 with at most a high school education in all states from $-11 \leq \tau \leq 6$. Columns (6)-(7) include state fixed effects. Column (7) includes year fixed effects.
*p<0.1; **p<0.05; ***p<0.01

Table A.8: Effect of family caps on the number of hours usually worked per week, full regression results

	Number of hours worked per week						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	33.784***	33.912***	22.363***	24.476***	24.655***		
	(0.187)	(0.186)	(0.483)	(0.768)	(0.801)		
IW Estimate	0.569**	0.590**	0.510*	0.399	0.393	0.380	0.240
	(0.277)	(0.277)	(0.266)	(0.275)	(0.279)	(0.264)	(0.293)
Birth		-4.36***	-2.12***	-2.11***	-2.11***	-2.09***	-2.09***
		(0.411)	(0.418)	(0.418)	(0.418)	(0.418)	(0.421)
Age			0.395***	0.396***	0.396***	0.395***	0.395***
			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Race							
Hispanic			-0.343	-0.005	-0.012	-0.128	-0.095
			(0.736)	(0.748)	(0.750)	(0.727)	(0.735)
White			0.197	0.286	0.277	0.478**	0.476**
			(0.216)	(0.243)	(0.244)	(0.229)	(0.231)
Other			0.101	0.454	0.439	0.430	0.446
			(0.485)	(0.453)	(0.457)	(0.507)	(0.505)
Number of children			-0.355***	-0.364***	-0.364***	-0.346***	-0.346***
			(0.063)	(0.062)	(0.062)	(0.065)	(0.065)
Female unemployment rate, t				-0.028	0.041	-0.052	-0.120*
				(0.064)	(0.063)	(0.065)	(0.071)
Median female weekly wage, t				-0.003**	-0.003	-0.006**	-0.008**
				(0.001)	(0.002)	(0.002)	(0.003)
Maximum monthly benefit for a family of 3, t				-0.002***	-0.004*	-0.005*	-0.005*
				(0.0008)	(0.002)	(0.003)	(0.003)
Welfare time limit				0.296	0.282	0.197	0.342
				(0.298)	(0.308)	(0.216)	(0.282)
Female unemployment rate, t-1					-0.091	-0.245***	-0.166**
					(0.065)	(0.061)	(0.080)
Median female weekly wage, t-1					0.0006	0.002	0.002
					(0.003)	(0.003)	(0.003)
Maximum monthly benefit for a family of 3, t-1					0.001	0.001	0.002
					(0.002)	(0.002)	(0.003)
State fixed effects						X	X
Year fixed effects							X
Observations	62,823	62,823	62,823	62,823	62,823	62,823	62,823
R^2	0.004	0.007	0.064	0.065	0.065	0.069	0.069
Within R^2						0.06198	0.06105

Notes: Standard errors clustered at the state level shown in parentheses. Effect of family caps on the number of hours worked per week by unmarried women aged 18-45 with at most a high school education in all states except New Jersey from $-7 \leq \tau \leq -6$. Columns (6)-(7) include state fixed effects. Column (7) includes year fixed effects.
*p<0.1; **p<0.05; ***p<0.01

Table A.9: Effect of family caps on an alternative proxy

	(1)	(2)	(3)
	Probability of birth	Probability of being employed	
	A: All states		
IW estimate	-0.012	0.006	0.004
	(0.008)	(0.014)	(0.013)
R^2	0.097	0.058	0.063
TWFEDD Estimate	0.012	0.008	0.008
	(0.013)	(0.025)	(0.025)
R^2	0.093	0.055	0.059
Observations	39,859	42,891	42,891
	B: Excluding implicit caps and marginal benefit increases (cash)		
IW estimate	-0.011	0.004	0.001
	(0.008)	(0.013)	(0.013)
R^2	0.097	0.061	0.066
TWFEDD Estimate	0.013	-0.003	0.000
	(0.014)	(0.026)	(0.003)
R^2	0.094	0.058	0.062
Observations	36,760	39,565	39,565
	C: Excluding implicit caps and marginal benefit increases (cash and in-kind)		
IW estimate	-0.015	0.002	-0.0002
	(0.008)	(0.013)	(0.013)
R^2	0.097	0.061	0.066
TWFEDD Estimate	0.013	-0.004	-0.001
	(0.014)	(0.027)	(0.027)
R^2	0.098	0.058	0.062
Observations	35,954	38,670	38,670
Birth in the past year dummy			X
Individual controls	X	X	X
State controls	X	X	X
State fixed effects	X	X	X
Year fixed effects	X	X	X

