

Strathclyde

Discussion Papers
in Economics



Does Paid Family Leave Affect Abortion? Evidence from New York

Rory Allanson, Otto Lenhart, Agnese Romiti

No. 24 – 1

March 2024

Department of Economics
University of Strathclyde, Glasgow

Does Paid Family Leave Affect Abortion?

Evidence from New York

Rory Allanson [†] Otto Lenhart [‡] Agnese Romiti ^{‡§}

March 22, 2024

Abstract

As state Paid Family Leave (PFL) programs continue to roll out across the United States, previous work has sought to investigate their impacts on economic, child, and maternal outcomes, including fertility. The impact they may have on abortion is however still unexplored. We employ the Synthetic-Difference-in-Differences estimator developed by Arkhangelsky et al. (2021) to estimate the effect of New York's PFL program (NY-PFL) on abortion rates. Using abortion data from the Centers for Disease Control and Prevention, we find that the launch of NY-PFL in 2018 led to a 13.6% decrease in abortion rates per 1,000 women for the 20-39 age group, with smaller effects observed for older women. Event-study estimates reveal that this decrease intensified from an initial 7.1% decline in 2018-19 to 13.6% in 2021, while robustness checks underline the significance of our findings. Our exercise contributes further evidence towards the deliberation of state PFL programs.

Keywords: Abortion, Paid Family Leave, Maternity Leave, Synthetic Difference-in-Differences
JEL-Codes: I18, I38, J13, J18

1 Introduction

Paid Family Leave (PFL) policies - which provide paid time off work for new parents at some level of wage replacement - are currently offered at the national level by all 38 OECD members states with the exception of the United States (OECD, 2022). It is hence up to each U.S. state whether to establish their own family leave policies in excess of the limited federal foundation. While many offer some package of medical and disability benefits (Byker, 2016; Bullinger, 2019), ten states have now launched their own PFL programmes, including New York, whose benefits began on 2018/1/1. Previous work has examined the ability of these programmes to affect child, parental, and economic outcomes (e.g. Rossin-Slater et al., 2013; Bütikofer et al., 2021; Lenhart, 2021). Abortion however, despite its predominance in the modern American consciousness following *Dobbs vs. Jackson Women's Health Organization* (Supreme Court, 2022; Grossman et al., 2022; Rader et al., 2022) remains unexplored with regards to PFL, a gap which this paper addresses.

[†]Corresponding Author. Department of Economics, University of Strathclyde. Email: rory.allanson@strath.ac.uk.

[‡]Department of Economics, University of Strathclyde.

[§]This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We are grateful to Caitlin Knowles Myers, Kelly Jones, and Mayra Pineda-Torres for their assistance in sourcing abortion data. Special thanks are due to Cristina Tealdi, Daniel I. Rees, and Tanya Wilson for their helpful comments, which have improved our paper substantially. We also thank participants at the 2023 University of Strathclyde PGR Conference and 2024 SGPE Conference, as well as members of the Erasmus University Department of Applied Economics, for their feedback. We have no conflicts of interest to disclose. All errors remain our own.

We propose that PFL programmes can influence abortion decisions through an income effect. More specifically, research into why women pursue abortions has revealed that perceived financial inability to provide for a child is one of the most common (Finer et al., 2005; Kirkman et al., 2009), if not the most common reason given (Larsson et al., 2002; Santelli et al., 2006; Biggs et al., 2013). As PFL offers workers direct financial assistance through the post-birth period as well as indirect assistance through job protections that aid them in retaining employment, it may decrease abortion rates by alleviating some of the financial pressure faced by prospective parents, alongside associated stress (Bullinger, 2019). Secondly, it has already been demonstrated that PFL can impact fertility: Thunell (2017) and Golightly and Meyerhofer (2022) find a positive, heterogeneous effect on fertility rates following the launch of California's PFL programme, findings reinforced by supporting evidence from other OECD countries (e.g. Shim, 2014; Bassford and Fisher, 2016; Raute, 2019). It follows naturally to enquire whether PFL could similarly affect other reproductive choices, including abortion. Moreover, empirical evidence (e.g. Prifti and Vuri, 2013; Clark and Lepinteur, 2022; Nieto, 2022) outlines that increased job security can positively affect fertility and family size, suggesting that the job protection offered by NY-PFL could impact reproductive choices, and more specifically abortion.

To investigate whether NY-PFL did impact the abortion rates of women in New York, we employ the Synthetic Difference-in-Differences (SDID) framework recently developed by Arkhangelsky et al. (2021). The SDID estimator is ideal for small sample comparative studies like ours as it combines two of the most widely used methods in causal inference, namely Difference-in-Differences (DID) and Synthetic Control, to identify robust estimates of treatment effects (Clarke et al., 2023). Our exercise is conducted by constructing a synthetic counterfactual of the treated state from our pool of untreated 'donor' states, before implementing a standard DID procedure between treated and synthetic units (Arkhangelsky et al., 2021). This method provides a series of desirable properties, introducing double robustness and invariance to additive unit-level shifts, while weakening our reliance on parallel trends type assumptions (Clarke et al., 2023; Melo et al., 2023).

This work makes two main contributions to the relevant literature. First, it is the first to provide evidence of the effect of state PFL programmes on abortion rates, furthering our understanding of the impact of such policies on reproductive choices (Golightly and Meyerhofer, 2022). When acknowledging the prevalence of abortion discourse in today's America (Chang et al., 2023), consideration of the indirect effects of PFL policies on this outcome is essential towards the full evaluation of such policies. Second, our exercise provides additional evidence towards the deliberation of the numerous state PFL programmes currently forthcoming, including those of Maryland, Delaware, and Minnesota, by highlighting another channel of effect these programmes may possess. Unlike much current work - which considers California's PFL policy (Coile et al., 2023) - our evidence stems through examination of NY-PFL, hence serving to expand the nascent literature evaluating this programme.¹ Moreover when considering the distinctions between CA-PFL and the other state PFL programmes, as outlined by Table 5, providing evidence from outside California is of notable value. Finally, our exercise expands the body of work that implements the SDID method empirically,

¹Joining previous work by Bartel et al. (2021), Coile et al. (2023), and J. Kim and Lenhart (2024), among a few others. A select number of works also consider New York in addition to other states, including Winston et al. (2019) and Boyens et al. (2022).

following Melo et al. (2023) and Clifton-Sprigg et al. (2023). Arkhangelsky et al. (2021) outline the ability of this method to dominate both the DID and Synthetic Control methods in settings like ours.²

Using abortion data from the Centers for Disease Control and Prevention (CDC)'s 'Abortion Surveillance' report series,³ we achieve accurate synthetic matching for most age groups, and find that abortion rates in New York decreased by 13.6% for our core 20-39 age group following the introduction of NY-PFL in 2018. Subgroup analysis reveals that this effect declines from a 13.3% reduction for the 25-29 age group to respective 10.4% and 12.5% reductions for the 30-34 and 35-39 age groups; unsatisfactory matching limits inference for the 20-24 age group. We also present event-study style output from our SDID exercise, following Clarke et al. (2023), which reveals a dynamic effect that intensifies significantly in 2020 after an initial 7.1% decrease in 2018. Results remain significant despite the introduction of two policies introduced in New York in 2019 that affected abortion positively. In addition, the increasing uptake and generosity of NY-PFL benefits through the post-treatment period likely leads to the larger effects we observe in latter years. Sensitivity analysis then confirms that our main 20-39 result remains qualitatively robust and significant when excluding 2019, 2020, and 2021 from the sample. No notable results are found when testing for anticipation effects in 2017 before the launch of NY-PFL, nor in any of the placebo tests we conduct in additional years prior. In sum our results shed light on a previously unexplored channel of impact for PFL programmes, one which aligns well with the corresponding rise in fertility observed with CA-PFL (Thunell, 2017; Golightly and Meyerhofer, 2022).

We conclude by remarking that when decreases in abortion rates are observed, it is often in the wake of policies which restrict the rights of women to make their own reproductive choices (Myers, 2021a; Myers, 2021b). Our findings contrast with such settings as the declines we observe result from women who may have previously aborted due to perceived economic necessity now feeling sufficiently secure to pursue their preferred outcome of pregnancy. As such we depart from past contributions by outlining a socially desirable fall in abortion rates, created through the alleviation of the burdens faced by prospective parents. This effect should be most pronounced for younger workers due to their lower average earnings (Lazear, 1981; Van Ours and Stoeldraijer, 2011), and indeed these are the groups for whom we find the largest effects.

2 Policy Background

Mothers in OECD countries are entitled to 50.8 weeks of total paid leave on average, combining maternity and parental home care leave (OECD, 2022). In contrast, U.S. family leave provision consists of only 12 weeks of job-protected *unpaid* leave available through the 1993 Family and Medical Leave Act (FMLA), which as of 2020-21 covered just 56% of employees (Brown et al., 2020). While recent federal legislative efforts have continually failed to expand coverage further (Bullinger, 2019; Schnake-Mahl et al., 2023), the importance of lacking PFL provision has been continually underlined by empirical evidence demonstrating how these

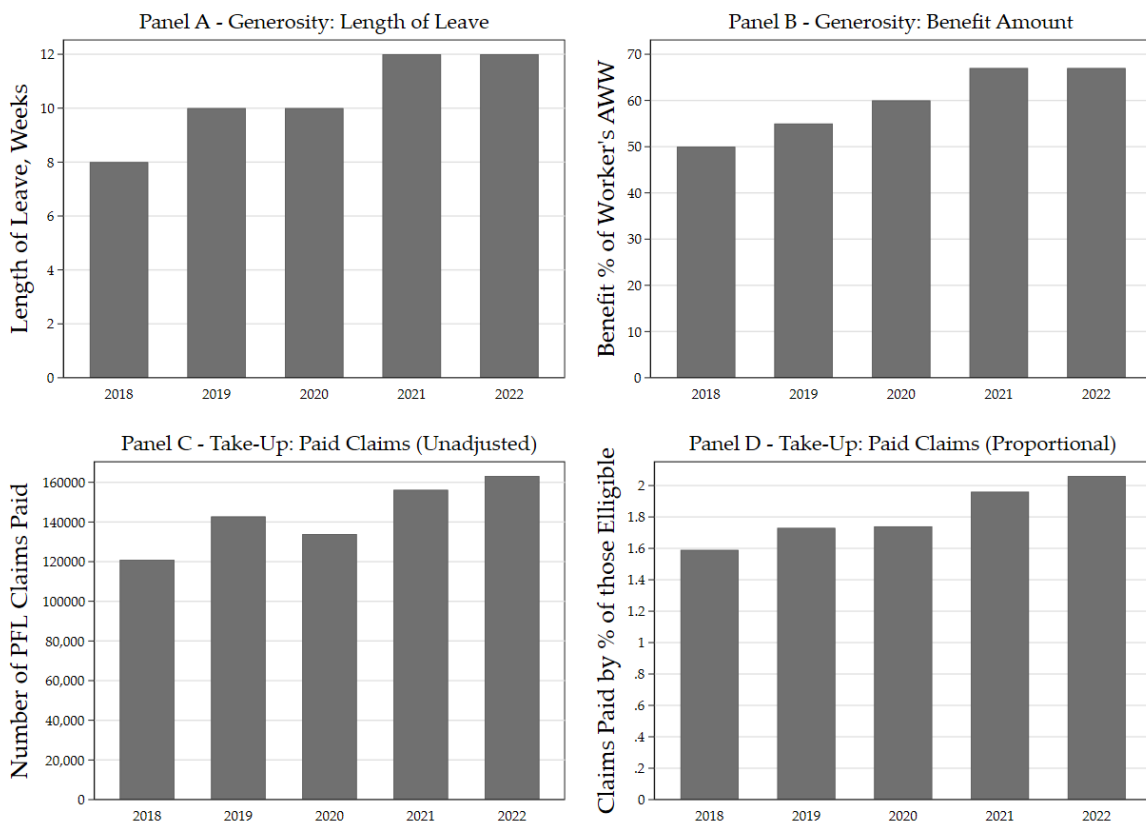
²We discuss throughout Section 4 and Appendix A.2 how our results support this position.

³Abortion data is provided as counts, which we transform into rates per 1,000 women for each subgroup using Bridged-Race (2000-2020) and Single Race (2021, after the discontinuation of the Bridged-Race dataset) population data from the CDC's WONDER platform (accessed 2023/1/4). URL: <https://wonder.cdc.gov/.html>.

programmes can improve the health outcomes of children (e.g. Lichtman-Sadot and Bell Pillay, 2017; Jou et al., 2018; Pihl and Basso, 2019), and increase fertility by reducing the child career ‘cost’ (Lalive and Zweimüller, 2009). Moreover, insofar as Americans have unequal access to paid and unpaid parental leave across states - with some having several supporting programmes and some having none (Winston et al., 2019) - as well as occupations and incomes (Bartel et al., 2023), the lack of national PFL provision also serves to reinforce socioeconomic inequalities.

There are however several specific circumstances under which American parents can qualify for PFL. One of the most common is to live in one of the select states that operated a PFL programme within our sample period (2000-2021), the characteristics of which are compared in Table 5 of Appendix A.1.⁴ NY-PFL was the fourth of these programmes to launch, enacted in 2017 with benefits beginning 2018/1/1. As Figure 1 details, these benefits did not begin at their ultimate values: maximum length of leave was initially capped at 8 weeks, increasing to 10 in 2019/20, before reaching its current and final limit of 12 weeks in 2021 (New York State Department of Financial Services, 2023). Similarly, the maximal weekly benefit amount increased from 50% of the worker’s average weekly wage in 2018, to 55% in 2019, to 60% in 2020, and up to its present cap of 67% in 2021.⁵ As of July 2023, the maximum benefit amount was \$1,131.08 per week (A Better Balance, 2023).

Figure 1: Generosity and Take-Up of NY-PFL Benefits, 2018-2022



Notes: AWW = Average Weekly Wage. Figures sourced from New York State Department of Financial Services (2023).

⁴The other qualifiers are to form part of the 24% of U.S. employees with PFL access through their employer (Bartel et al., 2023) or to live in one of the five states (CA, HI, NJ, NY, RI) that, since 1978, have offered temporary disability insurance (TDI) that covers pregnancy and childbirth (Stearns, 2015). TDI does not factor into our analysis as NY-TDI is constant throughout our sample period (2000-2021), and because the other 4 TDI states are each otherwise ineligible to enter our pool of control states; see Section 3.

⁵Section 5.2 discusses the role of the ramp-up in informing our results.

The coverage and eligibility criteria for inclusion in the programme have remained largely constant through its lifespan, with only some minor increases to eligibility in 2023 (New York State, 2023b). To first consider coverage, the vast majority of private-sector workers are covered under NY-PFL,⁶ while the self-employed can opt-in to the policy (A Better Balance, 2023). Public employers and unions covering public-sector workers can also opt-in to coverage through a collective bargaining process, with domestic workers who work ≥ 20 hours per week for a single employer also covered. Overall, while slightly less inclusive than NY-TDI (Worker's Compensation Board, 2023), NY-PFL still covers approximately 2 million additional employees not covered by FMLA (New York State, 2019). To be eligible to take leave, a worker must have been at their current employer for at least 26 consecutive weeks; those who work less than 20 hours per week must have worked for at least 175 days at their current employer (A Better Balance, 2023). Lastly, families with members in the military can also utilise NY-PFL (New York State, 2023d), although this summed to only 427 claims total from 2018-21 (New York State Department of Financial Services, 2023).

The policy offers both mothers and fathers the opportunity to bond with newborn, adopted, or foster children at any point during their first 12 months with the child, or to care for a sick family member (New York State, 2023b). There is no unpaid waiting period (A Better Balance, 2023), and leave can be taken by both parents simultaneously or separately in either continuous or intermittent increments of at least one full day (New York State, 2023a). Importantly, NY-PFL also offers job protection to employees who take leave, unlike FMLA, which workers cannot claim at the same time as NY-PFL (New York State, 2023c). Take-up of the policy has been strong: the New York State Department of Financial Services (2023) report 120,910 claims made in 2018 (1.60% of the 7,571,312 eligible), increasing up to 163,124 by 2022 (2.06% of 7,936,258). The majority of total claims from 2018-21 were for newborn bonding (71.4%) by female employees (64.9%). Notably, while absolute take-up of NY-PFL benefits declined from 2019 to 2020, proportional take-up increased marginally, suggesting that widespread loss of employment (Collyer et al., 2020) in New York in 2020 (and hence NY-PFL eligibility) caused raw take-up to decline, rather than decreasing engagement with the programme.

3 Data

3.1 Abortion

We employ abortion data sourced from the 'Abortion Surveillance' report series published annually by the CDC.⁷ Beginning in 1969, these reports provide a range of information about the prevalence of abortion in the U.S., breaking down abortions by state, age, ethnicity, procedure, weeks of gestation, and more. Age-disaggregation is of particular interest for our exercise due to the disparity of abortion rates across ages in New York and elsewhere (Kortsmitt et al., 2022). As the CDC only provide age-disaggregated abortion counts however, we employ data from CDC WONDER, which provides by-state-age-gender population counts from

⁶For a list of exceptions see Worker's Compensation Board (2023).

⁷The Guttmacher Institute also provide an established national panel of abortion in the U.S., however as it is not published annually we cannot use it for an SDID exercise.

1990 onwards, to transform counts into rates per 1,000 women for each age group. Rates are calculated for eight age groups for each state: three - 20-39, 20-44, and 25-39 - are composites, alongside the five subgroups of 20-24, 25-29, 30-34, 35-39, and 40-44.⁸ Of these eight, the three composites and the younger age groups of 20-24 and 25-29 are perhaps the most interesting, as these groups are more abortion prone than their older counterparts as clearly shown in Table 1. The upper panel of Table 1 provides descriptive statistics for the abortion outcomes of each group, and reveals that New York has abnormally high abortion rates relative to the control states included in our exercises (particularly for younger women), although the gap declines into the post-treatment period. While the CDC rate additionally report abortions performed on those aged ≤ 15 and 15-19, we exclude these groups⁹ as a significant majority of U.S. teen pregnancies are unintended (Marseille et al., 2018) meaning the profile of reproductive choices faced by teenagers does not well equate to that faced by our adult groups

While the data provided by the Abortion Surveillance series is thorough, state reporting to the CDC is voluntary, which affects our creation of the balanced panel SDID requires. Indeed, ten states¹⁰ have missing year(s) between 2000-2021 and so must be dropped from the donor pool,¹¹ while six additional states (HI, ID, ME, MT, RI, SD) have suppressed (often due to small counts) or missing data for the 40-44 age group specifically and so must be dropped from our 20-44 and 40-44 exercises. We select the period between 2000 and 2021 for several reasons. On the upper end, 2021 is the most recent data published by the CDC. On the lower, state reporting is increasingly poor in the mid-nineties, and 2000 specifically allows for the inclusion of Oklahoma, which has no data before this date. Establishing this cutoff still leaves us with 18 years of pre-treatment data upon which to construct our synthetic approximation of New York, along with four years of post-treatment data. In sum, as we also drop the other six states with PFL programmes in this period, we are left with 32 donors for the 20-44 and 40-44 age groups, and 37 for the other six. It is the need to drop these six additional states that motivates us to consider 20-39 and 25-39 alongside the complete 20-44 distribution; 20-39 may be considered the main result of these three, as it retains all 37 available donor clusters while covering a wider distribution than 25-39, including the most abortion-prone subgroup (20-24). 25-39 then provides an additionally robust estimate, omitting the 20-24 subgroup which has an inferior synthetic match than 25-29, 30-34, or 35-39.

⁸The latter five, along with ≤ 15 and 15-19, are the standard reporting categories used by the CDC (Kortsmit et al., 2022).

