The Effects of Paid Parental Leaves on the Labor Force Participation of Mothers

Sara Vargas

Undergraduate Honors Thesis in Economics

Abstract:

This paper investigates the impact of paid parental leave policies on the labor force participation (LFP) of mothers in the United States. Using a dynamic difference-in-differences (DiD) framework applied to Survey of Income and Program Participation (SIPP) data from 1996 to 2013, I examine the effects of state-level paid family leave programs implemented in California and New Jersey. The analysis traces changes in maternal LFP before and after childbirth, comparing states in which the policy was implemented and states without policy effect. I explore heterogeneity by race and marital status, finding that paid leave policies significantly increase LFP around childbirth, with the strongest effects concentrated among people of color (POC) and unmarried women. The results suggest that paid leave can reduce short-term labor market detachment and promote more equitable outcomes across different demographic groups.

1. Introduction

The increased participation of women in the labor market has been one of the most significant economic shifts in developed countries over recent decades. Despite these advancements, persistent structural barriers continue to drive a substantial gender wage gap, with childbearing identified as a critical factor exacerbating labor market inequalities (Kleven et al.; Goldin; Anderson et al.). The transition to parenthood often marks a turning point, where women disproportionately bear the economic costs associated with caregiving responsibilities, leading to what is widely known as the motherhood penalty. This penalty resulted in reduced earnings, diminished career advancement opportunities, and increased part-time employment among mothers, stemming from career interruptions and limited access to supportive workplace policies.

Claudia Goldin underscores "the portion of the difference in earnings by gender that was once due to differences in productive characteristics has largely been eliminated" (Goldin 1116) , reaffirming her later claim that women's wage penalty is due to the inflexibility of jobs in the current job market disproportionately penalizing women despite their demonstrated convergence with and, in some cases, outperformance of men as labor force participants (1118). There is also evidence of compounding long-term earnings losses resulting from even brief periods out of the labor force. Anderson, Binder, and Krause found that the wage penalty varies significantly by education level: "Mothers who did not complete high school do not earn less than their childless counterparts, while high-school and college graduates earn about 10 percent less per child" (Anderson et al. 357). These findings highlight the heterogeneity of the penalty's effects across subgroups of mothers, a disparity further exacerbated by differences in human capital. As the authors show, “years out of the workforce alone explain 40 percent of the gap” (357). This conclusion corroborates Goldin’s claim that time out of the workforce, often driven by the need for temporal flexibility following childbirth, is a major contributor to the gender earnings gap. As Goldin explains, “lower hours mean lower earnings in a nonlinear (convex) fashion,” and for many women, particularly those with higher-earning spouses, this results in “lower labor force participation” (Goldin 1109). Together, these findings underscore how structural features of the labor market amplify wage penalties for mothers, particularly those with higher levels of education and earning potential. These challenges are particularly salient in the United States, where federal paid family leave policies are absent, leaving many workers without adequate support.

The only federal protection for workers balancing work and family responsibilities is the Family and Medical Leave Act (FMLA), passed in 1993. The FMLA allows eligible employees to take up to 12 weeks of unpaid leave per year for specific family and medical reasons, including childbirth, adoption, or caring for a seriously ill family member. While this was a landmark policy for its time, offering job protection during leave, it does not provide wage replacement, limiting its accessibility for low-income workers who cannot afford to take unpaid time off. Moreover, many workers are excluded from the FMLA due to eligibility requirements, which mandate that employees must work for a company with at least 50 employees and meet certain tenure and hours-worked thresholds. These gaps in coverage leave millions of U.S.

workers, particularly those in low-wage jobs or employed by small businesses, without any guaranteed family leave ("FMLA") .

In the absence of a federal paid leave policy, several states have taken the initiative to implement their own paid family leave programs, aiming to address the limitations of the Family and Medical Leave Act (FMLA). California and New Jersey were among the first to adopt such policies, building on their long standing Temporary Disability Insurance (TDI) systems to offer wage replacement for workers taking leave to bond with a new child or care for a seriously ill family member. California's Paid Family Leave (C-PFL), introduced in 2004, initially provided six weeks of leave at 55% of a worker's wages, and has since been expanded to cover up to eight weeks at 60–70% wage replacement. New Jersey followed with its Family Leave Insurance (NJ-FLI) in 2009, offering up to 12 weeks of leave at 85% wage replacement.

In both states, paid family leave benefits are stacked on top of existing TDI benefits, especially for birth mothers. In California, birth mothers could receive 6–8 weeks of paid TDI leave for pregnancy-related disability, followed by 6 weeks of PFL bonding leave, totaling up to 12–14 weeks of paid leave. Non-birth parents were eligible for 6 weeks of PFL. Similarly, in New Jersey, eligible mothers could receive about 4 weeks of pre-birth and 6 weeks of postpartum TDI leave, followed by 6 weeks of NJ-FLI, for a combined 12 weeks of paid leave. Non-birth parents were also entitled to 6 weeks of bonding leave. These benefits do not guarantee job protection and are meant to be used at the same time as FMLA. The policies are meant to reduce career interruptions for caregivers and support the labor force attachment of mothers after childbirth.

In recent years, additional states—including New York, Washington, Massachusetts, Oregon, and Colorado—have enacted paid parental leave policies, and several more plan to implement similar programs in the near future. This growing trend signals a broader national movement toward more comprehensive family leave coverage.

However, empirical evidence shows that even brief interruptions in labor force participation can lead to long-term reductions in women's earnings (Anderson et al.). This raises an important question: do newly implemented paid leave policies help mitigate these disruptions and promote sustained employment among mothers, or might they inadvertently reinforce patterns of career detachment? Importantly, the effects of such policies are unlikely to be uniform. Women's responses to paid leaves may vary significantly depending on factors such as educational attainment, occupation, race and ethnicity, income level, and spousal earnings. As a result, evaluating the heterogeneous effects of state-level paid family leave programs is essential to understanding both their equity and effectiveness.

This paper addresses the central question: How do paid parental leave policies affect mothers' labor force attachment, and how do these effects differ across demographic subgroups? Using data from the Survey of Income and Program Participation (SIPP), a nationally representative panel survey well-suited for evaluating state-level policy changes, I apply a dynamic Difference-in-Differences (DiD) framework. The analysis compares labor force outcomes in California and New Jersey, where paid leave policies were implemented, to outcomes in control states such as Texas, Florida, and New York. By incorporating event-study methods and interaction terms for race and marital status, this study examines not only whether paid leave increases overall labor force participation, but also whether it contributes to reducing disparities in maternal employment outcomes. In doing so, this research aims to identify whether and for whom paid leave policies effectively promote equity in the labor market.

2. Literature review

Tanya Byker's 2016 article, *"Paid Parental Leave Laws in the United States: Does Short-Duration Leave Affect Women's Labor-Force Attachment?"* serves a suitable starting point for my research. Byker investigates the introduction of paid family leave policies in California and New Jersey using a Difference-in-Differences (DiD) methodology; she finds that even a relatively short leave duration can increase women's labor force attachment by 5–8 percentage points around childbirth, particularly for less-educated women. This indicates that the availability of paid leave has significant positive effects on workforce attachment. She further analyzes the effect on a subgroup of mothers who have less than a bachelor's degree to find that mothers with less than a bachelor's degree are the ones that see a greater difference in labor force participation. Therefore she decomposes the main labor force participation outcome into two separate regressions: one for the probability of being employed ("working") and another for the probability of actively searching for a job ("looking for work"). This decomposition, limited to the less-educated subgroup, shows that paid leave policies reduce job search in the year following childbirth and increase the likelihood that mothers remain attached to an employer, indicating the paid leave is successfully encouraging women with less than a bachelor's degree to remain in the labor force.

Building on Byker's (2016) framework, this study revisits her findings to examine whether the effects of paid family leave on women's labor force attachment differ across various subgroups of mothers. It focuses on heterogeneity in outcomes such as labor force participation, employment status, and job-search activity. By re-evaluating Byker's central question, how access to paid leave influences women's engagement with the labor market, this research introduces a more granular demographic analysis, specifically examining the experiences of Black and Asian mothers. Both of these groups are likely to face a disadvantage in the labor force market due to their racial background and caregiving responsibilities. This study will analyze the experience of married and unmarried women as previously illustrated in Goldin's conclusions, mothers who are not married and therefore have a different support system than that of their married counterparts and experience a different labor force market. Existing research highlights that mothers of color often face both racial wage gaps and motherhood penalties, resulting in disproportionately severe earnings losses compared to White mothers (Rose et al.). This paper seeks to determine whether paid parental leave policies mitigate or reinforce those disparities. In doing so, it contributes to a growing literature on the distributional equity of family policies and whether their implementation yields equitable benefits across racial groups and single parents, or unintentionally exacerbates existing inequalities.

3. Data, Methodology and Results:

A. Data: SIPP Panels 1996–2023

The primary dataset I will use to obtain employment and household relationship data for women who gave birth during the pre- and post-policy periods is the *Survey of Income and Program Participation (SIPP)*, a dataset produced by the U.S. Census Bureau.SIPP is a nationally representative longitudinal survey that provides detailed information on income dynamics, labor market activity, and household composition over time . This dataset is particularly well-suited for my research because it allows for both cross-sectional and longitudinal analysis. Specifically, SIPP facilitates comparisons across states that implemented paid parental leave policies and those that did not, while also offering panel data that enables tracking individuals' labor force participation trajectories over time. The SIPP dataset follows individuals and households over a 48-month period making this availability of multiple observations per individual permit a more robust analysis of employment dynamics surrounding childbirth, enhancing the credibility of causal inferences regarding policy effects. I will be drawing data from 4 panels that provide data before the policy and after. These panels will be the 1996, 2000, 2004 and 2008 panels. This window captures periods before and after the implementation of paid family leave laws in California (2004) and New Jersey (2009). The states that constitute the treatment group are California and New Jersey whereas the control group are women in states without parental leave policy, including Texas, Florida and New York. The sample is composed of 103, 624 women aged 24 to 45 who give birth during one of the panels. To construct this dataset I used the household identification number to match children born into a household to their mothers.