Notes: Standard errors clustered at the state level shown in parentheses. Effect of family caps on the probabilities of birth and employment on low-income women aged 18 to 45. Column (1) reports the IW and TWFEDD estimates for the effect on the probability of birth. Columns (2) and (3) report the IW and TWFEDD estimates for the effect on the probability of employment. Columns (1)-(3) include individual controls, state controls, lagged state controls, state fixed effects, and year fixed effects. Column (3) includes a dummy for whether the individual gave birth in the previous year. Panel A is estimated with all capped states. Panel B excludes capped states with marginal cash benefit increases and implicit family caps. Panel C excludes capped states with marginal benefit increases of any kind and implicit family caps.
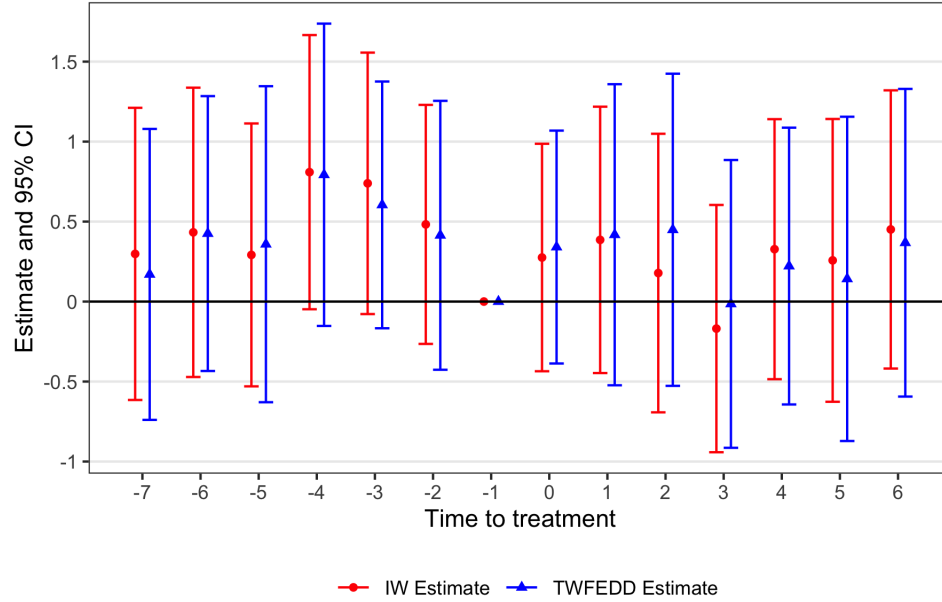*p<0.1; **p<0.05; ***p<0.01

B   Supplementary figures

Figure B.1: TWFEDD and IW event study estimates for the effects of family caps on the number of hours usually worked per week.