⁹Teens are less likely to be treated as many are still completing their education. However as they are not specifically excluded from NY-PFL, they are not fully untreated and so are problematic to use as a placebo test.

¹⁰AK, CA, D.C., FL, MD, NH, NJ, VT, WV, & WY.

¹¹Data may be missing from the surveillance reports but available through the respective state's department of health. For 2005 & 2006 Louisiana (Louisiana Department of Health, 2023), 2009 Delaware (Division of Public Health, 2013), 2012 D.C. (State Center for Health Statistics, 2014), 2012 Maine (Maine Department of Health and Human Services, 2023), 2014 Texas (Texas Department of Health and Human Services, 2023), and 2020 Tennessee (Tennessee Department of Health, 2023) this was the case. For the 40-44 age group, 2006 Arkansas (Arkansas Department of Health, 2007), 2016 Nebraska (Nebraska Department of Health and Human Services, 2017), 2016-2018 and 2021 Delaware (Delaware Health Statistics Center, 2023), 2021 Missouri (Missouri Department of Health and Senior Services, 2023), and 2010, 2011, 2014, 2017, 2018, 2020, & 2021 North Dakota (Division of Vital Records, 2022) were also available. Additionally, remark that the 2017 surveillance report is missing but its tables are provided by Kortsmit et al. (2020).

Table 1: Descriptive Statistics, Outcomes and Covariates

	Pre-NY-PFL: 2000-2017				Post-NY-PFL: 2018-2021			
	New York		Control States		New York		Control States	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Outcome Variables								
<i>Composites</i>								
20-44 Abortion Rate	27.815***	(2.521)	11.071	(4.350)	19.466***	(2.242)	9.752	(4.583)
20-39 Abortion Rate	33.608***	(3.650)	13.255	(5.450)	22.632***	(2.538)	11.093	(5.475)
25-39 Abortion Rate	27.759***	(2.115)	10.523	(4.259)	20.700***	(2.258)	9.866	(5.019)
<i>Subgroups</i>								
20-24 Abortion Rate	51.657***	(9.169)	21.169	(9.519)	28.877***	(3.488)	14.877	(7.061)
25-29 Abortion Rate	40.232***	(4.868)	16.147	(6.799)	27.555***	(2.902)	14.354	(7.318)
30-34 Abortion Rate	27.135***	(2.169)	9.927	(4.079)	20.694***	(2.037)	9.693	(5.006)
35-39 Abortion Rate	15.932***	(0.688)	5.655	(2.256)	13.037***	(1.572)	5.415	(2.746)
40-44 Abortion Rate	5.723***	(0.276)	2.026	(0.922)	5.066***	(0.643)	1.983	(0.992)
Covariates								
<i>Demographic</i>								
Share Female aged 15-44	0.211***	(0.007)	0.202	(0.010)	0.199***	(0.001)	0.192	(0.008)
% Reported as White	54.185***	(2.305)	70.194	(15.627)	50.644***	(0.715)	66.316	(16.286)
% Reported as Black	16.552***	(0.262)	12.376	(10.869)	16.186***	(0.241)	12.501	(10.788)
% Reported as Hispanic	19.432***	(1.317)	10.599	(10.669)	21.132***	(0.101)	12.767	(11.301)
% Masters+ Best Qual.	12.243***	(2.958)	7.396	(3.408)	17.626***	(2.408)	10.854	(4.328)
% Bachelors Best Qual.	19.847	(2.749)	18.771	(4.650)	27.461***	(2.566)	22.870	(5.892)
% Some College Best Qual.	26.370***	(2.299)	29.615	(4.648)	21.943***	(1.889)	27.815	(5.168)
% High School Best Qual.	22.044**	(3.978)	24.247	(5.633)	17.289***	(1.483)	20.824	(5.169)
<i>Economic</i>								
% No Health Insurance	11.976*	(3.414)	13.311	(4.546)	5.017***	(0.363)	8.465	(3.028)
State Unemployment Rate	6.094	(1.520)	5.651	(1.986)	6.250	(2.838)	4.762	(1.851)
State Poverty Rate	14.533***	(1.166)	13.042	(3.337)	11.975	(0.670)	11.400	(2.953)
<i>Legislative</i>								
Medicaid Expanded	0.222	(0.428)	0.110	(0.313)	1***	(0)	0.635	(0.483)
Governor is Democrat	0.611*	(0.502)	0.402	(0.491)	1***	(0)	0.395	(0.490)
Trips to Abortion Provider	0***	(0)	0.805	(0.761)	0***	(0)	0.946	(0.816)

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. ‘Control States’ excludes other treated states and those with imbalanced data. The pre-NY-PFL period contains 684 observations for all variables except the 20-44 and 40-44 abortion rates, while the post-NY-PFL period contains 154 under the same condition. The 20-44 and 40-44 abortion rate variables must drop a further 6 control states, and so use 594 and 132 observations respectively. All abortion rates are per 1,000 women. The earliest date Medicaid could be expanded was 2014/1/1; New York did so then (Kaiser Family Foundation, 2023). **Data Sources:** Education data is sourced from the IPUMS-CPS database (Flood et al., 2023), with the other demographic controls sourced from the Bridged-Race (2000-2020) and Single-Race (2021) population estimates provided by CDC WONDER. Health insurance data is sourced from the U.S. Census Bureau, more specifically HIC series American Community Survey data is used for 2008-2021 (U.S. Census Bureau, 2023a), with HIA series Community Population Survey data used for 2000-2007 (U.S. Census Bureau, 2023b). Unemployment and poverty rates are retrieved from the University of Kentucky Center for Poverty Research (2023). Medicaid expansion times are available from Kaiser Family Foundation (2023), governor party from the University of Kentucky Center for Poverty Research, and abortion visit requirements from Myers (2021a).

3.2 Covariates

We include 13 covariates in our SDID regressions, encompassing demographic, economic, and legislative variables.¹² Demographic controls include the percentage of the state population which is female and

¹²Note that while the applied examples of Arkhangelsky et al. (2021) do not include any covariate controls, the authors discuss how the SDID estimator can be naturally expanded to allow for their inclusion. Dealt with as a bias-removing pre-processing task (Clarke et al., 2023), Appendix A.3 underlines the benefit of controlling for our covariates by demonstrating that their omission worsens our pre-treatment matching, with no effect on unit or time weighting.

of reproductive age (15-44),¹³ the percentages of the state’s female population of reproductive age that is identified as (non-Hispanic) White,¹⁴ (non-Hispanic) Black, and Hispanic,¹⁵ and then four variables capturing the proportion of the state’s female population of reproductive age with a masters, a bachelors, some college, or a high-school degree as their highest educational attainment.¹⁶ Our economic covariates capture the percentage of the state population without health insurance, the state unemployment rate, and the state poverty rate. Finally we include three legislative controls, namely whether or not Medicaid eligibility had been expanded in the state post-2014 under the Affordable Care Act (see Congressional Research Service, 2021), whether the state governor was a democrat or not,¹⁷ and how many trips were required to a abortion provider before an abortion could be performed.¹⁸ The lower panel of Table 1 details that New York is on average better educated than the average of our control states, and has a distinct ethnic and legislative profile. As such, assuming parallel trends between the two without constructing a synthetic control unit, as in our method, may be unreasonable.

4 Empirical Strategy

4.1 Main Estimation

To examine the causal effect of NY-PFL on age-disaggregated abortion rates in New York, we employ the SDID method first proposed by Arkhangelsky et al. (2021). This method provides accurate estimation within a small sample comparative study like ours with non-random treatment assignment (Clarke et al., 2023). We employ SDID over the more established Difference-in-Differences (DID) and Synthetic Control methods as Arkhangelsky et al. (2021) demonstrate how SDID can outperform both in situations wherein each would be respectively favoured conventionally by combining attractive features of both, while preserving desirable asymptotic properties including efficiency.

In providing an overview of the SDID procedure, we draw on the introductions provided by Arkhangelsky et al. (2021) and Clarke et al. (2023). To begin, the SDID estimator requires a balanced panel of N units observed over T periods. An outcome, Y_{it} , is observed for each unit i in each time period t . Our interest is in the effect of a binary treatment, W_{it} , which some ($W_{it} = 1$) but not all ($W_{it} = 0$) units receive. Denote the treated unit(s) as N_{tr} and the never-treated control units as N_{co} , such that $N = N_{tr} + N_{co}$; in our setting only one unit, New York, is treated ($N_{tr} = 1$) by NY-PFL (W_{it}).¹⁹ We additionally require that treatment is irreversible, that there are no always treated units, and that there at least two pre-treatment periods. As New

¹³As defined by the CDC.

¹⁴Ethnicity and fertility data are both provided as counts, and so are transformed into rates using WONDER’s population data.

¹⁵From among five main ethnic groups, we include only White as the most populous, Black as the group with the highest abortion rate (Kortsmit et al., 2022), and Hispanic as the more fertile (Golightly and Meyerhofer, 2022).

¹⁶Less than High School is the omitted reference group.

¹⁷We use the mayor of D.C. as a proxy for governor, although as a treated state D.C. is dropped regardless.

¹⁸Either zero, one, or two trips, following Myers (2021a).

¹⁹The SDID estimator can also handle $N_{tr} > 1$, as well as staggered adoption (Arkhangelsky et al., 2021).

York becomes treated eighteen years into our 2000-2021 sample period, we have $T_{pre} = 18$ and $T_{post} = 4$, with $N_{co} = 37$ for six of our eight age groups.²⁰

The purpose of the SDID estimator is to consistently estimate the causal effect of receiving treatment on units who are treated, referred to as the Average effect of Treatment on the Treated (ATT). Importantly, the SDID estimator can do so if the parallel trends assumption between N_{tr} and N_{co} is implausible on average (Clarke et al., 2023), unlike a conventional DID. An ATT estimate, $\hat{\tau}^{sdid}$, is then retrieved using a two-way fixed effects regression:

$$\left(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta} \right) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (1)$$

Where β_t are time fixed-effects, α_i are unit fixed-effects,²¹ and $\hat{\omega}_i^{sdid}$ and $\hat{\lambda}_t^{sdid}$ are optimally chosen unit and time weights respectively. The purpose of $\hat{\omega}_i^{sdid}$ in (1) is similar to their purpose in the Synthetic Control estimator, namely to use a weighted average of the donor pool to construct a synthetic approximation of N_{tr} with as close to a parallel trend to N_{tr} as possible across the pre-treatment period. The additional inclusion of unit fixed effects α_i meanwhile is similar to the DID method, and allows the SDID estimator to match N_{tr} and N_{co} on pre-treatment trends only rather than pre-treatment trends and levels like in a Synthetic Control exercise. This grants the SDID estimator greater flexibility than the Synthetic Control method to work with outliers outside the convex hull of the donor pool. Fully unique to the SDID however is the inclusion of time weights $\hat{\lambda}_t^{sdid}$, which assign larger weights to pre-treatment periods with outcomes closer to - and hence more representative of - those post-treatment, in order to prevent periods with inferior matching from biasing our estimate $\hat{\tau}^{sdid}$.²²

We find our unit weights $\hat{\omega}_i^{sdid}$ by solving:

$$\left(\hat{\omega}_0, \hat{\omega}^{sdid} \right) = \arg \min_{\omega_0 \in \mathbb{R}, \omega \in \Omega} \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N Y_{it} \right)^2 + \zeta^2 T_{pre} \|\omega\|_2^2 \quad (2)$$

$$\text{where } \Omega = \left\{ \omega \in \mathbb{R}_+^N, \text{ with } \sum_{i=1}^{N_{co}} \omega_i = 1 \text{ and } \omega_i = \frac{1}{N_{tr}} \text{ for all } i = N_{co} + 1, \dots, N \right\}$$

and \mathbb{R}_+ denotes the positive real line. ζ is a regularisation parameter that increases the dispersion of weights across a larger set of control units relative to those assigned by the canonical Synthetic Control method (Arkhangelsky et al., 2021),²³ and is defined as:

$$\zeta = (N_{tr} T_{post})^{1/4} \hat{\sigma} \quad \text{with} \quad \hat{\sigma}^2 = \frac{1}{N_{co}(T_{pre} - 1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} (\Delta_{it} - \bar{\Delta})^2, \quad (3)$$

²⁰With $N_{co} = 32$ for the other two, as discussed in Section 3.1.

²¹One α_i and β_t are normalised to zero such to avoid perfect multicollinearity (Clarke et al., 2023).

²²Appendix A.2 compares the matching and weights each method produces for our 20-39 age group.

²³This consideration can be made as we are not required to construct an identical trend, only a parallel trend (Clarke et al., 2023). This is made possible by the inclusion of an intercept, ω_0 , itself possible due to the inclusion of α_i (Melo et al., 2023).

$$\text{where } \Delta_{it} = Y_{i(t+1)} - Y_{it} \quad \text{and} \quad \bar{\Delta} = \frac{1}{N_{co}(T_{pre} - 1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} \Delta_{it}$$

Our time weights $\hat{\lambda}_t^{sdid}$ are then selected by solving a similar optimisation problem where the regularisation penalty ζ is omitted (Arkhangelsky et al., 2021):

$$\left(\hat{\lambda}_0, \hat{\lambda}^{sdid} \right) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\operatorname{argmin}} \sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T Y_{it} \right)^2 \quad (4)$$

$$\text{where } \Lambda = \left\{ \lambda \in \mathbb{R}_+^T, \text{ with } \sum_{t=1}^{T_{pre}} \lambda_t = 1 \text{ and } \lambda_t = \frac{1}{T_{post}} \text{ for all } t = T_{pre} + 1, \dots, T \right\}$$

4.2 Event Study Exercise

Alongside our main exercise, we additionally provide event-study style output which dissects the impact of NY-PFL on abortion rates into annual point estimates. This exercise, proposed by Clarke et al. (2023) and first applied empirically by Clifton-Sprigg et al. (2023), allows us to better evaluate the validity of our parallel trends assumption between the treated and synthetic units relative to default SDID output. Moreover, it allows us to quantify the dynamics of treatment effects that may be obscured within the point estimates our static SDID exercise produces. This extension functions by calculating the difference in outcomes between the treated and synthetic units relative to their baseline pre-treatment means for each period, in kind with a regular DID event study:

$$\left(\hat{Y}_t^{Tr} - \hat{Y}_t^{Co} \right) - \left(\bar{Y}_{baseline}^{Tr} - \bar{Y}_{baseline}^{Co} \right) \quad (5)$$

However in contrast to a DID event study, where an arbitrary period is chosen as the baseline period,²⁴ the baseline pre-treatment means $\bar{Y}_{baseline}^{Tr}$, $\bar{Y}_{baseline}^{Co}$ in an SDID event study are chosen optimally according to our unit weights:

$$\bar{Y}_{baseline}^{Tr} = \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} \bar{Y}_t^{Tr}, \quad \bar{Y}_{baseline}^{Co} = \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} \bar{Y}_t^{Co} \quad (6)$$

Following Clarke et al. (2023), we perform this exercise by calculating (5), the set of point estimates for each period t , from the results of 1,000 stored repetitions of our main SDID exercise with placebo inference. Confidence intervals are then constructed for these estimates using the variance of 500 block-bootstrapped resamples, which re-calculate (5) from each t by replacing our observations with state clusters.²⁵

²⁴Most commonly the period immediately before treatment (Schmidheiny and Siegloch, 2019).

²⁵This is the same block bootstrap process as is conducted internally within the estimation of our main SDID estimates (Arkhangelsky et al., 2021), drawn out such that it can be dynamically plotted (Clarke et al., 2023).

4.3 Inference

We must make two choices in order to infer from our estimation: which form of variance estimation to employ, and how to incorporate our covariates. Although three options are proposed by Arkhangelsky et al. (2021) for variance estimation, namely bootstrap, jackknife, and placebo inference, only the latter is possible in our context where $N_{tr} = 1$. This procedure is similar to that used in the Synthetic Control exercise (Melo et al., 2023), and operates by assigning the treatment structure of N_{tr} onto the set of untreated control units. We assign this placebo treatment over 1,000 resampling repetitions, collect the vector of estimates $\hat{\tau}_{(p)}^{sdid}$, and calculate its variance $\hat{V}_\tau^{placebo}$, from which we compute our confidence intervals.²⁶

Covariates are then incorporated into our estimation using the projected optimisation procedure proposed by Kranz (2022),²⁷ rather than ‘sdid’ STATA package’s default use of the optimised Hirshberg method (Clarke et al., 2023). Kranz (2022) demonstrates that this method can outperform the Hirshberg method in settings with heterogeneously time-variant relationships between the outcome and covariates, as is likely the case in our setting with abortion rates and state unemployment and poverty rates.²⁸

5 Results

5.1 Main Results

Figures 2 and 3 provide the results of our SDID matching for our eight age groups: in solid grey is the trend of our treated state, New York, while the trend of the synthetic counterfactual, which we construct by solving (3), (2), (4), and (1) is shown in dashed grey. The optimal time weights we compute to solve (4) are shown at the bottom of each panel, while in solid red is the time of reform, placed at 2017 as NY-PFL benefits were available from 2018/1/1, making 2018 a treated year. Overall, we achieve close matches in trends between New York and the synthetic control group across all panels, particularly for the middle age groups of 25-29, 30-34, and 35-39. With the general decline in U.S. abortion rates since their peak in 1990 (R. Jones et al., 2022), the periods immediately before 2017 have outcomes more representative of those post-treatment, explaining why these periods are weighted so highly for most groups.²⁹ Furthermore, Figures 2 and 3 not only show how New York’s abortion rates fell after 2017, but also that abortion rates in our synthetic control picked up slightly in the post-treatment period (see R. Jones et al., 2022).

Next, Table 2 provides the main SDID ATT estimates of our exercise for each of our eight age groups. As our main result, we find that abortions fell after the introduction of NY-PFL in 2018 by 4.583 abortions per

²⁶Clarke et al. (2023) notes that the assumption of homoskedasticity across units is required for this procedure as $\hat{V}_\tau^{placebo}$ is calculated only from the set of control units N_{co} , not N_{tr} .

²⁷Covariate adjustment is implemented as a pre-processing task in an SDID exercise, absorbing their impact before constructing the synthetic unit. This contrasts sharply to the canonical Synthetic Control method, where covariates are used as vectors in the RMSPE minimisation matrix upon which unit weights are assigned (Arkhangelsky et al., 2021).

²⁸Projected errors carry the additional benefit of calculating much faster than optimised errors.

²⁹Figures 8 and 9 of Appendix A.2 detail the optimal unit weights assigned to each control state in each regression, while Figure 10 illustrates how unit weights vary between the SDID, DID, and Synthetic Control methods for our main 20-39 regression.