The SIPP dataset has a detailed labelling of employment status amounting to 8 possible categories a subject can fall into. Taking a parental leave is encoded into SIPP as remaining in the labor force and working since the mother maintains a relationship with their employer. Furthermore the status in which a subject is in the labor force is noted by specifying if they are looking for a job or working. I will use this differentiation in participation to decompose the effect of the paid leave in the short term and specifically quantify its effects on working or looking rates after childbirth.

This dataset provides rich demographic information for each observation, which I will use to decompose the regression model by subgroups of mothers to assess potential heterogeneous effects of the policy on labor force participation. Specifically, I will examine differences by race (“ERACE”) and marital status (“EMS”). Existing literature has largely overlooked these demographic differences in evaluating the labor force effects of paid family leave policies. Therefore, my analysis contributes to the literature by investigating whether the policy has differential impacts across these groups of mothers.

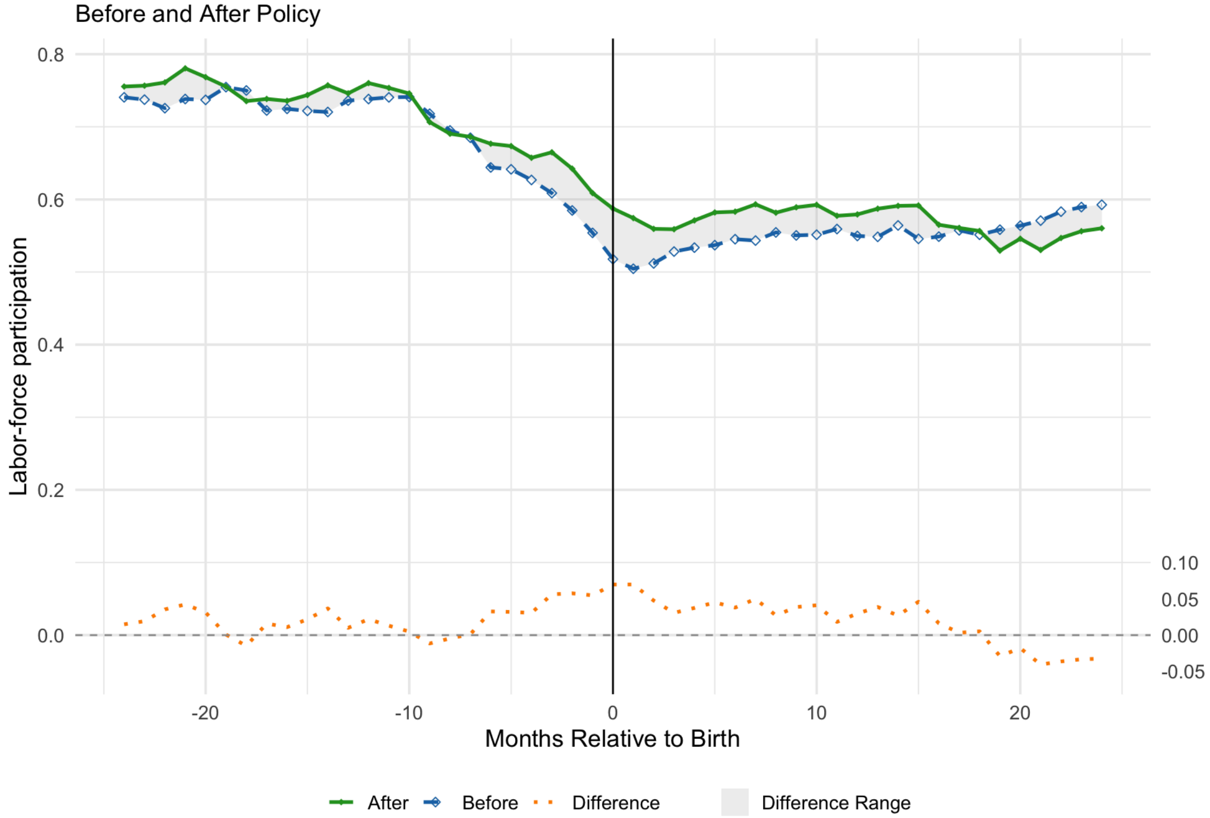
B. Empirical Model : Assessing LFP Effect on Women

The longitudinal structure of the SIPP allows me to track individuals over time, enabling comparisons of each woman’s labor-force status before and after childbirth, particularly I will be looking at 2 years before and after childbirth. To analyze the dynamic change of labor force participation around the date of childbirth I will use a dynamic difference in difference framework and I will index the event time to be the month relative to the childbirth date, which will span from -24 to +24. This way I aim to estimate the difference-in-difference coefficients that attribute to the effect of the policy on the labor force participation. I apply this framework into a fixed effects ordinary least squares linear model where I include fixed effects to target the time invariant unobserved heterogeneity and enable within individual comparisons across time.

Figure 1 illustrates why a dynamic difference-in-differences (event-study) approach is well-suited to evaluate the impact of paid parental leave policies on women's labor-force participation. The figure plots average labor-force participation (LFP) for women who gave birth before and after the policy was implemented in California and New Jersey, covering a window of 24 months before to 24 months after childbirth. For most of the observation window, particularly in the distant months from birth, the pre- and post-policy LFP trajectories appear similar, indicating parallel trends and supporting the validity of the identification strategy.

Figure 1. LFP Around Birth in CA and NJ

Before and After Policy



The figure shows labor force participation (LFP) trends in California and New Jersey around childbirth, comparing periods before and after the implementation of the paid parental leave policy. The solid green line represents LFP after the policy, while the dashed blue line shows LFP before. The dotted orange line reflects the difference between after and before.The shaded area represents the range of differences. A noticeable drop in LFP occurs around birth, with a slight improvement after policy implementation. However, the difference remains small and fluctuates within a small range.

A notable divergence emerges in the months surrounding childbirth, particularly between −6 and +6 months, where LFP dips for both groups. Formal slope comparisons indicate that the decline in labor-force participation between six months before and the month of birth was significantly steeper prior to the policy's implementation. Specifically, LFP fell by approximately 2.1 percentage points per month pre-policy versus 1.5 percentage points post-policy, supporting the interpretation that paid leave laws helped moderate labor-market exits around childbirth. This visual pattern is consistent with the interpretation that the policy increased labor-force attachment by allowing women to remain formally employed, even if temporarily away from work. These graphical patterns motivate formal estimation of month-by-month treatment effects using the event-study specification.

To estimate the effect of paid family leave policies on women's labor-force participation, I use a dynamic difference-in-differences (event-study) framework. Leveraging a series of event time dummies which interact with the treatment (policy) determinant dummy to obtain effects per month.

Simple event study framework

$$(0)\ LFP_{its} = a_i + \lambda_t + \theta_s \times \lambda_t + \sum_{m=-24}^{+24} \beta_j B_{it}^{m} \times PostPolicy_{ts} + \sum_{m=-24}^{+24} \theta_s \times B_{it}^{m} + \epsilon_{its}$$

Event Study with added controls

$$(1)\ LFP_{its} = a_i + \lambda_t + \theta_s \times \lambda_t + \sum_{m=-24}^{+24} \beta_j B_{it}^{m} \times PostPolicy_{ts} + \sum_{m=-24}^{+24} \delta_j B_{it}^{m}$$

$$+ \sum_{m=-24}^{+24} \pi_j B_{it}^{m} \times \theta_s + \sum_{m=-24}^{+24} \gamma_j B_{it}^{m} \times \lambda_t + \epsilon_{its}$$

The outcome variable, LFP_{its}, is a binary indicator equal to 1 if woman *i* in state *s* is in the labor force at time *t*, and 0 otherwise. Given the longitudinal structure of the data, fixed effects need to be specified to isolate the causal impact of the paid leave policy. Individual fixed effects a_i are included to account for time-invariant unobserved heterogeneity across individuals. By controlling for individual fixed effects , the estimation leverages individual variation over time and is able to remove bias from any confounding characteristics. Year fixed effects, λ_t capture nationwide time shocks common to all observations in that year. These shocks can include any macroeconomic fluctuations that could affect the economic state of the nation which might confound the state of the policy. The third fixed effect is the state-by-time fixed effect $\theta_s \times \lambda_t$ which absorb any time-varying factors at the state level. These include business cycle dynamics in state and other policy factors in that state which may influence the LFP in state.

I constructed a series of dummy variables, B_{it}^{m}, for the months around the child's birth excluding -24 to -18 months as the reference period. To select those who were treated by the policy I construct the binary dummy variable $PostPolicy_{ts}$ which is evaluated 1 for those observations which happened in California after July 2004 or in New Jersey after July 2009, any other state and time would evaluate post-policy to zero. The coefficients estimated as β_j for each interaction between event time j and the post-policy indicator capture the treatment effects of the paid leave laws for that month . Specifically, each coefficient captures the differential change in labor-force participation between the treated and control groups at event month j, relative to a pre-birth baseline, after accounting for fixed differences and time-varying state-level factors. These coefficients trace out how the policy impacted labor-force attachment month-by-month around childbirth, allowing us to visualize both the timing and duration of the effect. The interaction terms between relative month and either year or state serve as control variables that account for time-varying shocks to labor-force participation that may differ by state or calendar year. Specifically, these interactions allow each state and each year to have its own distinct labor-force trajectory around childbirth, ensuring that the estimated policy effects are not confounded by other trends unrelated to the paid leave laws. This specification allows me to compare within-individual changes in labor-force participation before and after childbirth across policy and non-policy periods, while flexibly controlling for both individual and aggregate time trends.