Notes: Figure shows TWFEDD (red) and IW (blue) estimates for the effect of family caps on number of hours usually worked per week by women at risk of welfare receipt. Points correspond to event study estimates and bars denote 95% confidence intervals. All estimates are relative to the period $\tau = -1$, the last full year before family cap implementation. Sample includes unweighted observations for unmarried women with at most a high school education (between ages 18 and 45) from all states except New Jersey and spans 1989-2010.
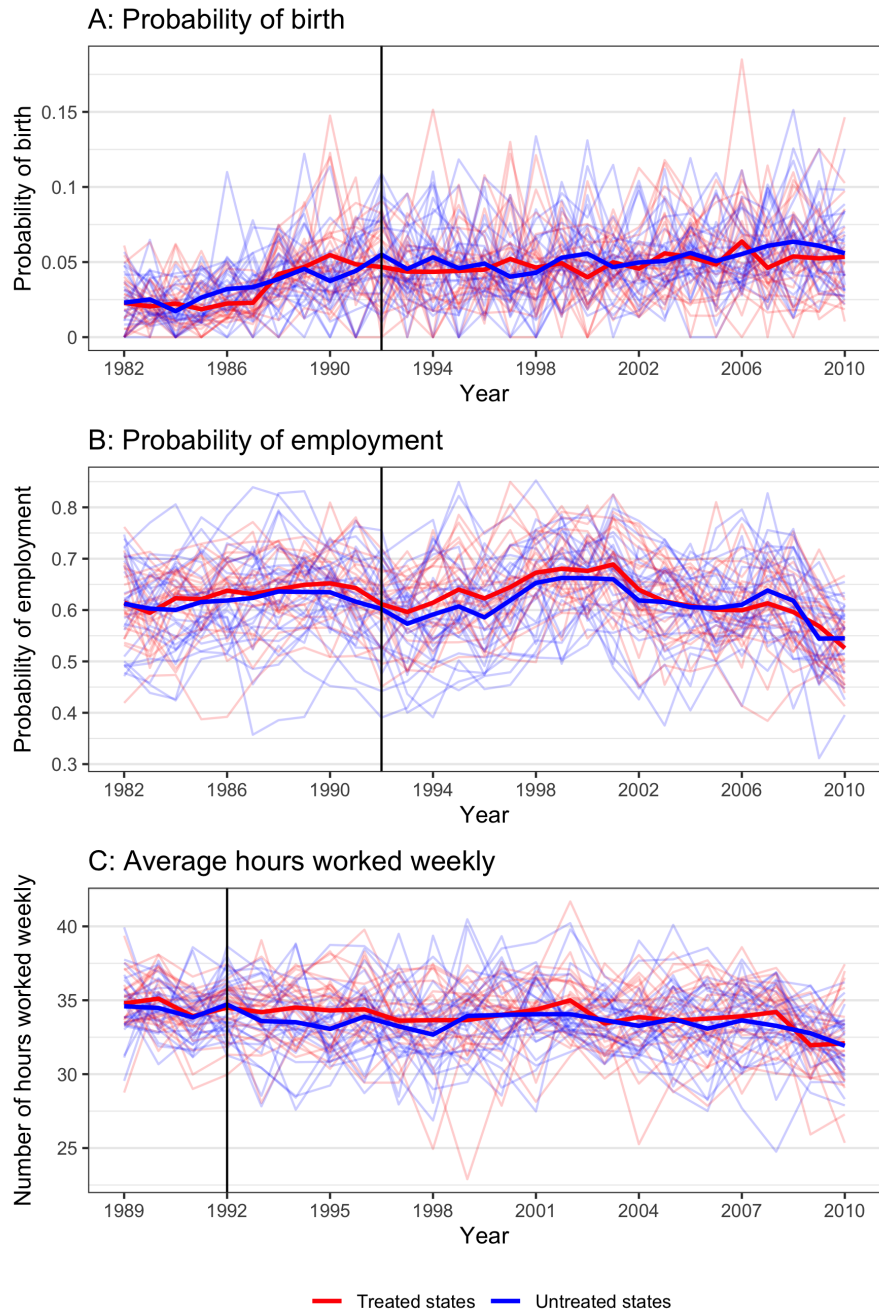
Figure B.2: Weighted outcomes by state and treatment group.

Notes: Figure shows survey weighted probabilities of birth (Panel A), probabilities of employment (Panel B), and number of hours worked weekly by women at risk of welfare receipt (Panel C) in the March CPS. Darker lines denote average outcomes by state treatment status. Lighter lines denote average outcomes by state. Black horizontal lines at 1992 indicate the first year a treated state implemented a family cap. Samples defined as unmarried women with at most a high school education (between ages 18 and 45).

C   Goodman-Bacon Decomposition

In this section I replicate Goodman-Bacon (2021)'s illustration of their DD decomposition theorem with data from the CPS to identify sources of variation and bias in the TWFEDD estimate of effects of family caps.

Because the decomposition requires a strongly balanced panel with the same number of observations per period, I compute state-level aggregates of the target group's birthrate to ensure one observation per state per year. Using person weights from the CPS, I calculate the number of unmarried women in each state and year aged 15-45 with at most a high school education. I then use the same weights to calculate the number of women in this group who have birth to a child in the past year and compute the birth rate for the target population in each state and year by dividing this value by the size of the target group. I include observations for all years before the first treated state repealed its policy, resulting in a study period from 1982 to 2003 and seven treatment times outlined in Table C.10.

Table C.10: Child exclusion policy roll out: Treatment times, group sizes, and shares of treatment & study periods

Group	Number of states	Share of states	Treatment share
Never-treated	27	0.53	0.00
1992	1	0.02	0.52
1994	2	0.04	0.43
1995	7	0.14	0.38
1996	7	0.14	0.33
1997	5	0.10	0.29
1998	1	0.02	0.24
2003	1	0.02	0.05

Notes: The table lists the number of states in each group, the shares they comprise of all states, and the share of the study period they spend treated from 1982 to 2003.