Figure 2: SDID Matching, First Four Age Groups, Abortion Rates, New York

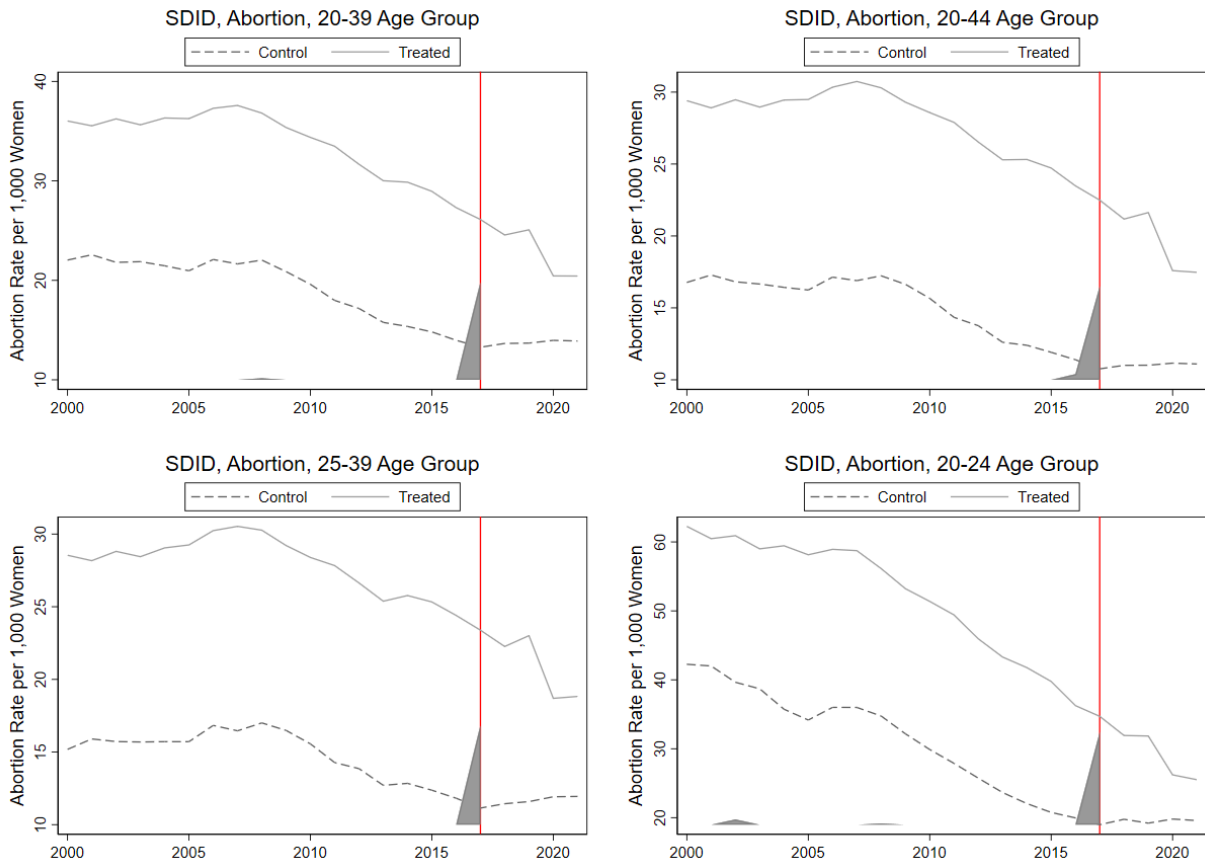
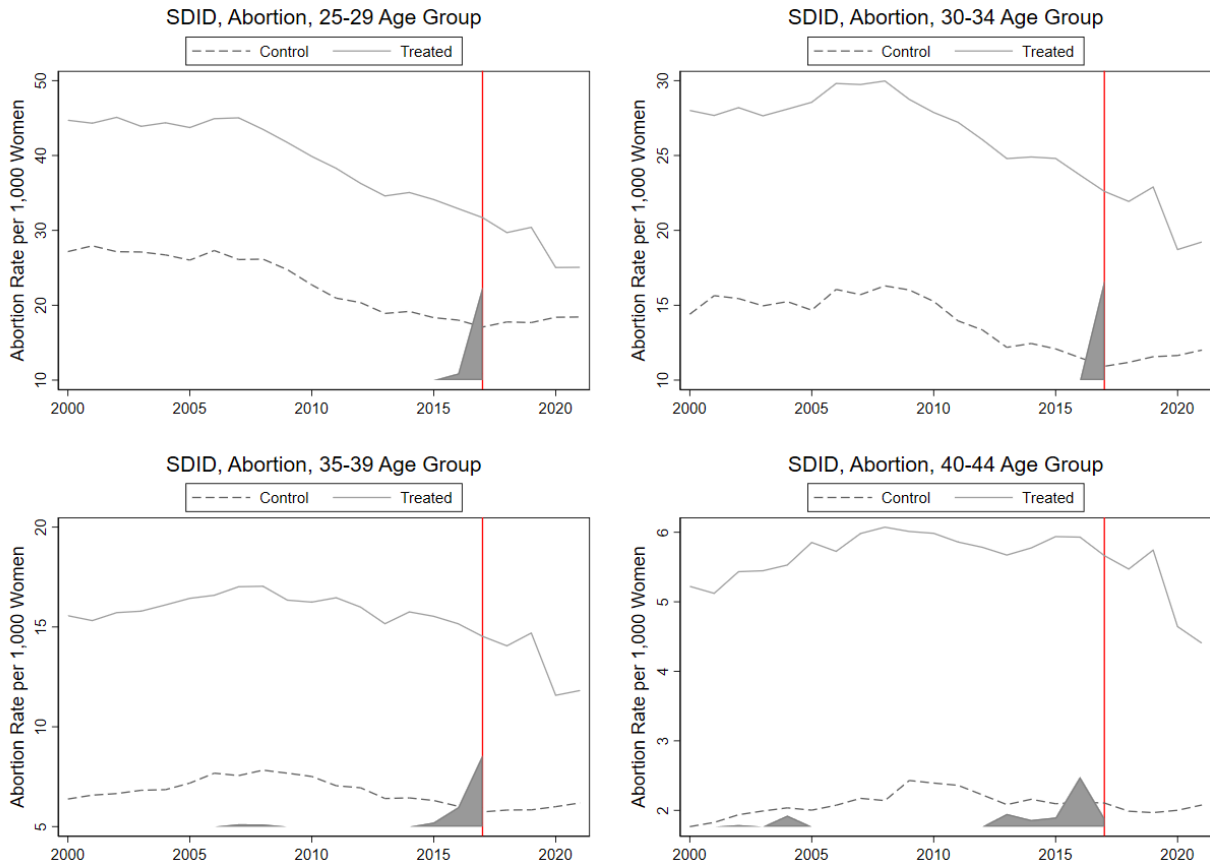


Figure 3: SDID Matching, Final Four Age Groups, Abortion Rates, New York



1,000 women for our primary 20-39 age group, a 13.6% reduction from the pre-treatment mean of 2000-2017, relative to our synthetic control. This result is highly significant, as are our estimates for several other age groups, although their magnitudes decline linearly with age. Overall, our estimated effects are consistently large, with the highest effect for the 20-24 age group (14.9%). The cause for these magnitudes is the sharp secondary decline in abortion rates for all groups in 2020, as illustrated by Figures 2 and 3. This effect is as large as the initial policy effect in 2018 following the initial launch of NY-PFL, and is likely driven by the increasing generosity and take-up of NY-PFL benefits between the programmes launch and its full realisation in 2021, as illustrated by Figure 1. Part of this decline may also be attributed to the impact of COVID-19 on abortion take-up (Wolfe and Meulen Rodgers, 2021). Yet, while it is difficult to fully rule out that COVID effects help drive our results, the DID aspect of our SDID estimation accounts for the fact that declines in abortion take-up occurred across the U.S. in 2020, not just New York.³⁰

Table 2: SDID Results, Age-Disaggregated Abortion Rates per 1,000 women, 2000-2021

	(1) 20-39 yr.	(2) 20-44 yr.	(3) 25-39 yr.	(4) 20-24 yr.	(5) 25-29 yr.	(6) 30-34 yr.	(7) 35-39 yr.	(8) 40-44 yr.
NY-PFL Active	-4.583***	-3.660***	-3.408***	-7.718***	-5.364***	-2.823**	-1.986***	-0.673
<i>S.E.</i>	(1.420)	(1.220)	(1.287)	(1.961)	(1.894)	(1.403)	(0.741)	(0.757)
<i>N</i>	836	726	836	836	836	836	836	726
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	33.608	27.815	27.759	51.657	40.232	27.135	15.932	5.723
Pre-Trt. Mean, Donors	19.189	15.054	14.740	31.134	23.446	14.232	6.868	2.107
Sample Outcome S.D.	6.218	5.095	5.024	10.420	7.696	4.887	2.769	1.097

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Payroll deductions began 2017/7/1, benefits become available 2018/1/1. *N* is the number of state-year-age group abortion rate observations. 'Clusters' refers to the number of states in each regression and includes the single treated state. Estimates are drawn from a restricted time sample from 2000 to 2021, and are based off 1000 placebo repetitions of this sample. 'Trt.' = Treatment. 'S.D.' = Standard Deviation of outcome for all clusters.

Moreover, as Section 5.2 addresses and Figure 7 of Appendix A.2 illustrates, while there was some pandemic-related disruption to abortion provision in New York across 2020/1, this disruption was felt similarly in many other states around this time (R. Jones et al., 2022). To further explore the robustness of our results, Section 5.4 performs a number of alternative and placebo specifications, and re-runs our main regression excluding 2019, 2020, and 2021 respectively from the sample.

5.2 Event Studies

Next we complement our main estimates of Table 2 by providing dynamic, event-study style output in Figures 4 and 5 which dissect our static effects into annual point estimates. Each point of these panels expresses the difference in trend between the treated and synthetic units of Figures 2 and 3, irrespective of their difference in levels. Overall, while we do generally observe a significant treatment effect in 2018 that then increases substantially in 2020, there are significant pre-trends in the pre-treatment periods for several of our figures.

³⁰Also see Figure 6.

Figure 4: Dynamic SDID Estimation, First Four Age Groups, Abortion Rates, New York

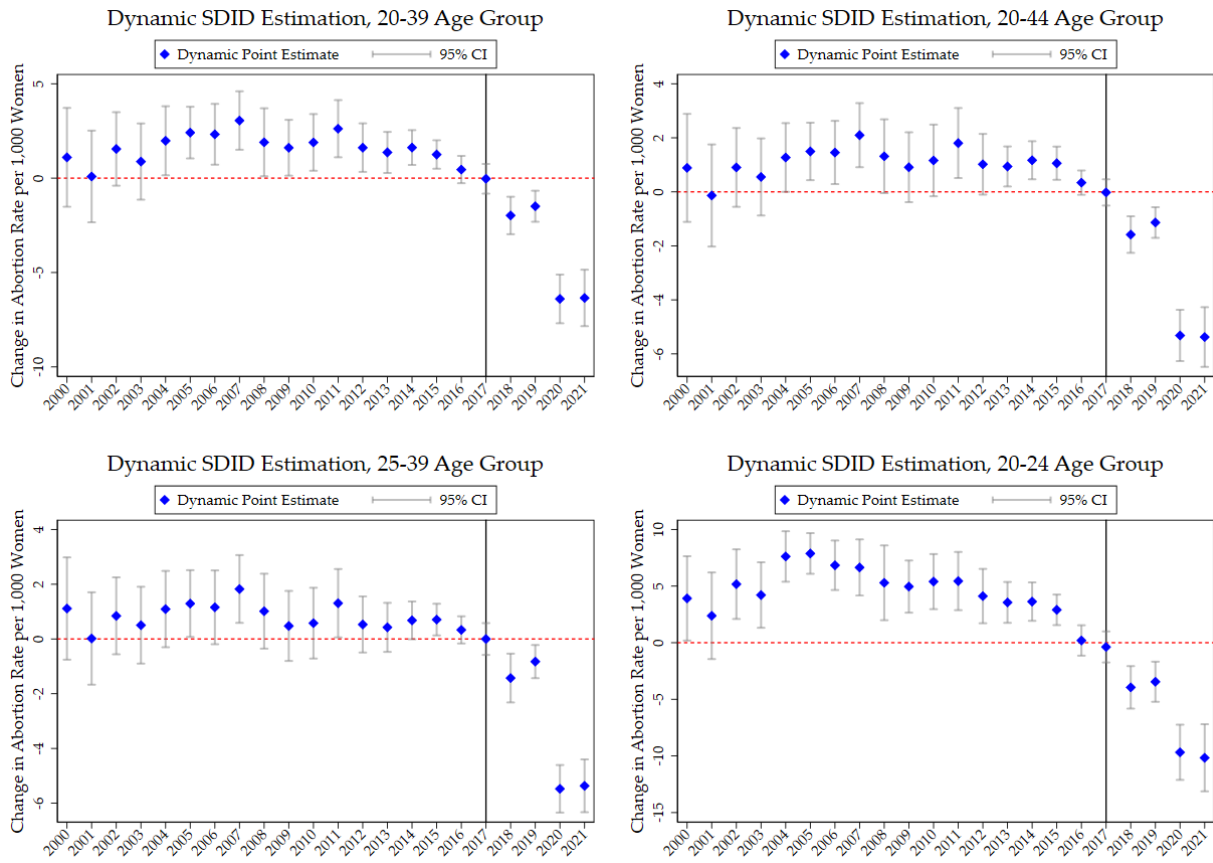
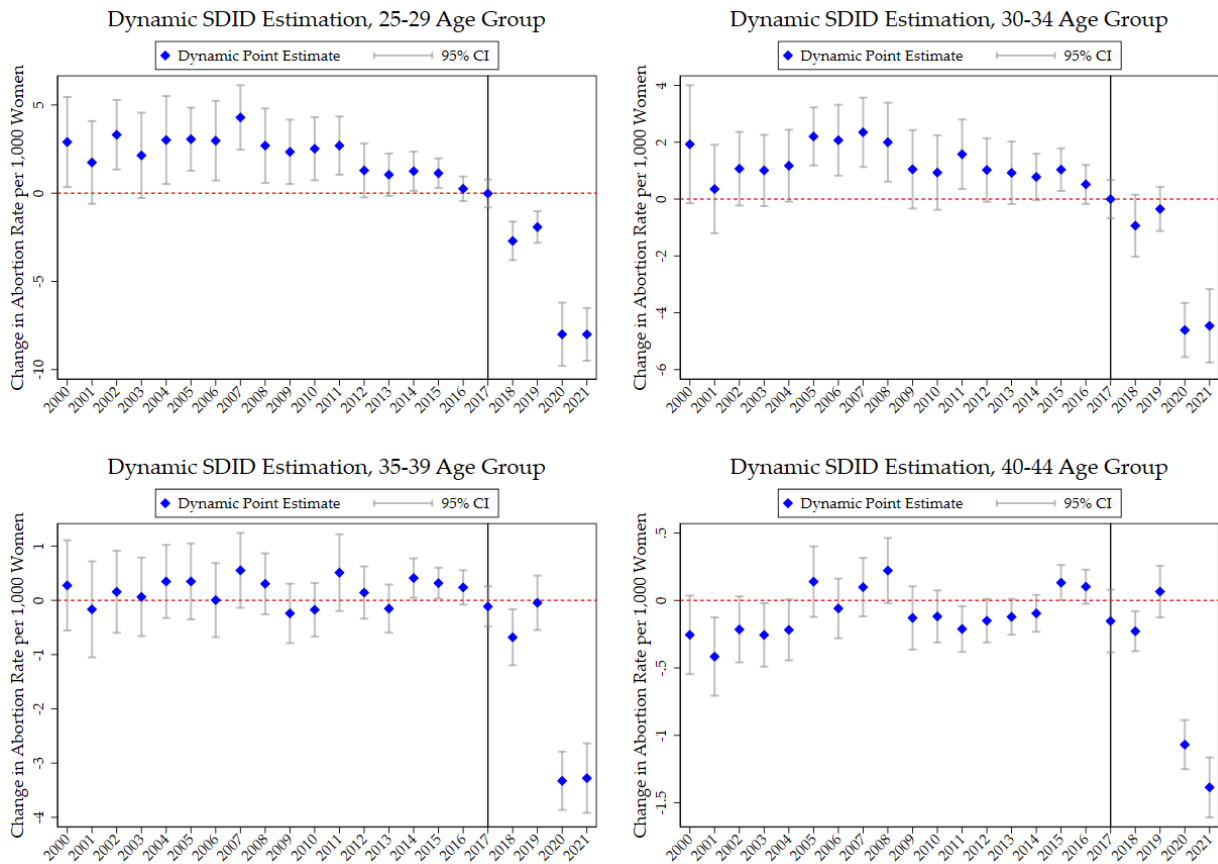


Figure 5: Dynamic SDID Estimation, Final Four Age Groups, Abortion Rates, New York



This issue is particularly prevalent before the great recession³¹ and for the 20-24 age group,³² however the former is no issue as the concentration of our time weights around the time of treatment ensures that subpar matching before 2010 does not affect our estimation. Overall, we contend that the seven groups other than 20-24 provide robust results, as while many pre-treatment point estimates lie above zero, few are statistically significant, particularly in the decade before treatment in 2018. For our main 20-39 age-group, point estimates switch from consistently above zero to significantly below it in 2018, a trait shared by the 20-44, 25-29, and 25-39 age groups. The latter of these groups contains almost zero significant pre-trend, and hence provides the most robust evidence of the impact of NY-PFL in reducing abortion rates in 2018.

Interestingly an uptick in abortion rates is observed for all groups in 2019, reversing some of the initial treatment effect and defying the subsequent decline in 2020. This is very likely caused by New York enacting two abortion-affecting policies in 2019: the Reproductive Health Act, which legally reclassified and increased access to abortion (see National Institute for Reproductive Health, 2022), and the Comprehensive Contraception Coverage Act, which aimed to reduce the number of unintended pregnancies and increase access to health insurance (New York Senate, 2019). The presence of these policies makes it difficult to distinguish whether the increasing generosity of NY-PFL benefits as the programme entered its second year influenced our observed effects.

All panels of Figures 4 and 5 then display a dramatic deviation in 2020, where the treatment effects approximately double before plateauing in 2021. These decreases feed into our main point estimates of Table 2, explaining their large magnitude. While these declines may immediately seem like a natural consequence of COVID-related disruption to abortion provision in New York in 2020, the DID aspect of our SDID method will absorb this decline as long as the disruption New York experienced was also felt in our donors. Indeed Figure 7 of Appendix A.1, using data from the Guttmacher institute (R. Jones et al., 2022),³³ reveals that while the number of abortion clinics in New York declined by approximately 8% in 2020 due to COVID restrictions (Zraick and Meko, 2022), this trend was repeated across many states in 2020 as the number of clinics fell by 11% in Florida and Massachusetts, 12% in Tennessee, and 25% in Delaware (R. Jones et al., 2022). Moreover, compared to the eight states assigned weight in our main SDID estimation,³⁴ New York was not alone in experiencing a decline in abortion clinics between 2017 and 2020 (113 to 104), as Delaware saw a decline from 4 to 3. With the number of clinics in Kansas and New Mexico also unchanged over this period, it is unlikely that changes to abortion provision explain the sharp 2020 falls we observe in Figures 4 and 5.

We instead interpret our 2020 and 2021 effects to result from the increasingly generosity and take-up of NY-PFL benefits throughout the post-treatment period, as detailed by Figure 1. While the presence of the

³¹We suspect that the strong reproductive shock the U.S. experienced at this time, with fertility rates falling significantly (Schneider, 2015), may explain why this is the case.

³²This group has a uniquely high abortion rate that also follows a uniquely steady trend around 2010 relative to any of our donors. The presence of New York City, which has a very high abortion rate and large/young population (U.S. Census Bureau, 2022) relative to many other units, may produce this speciality. Matching this trend is then difficult as two of the closest equivalents to NYC - Los Angeles (2nd largest city) and D.C. (very high abortion rate) - are treated and hence must be dropped.

³³As data on this variable is not provided annually between 2000 and 2021 we cannot include it as a covariate.

³⁴Displayed by Figures 8 and 9 of Appendix A.2.

pandemic and the two 2019 policies means we cannot observe the effect of NY-PFL's progression directly, it is worth remembering that the initial 2018 effect³⁵ stems from only the least generous version of the programme, with 8 weeks of leave and 50% of the worker's Average Weekly Wage (AWW) as benefits. We argue that by 2021, with 12 weeks of leave and 67% AWW benefits, the ability of NY-PFL to reduce abortion rates through income mechanisms would have increased markedly.