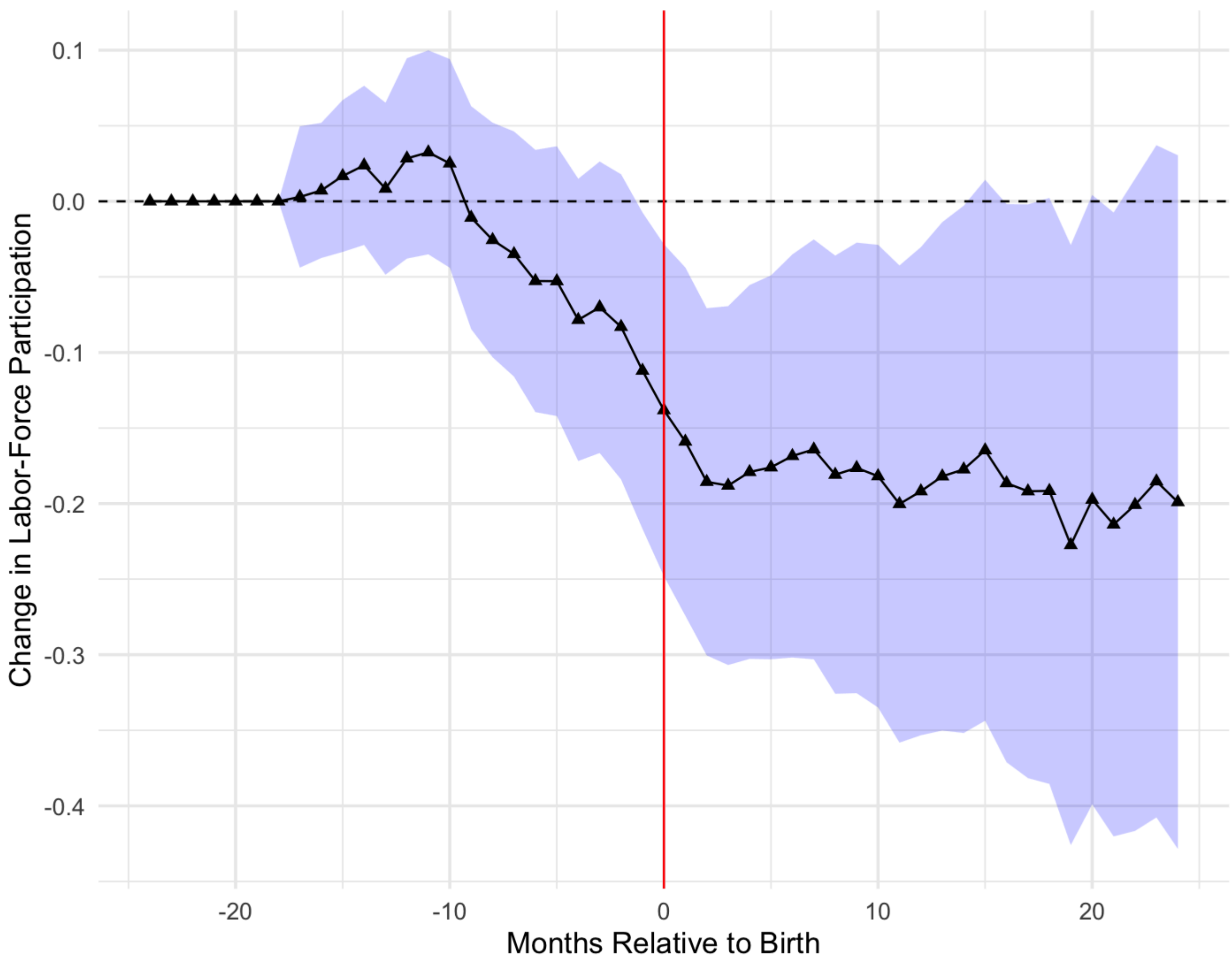
To determine the validity of the coefficients estimated in the dynamic DiD framework, I perform formal significance tests. While examining each event-time coefficient individually may help identify when effects arise, this approach fails to account for the multiple testing problem, which increases the likelihood of false positives due to random variation. Since the research question centers on whether there is a sustained effect over a post-treatment window, joint hypothesis testing is necessary as it allows for a formal evaluation of whether the treatment had an overall statistically significant impact across a defined time period, rather than relying on isolated significant coefficients that may reflect noise. To conduct this test, I use a Wald test, which assesses whether a group of coefficients, the post-treatment event-time indicators, are jointly different from zero.The Wald statistic follows a Chi-squared distribution under the null hypothesis and remains valid even in the presence of heteroskedasticity, which is commonly observed in panel data structures like the SIPP, where error variance can vary across individuals and over time. This test offers a more reliable basis for inference, particularly in models where individual standard errors may be inflated due to model complexity or correlated residuals. As a result, the joint test provides a clearer picture of the policy's cumulative effect on labor force participation around childbirth.

C. Result Effects of Policy on Overall Labor Force Participation of Mothers

Model 0 provides the basic two-way fixed effects ordinary least squares (OLS) model which displays the event study design estimating the effect of the policy on the labor force participation of the woman at that month relative to childbirth.

Figure 2: Model 0 – Basic Two-Way Fixed Effects Regression

Estimated Change in Labor Force Participation Around Childbirth



The figure shows estimated changes in labor force participation (LFP) from a basic two-way fixed effects model. The black line and points show the point estimates relative to birth month (red vertical line), while the shaded blue area represents the 95% confidence interval around the estimates. Estimates are not statistically significant where the confidence intervals cross zero. However we see tat around birth the estimates are statistically significant as well as during post birth months which indicate there is a persisten decrease in LFP after the policy.

Figure 2, shows a plot of the event-study estimates evaluating the change in labor force participation (LFP) among individuals relative to a pre-treatment baseline of 24 to 18 months before childbirth. The x-axis denotes time in months relative to birth (with birth at month 0, marked by a vertical red line), and the y-axis captures the estimated change in LFP. Prior to -4 months before the event (birth) the estimates are statistically indistinguishable from zero, supporting the validity of parallel trends, this can be observed in Table 1 where an excerpt of the model summary of Model 0 shows the estimate for the coefficients we are interested in, β_j in Regression (0). Starting at 4 months before birth to onwards we see a persistent decline in LFP reaching approximately 15 to 20 percentage point reduction within the first year post-birth. These effects remain negative and stable through 24 months after childbirth suggesting a sustained postnatal impact on labor force participation after the policy was enacted. This dynamic pattern indicates that the paid parental leave policy exerts a substantial and enduring influence on labor market attachment, with limited recovery in the two years following birth.

Table 0: Wald Tests: Post-Policy Effects (Model 0)

Window	Months Relative to Birth	# Terms	DF	Wald F-Stat	p-Value	VCOV
−2 to +3	-2 to 3	6	6 and 81784	3.780	0.0009131	Clustered (sippid)
−3 to +3	-3 to 3	7	7 and 81784	3.703	0.0005217	Clustered (sippid)
−6 to +6	-6 to 6	13	13 and 81784	2.667	0.0009545	Clustered (sippid)
Full Post (−17 to +24)	-17 to 24	42	42 and 81784	2.264	< 1e-04	Clustered (sippid)

Note: Wald tests show post-policy effects are jointly significant at the 95% level.

Table 1: Model 0 Post-Policy × Event Time Coefficients

Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	0.003	0.024	0.903		[-0.044, 0.05]
-16	0.007	0.023	0.752		[-0.038, 0.052]
-15	0.017	0.026	0.513		[-0.033, 0.067]
-14	0.024	0.027	0.375		[-0.029, 0.076]
-13	0.008	0.029	0.773		[-0.049, 0.065]
-12	0.028	0.034	0.401		[-0.038, 0.095]
-11	0.032	0.034	0.346		[-0.035, 0.1]
-10	0.025	0.035	0.476		[-0.044, 0.094]
-9	-0.011	0.038	0.774		[-0.085, 0.063]
-8	-0.026	0.040	0.519		[-0.103, 0.052]
-7	-0.035	0.041	0.399		[-0.116, 0.046]
-6	-0.053	0.044	0.234		[-0.139, 0.034]
-5	-0.053	0.046	0.246		[-0.142, 0.036]
-4	-0.078	0.048	0.100	*	[-0.172, 0.015]
-3	-0.070	0.049	0.154		[-0.167, 0.026]
-2	-0.083	0.052	0.107		[-0.184, 0.018]
-1	-0.112	0.053	0.036	**	[-0.217, -0.007]
0	-0.138	0.056	0.014	**	[-0.248, -0.028]
1	-0.159	0.059	0.007	***	[-0.274, -0.044]
2	-0.186	0.059	0.002	***	[-0.3, -0.071]
3	-0.188	0.061	0.002	***	[-0.307, -0.069]
4	-0.179	0.063	0.005	***	[-0.303, -0.055]
5	-0.176	0.065	0.007	***	[-0.303, -0.049]
6	-0.168	0.068	0.013	**	[-0.302, -0.035]
7	-0.164	0.071	0.021	**	[-0.303, -0.025]
8	-0.181	0.074	0.015	**	[-0.326, -0.036]
9	-0.176	0.076	0.020	**	[-0.325, -0.027]
10	-0.182	0.078	0.020	**	[-0.335, -0.029]
11	-0.200	0.081	0.013	**	[-0.358, -0.042]
12	-0.192	0.082	0.020	**	[-0.353, -0.03]
13	-0.182	0.086	0.034	**	[-0.35, -0.014]
14	-0.177	0.089	0.046	**	[-0.352, -0.003]
15	-0.165	0.091	0.071	*	[-0.344, 0.014]
16	-0.186	0.094	0.048	**	[-0.371, -0.002]
17	-0.192	0.097	0.048	**	[-0.382, -0.002]
18	-0.192	0.099	0.053	*	[-0.385, 0.002]
19	-0.227	0.101	0.025	**	[-0.426, -0.029]
20	-0.197	0.103	0.055	*	[-0.399, 0.004]
21	-0.214	0.105	0.042	**	[-0.42, -0.007]
22	-0.201	0.110	0.068	*	[-0.417, 0.015]
23	-0.185	0.113	0.103		[-0.408, 0.037]
24	-0.199	0.117	0.089	*	[-0.428, 0.03]

Note: Estimates are derived from the simple dynamic DiD specification in regression 0. Standard errors are clustered at the individual level.Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