I conduct an event study with the TWFEDD estimator with the specification

$$Y_{st} = \sum_{\tau \neq 0} \beta_\tau YearRelativeCap_\tau + \alpha_d + \alpha_t \tag{7}$$

where y_{st} is the target group's birth rate in state s and calendar year t. α_s are state fixed effects, α_t are year fixed effects. τ refers to the year relative to family cap implementation. For each state, $\tau = 0$ represents the first full year of family cap implementation. $YearRelativeCap_\tau$ are a set of dummy variables equal to one if a

state had a family cap for the whole year at event time τ. The effect at $\tau = 0$ is normalised to 0 so each coefficient is interpreted as the effect of family caps at each event time relative to the full year of treatment. I do not include any time-varying controls because this event study serves as a comparison for the TWFEDD estimator and the latter is sensitive to additional identifying variation through the inclusion of covariates (Goodman-Bacon, 2021).

I then conduct an event study with the IW estimator with a similar specification to Equation (4), but without state and individual controls and with a wider event-time window $-22 \leq \tau \leq 11$. The effect at $\tau = 0$ is normalised to 0 so each coefficient is interpreted at the effect of family caps at each event time relative to first full year of treatment.

The TWFEDD estimate of the effect of family caps is 0.97 (SE = 2.70) whereas the IW estimate is 3.42 (SE = 5.84). Although this indicates bias, neither effect is statistically significant so it stands to reason that family caps had no effect on birth rates in the target group. Figure C.3 illustrates the results of both event studies and the overall treatment effect. Both methods produce fairly consistent estimates for $\tau \geq -15$ but diverge before this period: the IW estimates suggest that treated states had significantly lower birth rates in the target group relative to $\tau = 0$ whereas the TWFEDD estimates indicate significantly higher birth rates for $\tau = 20$. These differences may contribute to divergences in the overall treatment effect estimates.

To demonstrate the Goodman-Bacon decomposition for the TWFEDD estimator, I estimate

$$y_{st} = \delta^{DD} D_{st} + \alpha_d + \alpha_t \tag{8}$$

where y_{st} is the target group's birth rate in state s and year t, D_{st} is a dummy indicating whether a state s has an active family cap in year t, α_s are state fixed effects, α_t are year fixed effects and δ^{DD} is the TWFEDD estimator.

The TWFEDD estimate is the weighted average of 49 $2 \times 2DD$ estimates. Table C.11 shows the number, average effect and weight of comparisons by type. The average estimate is positive for treated vs. untreated comparisons and accounts for 88.3% of the final estimate. This is driven by the large number of untreated states (27) relative to treated states (at most 7 in any given comparison group). Timing variation accounts for only 11.7% of the final estimate. Figure C.4 plots $2 \times 2DD$ estimates by their weight, showing that the ten largest estimates account for 91.4% of the final treatment effect.
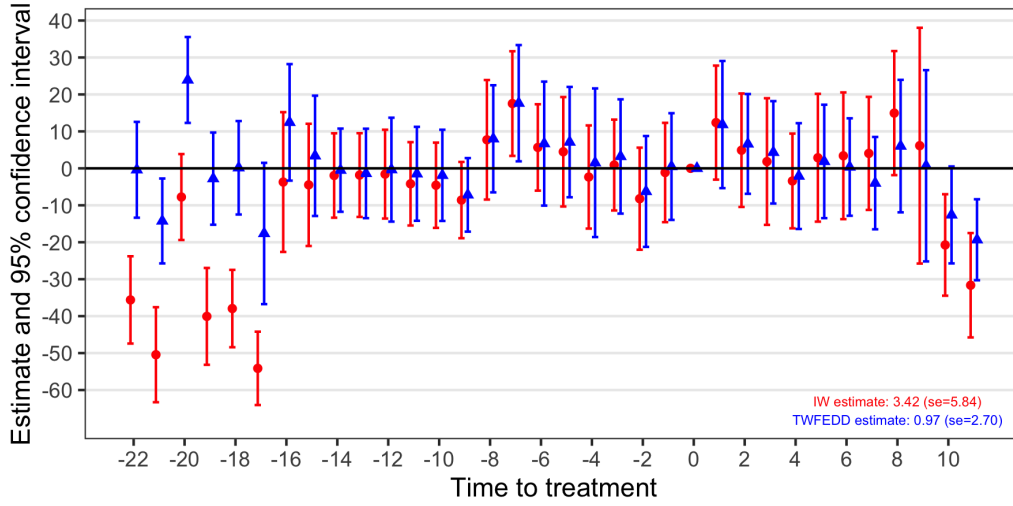
55