5.3 Mechanisms

To best explore the income mechanism we propose would require long-running, national survey data on the fertility intentions of U.S. women, and more specifically whether PFL or any other public assistance programme played a part towards their feelings of financial security in having a child or terminating a pregnancy. Unfortunately, we are unaware of any survey that could fulfill this function within our sample period.³⁶ We can however consider our income mechanism within the wider literature of works considering the relationship between welfare and seemingly non-economic decisions, within which abortion has been previously considered (Hussey, 2011; Ressler et al., 2022). Unlike fertility, where the links between various welfare programmes³⁷ and fertility are well established (e.g. Luci-Greulich and Thévenon, 2013; Raute, 2019; W. Kim, 2023), evidence on the welfare effects on abortion is mixed. While economic theorists have posited that public assistance policies like PFL should reduce abortion rates by alleviating some of the costs involved in having children (Medoff, 1988; Gohmann and Ohsfeldt, 1993), authors such as Kelly and Grant (2007) find very limited effects on the abortion rates of both teenage and adult women following the 1996 introduction of the Temporary Assistance for Needy Families (TANF) welfare programme in the U.S.

The most complete platform upon which to discuss the income effects of welfare provision like NY-PFL on abortion rates however is the works of Hussey (2010; 2011). To begin, Hussey (2010) considers the effects of several welfare measures including TANF, Medicaid, and most notably the introduction of FMLA in 1993. While no significant effects are found for the former two programmes, she finds that states which introduced unpaid family leave through FMLA saw statistically significant reductions in abortion rates (Hussey, 2010). This result corroborates our findings and income mechanism, as if a significant reduction in abortion can be observed following the introduction of *unpaid* family leave, it is likely that the launch of PFL programmes like NY-PFL can produce at least a similar effect. Expanding on this first conclusion, Hussey (2011) reconsiders the introduction of TANF and its effect on abortion, finding that there were reductions in abortion rates but only in states with restrictive, pro-life legislative attitudes towards abortion. Our results argue that NY-PFL defied this trend by affecting abortion in a liberal state like New York, although we note our exercise considers a different welfare measure than Hussey (2011), and only one treated state.

³⁵Remark that to the best of our knowledge there were no systematic influences or changes to abortion provision in New York in either 2017 (Longbons, 2019) or 2018 (Longbons, 2021). Additionally, Rice et al. (2022) details that from 2006-2018, the only one of 20 expansive and 3 restrictive contraceptive policies in change in New York was the rollout of Medicaid, which we already control for.

³⁶Perhaps closest is the American Community Survey, which records whether female respondents reported giving birth in the last twelve months and whether they had received any public assistance in the same time period. However this only covers fertility outcomes, not intentions, and does not isolate PFL as a factor.

³⁷Notably including CA-PFL (Thunell, 2017; Golightly and Meyerhofer, 2022)

There are other additional elements of the income effect of welfare on abortion rates addressed in the existing literature that also relate to our findings. In contrast to our study, which considers the totality of abortions in a state across the whole income distribution, Hussey (2010; 2011) and Ressler et al. (2022) focus solely on low-income women. Their focus is a product of considering TANF - only available to those on low-incomes - but may suggest that our income mechanism is likely most relevant to women on lower incomes for whom NY-PFL benefits represent a higher portion of their incomes.³⁸ Future research could explore further whether this is true for NY-PFL, or the other state PFL programmes active today. Moreover, future work may also investigate whether statewide attitudes towards reproductive choices can interact with the income effect of PFL on abortion, as Hussey (2011) finds with TANF, or whether the launch of a state PFL programme can have symbolic importance towards the reproductive climate in a state (Gauthier, 2007; Hussey, 2011) alongside the income effect outlined here.

5.4 Sensitivity

To reinforce our findings from Sections 5.1 and 5.2 we now conduct an array of sensitivity analysis, before discussing other relevant factors in Section 6. As a first set of robustness tests, we compare our main results of Table 2 to those of Table 3, where we restrict the time period to end at 2020, 2019, and 2018 respectively. The results of Panel A, which restrict the sample to exclude 2021, are very similar to that of Table 2. This lends further credence to the belief that the increased effect we observe in 2020 (see Figures 4 and 5) cannot result from pandemic-related disruption to abortion provision, as New York's experience of the pandemic in 2021 was less acute relative to the rest of the country compared to 2020 (Centers for Disease Control and Prevention, 2023).

For Panels B and C, the result for our main 20-39 age group remains significant at 5% level in both, demonstrating that the effect of NY-PFL is already significant when considering only the first two post-treatment periods of 2018 and 2019. In addition, our estimates for the 25-29 age group, perhaps the second most policy relevant subgroup after 20-24, also remain large (6.4%, 7.3%) and significant in both latter panels despite the exclusion of 2020 and 2021. Note however that the effects of several other age groups are now imprecisely estimated in Panels B and C, although they remain qualitatively similar. The magnitude of these estimates also falls noticeably from those of Panel A and Table 2, with the decrease in abortion rates for our 20-39 age group now estimated at 7.1% in 2018 rather than the previous 13.6% reduction, likely due to the reduced take-up and generosity of NY-PFL benefits in their first two years (New York State Department of Financial Services, 2023).

Next, Table 4 repeats our main analysis of Table 2 using a number of alternative treatment dates, namely 2017, 2014, 2012, and 2010; the former tests for anticipation effects - as payroll deductions began up to six months before benefits were available (New York State, 2019) - while the latter three conduct additional

³⁸Discussed in more detail in Section 6.

Table 3: SDID Results, Age-Disaggregated Abortion Rates per 1,000 women, Restricted Time Samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	20-39 yr.	20-44 yr.	25-39 yr.	20-24 yr.	25-29 yr.	30-34 yr.	35-39 yr.	40-44 yr.
Panel A: 2000-2020								
NY-PFL Active	-3.897***	-2.953***	-2.739**	-6.841***	-4.505**	-2.223*	-1.425**	-0.380
<i>S.E.</i>	(1.149)	(1.085)	(1.107)	(1.805)	(1.911)	(1.314)	(0.663)	(0.711)
<i>N</i>	798	693	798	798	798	798	798	693
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	33.608	27.815	27.759	51.657	40.232	27.135	15.932	5.723
Pre-Trt. Mean, Donors	19.381	15.049	14.847	31.374	23.689	14.490	6.921	2.125
Sample Outcome S.D.	6.242	5.114	5.022	10.500	7.710	4.874	2.768	1.101
Panel B: 2000-2019								
NY-PFL Active	-2.254**	-1.468*	-1.226	-4.850***	-2.588*	-0.838	-0.277	-0.0375
<i>S.E.</i>	(1.004)	(0.858)	(0.845)	(1.481)	(1.333)	(0.937)	(0.616)	(0.624)
<i>N</i>	760	660	760	760	760	760	760	660
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	33.608	27.815	27.759	51.657	40.232	27.135	15.932	5.723
Pre-Trt. Mean, Donors	19.347	14.967	14.757	31.200	23.638	14.540	6.851	2.123
Sample Outcome S.D.	6.257	5.124	5.017	10.560	7.716	4.866	2.768	1.105
Panel C: 2000-2018								
NY-PFL Active	-2.386**	-1.670*	-1.500	-4.898***	-2.943**	-1.081	-0.630	-0.175
<i>S.E.</i>	(0.999)	(0.949)	(0.920)	(1.442)	(1.265)	(1.032)	(0.788)	(0.456)
<i>N</i>	722	627	722	722	722	722	722	627
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	33.608	27.815	27.759	51.657	40.232	27.135	15.932	5.723
Pre-Trt. Mean, Donors	19.492	15.082	14.872	31.765	23.831	14.654	6.952	2.102
Sample Outcome S.D.	6.283	5.140	5.024	10.620	7.743	4.870	2.768	1.107

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Payroll deductions began 2017/7/1, benefits become available 2018/1/1. *N* is the number of state-year-age group abortion rate observations. 'Clusters' refers to the number of states in each regression and includes the single treated state. Estimates are drawn from restricted time samples from 2000 to 2020/2019/2018, and are based off 1000 placebo repetitions of each sample. 'Trt.' = Treatment. 'S.D.' = Standard Deviation of outcome for all clusters.

placebo tests.³⁹ The four panels display that, other than some marginally significant effects at 10% in 2012,⁴⁰ there are no significant effects associated with any of our placebo treatments. The sparse significance of our estimates here contrasts sharply with that observed in Table 3, where even a single post-treatment period reveals a significant effect for several age groups. Perhaps most interesting in Table 4 are the results of Panel D, which confirm that even though payroll deductions began for some employees as early as 2017/7/1, there was no statistically significant change in abortion rates until benefits began in 2018. Our estimates in Panels F and G meanwhile are sizable, however it is highly probable that their magnitude derives from the wider contemporary reversion of New York's uniquely high abortion rates before the financial crisis to rates closer to the national average, rather than any particular event. In sum, Tables 3 and 4 give us greater confidence that our main results of Table 2 and Figures 4 and 5 are unlikely to be driven by spurious factors.

As further robustness tests, we first demonstrate in Appendix A.3 how results become marginally worse when our covariates are excluded from the main SDID exercise, before discussing the alternative use of optimised errors over our projected errors. Appendix A.4 then generates the results of placebo treatments

³⁹In all four regressions the true post-treatment period (2018-2021) is omitted in order to avoid confounding.

⁴⁰Figures 4 and 5 offer strong evidence that these result from an outlier uptick in abortion rates in 2011, with a return to baseline in 2012. As such they are not indicative of any factor acting to substantially decrease abortion rates as NY-PFL does after 2017.

in 2018 for each control state to compare against our New York result⁴¹ and demonstrates that the effect estimated for New York is unique among the distributions, amongst other additional considerations.

Table 4: SDID Results, Age-Disaggregated Abortion Rates per 1,000 women, Placebo Treatments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	20-39 yr.	20-44 yr.	25-39 yr.	20-24 yr.	25-29 yr.	30-34 yr.	35-39 yr.	40-44 yr.
Panel D: 2017 Treatment								
NY-PFL Active	-0.921	-0.719	-0.671	-1.457	-0.695	-0.808	-0.575	-0.377
<i>S.E.</i>	(0.765)	(0.804)	(0.723)	(1.269)	(1.390)	(0.712)	(0.429)	(0.871)
<i>N</i>	684	594	684	684	684	684	684	594
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	34.048	28.128	28.016	52.654	40.734	27.401	16.014	5.726
Pre-Trt. Mean, Donors	19.379	15.043	14.802	31.598	23.751	14.412	6.802	2.069
Sample Outcome S.D.	6.315	5.162	5.040	10.68	7.777	4.891	2.771	1.110
Panel E: 2014 Treatment								
NY-PFL Active	-1.420	-0.682	-0.752	-3.506	-1.430	-1.460	0.0672	0.125
<i>S.E.</i>	(1.409)	(1.178)	(1.337)	(2.230)	(1.885)	(1.415)	(0.749)	(0.502)
<i>N</i>	684	594	684	684	684	684	684	594
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	35.191	28.904	28.626	55.523	42.170	28.030	16.128	5.693
Pre-Trt. Mean, Donors	18.736	14.866	14.598	30.734	22.676	13.812	6.734	2.105
Sample Outcome S.D.	6.315	5.162	5.040	10.680	7.777	4.891	2.771	1.110
Panel F: 2012 Treatment								
NY-PFL Active	-3.143*	-1.849	-1.926	-5.275*	-3.215	-2.289	-0.694	0.131
<i>S.E.</i>	(1.757)	(1.359)	(1.361)	(2.984)	(2.361)	(1.422)	(0.747)	(0.501)
<i>N</i>	684	594	684	684	684	684	684	594
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	35.915	29.402	29.062	57.339	43.292	28.464	16.219	5.687
Pre-Trt. Mean, Donors	15.896	13.053	12.603	29.563	20.447	11.725	6.297	2.113
Sample Outcome S.D.	6.315	5.162	5.040	10.680	7.777	4.891	2.771	1.110
Panel G: 2010 Treatment								
NY-PFL Active	-3.028	-1.616	-1.618	-5.213	-3.893	-2.058	-0.291	0.009
<i>S.E.</i>	(2.692)	(2.343)	(2.125)	(4.729)	(3.622)	(2.099)	(1.105)	(0.447)
<i>N</i>	684	594	684	684	684	684	684	594
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	36.312	29.635	29.253	58.726	44.132	28.647	16.192	5.640
Pre-Trt. Mean, Donors	15.196	12.897	12.269	28.312	19.057	11.221	6.373	2.041
Sample Outcome S.D.	6.315	5.162	5.040	10.680	7.777	4.891	2.771	1.110

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: *N* is the number of state-year-age group abortion rate observations. 'Clusters' refers to the number of states in each regression and includes the single treated state. Estimates are drawn from restricted time samples from 2000 to 2017/14/12/10, and are based off 1000 placebo repetitions of each sample. 'Trt.' = Treatment. 'S.D.' = Standard Deviation of outcome for all clusters.

6 Discussion

In discussing the validity of our results, we first consider a few further elements of the reproductive climate in New York around the time of treatment beyond those detailed in Section 5.2. Regarding fertility, CDC WONDER data reveals that the fertility rate of New York women of reproductive age was almost constant from 2016-2019, at 58.29, 57.14, 57.63, and 56.89 births per 1,000 women. In 2020 this dipped to 54.30, likely impacted by the pandemic, and remained low in 2021 at 53.79.⁴² In contrast, the Abortion Surveillance

⁴¹Following Melo et al. (2023).

⁴²Appendix A.4 demonstrates that replacing abortion rates with fertility rates in our main regressions produces estimates of little significance, with inferior synthetic matching than that which we achieve for abortion. Notably however the small, positive effects

reports describe that the percentage of abortions in New York performed on out-of-state residents was largely unaffected by COVID, progressing from 5.6% in 2016, 6.1% in 2017 and 2018, 5.3% in 2019, to 5.8% in 2020 and 5.9% in 2021.⁴³ As this percentage remains roughly constant across the treatment period, we understand that the introduction of NY-PFL benefits did not incur significant migration into the state as the composition of those accessing abortion did not shift dramatically in favour of in-state residents. Ideally, we could further explore the migration question by considering age-disaggregated abortion rates reported by the mother's state of residence, alongside the standard measure of state of occurrence (Kortsmit et al., 2020). Unfortunately however the CDC only report abortion rates by residence at the state level, with all disaggregation, including by age, only reported by state of occurrence. As such we instead explore in Appendix A.4 whether the number of New York state residents who lived in a different state one year prior⁴⁴ was measurably affected by NY-PFL using our SDID method, finding no significant results.

Also approximately constant from 2016-2020 in the Abortion Surveillance data was the percentage of abortions performed on unmarried mothers in New York City⁴⁵ (83.7%, 82.7%, 81.6%, 81.5%, 82.7%), with no reporting in 2021. For this reason it is unlikely that New York's inverse trend in the share of births to unmarried mothers relative to our control states - as outlined by Table 1 - can be attributed to NY-PFL, as we might then expect the abortion profile to also change with regards to marital status. Where we do see significant change in the pre-treatment period is in the relative popularity of different abortion procedures. The percentage of all New York abortions performed using surgical methods declined from 74.5% in 2016, to 70.2% in 2017, 67.1% in 2018, 62.7% in 2019, to reach 59.0% and 58.4% in 2020 and 2021. This decrease results from the rise of medical abortion methods, employing pills and medications to terminate pregnancy, whose usage increased linearly from 19.2% in 2016 to 33.4% in 2020 for abortions performed in or before the ninth week of gestation. Note that this trend is not unique to New York however, as for the 37 areas who reported annually on this variable between 2011-2020, the share of early abortions using medical procedures more than doubled from 19.7% to 50.0% (Kortsmit et al., 2022).⁴⁶ Accordingly, the DID aspect of our SDID method incorporates the similar increase in presence of medical abortions in many other U.S. states.⁴⁷

We now address the policy implications of our results. Unlike previous work like Lenhart (2021), who studies an outcome (food insecurity) whose minimisation is of clear benefit to society, it is not immediately apparent whether the decreases we detail in abortion rates following the launch of NY-PFL are socially

we find for all age groups line up well with those of Thunell (2017) and Golightly and Meyerhofer (2022) following the launch of CA-PFL, granting further confidence in our methods and mechanisms.

⁴³Due to changing reporting standards and general data missingness, we discuss these variables only around the time of treatment, rather than the full 2000-2021 sample range. For example, reporting on abortions on out-of-state residents is reported for New York City and New York State separately (but not the state as a whole) from 2000 until 2008, and then missing for both in 2012, 2013, 2015, and 2016 (but reported at the state level).

⁴⁴Using American Community Survey Data provided by the U.S. Census Bureau (2023d).

⁴⁵Data for New York State, but not New York City, is missing around the treatment period for several of the abortion disaggregations the CDC perform, including by number of previous abortions, number of previous live births, and marital status.

⁴⁶Surgical methods still dominate terminations performed at >13 weeks of gestation however, employed in more than 90% of such cases across the U.S. in 2020 (Kortsmit et al., 2022).

⁴⁷It is worth noting that the Abortion Surveillance reports reveal that that the distribution of abortions performed in New York by weeks of gestation was largely unchanged from 2016-2021, while racial or ethnic disaggregations of abortions were not reported for New York, New York City, or New York State in 2018, 2019, and 2020, and only New York City in 2021.

desirable. We argue that they can be seen as beneficial insofar as through an income mechanism, that the financial support of NY-PFL may assist women who previously sought an abortion due to a perceived inability to ‘afford’ their child, NY-PFL can reduce the number of terminations brought about by economic necessity. As such, in contrast to TRAP laws which decrease abortion rates by inhibiting abortion-seeking women from accessing provision (see K. Jones and Pineda-Torres, 2022; Myers, 2021a; Myers, 2021b), the decrease we observe may operate in the opposite direction whereby NY-PFL can alleviate abortive coercion, rather than creating it. We find evidence in line with this conjecture in our results. Specifically, due to the cap on NY-PFL benefits at a certain percentage of the average statewide weekly wage, the financial support they offer will be proportionally larger, and hence likely more impactful, for workers on lower wages. As younger workers are paid less on average than senior counterparts (Lazear, 1981; Van Ours and Stoeldraijer, 2011), we could expect the decline in abortion rates to be proportionally larger for younger women; indeed we observe this in our main results of Tables 2 and 3. The interactions between age, income, and abortion hence allow NY-PFL to combat inequality while still covering parents of all incomes by particularly benefiting young, low-income parents who face relative financial hardship around childbirth.

An interesting avenue for further research could be to further explore the mechanisms by which NY-PFL can decrease abortion rates, whether through reducing poverty rates, or by increasing food security or feelings of financial security. Such work would contribute to the limited literature considering abortion and PFL, as we are only aware of the welfare literature surveyed in Section 5.3. Outside of abortion, notably comparable to our study is Golightly and Meyerhofer (2022), who use a synthetic control exercise to find a mean 2.8% increase in fertility for women of reproductive age associated with the introduction of CA-PFL; as with our exercise they find an age-heterogeneous effect, with larger effects for women in their thirties.

These complementary results are not the only ones to be found in the study of fertility. Sobotka et al. (2011), reviewing literature from the developed world, find that fertility increases with better economic conditions, in line with the modest fertility increases we observe in Appendix A.4 following the launch of NY-PFL. Closer still to our exercise is the work of Cavallini (2024), who finds an increase in abortions and decline in fertility following rises in unemployment. Such a case represents the inverse of our own, in which NY-PFL creates an economic uplift for prospective parents, leading abortions to decline. Cavallini (2024) also hints towards the interplay between contraception, abortion, and realised fertility: future work could explore this theme further as to if the declines in abortion we observe with NY-PFL interact with contraceptive behaviours,⁴⁸ or stem primarily from women who were already pregnant upon exposure to NY-PFL.