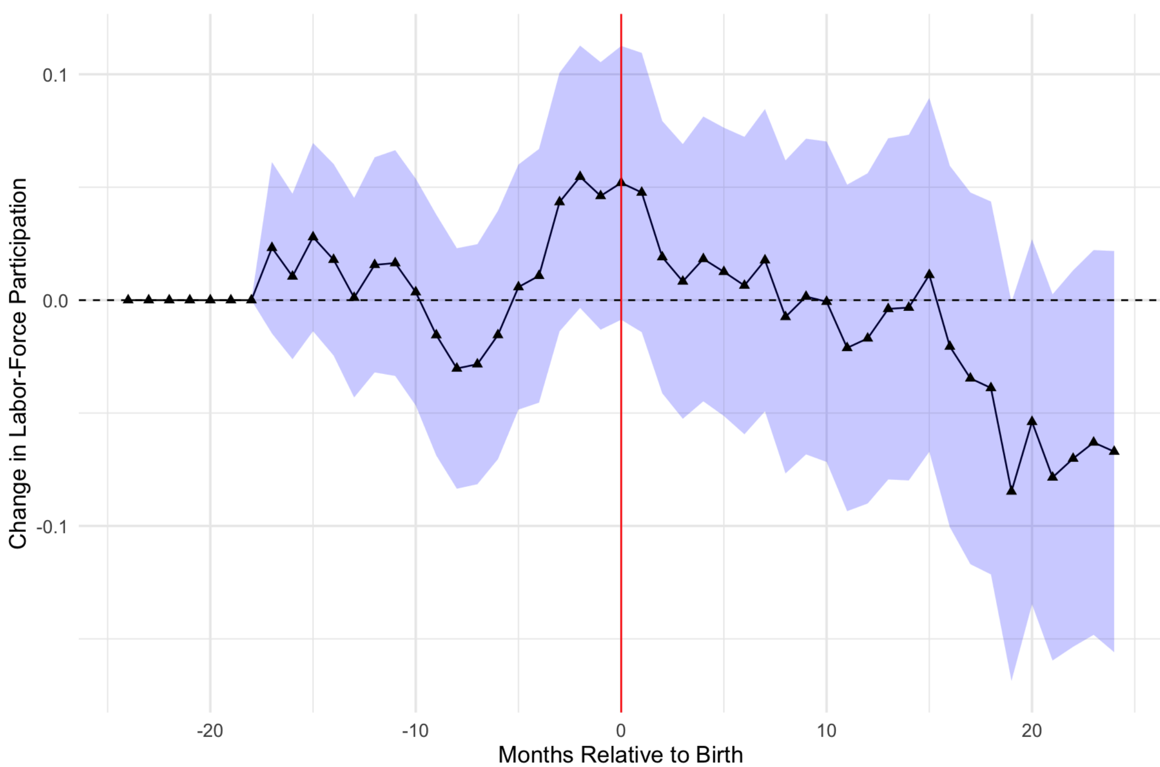
To assess the overall impact of the policy, Table 0 explains the results of a series of joint significance tests on different time windows for the event-time coefficients. The window from -3 to +3 months relative to birth captures the period directly surrounding childbirth, when labor force attachment is most likely to be affected. To evaluate whether the cumulative treatment effect in this window is statistically significant I ran a Wald test on the time window which resulted in a statistically significant result with a p-value of 0.0005217 indicating we can reject the null hypothesis. The other time windows also yield low p-values, indicating statistically significant cumulative effects. Among these, the −3 to +3 window exhibits the strongest joint significance. These results reinforce the conclusion that the policy's impact is not driven by noise in individual coefficients but reflects a systematic and statistically meaningful effect over time and estimates the persistent decrease of labor force participation during the months after birth.

Model 0 estimates the dynamic treatment effect around childbirth while controlling for time-invariant unobserved heterogeneity through individual and state fixed effects, as well as common shocks through time fixed effects. These fixed effects assume that any differences across individuals, states, or calendar months are constant and additive, but they do not account for state-specific time-varying confounders that could bias the dynamic treatment effect estimates, which are likely confounding factors. To address this potential limitation I consider Model 1 which addresses this assumption by interacting the month event-time dummies with both state and time fixed effects, allowing for more flexible pre-treatment dynamics and controlling for time-varying shocks that may differentially affect states. This helps identify whether observed effects are due to the treatment or instead reflect state or time specific trends that align with the event.

In Figure 3, the event study coefficients estimated from regression equation (1) are displayed. The trend these coefficients follow closely resembles the raw difference in labor-force participation observed in Figure 1, suggesting that the regression is capturing the underlying treatment effect in a more rigorous and controlled way. The interaction terms added provide baseline dynamics around childbirth to vary across states and different months in a year. These controls help isolate the true impact of the policy by accounting for time-varying state-level and temporal confounders that might otherwise bias the event-study estimates. However, The figure shows large confidence intervals including zero and therefore indicating they are not statistically significant.

Figure 3: Model 1 TWFE Including Controls

Estimated change in LFP due to paid parental leave policy relative to the baseline in months relative to childbirth



Estimates come from a TWFE model with dynamic controls for months in a year and for state.
The estimated trend shows an increase in LFP during the immediate months around childbirth and a decrease in the long term.
However, the estimates exhibit large 95% confidence intervals which include 0 and therefore do not provide statistically significant evidence.

Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	0.023	0.025	0.359		[-0.026, 0.073]
-16	0.010	0.024	0.667		[-0.037, 0.058]
-15	0.028	0.028	0.315		[-0.027, 0.082]
-14	0.018	0.028	0.526		[-0.037, 0.073]
-13	0.001	0.029	0.970		[-0.057, 0.059]
-12	0.016	0.032	0.623		[-0.047, 0.078]
-11	0.016	0.033	0.622		[-0.049, 0.082]
-10	0.004	0.033	0.916		[-0.062, 0.069]
-9	-0.015	0.036	0.663		[-0.085, 0.054]
-8	-0.030	0.035	0.394		[-0.1, 0.039]
-7	-0.028	0.035	0.423		[-0.098, 0.041]
-6	-0.015	0.037	0.674		[-0.087, 0.056]
-5	0.006	0.036	0.873		[-0.065, 0.077]
-4	0.011	0.037	0.773		[-0.063, 0.084]
-3	0.043	0.038	0.255		[-0.031, 0.118]
-2	0.055	0.039	0.158		[-0.021, 0.131]
-1	0.046	0.039	0.243		[-0.031, 0.124]
0	0.052	0.040	0.199		[-0.027, 0.131]
1	0.048	0.041	0.248		[-0.033, 0.128]
2	0.019	0.040	0.635		[-0.06, 0.098]
3	0.008	0.041	0.838		[-0.071, 0.088]
4	0.018	0.042	0.664		[-0.064, 0.101]
5	0.013	0.043	0.768		[-0.071, 0.096]
6	0.006	0.044	0.883		[-0.08, 0.093]
7	0.018	0.045	0.692		[-0.07, 0.105]
8	-0.007	0.046	0.872		[-0.098, 0.083]
9	0.002	0.047	0.973		[-0.09, 0.093]
10	-0.001	0.047	0.989		[-0.093, 0.092]
11	-0.021	0.048	0.661		[-0.116, 0.073]
12	-0.017	0.049	0.728		[-0.112, 0.079]
13	-0.004	0.050	0.939		[-0.103, 0.095]
14	-0.003	0.051	0.948		[-0.103, 0.097]
15	0.011	0.052	0.830		[-0.091, 0.114]
16	-0.021	0.053	0.700		[-0.125, 0.084]
17	-0.035	0.055	0.528		[-0.142, 0.073]
18	-0.039	0.055	0.480		[-0.147, 0.069]
19	-0.085	0.056	0.130		[-0.194, 0.025]
20	-0.054	0.054	0.318		[-0.159, 0.052]
21	-0.078	0.054	0.147		[-0.184, 0.027]
22	-0.070	0.056	0.207		[-0.179, 0.039]
23	-0.063	0.057	0.267		[-0.174, 0.048]
24	-0.067	0.059	0.258		[-0.183, 0.049]

Table 2: Model 1 Post-Policy × Event Time Coefficients

Note: Table 2 reports monthly estimates from Model 1, showing the effect of the post-policy × treatment interaction relative to birth. Pre-policy coefficients support parallel trends; post-policy estimates reflect the policy's dynamic impact on labor force attachment. Significance levels: $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***)..

In Figure 3 the estimated treatment effect of the paid leave policy reaches a peak of approximately 5.2 percentage points in month 0 relative to birth, suggesting a positive and economically meaningful impact on LFP immediately post-birth. In Table 2, we can see that this point estimate is not statistically significant at the conventional alpha levels, since the p-value is higher ($p = 0.199$). Despite the individual coefficients not showing significance, the surrounding months (−1 to +3) also show consistently positive effects, reinforcing the hypothesis that short-term paid leaves support LFP attachment during the short-term postnatal period. The trend corroborates previous literature (Byker 2016) which indicates the policy had a positive short term effect on labor force participation, however since the estimates are not statistically significant this conclusion cannot be drawn from these results. In contrast with Model 0, the individual-month coefficients remain statistically imprecise, likely due to the inclusion of high-dimensional interaction controls, which improve model credibility but inflate standard errors. To assess the cumulative effect of the policy, I conduct a series of Wald joint significance tests which can be found in Table 3.

Table 2.5 : Wald Tests Joint Significance of post_policy × lBirth Coefficients

Window	Months	Terms	DF	Wald F	p-value
−6 to +6	−6 ... +6	13	13 and 82623	1.148	0.3121
−4 to +3	−4 ... +3	8	8 and 82623	1.342	0.2171
−3 to +3	−3 ... +3	7	7 and 82623	0.853	0.5429
0 to +6	0 ... +6	7	7 and 82623	0.810	0.5789
0 to +12	0 ... +12	13	13 and 82623	1.072	0.3786
+3 to +8	+3 ... +8	6	6 and 82623	0.566	0.7577
Full −17 to +24	−17 ... +24	42	42 and 82623	1.704	0.0030

Note: This table reports Wald tests of the joint significance of the post-policy × treatment interaction coefficients over selected event-time windows. Each test evaluates the null hypothesis that the policy had no effect within the specified window. Reported are the number of coefficients tested, F-statistics, p-values, and degrees of freedom. Standard errors are clustered at the individual level (sippid).