Figure C.3: TWFEDD and IW event study estimates of the effect of family cap policies on birth rates

Notes: Figure shows IW (red) and TWFEDD (blue) event study estimates for the effect of family caps on birth rates among women at risk of welfare receipt. Points denote event study estimates and lines denot 95% confidence intervals. All event study estimates are expressed relative to $\tau = 0$, the first full year treated states had a family cap in place.

Table C.11: Goodman-Bacon Decomposition Summary

Type	Average effect	Number of 2×2 estimators	Total weight
Earlier vs Later Treated	0.83	21	0.084
Later vs Earlier Treated	6.41	21	0.033
Treated vs Untreated	0.78	7	0.883

Although all cross-group treatment effects have positive weights, there is evidence of downward bias from time-varying treatment effects (Table C.11). The average post-treatment event study estimate is -0.57 whereas the the TWFEDD estimate is 0.97. This is likely driven by comparisons that use earlier treated observations as controls and later treated observations as treated. These comparisons result in large, positive average treatment effect estimate. Though this only comprises 3% of the final estimate, excluding it from the VWATT reduces the effect size to 0.76, which is marginally closer to the post-treatment event study estimate. Although this is evidence of bias in the estimator, it ultimately has no bearing on the final interpretation of the VWATT because it is not statistically significant.
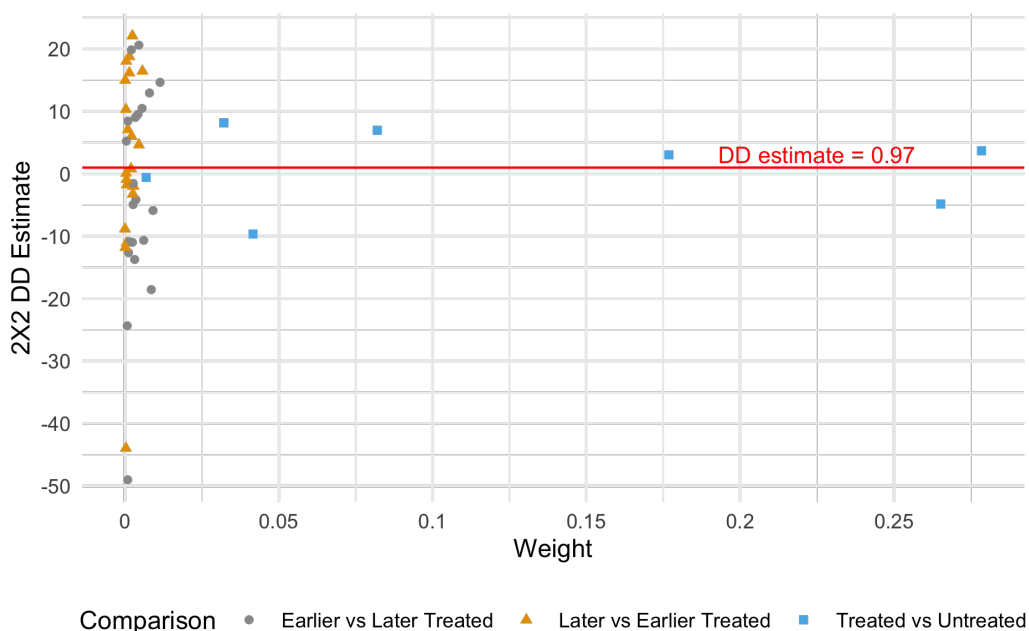
Figure C.4: 2×2DD estimators with weights

Notes: Figure plots 2×2DD estimators by their assigned weights. Grey circles denote 2×2DD estimates from earlier vs later treated comparisons. Yellow triangles denote 2×2DD estimates from later vs earlier treated comparisons. Blue squares denote 2×2DD estimates from treated vs untreated comparisons. The red horizontal line denotes the aggregate TWFEDD estimate for the effect of family caps on birth rates among women at risk of welfare receipt.