7 Conclusion

An increasing number of U.S. states have enacted Paid Family Leave programmes, making proper understanding of their effects essential. While a range of research has explored the impact of these programmes on

⁴⁸We hint towards this possibility in our discussion of the 2019 upticks observed in Section 5.2, which occur concurrently with legislation increasing access to contraception across New York.

many important outcomes, including reproductive choices like fertility, no existing studies have considered the possible effect of state PFL policies on abortion rates. Our investigation into this effect explores an important new channel of impact for these programmes, one that is highly policy-relevant considering the importance of abortion in the modern American consciousness. Moreover, we highlight that NY-PFL may reduce health inequalities by increasing the access of disadvantaged mothers to reproductive healthcare (Bartel et al., 2021; Bailey et al., 2021), and could serve to reduce the number of abortions undertaken due of financial pressures rather than true individual preference.

Using the recently developed Synthetic-Difference-in-Differences method, we derive significant results that the launch of New York's PFL benefits on 2018/1/1 led to a decrease of 4.583 abortions per 1,000 women for the 20-39 year old age group, an 13.6% decrease relative to the pre-treatment average. Similarly significant decreases are found for smaller subgroups, with the impact diminishing for progressively older women. Event-study output then reveals dynamic effects, with significant effects starting in 2018 (7.1%) and intensifying in 2021 (13.6%), following the increasing generosity and take-up of NY-PFL benefits. We perform a number of robustness tests to confirm the validity of our findings, demonstrating that there are no significant anticipation effects, that New York's results are unique among all placebos, and that they remain significant for our main age group when 2019, 2020, and 2021 are excluded from the sample. Additionally, we discuss that our 2018 treatment effect likely cannot be explained by any unobserved shock to abortion behaviours, and how the income mechanism we propose through which NY-PFL can affect abortion may be reflected in the proportional magnitudes we observe across age groups.

Our exercise makes several contributions to the current literature, and provides empirical evidence towards the deliberation of future public policy. First, we further the limited literature evaluating NY-PFL, an important contribution considering how much current work considers only the Californian programme. Second, we believe our findings have external validity for other states and high-income countries, meaningful as the ongoing state and national discussions over PFL fail to account for the indirect effects of PFL programmes on abortion. In contrast to the increasing presence of U.S. policies restricting access to abortion, PFL can make a positive contribution to reproductive choice as its financial assistance can support women in only seeking abortions they actually want, rather than economically require. Finally, this work provides an empirical application of the SDID method, and demonstrates in Appendix A.2 how it outperforms both the Differences-in-Differences and Synthetic Control techniques in our small-sample setting. Further work could seek to examine in detail the specific mechanisms driving the effects of PFL on abortion, or explore heterogeneity along different dimensions including income or ethnicity.

References

- A Better Balance (2023). *Overview of Paid Family & Medical Leave Laws in the United States*. A Better Balance.org. URL: <https://www.abetterbalance.org/wp-content/uploads/2021/12/CombinedPFMLChart-October-2023.pdf>.
- Abadie, A., A. Diamond, and J. Hainmueller (2010). “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program”. In: *Journal of the American Statistical Association* 105(490), pp. 493–505.
- Arkansas Department of Health (2007). *Vital Statistics*. URL: <https://www.healthy.arkansas.gov/programs-services/topics/vital-statistics>.
- Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021). “Synthetic difference-in-differences”. In: *American Economic Review* 111(12), pp. 4088–4118.
- Arora, K. and D. A. Wolf (2018). “Does Paid Family Leave Reduce Nursing Home Use? The California Experience”. In: *Journal of Policy Analysis and Management* 37(1), pp. 38–62.
- Bailey, Z. D., J. M. Feldman, and M. T. Bassett (2021). “How Structural Racism Works — Racist Policies as a Root Cause of US Racial Health Inequities”. In: *New England Journal of Medicine* 384(8), pp. 768–773.
- Bartel, A., M. Rossin-Slater, C. Ruhm, M. Slopen, and J. Waldfogel (2021). “The impact of paid family leave on employers: Evidence from New York”. In: *NBER Working Papers Series*. 28672. National Bureau of Economic Research.
- Bartel, A., M. Rossin-Slater, C. Ruhm, M. Slopen, and J. Waldfogel (2023). “The Impacts of Paid Family and Medical Leave on Worker Health, Family Well-Being, and Employer Outcomes”. In: *Annual Review of Public Health* 44, pp. 429–443.
- Bassford, M. and H. Fisher (2016). “Bonus babies? The impact of paid parental leave on fertility intentions”. In: *University of Sydney Working Papers* 2016(04). URL: <https://esacentral.org.au/images/FisherHBassfordM.pdf>.
- Ben-Michael, E., A. Feller, and J. Rothstein (2021). “The Augmented Synthetic Control Method”. In: *Journal of the American Statistical Association* 116(536), pp. 1789–1803.
- Biggs, M. A., H. Gould, and D. G. Foster (2013). “Understanding why women seek abortions in the US”. In: *BMC Women’s Health* 13, pp. 1–13.
- Boyens, C., M. Karpman, and J. Smalligan (2022). *Access to Paid Leave Is Lowest among Workers with the Greatest Needs*. Washington, DC: Urban Institute. URL: <https://www.urban.org/research/publication/access-paid-leave-lowest-among-workers-greatest-needs>.
- Brown, S., J. Herr, R. Roy, and J. A. Klerman (2020). *Employee and Worksite Perspectives of the FMLA: Who Is Eligible?* Abt Associates for the U.S. Department of Labor. URL: https://www.dol.gov/sites/dolgov/files/OASP/evaluation/pdf/WHD_FMLA2018PB1WhoIsEligible_StudyBrief_Aug2020.pdf.
- Bullinger, L. (2019). “The Effect of Paid Family Leave on Infant and Parental Health in the United States”. In: *Journal of Health Economics* 66, pp. 101–116.

- Bütikofer, A., J. Riise, and M. M. Skira (2021). “The Impact of Paid Maternity Leave on Maternal Health”. In: *American Economic Journal: Economic Policy* 13(1), pp. 67–105.
- Byker, T. S. (2016). “Paid parental leave laws in the United States: Does short-duration leave affect women’s labor-force attachment?” In: *American Economic Review* 106(5), pp. 242–46.
- California Employment Development Department (2023). *Paid Family Leave Benefit Payment Amounts*. URL: https://edd.ca.gov/en/Disability/Calculating_PFL_Benefit_Payment_Amounts.
- Cavallini, F. (2024). “Not the right time for Children: Unemployment, fertility, and abortion”. In: *Labour Economics* 87, p. 102492.
- Center for Poverty Research (2023). *National Welfare Data*. University of Kentucky. URL: <https://cpr.uky.edu/resources/national-welfare-data>.
- Centers for Disease Control and Prevention (2023). *COVID-19 Mortality by State*. URL: https://www.cdc.gov/nchs/pressroom/sosmap/covid19_mortality_final/COVID19.htm.
- Chang, R.-C., A. Rao, Q. Zhong, M. Wojcieszak, and K. Lerman (2023). “#RoeOverturned: Twitter Dataset on the Abortion Rights Controversy”. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 17, pp. 997–1005.
- Clark, A. E. and A. Lepinteur (2022). “A Natural Experiment on Job Insecurity and Fertility in France”. In: *Review of Economics and Statistics* 104(2), pp. 386–398.
- Clarke, D., D. Pailañir, S. Athey, and G. Imbens (2023). “Synthetic Difference-in-Differences Estimation”. In: *IZA Discussion Papers Series*. 15907. IZA Institute of Labor Economics.
- Clifton-Sprigg, J., I. Homburg, J. James, and S. Vujic (2023). “A Bad Break-Up? Assessing the Effects of the 2016 Brexit Referendum on Migration”. In: *IZA Discussion Papers Series*. 16468. IZA Institute of Labor Economics.
- Coile, C., M. Rossin-Slater, and A. Su (2023). “The Impact of Paid Family Leave on Families with Health Shocks”. In: *IZA Discussion Papers Series*. 15783. IZA Institute of Labor Economics.
- Collyer, S., S. Huq, K. Washington, and C. Wimer (2020). *Poverty Tracker: Spotlight on COVID-19*. Robin Hood Foundation. URL: <https://www.povertycenter.columbia.edu/news-internal/2020/covid-lost-income-new-yorkers>.
- Commonwealth of Massachusetts (2023). *Massachusetts law about parental leave*. URL: <https://www.mass.gov/info-details/massachusetts-law-about-parental-leave>.
- Congressional Research Service (2021). *Overview of the ACA Medicaid Expansion*. Congress of the United States. URL: <https://crsreports.congress.gov/product/pdf/IF/IF10399>.
- Delaware Health Statistics Center (2023). *Vital Statistics Reports*. Delaware Department of Health and Social Services. URL: <https://dhss.delaware.gov/dhss/dph/hp/annrepvs.html>.
- Division of Public Health (2013). *Delaware Reported Pregnancies, 2009 & 2010*. Delaware Department of Health and Social Services. URL: <https://dhss.delaware.gov/dph/hp/files/rp0910.pdf>.
- Division of Vital Records (2022). *Vital Records*. North Dakota Department of Health and Social Services. URL: <https://www.hhs.nd.gov/vital>.

- Finer, L. B., L. F. Frohworth, L. A. Dauphinee, S. Singh, and A. M. Moore (2005). "Reasons US women Have Abortions: Quantitative and Qualitative Perspectives". In: *Perspectives on Sexual and Reproductive Health* 37(3), pp. 110–118.
- Flood, S., M. King, R. Rodgers, J. R. Warren, and M. Westberry (2023). *Integrated Public Use Microdata Series, Current Population Survey: Version 11.0 [dataset]*. Minneapolis: University of Minnesota.
- Gauthier, A. (2007). "The Impact of Family Policies on Fertility in Industrialized Countries: A Review of the Literature". In: *Population research and policy review* 26, pp. 323–346.
- Gohmann, S. F. and R. L. Ohsfeldt (1993). "Effects of price and availability on abortion demand". In: *Contemporary Economic Policy* 11(4), pp. 42–55.
- Golightly, E. and P. Meyerhofer (2022). "Does paid family leave cause mothers to have more children? Evidence from California". In: *Journal of Labor Research* 43(2), pp. 203–238.
- Grossman, D., J. Perritt, and D. Grady (2022). "The Impending Crisis of Access to Safe Abortion Care in the US". In: *JAMA Internal Medicine* 182(8), pp. 793–795.
- Howe, L. (2020). *Massachusetts Paid Family and Medical Leave: Latest Updates Ahead of January 1, 2021, Effective Date*. URL: <https://ogletree.com/insights/massachusetts-paid-family-and-medical-leave-latest-updates-ahead-of-january-1-2021-effective-date/>.
- Hussey, L. (2010). "Welfare Generosity, Abortion Access, and Abortion Rates: A Comparison of State Policy Tools". In: *Social Science Quarterly* 91(1), pp. 266–283.
- Hussey, L. (2011). "Is Welfare Pro-life? Assistance Programs, Abortion, and the Moderating Role of States". In: *Social Service Review* 85(1), pp. 75–107.
- Jones, K. and M. Pineda-Torres (2022). "TRAP'd Teens: Impacts of Abortion Provider Regulations on Fertility & Education". In: *IZA Discussion Papers Series*. 14837. IZA Institute of Labor Economics.
- Jones, R., M. Kirstein, and J. Philbin (2022). "Abortion incidence and service availability in the United States, 2020". In: *Perspectives on Sexual and Reproductive Health* 54(4), pp. 128–141.
- Jou, J., K. B. Kozhimannil, J. M. Abraham, L. A. Blewett, and P. M. McGovern (2018). "Paid Maternity Leave in the United States: Associations with Maternal and Infant Health". In: *Maternal and Child Health Journal* 22(2), pp. 216–225.
- Kaiser Family Foundation (2023). *Status of State Medicaid Expansion Decisions: Interactive Map*. URL: <https://www.kff.org/medicaid/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/>.
- Kelly, K. and L. Grant (2007). "State Abortion and Nonmarital Birthrates in the Post—Welfare Reform Era: The Impact of Economic Incentives on Reproductive Behaviors of Teenage and Adult Women". In: *Gender & Society* 21(6), pp. 878–904.
- Kim, J. and O. Lenhart (2024). "Paid family leave and the fight against hunger: Evidence from New York". In: *Health Economics* 2024;1-25.
- Kim, W. (2023). *Baby Bonus, Fertility, and Missing Women*. DOI: [10.2139/ssrn.3704188](https://doi.org/10.2139/ssrn.3704188).

- Kirkman, M., H. Rowe, A. Hardiman, S. Mallett, and D. Rosenthal (2009). “Reasons women give for abortion: a review of the literature”. In: *Archives of Women’s Mental Health* 12, pp. 365–378.
- Korttsmit, K., T. Jatlaoui, M. Mandel, et al. (2020). “Abortion 2018 ss6907 - Supplementary Tables 1-17 with 2017 data”. In: *Morbidity and Mortality Weekly Report Supplementary Material*. URL: <https://stacks.cdc.gov/view/cdc/96608>.
- Korttsmit, K., A. T. Nguyen, M. G. Mandel, E. Clark, L. M. Hollier, J. Rodenhizer, and M. K. Whiteman (2022). “Abortion Surveillance—United States, 2020”. In: *MMWR Surveillance Summaries* 71(10), pp. 1–27.
- Kranz, S. (2022). *Synthetic Difference-in-Differences with Time-Varying Covariates*. Tech. rep. URL: https://github.com/skranz/xsynthdid/blob/main/paper/synthdid_with_covariates.pdf.
- Lalive, R. and J. Zweimüller (2009). “How does parental leave affect fertility and return to work? Evidence from two natural experiments”. In: *The Quarterly Journal of Economics* 124(3), pp. 1363–1402.
- Larsson, M., G. Aneblom, V. Odland, and T. Tydén (2002). “Reasons for pregnancy termination, contraceptive habits and contraceptive failure among Swedish women requesting an early pregnancy termination”. In: *Acta obstetrica et gynecologica Scandinavica* 81(1), pp. 64–71.
- Lazear, E. P. (1981). “Agency, Earnings Profiles, Productivity, and Hours Restrictions”. In: *The American Economic Review* 71(4), pp. 606–620.
- Lenhart, O. (2021). “The effects of paid family leave on food insecurity—evidence from California”. In: *Review of Economics of the Household* 19(3), pp. 615–639.
- Lichtman-Sadot, S. and N. Bell Pillay (2017). “Child Health in Elementary School Following California’s Paid Family Leave Program”. In: *Journal of Policy Analysis and Management* 36(4), pp. 790–827.
- Logue-Conroy, R., D. Lancaster, and L. Daux (2021). *Implementation of Paid Family Leave in New Jersey: A Promising Work in Progress*. Centre for Women, Work, Rutgers University School of Management, and Labor Relations. URL: https://smlr.rutgers.edu/sites/default/files/Documents/Implementation_PaidFamilyLeave_Oct2021.pdf.
- Longbons, T. (2019). *Abortion Reporting: New York (2017)*. Charlotte Lozier Institute. URL: <https://lozierinstitute.org/abortion-reporting-new-york-2017/>.
- Longbons, T. (2021). *Abortion Reporting: New York (2018)*. Charlotte Lozier Institute. URL: <https://lozierinstitute.org/abortion-reporting-new-york-2018/>.
- Louisiana Department of Health (2023). *Induced Termination of Pregnancy (ITOP) Data*. URL: <https://ldh.la.gov/page/709>.
- Luci-Greulich, A. and O. Thévenon (2013). “The Impact of Family Policies on Fertility Trends in Developed Countries”. In: *European Journal of Population* 29, pp. 387–416.
- Maine Department of Health and Human Services (2023). *Table A-4: Induced Abortions by Gestation, Age of Patient and Race. Maine State Totals (2012), Occurrence Data*. Maine Center for Disease Control, Prevention, Data, Research, and Vital Statistics. URL: https://www.maine.gov/dhhs/mecdc/public-health-systems/data-research/data/documents/pdf/ABO%20VSTAB%20Tables_2012.pdf.

- Marseille, E., A. Mirzazadeh, M. A. Biggs, A. Miller, H. Horvath, M. Lightfoot, M. Malekinejad, and J. Kahn (2018). “Effectiveness of School-Based Teen Pregnancy Prevention Programs in the USA: A Systematic Review and Meta-Analysis”. In: *Prevention Science* 19, pp. 468–489.
- Medoff, M. (1988). “An economic analysis of the demand for abortions”. In: *Economic inquiry* 26(2), pp. 353–359.
- Melo, V., E. Neilson, and D. Kemboi (2023). “Indoor Vaccine Mandates in US Cities, Vaccination Behavior, and COVID-19 Outcomes”. In: *Mercatus Working Papers*. NMercatus Center at George Mason University.
- Missouri Department of Health and Senior Services (2023). *Missouri Vital Statistics 2021*. URL: <https://health.mo.gov/data/vitalstatistics/mvs21/2021MissouriVitalStatistics.pdf>.
- Missouri Economic Research and Information Center (2023). *Cost of Living Data Series*. URL: <https://meric.mo.gov/data/cost-living-data-series>.
- Myers, C. (2021a). “Cooling off or Burdened? The Effects of Mandatory Waiting Periods on Abortions and Births”. In: *IZA Discussion Papers Series*. 14434. IZA Institute of Labor Economics.
- Myers, C. (2021b). “Measuring the burden: The effect of travel distance on abortions and births”. In: *IZA Discussion Papers Series*. 14556. IZA Institute of Labor Economics.
- National Institute for Reproductive Health (2022). *Passing and Protecting New York’s Reproductive Health Act*. URL: <https://nirhealth.org/our-impact/passing-protecting-new-yorks-reproductive-health-act/>.
- Nebraska Department of Health and Human Services (2017). *Vital Records and Statistics*. URL: <https://dhhs.ne.gov/Pages/vital-records-and-statistics.aspx>.
- New York Senate (2019). *Senate Bill S659A, 2019-2020 Legislative Session*. URL: <https://www.nysenate.gov/legislation/bills/2019/S659>.
- New York State (2019). *Paid Family Leave: 2018 Year in Review*. URL: https://www.governor.ny.gov/sites/default/files/atoms/files/PFL_EOYReport_2018_FINAL.pdf.
- New York State (2023a). *Eligibility: Who is eligible for Paid Family Leave?* URL: <https://paidfamilyleave.ny.gov/eligibility>.
- New York State (2023b). *New York Paid Family Leave Updates for 2023*. URL: <https://www.govdocs.com/new-york-state-paid-family-leave-update-2023/>.
- New York State (2023c). *Paid Family Leave and Other Benefits*. URL: <https://paidfamilyleave.ny.gov/paid-family-leave-and-other-benefits>.
- New York State (2023d). *What is Military Family Support Leave?* URL: <https://paidfamilyleave.ny.gov/paid-family-leave-military-families>.
- New York State Department of Financial Services (2023). *New York State Paid Family Leave Report 2018 – 2022*. URL: https://www.dfs.ny.gov/reports_and_publications/pfl.
- Nieto, A. (2022). “Can subsidies to permanent employment change fertility decisions?” In: *Labour Economics* 78, p. 102219.