Narrow windows such as event time −3 to +3 or -6 to +6 as well fail to reject the null of no treatment effect. Postpartum windows also show a higher than expected p-value indicating the cumulative effect of the policy was not significant. These results were unexpected as the trend in figure 3 resembled the raw difference in LFP of Figure 1, suggesting the effect of the policy was extracted. Despite the joint tests not confirming the cumulative significance of these time windows, a global Wald test over the full dynamic event time dummies window (event time −17 to +24) strongly rejects the null hypothesis of joint nullity across all post-policy coefficients ($F = 1.704$, $p = 0.002989$), indicating the overall signal of the trend is significant. This suggests that while monthly impacts may be individually noisy, the aggregate pattern indicates there is meaningful policy effect on maternal LFP, despite results not corroborating an increase in LFP around birth.

D. Heterogeneous Effects Analysis:

A. Ethnicity

To assess whether the paid leave policy had differential effects across racial groups, I begin by estimating Model 1 separately for racial subgroups of the sample. This preliminary analysis helps uncover whether pre-treatment dynamics differ meaningfully across subgroups. I estimate event-study regressions separately for three mutually exclusive racial categories: White (*erace* = 1), Black (*erace* = 2), and Asian (*erace* = 3), and construct group-specific confidence bands around each dynamic treatment effect.

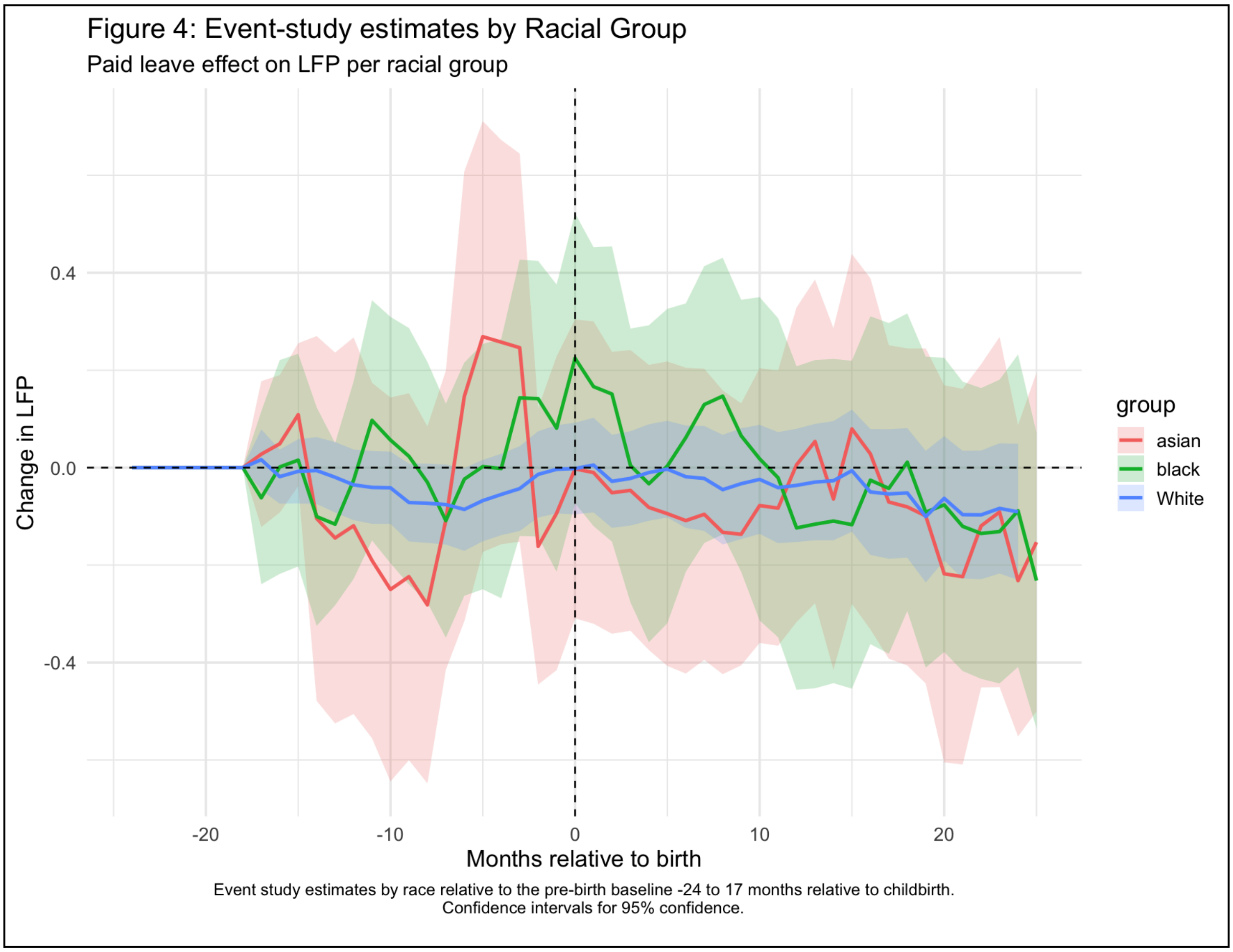
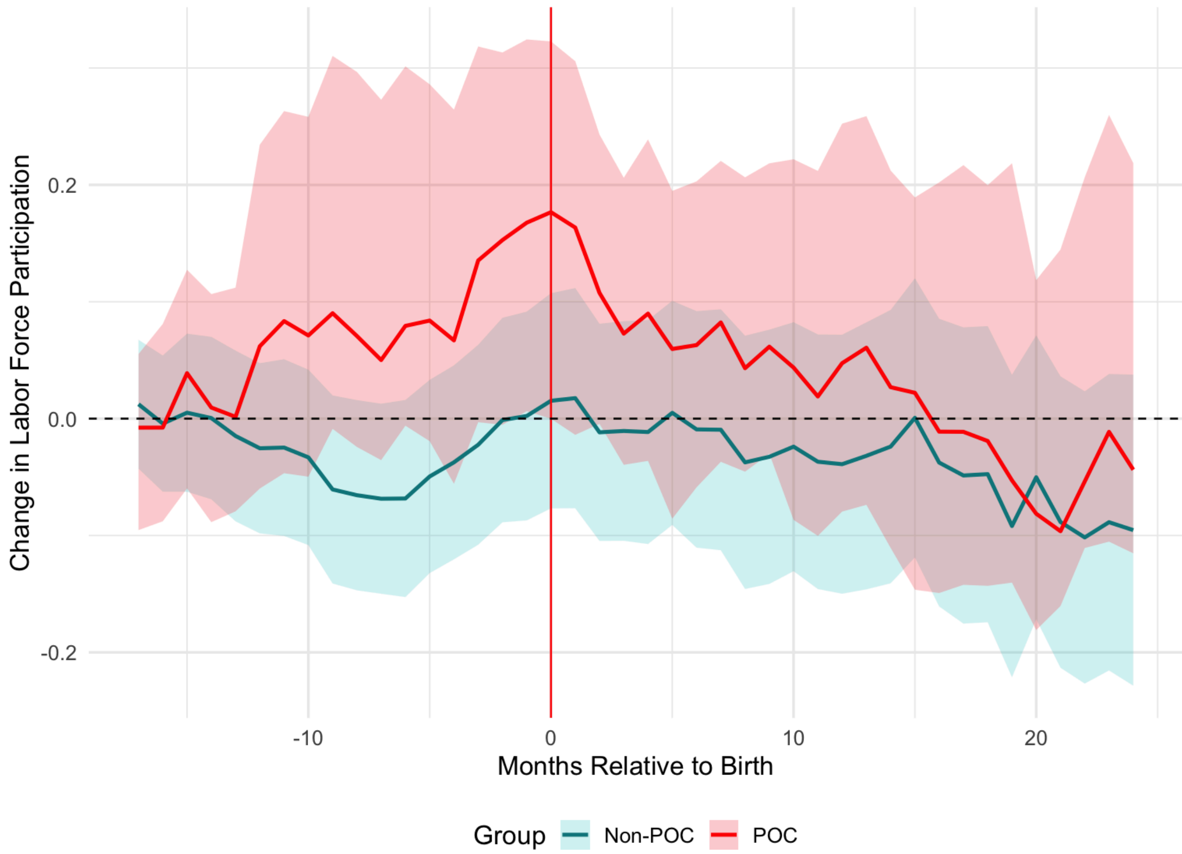


Figure 4 illustrates racial heterogeneity in dynamic treatment effects of paid leave on labor force participation. Black mothers exhibit the clearest positive response post-birth, with a peak around month 0 indicating they faced an increase in LFP around childbirth and another peak around 8 months after, however confidence intervals are wide and include zero. White mothers show stable estimates near zero, indicating no policy-induced change, while Asian mothers show noisier estimates, with a peak around 3 months pre birth and a slow descent after. Although individual-month estimates are not statistically significant, the divergence in trajectories suggest potential racial disparities in policy responsiveness, motivating the inclusion of formal interaction models in subsequent regressions. Because the number of observations for Black and Asian mothers was substantially smaller than for White mothers, direct comparisons at the individual group level suffer from wide confidence intervals and lower precision. To address this, I grouped Black and Asian mothers into a single “Person of Color” (POC) category and re-estimated treatment effects for POC versus White mothers. This aggregation improves statistical power and provides a cleaner test of group-level heterogeneity while still preserving the relevant racial distinction motivating the analysis.

Figure 5 plots dynamic treatment effect estimates from Model (1), separately for White and POC mothers. The estimates reflect changes in labor force participation (LFP) relative to a pre-birth baseline, defined as the average LFP between months −24 and −18. The red line shows the trajectory for POC mothers, who experience a pronounced and sustained increase in LFP beginning around childbirth. This effect peaks between approximately −6 and +3 months and remains elevated well into the post-birth period. In contrast, the trend for White mothers (blue line) is relatively flat, with only modest and transitory fluctuations in LFP. The divergence between the two subgroups suggests that the paid leave policy generated stronger and more persistent impacts for POC mothers. This could reflect differential baseline vulnerabilities for POC women. While confidence intervals overlap at times, the persistent gap in point estimates points to meaningful differences in policy responsiveness across racial groups.

Figure 5: Event-Study Estimates by Race (POC vs Non-POC)

Difference-in-Differences Estimates Relative to Pre-Birth Baseline (−24 to −18 months)



Event-study estimates of the change in labor force participation (LFP) relative to childbirth, by race (POC vs. Non-POC), based on a difference-in-differences model. The solid lines represent point estimates for each group, and the shaded areas represent the 95% confidence intervals. The red vertical line marks the month of birth, and the horizontal dashed line indicates no change in LFP. The estimates are relative to the pre-birth baseline period (−24 to −18 months).

Table 3: Wald Test Results: Post-Policy Effects for POC vs. White Subgroups

Subgroup	Event-Time Window	Degrees of Freedom	Wald Statistic	p-Value
Non-POC	−2 to +3	6 and 61992	0.776	0.588300
Non-POC	−3 to +3	7 and 61992	0.704	0.668400
Non-POC	−6 to +6	13 and 61992	0.990	0.457600
Non-POC	Full time-frame	42 and 61992	1.268	0.114000
POC	−2 to +3	6 and 16929	2.491	0.020710
POC	−3 to +3	7 and 16929	2.830	0.006012
POC	−6 to +6	13 and 16929	1.945	0.021230
POC	Full time-frame	42 and 16929	1.595	0.008526

Note: This table reports Wald tests of the joint significance of the post-policy × treatment interaction coefficients for POC and Non-POC subgroups over specified event-time windows. Each test evaluates the null hypothesis of no policy effect within the window. Standard errors are clustered at the individual level (sippid).

These visual patterns are reinforced by formal statistical evidence in Table 3 summarizing results from a series of Wald tests assessing whether the post-policy event-time coefficients are jointly equal to zero within each group. Across all windows, from narrow intervals (−3 to +3 months) to the full post-birth horizon (−17 to +24 months), POC mothers show statistically significant cumulative effects. For example, the test over months −3 to +3 results in an F-statistic of 2.83 with a p-value of 0.006 for POC, compared to an insignificant result for non-POC mothers ($F = 0.704$, $p = 0.668$). Even over the full dynamic window, the cumulative treatment effect remains significant for POC mothers ($F = 1.595$, $p = 0.0085$), but not for White mothers ($p = 0.114$). These results underscore the conclusion that the paid leave policy had a disproportionately impact on POC and non-POC mothers with a significant positive impact on POC mothers' labor force attachment following childbirth.

To formally test the significance of the difference between POC and non-POC labor force participation (LFP) responses to the policy, I estimate Model (2), which includes a three-way interaction between a binary indicator for being a person of color (POC), the event-time dummies (months relative to birth), and the post-policy dummy variable. This three-way interaction structure allows the model to estimate how the dynamic treatment effect of the policy varies by racial category (a difference in difference estimation). Specifically, the coefficients on the interaction terms, δ as seen in the Regression (2) below, capture the difference in policy effects between POC and non-POC individuals at each event month, while the baseline two-way interactions, β in Regression (2) represent the policy effect for the non-POC group.

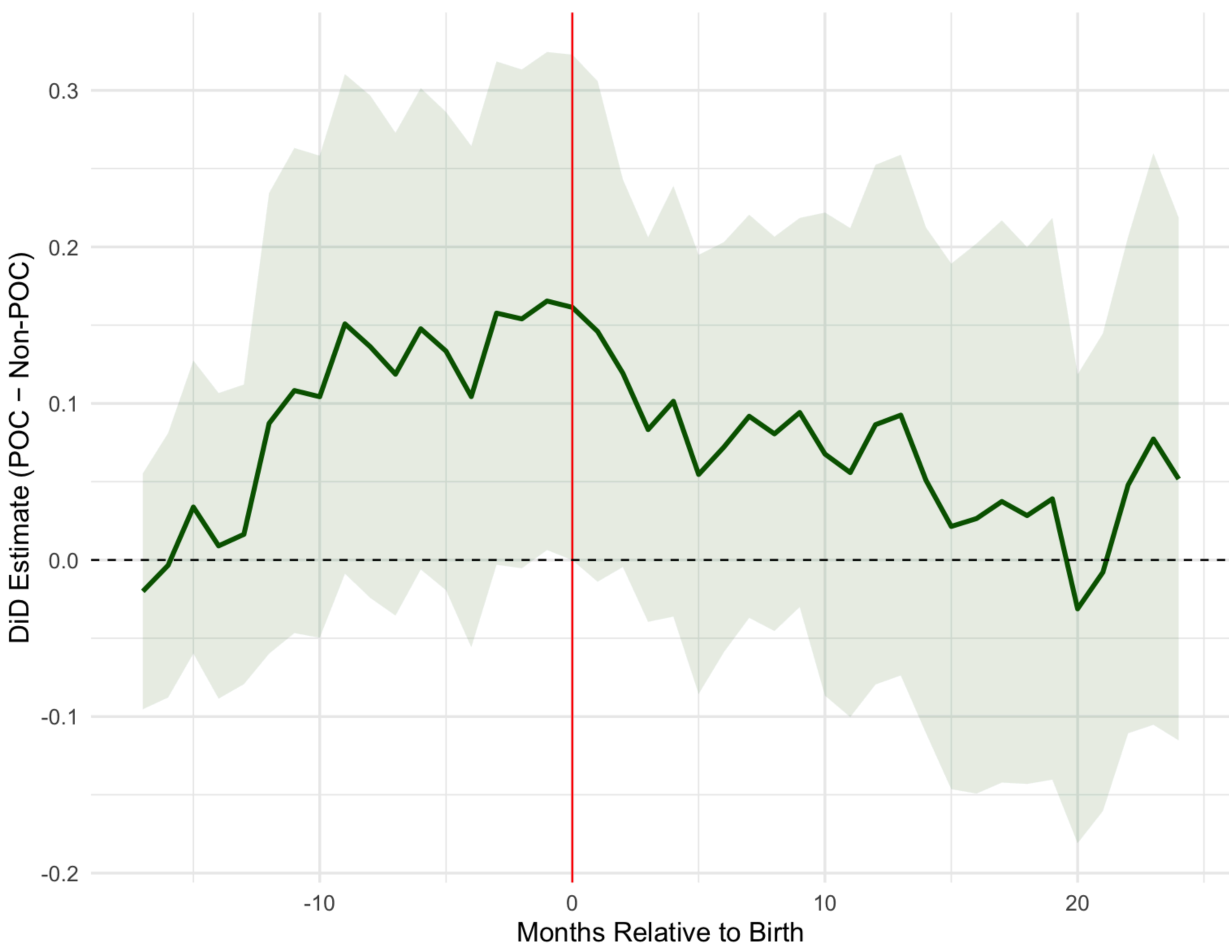
Model 2: FEOLS Main interaction term including POC dummy variable and baseline interaction terms

$$(2)\ \ LFP_{its} = \sum_{m=-24}^{+24} \delta_m B_{it}^{m} \times PostPolicy_{ts} poc \ + \sum_{m=-24}^{+24} \beta_m B_{it}^{m} \times PostPolicy_{ts}$$

$$+ \sum_{m=-24}^{+24} \lambda_t \times \ B_{it}^{m} \ + \sum_{m=-24}^{+24} \theta_s \times \ B_{it}^{m} + a_i + \lambda_t + \epsilon_{it}$$

Figure 6 Dynamic Difference-in-Differences (POC vs Non-POC)

POC × Policy × EventTime Interaction Coefficients



Event-study plot illustrating the difference-in-differences estimates of the policy effect on mothers of color compared to non-POC mothers, relative to the month of birth.

Table 4: POC × Post-Policy × Event Time Interaction Coefficients

Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	-0.02	0.038	0.602		[-0.095, 0.055]
-16	-0.003	0.043	0.937		[-0.088, 0.081]
-15	0.034	0.048	0.479		[-0.06, 0.127]
-14	0.009	0.05	0.856		[-0.089, 0.107]
-13	0.016	0.049	0.737		[-0.079, 0.112]
-12	0.087	0.075	0.245		[-0.06, 0.234]
-11	0.108	0.079	0.171		[-0.047, 0.263]
-10	0.104	0.079	0.184		[-0.05, 0.258]
-9	0.151	0.081	0.064	•	[-0.009, 0.31]
-8	0.136	0.082	0.096	•	[-0.024, 0.297]
-7	0.119	0.079	0.131		[-0.035, 0.273]
-6	0.148	0.078	0.06	•	[-0.006, 0.301]
-5	0.133	0.078	0.087	•	[-0.019, 0.286]
-4	0.104	0.082	0.201		[-0.056, 0.264]
-3	0.158	0.082	0.055	•	[-0.003, 0.318]
-2	0.154	0.081	0.058	•	[-0.005, 0.313]
-1	0.165	0.081	0.042	**	[0.006, 0.324]
0	0.161	0.082	0.05	•	[0, 0.323]
1	0.146	0.082	0.074	•	[-0.014, 0.306]
2	0.119	0.063	0.059	•	[-0.004, 0.243]
3	0.083	0.063	0.184		[-0.039, 0.206]
4	0.101	0.07	0.149		[-0.036, 0.239]
5	0.055	0.072	0.445		[-0.086, 0.195]
6	0.072	0.067	0.28		[-0.059, 0.203]
7	0.092	0.066	0.162		[-0.037, 0.221]
8	0.081	0.064	0.21		[-0.045, 0.206]
9	0.094	0.063	0.137		[-0.03, 0.218]
10	0.068	0.079	0.39		[-0.087, 0.222]
11	0.056	0.08	0.483		[-0.1, 0.212]
12	0.086	0.085	0.307		[-0.08, 0.252]
13	0.093	0.085	0.276		[-0.074, 0.259]
14	0.051	0.082	0.537		[-0.111, 0.212]
15	0.021	0.086	0.802		[-0.146, 0.189]
16	0.026	0.09	0.768		[-0.149, 0.202]
17	0.037	0.092	0.683		[-0.142, 0.217]
18	0.028	0.087	0.746		[-0.143, 0.2]
19	0.039	0.092	0.67		[-0.14, 0.218]
20	-0.031	0.076	0.683		[-0.181, 0.118]
21	-0.008	0.078	0.92		[-0.16, 0.145]
22	0.048	0.081	0.554		[-0.111, 0.206]
23	0.077	0.093	0.407		[-0.105, 0.26]
24	0.052	0.085	0.544		[-0.115, 0.219]

Note: This table reports the estimated coefficients on the triple interaction term POC × post-policy × event time. Each coefficient reflects the differential effect of the policy over time for POC compared to Non-POC individuals. Standard errors are clustered at the individual level (sippid). Significance levels: *p* < 0.1 (*), *p* < 0.05 (**), *p* < 0.01 (***).