- OECD (2022). *OECD Family Database PF2.1.: Parental Leave Systems*. URL: https://www.oecd.org/els/soc/PF2_1_Parental_leave_systems.pdf.
- Pihl, A. M. and G. Basso (2019). “Did California Paid Family Leave Impact Infant Health?” In: *Journal of Policy Analysis and Management* 38(1), pp. 155–180.
- Prifti, E. and D. Vuri (2013). “Employment protection and fertility: Evidence from the 1990 Italian reform”. In: *Labour Economics* 23, pp. 77–88.
- Rader, B., U. D. Upadhyay, N. K. Sehgal, B. Y. Reis, J. S. Brownstein, and Y. Hswen (2022). “Estimated Travel Time and Spatial Access to Abortion Facilities in the US Before and After the Dobbs v Jackson Women’s Health Decision”. In: *Journal of the American Medical Association* 328(20), pp. 2041–2047.
- Raute, A. (2019). “Can Financial Incentives Reduce the Baby Gap? Evidence from a Reform in Maternity Leave Benefits”. In: *Journal of Public Economics* 169, pp. 203–222.
- Ressler, R., M. S. Waters, and C. Hill (2022). “The Impact of Welfare Reform on Rates of Abortion”. In: *Southern University College of Business E-Journal* 12(1), p. 2.
- Rhode Island Senate (2023). *Rhode Island Senate Bill 0139 (Lawson)*. State of Rhode Island. URL: <http://webserver.rilegislatore.gov/BillText/BillText23/SenateText23/S0139.pdf>.
- Rice, W. S., S. Redd, A. Luke, K. Komro, K. J. Arriola, and K. S. Hall (2022). “Dispersion of contraceptive access policies across the United States from 2006 to 2021”. In: *Preventive Medicine Reports* 27, p. 101827.
- Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (2013). “The Effects of California’s Paid Family Leave Program on Mothers’ Leave-Taking and Subsequent Labor Market Outcomes”. In: *Journal of Policy Analysis and Management* 32(2), pp. 224–245.
- Santelli, J. S., I. S. Speizer, A. Avery, and C. Kendall (2006). “An Exploration of the Dimensions of Pregnancy Intentions Among Women Choosing to Terminate Pregnancy or to Initiate Prenatal Care in New Orleans, Louisiana”. In: *American Journal of Public Health* 96(11), pp. 2009–2015.
- Schmidheiny, K. and S. Siegloch (2019). “On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications”. In: *IZA Discussion Papers Series*. 12079. IZA Institute of Labor Economics.
- Schnake-Mahl, A. S., J. L. Pomeranz, N. Sun, I. Headen, G. O’Leary, and J. L. Jahn (2023). “Forced Birth and No Time off Work: Abortion Access and Paid Family Leave Policies”. In: *American Journal of Preventive Medicine*.
- Schneider, D. (2015). “The Great Recession, Fertility, and Uncertainty: Evidence From the United States”. In: *Journal of Marriage and Family* 77(5), pp. 1144–1156.
- Shim, J. (2014). “Family Leave Policy and Fertility Rates in OECD Countries including East Asia”. In: *PAA 2014 Annual Meeting*. Population Association of America.
- Sobotka, T., V. Skirbekk, and D. Philipov (2011). “Economic Recession and Fertility in the Developed World”. In: *Population and Development Review* 37(2), pp. 267–306.
- State Center for Health Statistics (2014). *Reported Pregnancies and Pregnancy Rates in the District of Columbia, 2008-2012*. Center for Policy, Planning, and Evaluation, D.C. Department of Health. URL: <https://>

doh.dc.gov/sites/default/files/dc/sites/doh/publication/attachments/Reported%20Pregnancy%20Rates%20in%20DC%202008-2012%20Final%20%289%2025%2014%29.pdf.

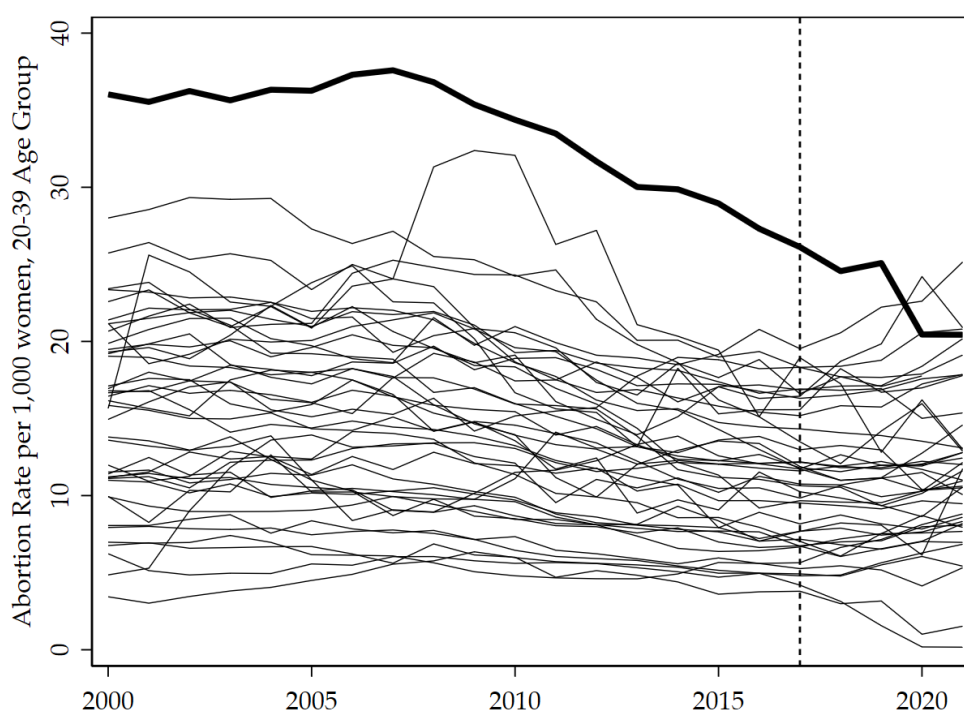
- Stearns, J. (2015). “The effects of paid maternity leave: Evidence from Temporary Disability Insurance”. In: *Journal of Health Economics* 43, pp. 85–102.
- Supreme Court (2022). *Dobbs v. Jackson Women’s Health Organization* (06/24/2022). Opinions of the Supreme Court of the United States, October Term 2021. URL: https://www.supremecourt.gov/opinions/21pdf/19-1392_6j37.pdf.
- Tennessee Department of Health (2023). *Induced Termination of Pregnancy Statistics*. URL: <https://www.tn.gov/health/health-program-areas/statistics/health-data/itop.html>.
- Texas Department of Health and Human Services (2023). *2014 ITOP Statistics*. URL: <https://www.hhs.texas.gov/about/records-statistics/data-statistics/itop-statistics>.
- Thunell, J. (2017). “Paid Family Leave and Women’s Childbearing Decisions in the United States”. In: *PAA 2017 Annual Meeting*. Population Association of America.
- U.S. Census Bureau (2022). *American Community Survey Data via FTP*. URL: <https://www.census.gov/programs-surveys/acs/data/data-via-ftp.html>.
- U.S. Census Bureau (2023a). *Health Insurance Historical Tables - HHI Series*. URL: <https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-series/hic.html>.
- U.S. Census Bureau (2023b). *Health Insurance Historical Tables - HIA Series*. URL: <https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-series/hia.html>.
- U.S. Census Bureau (2023c). *State Population Totals and Components of Change: 2020-2022*. URL: <https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html>.
- U.S. Census Bureau (2023d). *State-to-State Migration Flows*. URL: <https://www.census.gov/data/tables/time-series/demo/geographic-mobility/state-to-state-migration.html>.
- Van Ours, J. C. and L. Stoeldraijer (2011). “Age, Wage and Productivity in Dutch Manufacturing”. In: *De Economist* 159(2), pp. 113–137.
- Washington State (2023). *History: Washington Paid Family & Medical Leave*. URL: <https://paidleave.wa.gov/history/>.
- Winston, P, E. Coombs, R. Bennett, L. Antelo, P. Landers, and M. Abbott (2019). “Paid family leave: supporting work attachment among lower income mothers”. In: *Community, Work & Family* 22(4), pp. 478–511.
- Wolfe, T. and Y. van der Meulen Rodgers (2021). “Abortion During the COVID-19 Pandemic: Racial Disparities and Barriers to Care in the USA”. In: *Sexuality Research and Social Policy* 19, pp. 1–8.
- Worker’s Compensation Board (2023). *Is Disability and Paid Family Leave Benefits Coverage Required?* New York State. URL: <https://paidfamilyleave.ny.gov/eligibility>.
- Zraick, K. and H. Meko (2022). *What’s Next for Abortion Rights in New York?* The New York Times. URL: <https://www.nytimes.com/2022/07/08/nyregion/abortion-new-york-roe.html>.

A Appendix

A.1 Further Background and Abortion Provision

We illustrate in Figure 6 how the abortion rate of our core 20-39 age group varied between 2000-2021 across each of the 38 states included in our main analysis, with New York's outcome shown in bold:

Figure 6: Abortion Rate per 1,000 women of 20-39 age group, 38 state sample, 2000-2021



Clearly, New York has the highest pre-treatment abortion rate of all states with complete abortion data. This fact is a strong motivator to utilise the SDID method rather than the canonical Synthetic Control, as there is no combination of non-negative weights we can assign to our 37 donors that can contain New York's abortion rate within its convex hull. SDID by comparison only requires us to approximate the trend for New York without having to match the uniquely high level, making construction of the synthetic approximation feasible. An alternative option would be to perform an Augmented Synthetic Control exercise, which employs the minimum degree of negative weighting to capture an outlier within the convex hull (Ben-Michael et al., 2021). However we avoid this method as it involves imputation into the negative dimension of observed data, a requirement which we find undesirable.

Next, we provide in Table 5 an extensive summary and comparison of the state PFL programmes active during our 2000-2021 sample period. Overall they are quite alike, and much closer to each other than they are to other national OECD PFL programmes (see OECD, 2022). Upon its launch NY-PFL was likely the most generous programme yet, offering the longest leave, job protections, and the second highest maximum weekly benefit. The more recent programmes however offer similar benefits to NY-PFL, with Washington's perhaps the most generous of all.

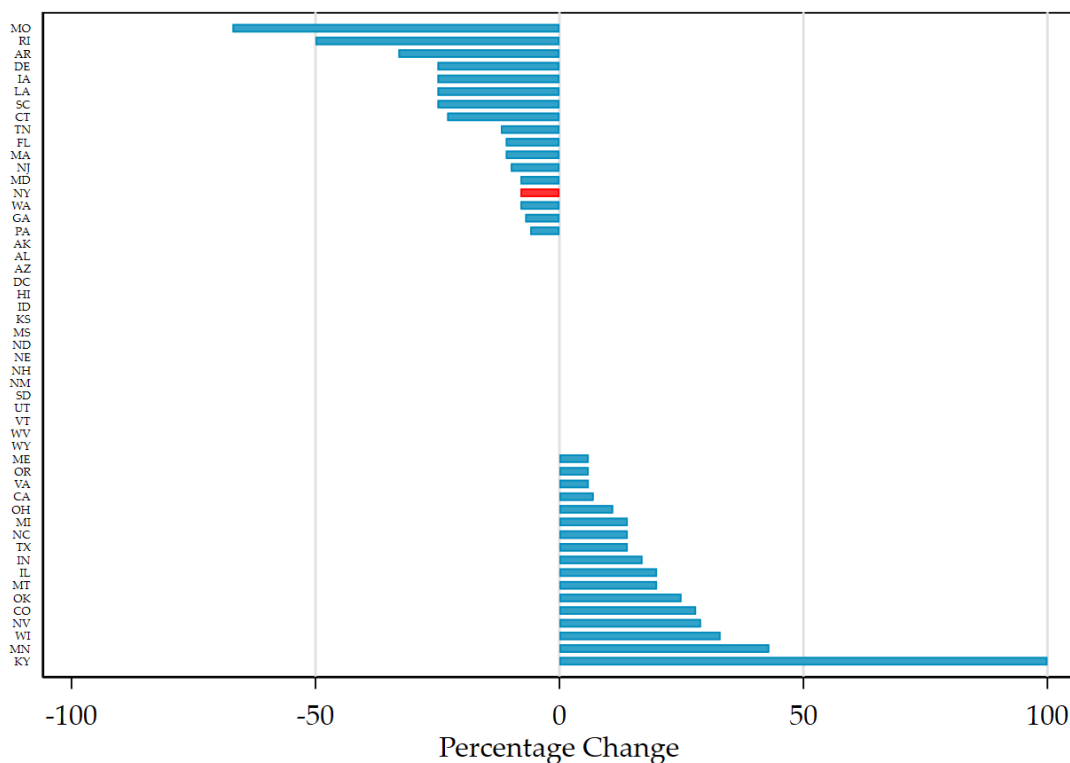
Table 5: Characteristics of selected state Paid Family Leave programmes, as of May 2023

	California	New Jersey	Rhode Island	New York	Washington	Massachusetts	D.C.
Effective From	2004/6/1	2009/1/1	2014/1/1	2018/1/1	2019/1/1	2019/1/1	2020/6/1
Max. Length	8 Weeks	8-12 Weeks	6 Weeks	12 Weeks	12 Weeks	12 Weeks	12 Weeks
Coverage	Employees covered by state unemployment insurance law, except for most public employees. Includes self-employed opt-in.	Employees covered by state unemployment insurance law, including public employees. No self-employed opt-in.	Employees covered under state unemployment insurance law, except for public employees. No self-employed opt-in.	Covers vast majority of private sector employees, with some exceptions. Includes self-employed opt-in.	All employees including public employees. Includes self-employed opt-in.	Employees covered by state unemployment insurance law, including some public employees. Includes self-employed opt-in.	Most private sector employees; no public employee opt-in. Includes self-employed opt-in.
Eligibility	Must have earned \geq \$300 during the base period (first 4 of the 5 most recently completed quarters).	Must have earned \geq 20 times the minimum wage (\$260) in at least 20 weeks or \geq 1,000 times the minimum wage (\$13,000) across the base period.	Must have earned \geq 200 times the minimum wage (\$2,600) in one quarter, and earned \geq 400 times the minimum wage (\$5,200) across the base period.	Must have been at current employer for \geq 26 consecutive weeks; those who work $<$ 20 hours per week must have worked \geq 175 days at current employer.	Must have worked \geq 820 hours in the qualifying period, that being the first 4 of the 5 most recently completed quarters or the 4 most recent quarters.	Must have earned \geq \$6,000 over the base period and meet an average earnings criterion that means the worker must have worked for several weeks.	Must have been employed by a covered D.C. employer during some of the 52 preceding weeks. Those employed for $<$ 1 year may receive a prorated amount.
Benefit Amount	60/70% of worker's AWW.	85% of worker's AWW.	\sim 60% of worker's AWW.	67% of worker's AWW.	50/90% of worker's AWW.	50/80% of worker's AWW.	50/90% of worker's AWW.
Max. Benefit	\$1,620p/w	\$1,025p/w	\$1,007p/w	\$1,131.08p/w	\$1,427p/w	\$1,129.82p/w	\$1,049p/w
Job Protections	No	Yes	Yes	Yes	Yes/No	Yes	No

Notes: AWW = Average Weekly Wage. Only programmes active during 2000-2021 are listed, hence excluding Connecticut (2022/1), Oregon (2023/1), Colorado (2024/1), Maryland (2026/1), Delaware (2026/1), Minnesota (2026/1), and Maine (2026/5). All otherwise unreferenced material is sourced from A Better Balance (2023). **Timing:** Employees began paying dues on 2009/1/1 in NJ although benefits were not available until 2009/6/1 (Arora and Wolf, 2018). Payroll deductions began in New York on 2017/7/1 although benefits were not available until 2018/1/1 (New York State, 2019). Washington's PFL programme began taking dues on 2019/1/1 but benefits could not be applied for until 2020/1/1 (Washington State, 2023). Massachusetts' PFL benefits could not be taken until 2021/1/1 (Commonwealth of Massachusetts, 2023), some not until 2021/6/1 (Howe, 2020). **Length:** Workers in New Jersey can take \leq 12 continuous weeks of PFL or \leq 8 weeks of intermittent leave. NJ's PFL programme was also expanded significantly on 2020/6/1, doubling the length of leave from 6 weeks to 12 (Logue-Conroy et al., 2021). Maximum leave length in RI increased from 4 weeks to 5 in 2021 to 6 in 2023 (Rhode Island Senate, 2023). New York's programme ramped-up from 8 weeks of leave in 2018, to 10 in 2019/20, to 12 in 2021 (New York State Department of Financial Services, 2023). Several states have limits on how many combined weeks of medical/family leave can be taken in a 12 month period, including D.C. (12, or 14 for pregnancy related health complications), Washington (16), Massachusetts (26), New York (26), and Rhode Island (30). **Coverage:** Many public employers can opt-in to coverage through a collecting bargaining process in CA, NJ (for own-health leave, PFL already covered), RI, NY, and MA. Domestic workers are covered in all states, subject to a low minimum earnings (CA, NJ, RI, D.C.) or hours (NY) requirement. **Eligibility:** Minimum income/hours requirements combine income across employers if a worker has multiple jobs in CA, NJ, RI, WA, and MA. The 'base period' is defined differently across states: either 3 or 4 quarters, and ignores quarters of unemployment in CA. RI also requires that income across the base period must be at least 150% of their single highest income quarter. New York eased some eligibility criteria on 2023/1/1 (New York State, 2023b). **Benefit Amount:** For California, earnings $<$ 1/3 state average: 70% AWW; earnings $>$ 1/3 state average: 60% AWW (California Employment Development Department, 2023). RI's benefit amount is precisely calculated as 4.62% of the worker's wages in their highest earning base year quarter. NY's benefit amount increased from 50% AWW in 2018, to 55% (2019), to 60% (2020), to 67% for subsequent years (New York State Department of Financial Services, 2023). For WA, MA, and D.C., the benefit amount is the high % up to a threshold (50% statewide AWW for MA and WA, 40x150% minimum wage for D.C.), the low % above. **Max. Benefit:** All figures are calculated as a certain % of statewide AWW, except D.C. which is a flat figure adjusted annually for inflation. **Job Protections:** In all states workers may have existing protections under FMLA or other state laws. Washington's protections only extend to certain workers, although the criteria for inclusion are similar to the state's PFL eligibility criteria.

Finally as introduced in Section 5.2, it is worth considering how abortion provision changed across the U.S. around the time of treatment. This is because the availability of abortion provision affects the ability of women to actually receive abortions that they desire (Myers, 2021b), hence feasibly confounding with our outcome of abortion rates. Unfortunately, consistent state-by-state panel data on the number of abortion providers is not available across our sample period (2000-2021), so we cannot include it as a covariate in our analysis. Instead we now display, using data from the Guttmacher Institute (R. Jones et al., 2022), how the number of abortion clinics⁴⁹ in each state changed between 2017 and 2020 in order to gauge if New York’s decline of 8% was unique. Figure 7 presents this information with New York’s outcome in red, amongst the set of all other states:

Figure 7: Change in Number of Abortion Clinics between 2017 and 2020, by State



While New York’s decrease of 9 clinics is the largest in absolute terms owing to the large number of clinics in the state, this decrease was proportionally similar to that of many other states, including other populous states like Florida, Georgia, and New Jersey. Furthermore, and as mentioned in Section 5.2, among the states given weight in our SDID estimation for the 20-39 age group, Delaware has a larger decrease than New York, while New Mexico and Kansas saw no change in their number of clinics. As such, by comparing New York to these other states in our synthetic control, we do not risk New York experiencing a unique fall in abortion provision that could explain some of decreases in abortion rates that we observe in Section 5. Moreover as Figure 7 shows that the other four states that compose our synthetic (NV, OR TX, VA) see modest rises in their number of abortion clinics between 2017 and 2020, our results are also not cherry picked in that they only compare to other states which saw no increase in abortion provision over the post-treatment period.

⁴⁹The Guttmacher Institute also share data on the number of abortion providers in each state, which encompasses a wider range of healthcare providers than just clinics. However the most recent year of data for this second variable is 2017, and so it cannot be used for our discussion of the post-treatment period here.

A.2 Optimal Unit Weights and Method Comparison

In addition to the SDID matching and time weights illustrated by Figures 2 and 3, we also provide the unique optimal unit weights we compute for each regression in order to solve (2).