Figure 6 plots the δ coefficients on the three-way interaction terms from Regression (2), which estimate the difference in policy effects on LFP between POC and non-POC individuals at each event-time month. A positive value indicates that the paid leave policy had a larger impact on POC individuals relative to their Non-POC counterparts. The trend shows a consistent upward shift around the month of birth, with positive estimates from approximately −10 through +6 months suggesting a persistent divergence in policy response.

Table 4 provides the corresponding coefficient estimates, standard errors, p-values, and confidence intervals for each event-time month. Notably, the interaction term for month −1 is statistically significant at the 5% level ($p = 0.042$), and several surrounding months (from −9 to +2) exhibit marginal significance at the 10% alpha level. This pattern is consistent with a local surge in differential impact around the time of childbirth. However, the wide confidence intervals, evident in both the plot and table, reflect substantial estimation uncertainty, likely due to the high dimensionality of the model and finite subgroup sample sizes.The individual months show suggestive evidence of racial heterogeneity in policy effects and the formal joint Wald tests' results in Table 5 partially corroborate these results.

Table 5: Wald Tests: POC × Post-Policy Interaction Effects (Model 2)

Window	Months Relative to Birth	# Terms	DF	Wald F-Stat	p-Value
−3 to +3	−3 to 3	7	7 and 71929	0.782	0.60240
−6 to +6	−6 to 6	13	13 and 71929	1.242	0.24170
0 to +10	0 to 23	24	24 and 71929	1.452	0.07083
Full Post (−17 to +24)	−17 to 24	42	42 and 71929	1.169	0.21020

Note: This table reports Wald tests of the joint significance of the POC × post-policy interaction coefficients from Model 2 over specified event-time windows. Each test evaluates the null hypothesis that the policy had no differential effect for POC within the window. Standard errors are clustered at the individual level (sippid).

Table 5 presents the results of joint Wald tests and tests evaluate whether the interaction coefficients within each window are jointly equal to zero, assessing whether the policy's effect differed significantly for POC compared to Non-POC. The results show that for the immediate post-birth window (0 to +10 months), the null hypothesis is rejected at approximately the 10% significance level ($p = 0.071$), suggesting some evidence of differential policy impact concentrated in the months following childbirth. In contrast, the shorter pre- and post-birth windows (−3 to +3 and −6 to +6) and the full time frame (−17 to +24) do not reject the null, indicating no statistically significant racial heterogeneity over those broader intervals. These findings are consistent with the notion that any differential effect of the policy for POC is most pronounced in the immediate postpartum period, though the evidence remains relatively weak and sensitive to the choice of window.

To further analyze the decomposition of the policy's effect on the labor force participation status of POC mothers I run the following Regression (3) and (4). These models have 2 dummy variables, *working* and *looking,* which are the dependent variables and regress them on the same covariates as those from Regression (1). These dummy variables are constructed from the *rmesr* variable in the SIPP dataset, *working*=1 when *rmesr*=1 and 0 otherwise, whereas *looking*=1 when *rmesr*={5,6,7} and 0 otherwise.

Model 3: Working status regressed on Dynamic DiD covariates

$$(3)\ working_{its} = a_i + \lambda_t + \theta_s \times \lambda_t + \sum_{m=-24}^{+24} \beta_j B_{it}^{m} \times PostPolicy_{ts} + \sum_{m=-24}^{+24} \delta_j B_{it}^{m}$$

$$+ \sum_{m=-24}^{+24} \pi_j B_{it}^{m} \times \theta_s + \sum_{m=-24}^{+24} \gamma_j B_{it}^{m} \times \lambda_t + \epsilon_{its}$$

Model 4: Working status regressed on Dynamic DiD covariates

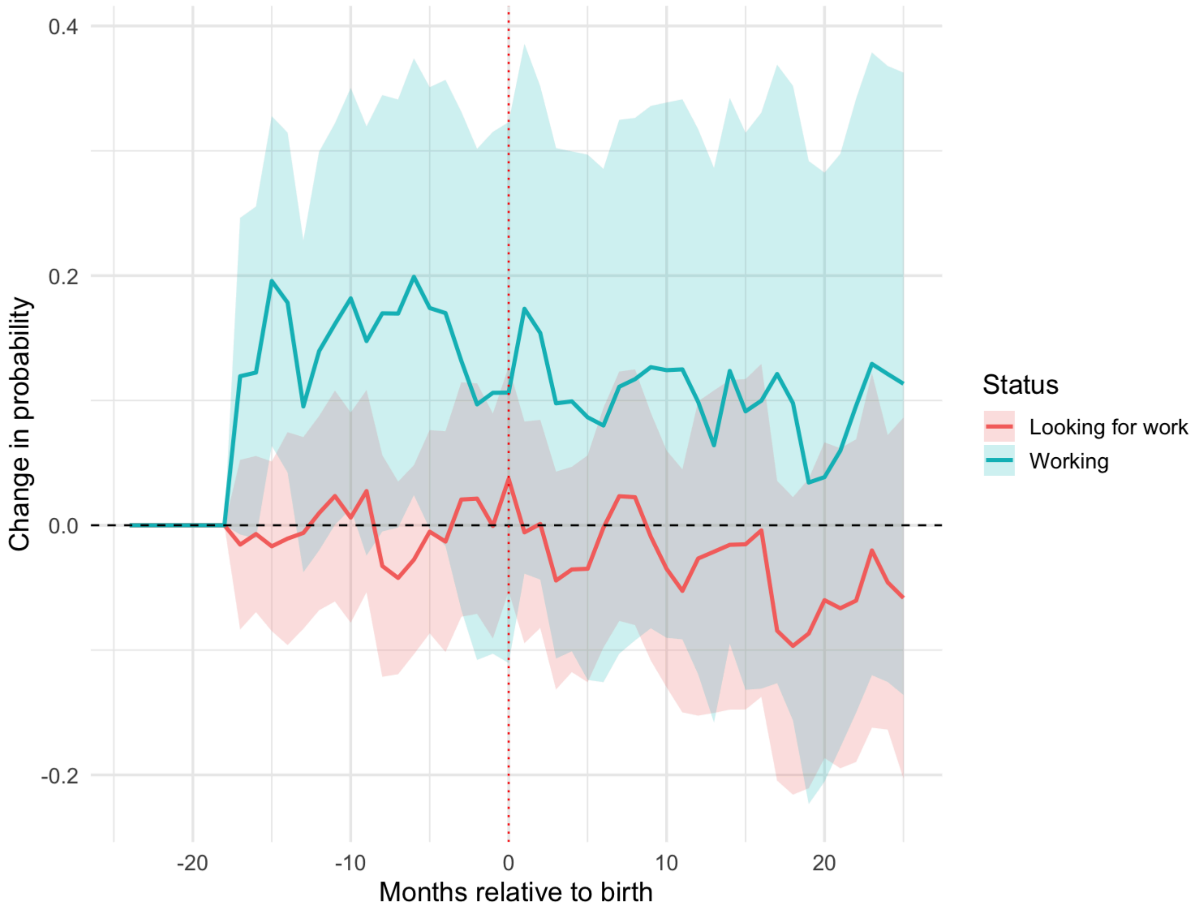
$$(4)\ looking_{its} = a_i + \lambda_t + \theta_s \times \lambda_t + \sum_{m=-24}^{+24} \beta_j B_{it}^{m} \times PostPolicy_{ts} + \sum_{m=-24}^{+24} \delta_j B_{it}^{m}$$

$$+ \sum_{m=-24}^{+24} \pi_j B_{it}^{m} \times \theta_s + \sum_{m=-24}^{+24} \gamma_j B_{it}^{m} \times \lambda_t + \epsilon_{its}$$

The results of Regression (3) and (4) are displayed in Figure 7, which plots the β estimates for each regression. In Figure 7 the plotted coefficients reveal that the positive treatment effect observed in overall LFP is largely driven by increases in actual employment, rather than job search activity, we see rates reaching 20% increase for working women around childbirth as an effect of the policy. Beginning several months before birth and persisting throughout the postpartum period, the working estimates show a pronounced upward trend, remaining elevated after the peak near childbirth. In contrast to the working estimates, the looking estimates exhibit a slow gradual decline close to zero across the event window Indicating the policy did not increase the rate of looking for work in mothers as they remain connected to

Figure 7. Impact of Paid Leave on Labor Force Status for POC women

Decomposition of LFP for POC women



Event-study decomposition of labor force participation for POC women, showing estimated changes in the probabilities of “Working” and “Looking for work” relative to the pre-birth baseline period (−24 to −17 months). Point estimates are shown with 95% confidence intervals.

their employment. Despite these positive effects, the 95%confidence intervals are very wide which indicates the conclusions drawn from this result might not be statistically significant.

Table 6 supports this visual impression with detailed estimates from the decomposed regressions. For the working outcome, coefficients from approximately −15 to −6 months show economically meaningful and occasionally statistically significant increases in employment probabilities, with estimates such as 0.196 at month −15 ($p = 0.004$) and 0.199 at month −6 (p =

0.026). However, the months immediately surrounding birth (−3 to +3) do not exhibit significance despite consistently positive point estimates. For the looking outcome, estimates are uniformly small, negative or near-zero, and never significant. This reinforces the interpretation that paid leave facilitates reattachment to work rather than stimulating job search.