Figure 8: Optimal Unit Weights, First Four Age Groups, Abortion Rates, New York

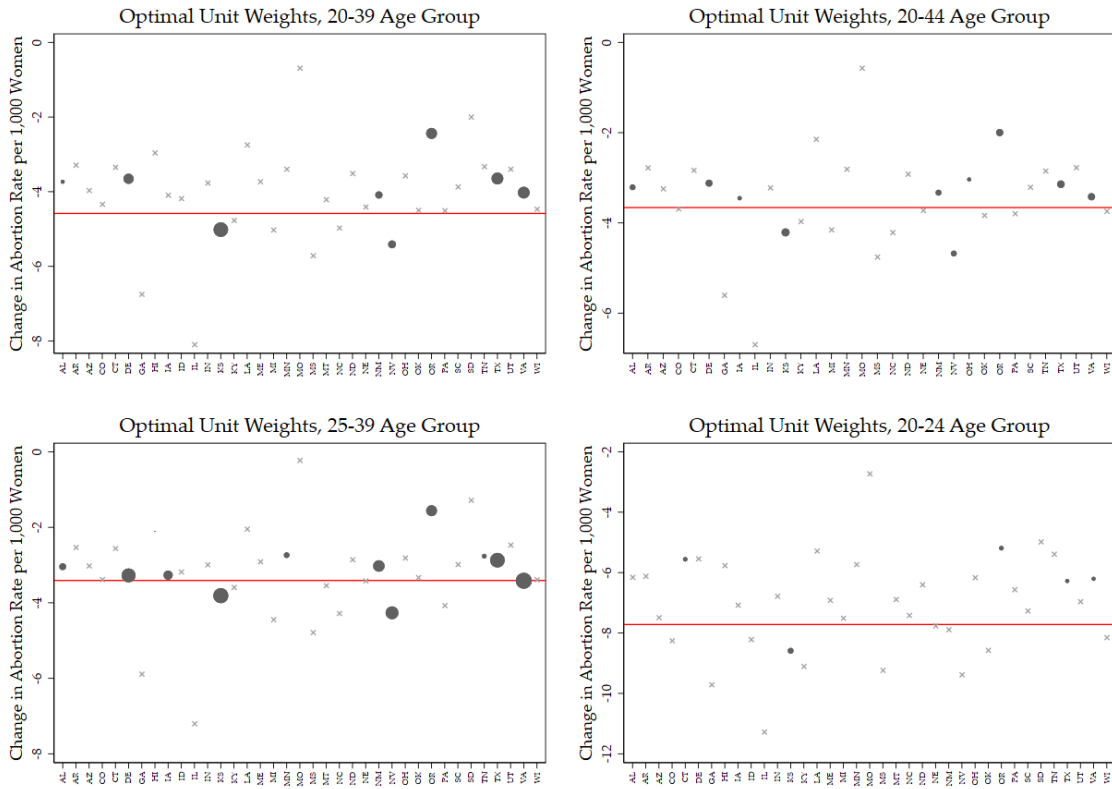
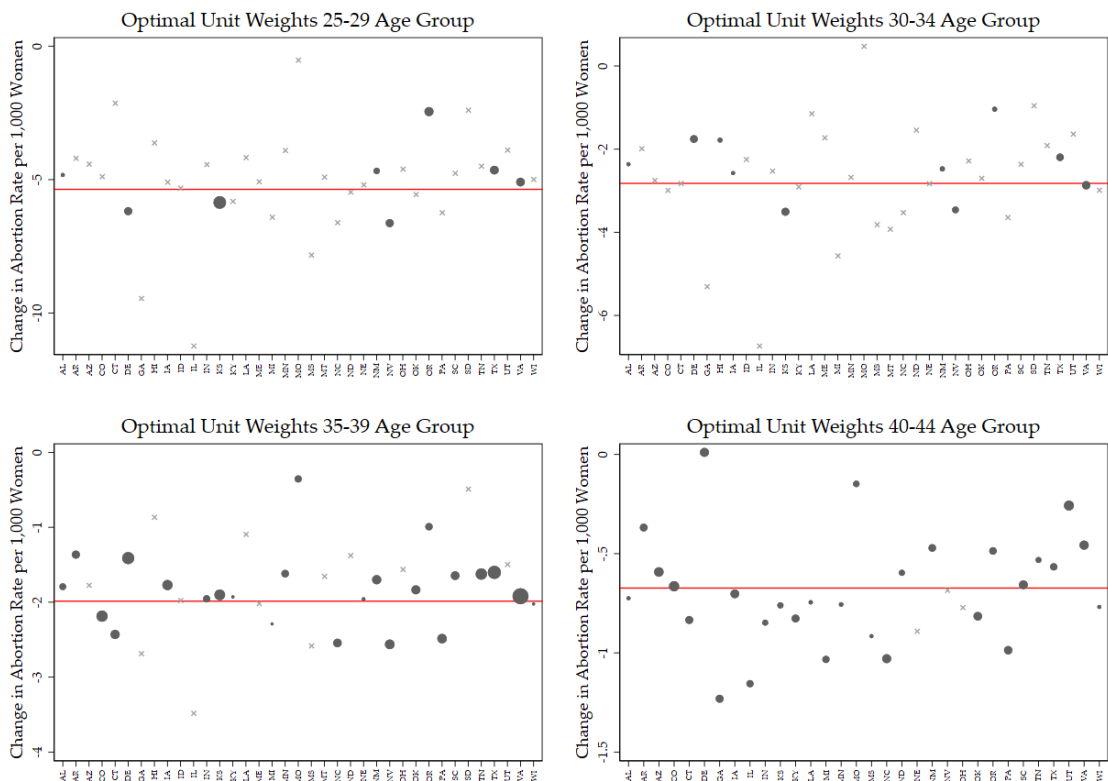


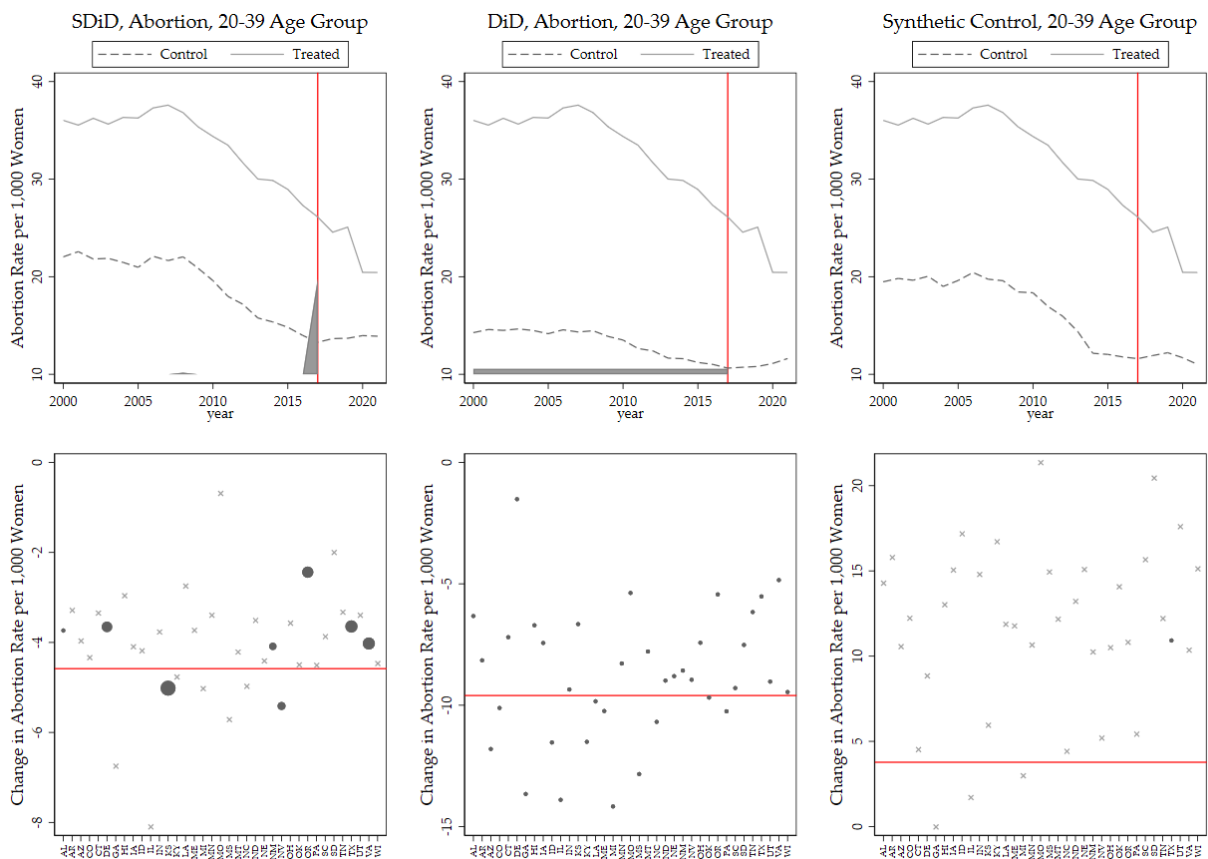
Figure 9: Optimal Unit Weights, Final Four Age Groups, Abortion Rates, New York



Donors are listed by postal code on the X-axis; there are 37 for six age groups and 32 for the two containing women aged 40-44, as discussed in Section 3.1. On the Y-axis is the change over time in abortion rates for each state, with New York's value shown as a red line. The size of the weight each control state is assigned is illustrated by the size of its dot, with states assigned zero weight denoted by an 'x'. Across the eight panels, an average of 13.1 states are assigned a positive weight in each regression, ranging from a minimum of 5 for the 20-24 age group to a maximum of 29 for those aged 40-44. This is due to New York's trend in abortion rates for older groups being less unique than its trend for younger groups, meaning more states are close of it and so are assigned a positive weight. For our main 20-39 age group, a diverse collection of eight states are assigned positive weights, namely Alabama, Delaware, Kansas, New Mexico, Nevada, Oregon, Texas, and Virginia. While it may initially seem strange that we compare New York to a synthetic that draws from rural, Western states like Kansas and New Mexico, in reality the need to SDID estimation to only match on trends and not levels enables us to draw on a more diverse pool of donors than an Synthetic Control exercise can. This is due to the latter needing to match both trends and levels, and hence is generally constrained to a set of comparators more homogeneous with the treated state than our exercise requires.

To evidence this claim, we now reproduce Figure 1 of Arkhangelsky et al. (2021) as Figure 10, which illustrates the matching and weighting of the SDID, DID, and Synthetic Control methods for our main 20-39 exercise. The 'sdid' STATA command of Clarke et al. (2023) allows us to run an identical regression for each respective method, making such a comparison easy to perform:

Figure 10: Method Comparison, 20-39 Abortion Rates, New York



The differences between the two-way fixed effects regressions each method employs (see Arkhangelsky et al., 2021) results in visibly different optimal time weights, unit weights, and trends matching. The SDID panels on the left of Figure 10 are those of our main results, where eight states are assigned a positive weight, and abortions are estimated to fall by 4.583 abortions per 1,000 women (red line, bottom panel). To first compare these results to those of the conventional DID, we see that this second method weights all states and all time periods equally. To then draw inference from their DID exercise, the researcher would be required to make an assumption of parallel trends between the treatment and control units on average, which the upper-middle panel of Figure 10 displays to be untrue. As such, the comparison between the two leads to an erroneous estimate of 9.600 abortions per 1,000 women reduced by the introduction of NY-PFL benefits in 2018, a 28.6% reduction relative to the pre-treatment mean of 33.608 abortions per 1,000 women. The use of optimal time and unit weights in the SDID thus prove their value here, as we can produce a synthetic unit with a trend that much better represents New York, leading to more reliable comparison and estimation.

To then compare our SDID results to the Synthetic Control results of the rightmost panels, we observe the fundamental reason why the latter fails in our setting: there is no combination of non-negative weights we can assign to our control units that can produce a representative synthetic trend near the level of New York's uniquely high abortion rates.⁵⁰ As such, the closest the Synthetic Control estimator can get is to assign all weight to Texas, another rich, populous state with a moderately high abortion rate and roughly comparable trend. However this match is not good enough to act as a valid counterfactual for New York in such an exercise. Such incompatibility leads the Synthetic Control estimator to produce a statistically insignificant result that NY-PFL increased abortion rates by 3.775 abortions per 1,000 women. We conclude that, due to this substandard matching, favouring the Synthetic Control method over SDID would be inopportune.

A.3 Alternative Specifications

Compared to a Synthetic Control exercise, where covariates are typically used alongside lagged outcomes as the vectors upon which the root mean squared prediction error is minimised in constructing the counterfactual (Abadie et al., 2010), covariates in SDID play a more minor role.⁵¹ Like a conventional DID or linear regression, covariates are used to capture out the presence of other confounding variables before the synthetic unit is constructed, thus improving the matching accuracy (Clarke et al., 2023). This is not to say however that their inclusion is unimportant. We evidence this statement by now re-running our exercise without any covariates and observing the matching quality relative to that of Figures 4 and 5.

The value of our covariates reveals itself here, as the matching achieved in Figures 11 and 12 is evidently worse than that we achieve in Figures 4 and 5. The positive pre-trends we minimised in our main event studies are substantially stronger here, as each point has shifted directly up with most now significantly above zero, neutering our ability to draw inference. Hence, while the inclusion or exclusion of individual

⁵⁰See Figure 6 of Appendix A.1.

⁵¹Indeed Melo et al. (2023) do not include any covariates in their analysis.

Figure 11: Dynamic SDID Estimation, First Four Age Groups, No Covariates, New York

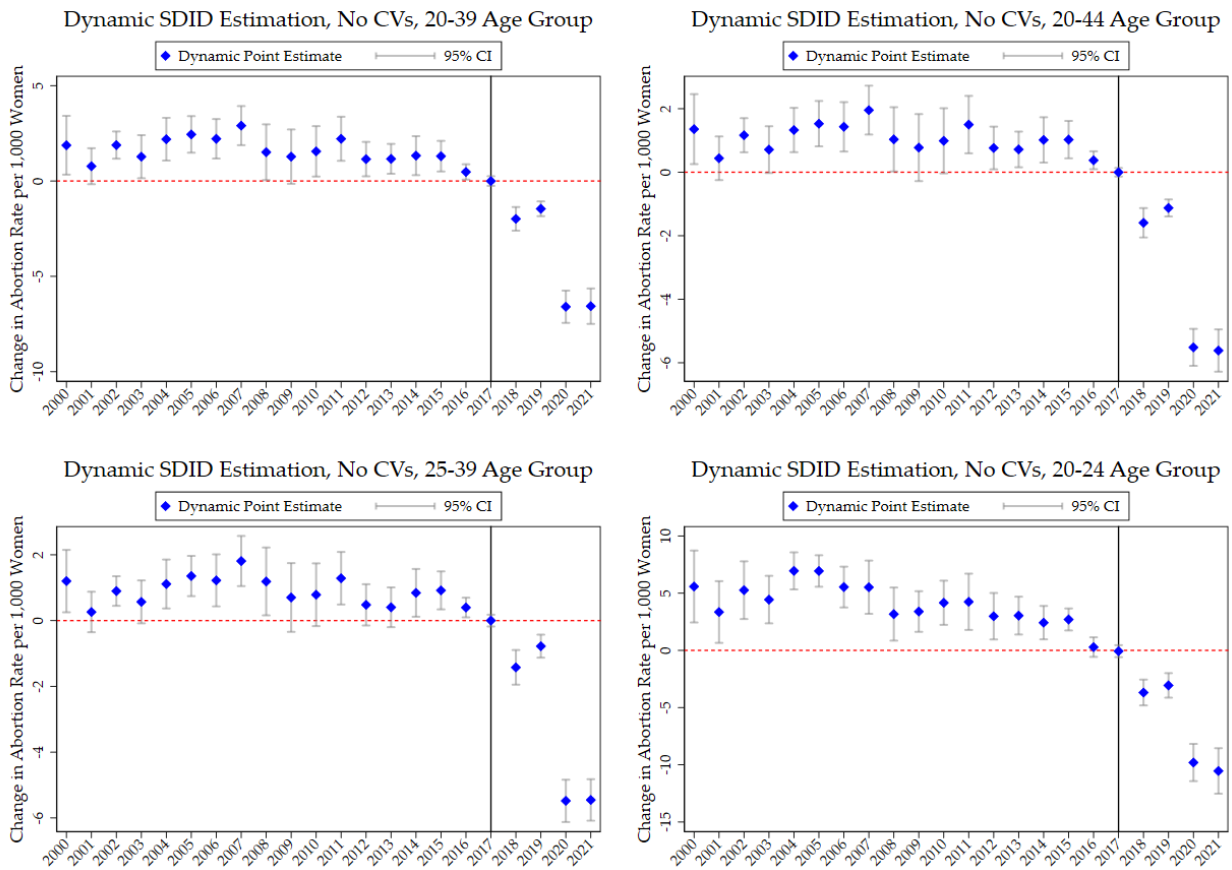
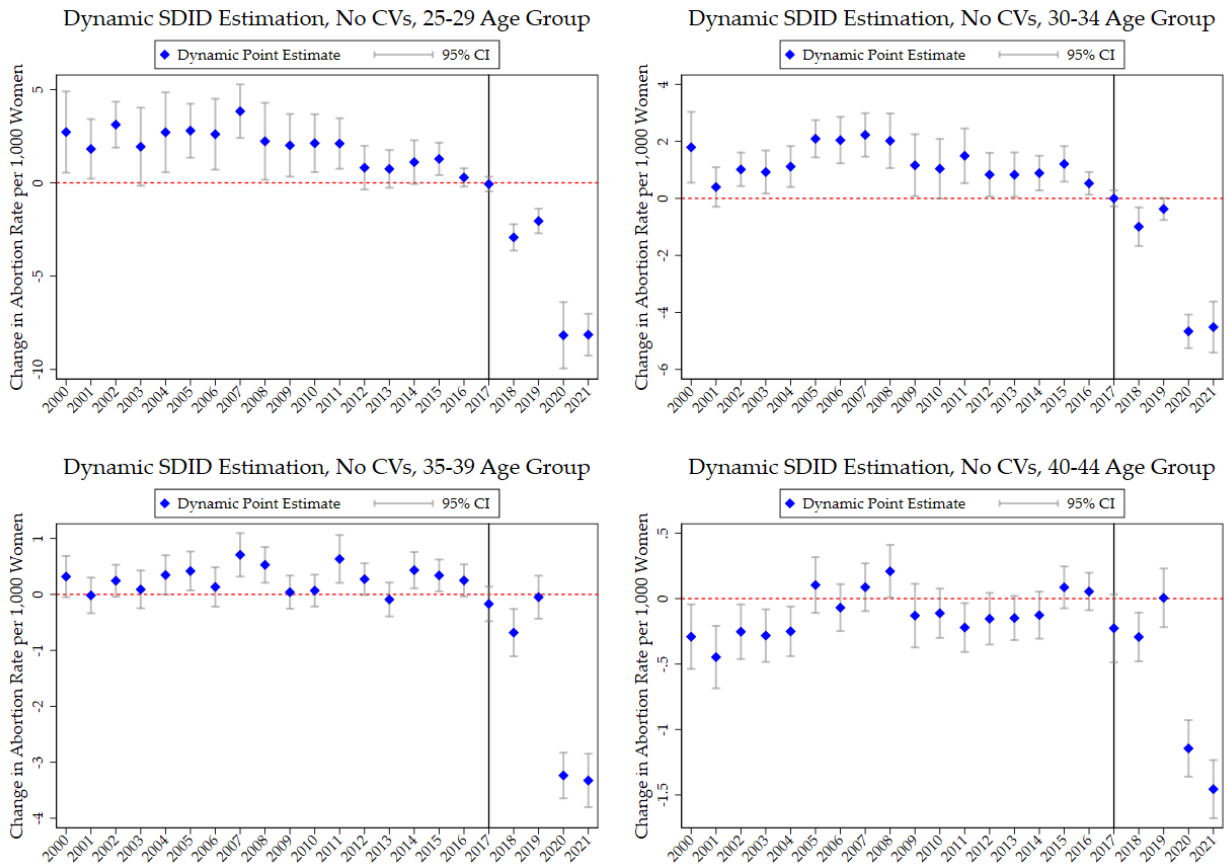


Figure 12: Dynamic SDID Estimation, Final Four Age Groups, No Covariates, New York



covariates may not shift our estimation by any significant amount,⁵² their sum contribution is to improve our estimation sufficiently such that we feel secure to infer from Figures 4 and 5.

When considering our covariates and their implementation into our exercise, we must decide between utilising the ‘projected’ and (default) ‘optimised’ procedures. As described in Section 4.3, we utilise projected errors as they better fit our covariates.⁵³ It is worth demonstrating however that this choice does not drastically affect our results, as the results of Table 2 are largely unchanged when optimised errors are employed instead, as shown in Table 6. The point estimates of this latter table are marginally more conservative, with a fall from 1% to 5% significance for the 25-39 age group particularly notable.

Table 6: SDID Results, Optimised Errors, Age-Disaggregated Abortion Rates per 1,000 women, 2000-2021

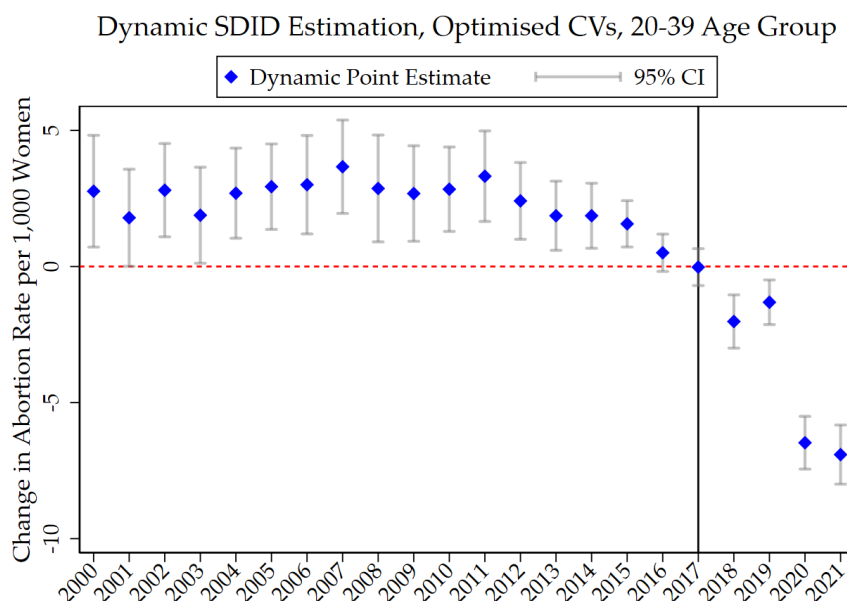
	(1) 20-39 yr.	(2) 20-44 yr.	(3) 25-39 yr.	(4) 20-24 yr.	(5) 25-29 yr.	(6) 30-34 yr.	(7) 35-39 yr.	(8) 40-44 yr.
NY-PFL Active	-4.346***	-3.713***	-3.258**	-7.433***	-5.167**	-2.562*	-1.781***	-0.601
<i>S.E.</i>	(1.432)	(1.210)	(1.288)	(1.902)	(2.058)	(1.333)	(0.660)	(0.417)
<i>N</i>	836	726	836	836	836	836	836	726
Clusters	38	33	38	38	38	38	38	33
Pre-Trt. Mean, N.Y.	33.608	27.815	27.759	51.657	40.232	27.135	15.932	5.723
Pre-Trt. Mean, Donors	17.869	14.044	13.284	30.222	21.802	11.561	5.705	1.872
Sample Outcome S.D.	6.218	5.095	5.024	10.420	7.696	4.887	2.769	1.097

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Payroll deductions began 2017/7/1, benefits become available 2018/1/1. *N* is the number of state-year-age group abortion rate observations. ‘Clusters’ refers to the number of states in each regression and includes the single treated state. Estimates are drawn from a restricted time sample from 2000 to 2021, and are based off 1000 placebo repetitions of this sample. ‘Trt.’ = Treatment. ‘S.D.’ = Standard Deviation of outcome for all clusters.

To further evidence why we prefer projected errors, Figure 13 provides an example of the matching we achieve under optimised errors for our main 20-39 age group:

Figure 13: Main SDID Estimation with Optimised Errors, 20-39 Age Group, New York



⁵²For this reason neither the possibility of ‘bad’ controls nor our inability to include other interesting covariates - such as Rice et al. (2022)’s contraceptive index (as this only goes back to 2006) - are pressing concerns. In addition, remark that we already exclude possible controls that could feasibly be affected by NY-PFL, such as mean state household income.

⁵³Incidentally they also calculate much faster.

Where, similar to excluding the covariates, the matching we achieve is inferior to the upper left panel of Figure 4, with a stronger positive pre-trend. This fact, a result of the theoretical justification outlined by Kranz (2022), pushes us to prefer projected errors.

A.4 Additional Robustness Tests

Alongside the placebo treatments considered in Section 5.4, where we showed that results became insignificant when New York's treatment date was moved to a previous year, we now additionally illustrate the distribution of placebo ATT estimates created by assigning treatment to each donor at the true time of treatment (2018), following Melo et al. (2023). Upon each unit we run our main regression with all covariates and projected errors, the only change being that we run our 1000 repetitions with no inference rather than placebo inference as this minimises crashes in calculation and does not affect the ATTs we now plot.

Figures 14 and 15 hence plot our ATT estimate for New York (shown in red) against those of each donor. They demonstrate that the large negative effect on abortion rates observed in 2018 for New York stands out among our placebos, located near or at the top of the distribution for seven of our eight age groups. Moreover Missouri, the only state that consistently produces comparably large negative ATT estimate across our age groups, only does so due to almost totally ceasing abortion provision during the post-treatment period (Kortsmit et al., 2022). While New York's outcome for the 40-44 age group is nested within the wider distribution, this is the group for whom we observe the smallest (and least significant) effect, so this result is not of much concern. All 8 distributions are approximately normally distributed around zero, although the small number of units in our dataset limits the shape of the distribution. Overall, Figures 14 and 15 provide further proof of the validity of our conclusion that age-disaggregated abortion rates fell in New York following the introduction of NY-PFL in 2018, as this effect is not observed for other states or for other time periods, decreasing the likelihood that our results are simply part of some overarching state or national trend.

Furthermore to our placebo tests using other states and other years, we can also check to see if a significant effect is observed on fertility rates, another key birth-related outcome. We hence next replace abortion rates in our main regressions with fertility rates, leaving all other aspects of the our specification constant:

Table 7: SDID Results, Age-Disaggregated Fertility Rates per 1,000 women, 2000-2021

	(1) 20-39 yr.	(2) 20-44 yr.	(3) 25-39 yr.	(4) 20-24 yr.	(5) 25-29 yr.	(6) 30-34 yr.	(7) 35-39 yr.	(8) 40-44 yr.
NY-PFL Active	0.858	1.202	0.324	3.788	-0.512	0.824	0.388	1.169**
<i>S.E.</i>	(1.350)	(1.426)	(1.425)	(2.455)	(2.577)	(2.251)	(1.310)	(0.566)
<i>N</i>	990	990	990	990	990	990	990	990
Clusters	45	45	45	45	45	45	45	45
Pre-Trt. Mean, N.Y.	80.555	66.747	84.164	69.938	92.717	102.08	58.403	13.770
Pre-Trt. Mean, Donors	85.875	69.700	83.096	82.491	111.97	97.184	46.060	10.936
Sample Outcome S.D.	9.278	7.951	9.916	23.830	17.840	12.600	8.901	2.499

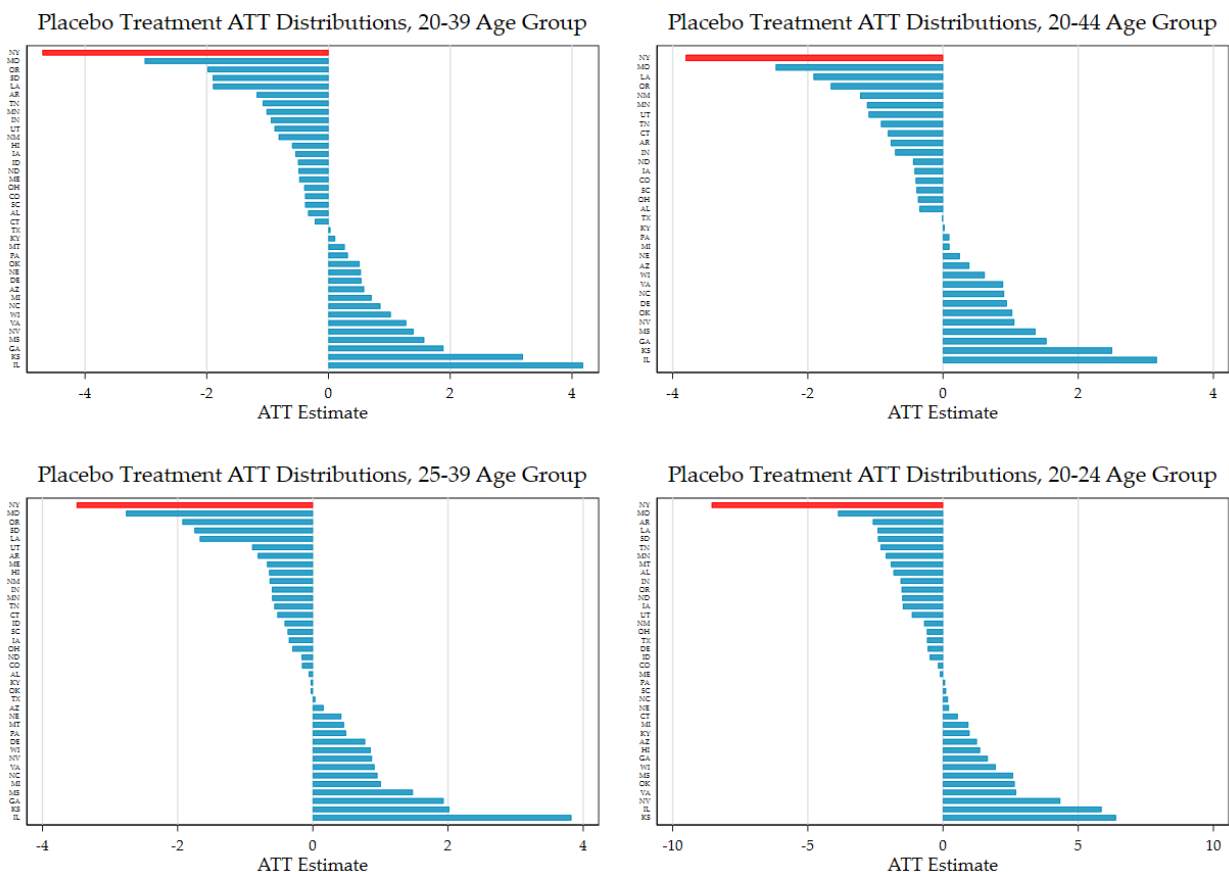
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Payroll deductions began 2017/7/1, benefits become available 2018/1/1. *N* is the number of state-year-age group fertility rate observations. 'Clusters' refers to the number of states in each regression and includes the single treated state. Estimates are drawn from a restricted time sample from 2000 to 2021, and are based off 1000 placebo repetitions of this sample. 'Trt.' = Treatment. 'S.D.' = Standard Deviation of outcome for all clusters.

We observe very limited significance here, with only the effect for our 40-44 age group found to be significant at the 5% level. Magnitudes and signs are what could be expected however, with a small increase in fertility accompanying our declines in abortion. Overall, these results align well with the small, positive effect observed by Thunell (2017) and Golightly and Meyerhofer (2022) following the introduction of CA-PFL, despite their limited significance. Repeating this exercise with monthly rather than annual fertility rates to increase our sample size - as Golightly and Meyerhofer (2022) do - may produce significant results, and so represents an interesting possible extension to our work.

Finally as raised in Section 6, consider the role of possible role of migration as a confounder to our analysis. More specifically, consider if the introduction of NY-PFL led to an increase in migration to New York as prospective parents sought to take advantage of the incoming benefits. We could be then observing a selected sample in New York that was less likely to terminate a pregnancy due to specifically moving to the state in order to have a child and utilise NY-PFL. Such a situation would also colour the control groups in our analysis as the outward migration would change the composition of their population in a selective way.

Figure 14: Placebo Treatment ATT Distributions, First Four Age Groups

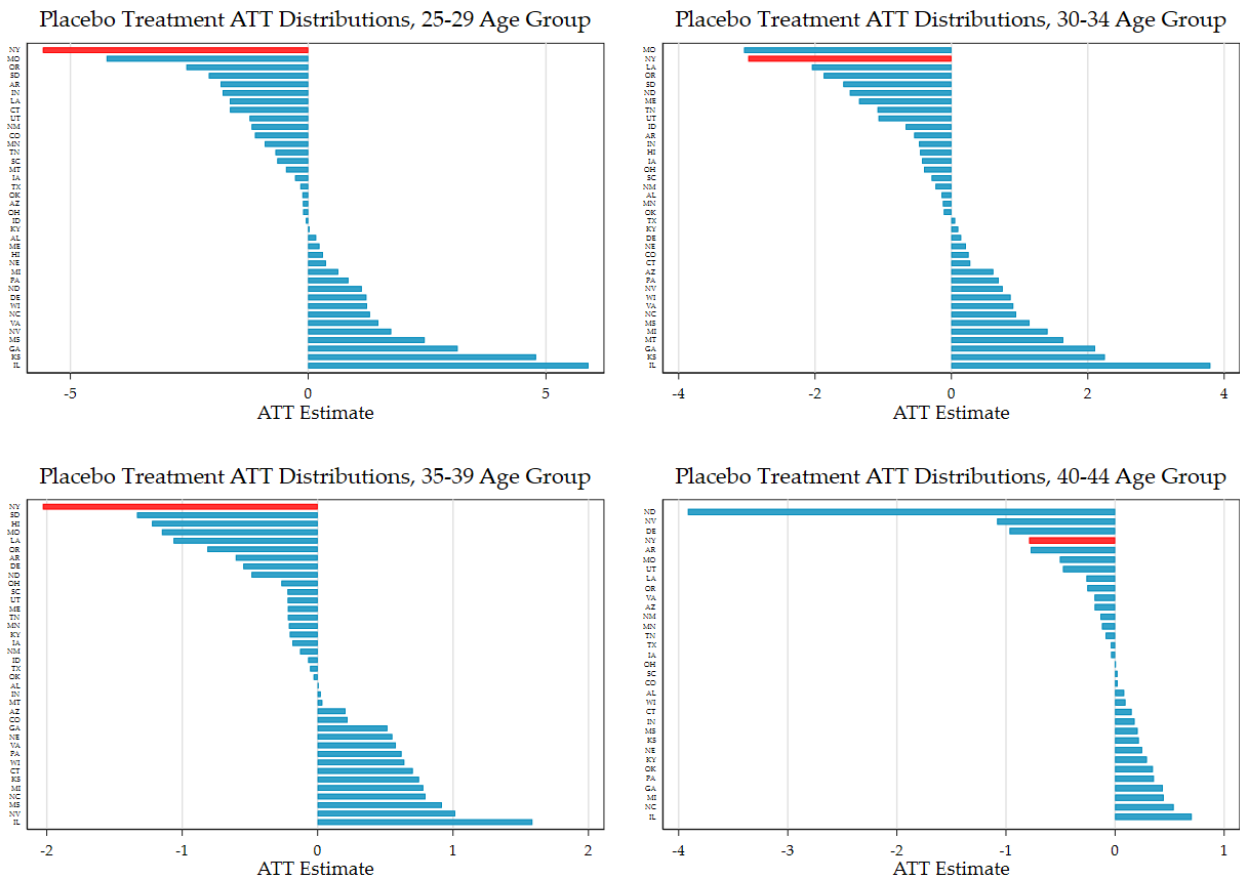


Notes: New York ATT shown in red. All states with complete abortion data are included except other treated states. Results based off 1000 repetitions of the full 2000-2021 sample period with covariates from our main specification using no inference option.

Some aspects of our exercise argue against this already. Firstly we already control for the percentage of the state population that is female and of reproductive age as a covariate, and this group additionally forms the denominator in our ethnicity and education variables. As such we already factor in the possible migration of fertile women throughout the sample period. Second, Section 6 addresses that there was no

decline in the share of New York abortions conducted on out-of-state residents in the post-treatment period, implying that women who previously lived close enough to New York to pursue abortions there did not take the chance to relocate in-state to take advantage of NY-PFL.⁵⁴ In addition, we now also run a direct test to see if there is any significant effect on the number of state residents who lived in a different state one year prior for New York relative to our control states, using data from the American Community Survey sourced from U.S. Census Bureau (2023d). We perform two regressions here, 2005-2019 and 2010-2019,⁵⁵ using our regular SDID method and covariates.

Figure 15: Placebo Treatment ATT Distributions, Final Four Age Groups



Notes: New York ATT shown in red. All states with complete abortion data are included except other treated states. Results based off 1000 repetitions of the full 2000-2021 sample period with covariates from our main specification using no inference option.

⁵⁴There are also intuitive reasons to suspect that widespread migration into New York for financial gain through NY-PFL is unreasonable. First, NY-PFL benefits are funded entirely through employee payroll deductions (A Better Balance, 2023), meaning that only the job protections offered by the programme could represent true financial gain. Second, several states nearby to New York either already had their own established PFL programmes (NJ, RI) or would launch them shortly after NY (MA, CT), diminishing the chance that others would be convinced to move to New York over their current residence by NY-PFL. Third, New York has one of the highest average cost of living rates in the U.S. (Missouri Economic Research and Information Center, 2023), likely offsetting any gain from NY-PFL benefit through more costly general expenditure. Hence while young couples may move to New York in search of financial betterment through higher salaries, there is no intuitive reason why the introduction of NY-PFL affected this pull.

⁵⁵Reporting of ACS migration statistics by the Census Bureau from 2010 sum the total of state residents who lived in a different state the year prior directly, while for 2005-2009 individual state totals must be summed manually. For the majority of states (including New York) the 2005-2009 and 2010-2019 totals are very comparable, but for some (e.g. Alaska) they are quite disparate. For this reason we include 2010-2019 alongside 2005-2019; both regressions end at 2019 as collection of the 2020 ACS was disrupted by the pandemic, leading the Census Bureau to not report ACS migration statistics for 2020.

Table 8: SDID Results, Number of State Residents who lived in a Different State the year prior, 2005/10-2019

	(1) 2005-2019	(2) 2010-2019
NY-PFL Active	-952.7	-10100.0
<i>S.E.</i>	(10403.3)	(10128.1)
<i>N</i>	675	450
Clusters	45	45
Pre-Trt. Mean, N.Y.	332998	269489
Pre-Trt. Mean, Donors	175019	117257
Sample Outcome S.D.	6.034	5.560

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: NY-PFL payroll deductions began 2017/7/1, benefits become available 2018/1/1. *N* is the number of state-year-age group abortion rate observations. 'Clusters' refers to the number of states in each regression and includes the single treated state. Estimates are based off 1000 placebo repetitions of each sample. 'Trt.' = Treatment. 'S.D.' = Standard Deviation of outcome for all clusters.

We find no significant effect for New York here, in fact observing negative point estimates rather than positive ones. These estimates align with the negative net migration for New York from 2020-2022 reported by the Census Bureau (2023c), and provide further evidence against a possible migration channel.