Table 6: Estimates for POC Working vs Looking Models

Event Time	Estimate (Working)	SE (Working)	p-value (Working)	Estimate (Looking)	SE (Looking)	p-value (Looking)
−17.000	0.119	0.065	0.066	−0.015	0.035	0.655
−16.000	0.122	0.068	0.072	−0.007	0.032	0.826
−15.000	0.196	0.067	0.004	−0.017	0.035	0.627
−14.000	0.178	0.069	0.010	−0.011	0.043	0.806
−13.000	0.095	0.068	0.161	−0.006	0.039	0.876
−12.000	0.140	0.081	0.087	0.010	0.040	0.805
−11.000	0.161	0.082	0.050	0.023	0.043	0.587
−10.000	0.182	0.086	0.035	0.006	0.043	0.886
−9.000	0.148	0.088	0.093	0.027	0.041	0.507
−8.000	0.170	0.089	0.057	−0.033	0.045	0.471
−7.000	0.170	0.087	0.053	−0.042	0.039	0.284
−6.000	0.199	0.089	0.026	−0.028	0.039	0.476
−5.000	0.174	0.090	0.054	−0.005	0.042	0.902
−4.000	0.170	0.095	0.075	−0.013	0.045	0.772
−3.000	0.132	0.102	0.196	0.021	0.048	0.667
−2.000	0.097	0.104	0.354	0.021	0.047	0.650
−1.000	0.106	0.107	0.320	−0.001	0.046	0.990
0.000	0.106	0.111	0.337	0.037	0.046	0.420
1.000	0.173	0.108	0.110	−0.006	0.045	0.901
2.000	0.154	0.101	0.127	0.001	0.042	0.979
3.000	0.098	0.104	0.349	−0.044	0.045	0.320
4.000	0.099	0.102	0.332	−0.035	0.042	0.399
5.000	0.087	0.107	0.421	−0.035	0.046	0.452
6.000	0.080	0.105	0.446	−0.002	0.049	0.968
7.000	0.111	0.109	0.309	0.023	0.051	0.648
8.000	0.117	0.107	0.273	0.022	0.052	0.667
9.000	0.127	0.107	0.236	−0.009	0.051	0.857
10.000	0.124	0.109	0.256	−0.035	0.048	0.471
11.000	0.125	0.110	0.258	−0.053	0.050	0.290
12.000	0.099	0.111	0.375	−0.027	0.064	0.679
13.000	0.064	0.113	0.572	−0.021	0.066	0.748
14.000	0.124	0.112	0.268	−0.016	0.067	0.816
15.000	0.091	0.114	0.423	−0.015	0.068	0.822
16.000	0.100	0.118	0.397	−0.004	0.068	0.952
17.000	0.121	0.126	0.338	−0.084	0.061	0.168
18.000	0.098	0.130	0.451	−0.097	0.061	0.112
19.000	0.034	0.131	0.794	−0.087	0.063	0.171
20.000	0.039	0.124	0.757	−0.060	0.065	0.353
21.000	0.060	0.121	0.622	−0.066	0.065	0.310
22.000	0.096	0.125	0.446	−0.060	0.066	0.360
23.000	0.129	0.127	0.310	−0.020	0.072	0.781
24.000	0.121	0.126	0.336	−0.046	0.060	0.448
25.000	0.113	0.127	0.373	−0.058	0.074	0.430

Table 7: Wald Tests for Working Outcome (POC)

Window	Months Relative to Birth	# Terms	Degrees of Freedom	Wald F-Stat	p-Value
−2 to +3	−2 to 3	6	6 and 16,929	1.029	0.40408
−3 to +3	−3 to 3	7	7 and 16,929	1.059	0.38746
−6 to +6	−6 to 6	13	13 and 16,929	0.897	0.55586
Full Post (−17 to +24)	−17 to 24	42	42 and 16,929	1.132	0.25811

Note: This table reports Wald tests of the joint significance of the POC × post-policy interaction coefficients for the 'working' outcome across specified event-time windows. Each test evaluates the null hypothesis that the policy had no differential effect on the likelihood of working for POC within the window. Standard errors are clustered at the individual level (sippid).

Table 8: Wald Tests for Looking Outcome (POC)

Window	Months Relative to Birth	# Terms	Degrees of Freedom	Wald F-Stat	p-Value
−2 to +3	−2 to 3	6	6 and 16,929	2.285	0.03310
−3 to +3	−3 to 3	7	7 and 16,929	1.959	0.05663
−6 to +6	−6 to 6	13	13 and 16,929	1.617	0.07254
Full Post (−17 to +24)	−17 to 24	42	42 and 16,929	1.076	0.34055

Note: This table reports Wald tests of the joint significance of the POC × post-policy interaction coefficients for the 'looking for work' outcome across specified event-time windows. Each test evaluates the null hypothesis that the policy had no differential effect on the likelihood of looking for work for POC within the window. Standard errors are clustered at the individual level (sippid).

To test the significance of the trend's results, Tables 7 and 8 report formal Wald tests for joint significance across various post-birth event windows. For the *working* outcome in Table 7, none of the test windows show significant cumulative effects, with p-values ranging from 0.26 to 0.56, suggesting that despite upward trends in employment, the model lacks power to detect aggregate significance. In contrast, the *looking* outcome in Table 8 exhibits joint significance over the −2 to +3 window ($p = 0.033$). This result appears inconsistent with the underlying pattern, as individual point estimates in this window are largely small and statistically insignificant, suggesting the joint result may reflect statistical noise rather than a robust effect. It is important to note both models suffer from reduced statistical power due to sample trimming (since only individuals with complete working and looking status in POC subgroup) and the high dimensionality of the interactions in the event-study regressions. This loss of precision is evident in large standard errors and wide confidence intervals shown in Table 6, limiting the strength of inference even when point estimates are relevant.

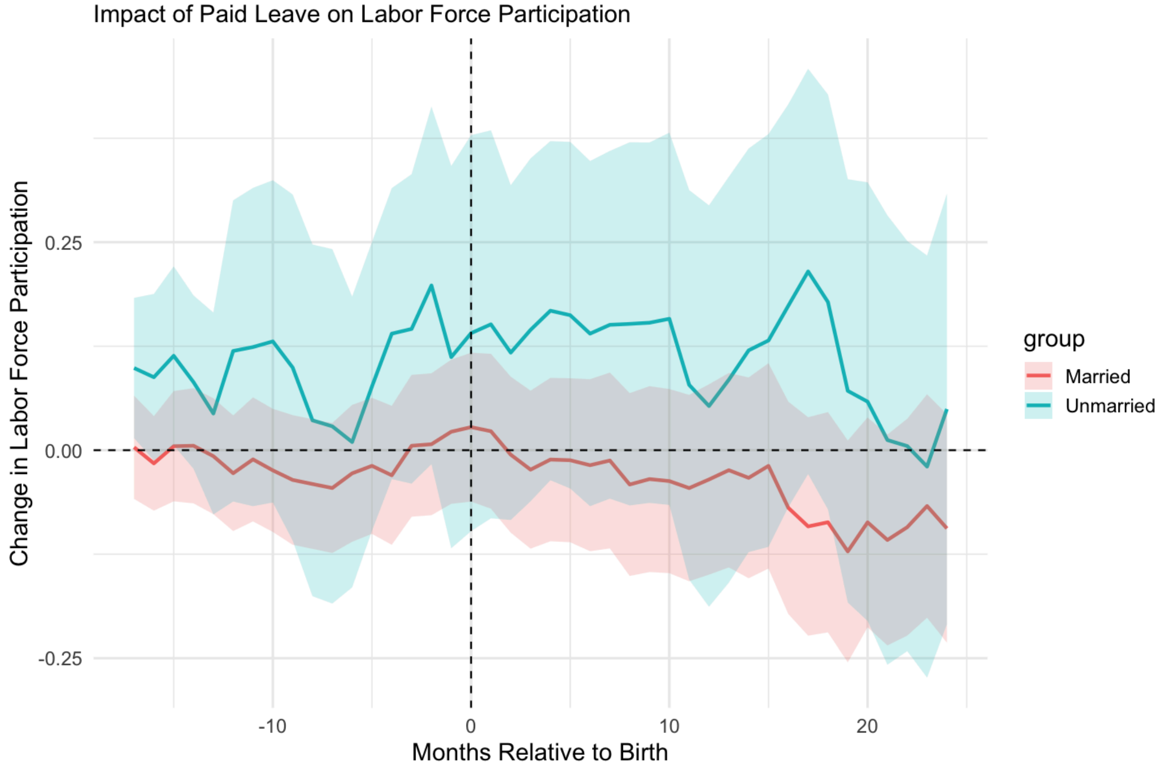
Considering these trends, the results suggest that the policy's primary channel of impact for POC mothers is through sustained employment. However, the inability to achieve formal statistical significance in more decomposed models likely reflects the limitations imposed by smaller subgroup sample sizes and increased model complexity, which reduce statistical power.

B. Marriage Status

To assess if the paid leave policy had differential effects by marriage status I use the same strategy as previously shown for the subgroup analysis by ethnicity and begin by separating mothers in the sample into subgroups of married and not married women and using each to

Figure 8: Event-Study Estimates by Marital Status

Impact of Paid Leave on Labor Force Participation



Event Study coefficient estimates per month relative to childbirth for unmarried and married women. relative to the pre-birth baseline period -24 to -17 months relative to childbirth.Point estimates are shown with 95% confidence intervals.

estimate Model 1. This subgroup classification is determined by the "ems" column of data in SIPPP. If the status of the mothers changed during the window they were not included.

Table 9: Joint Significance of Policy Effects by Marital Status

Group	Event-Time Window	DF (Model)	F Statistic	p-value
Unmarried	−3 to +3	23 and 16,496	4.441	< 1e-04
Unmarried	−6 to +6	23 and 16,496	4.441	< 1e-04
Unmarried	Full 0 to +24	23 and 16,496	4.441	< 1e-04
Married	−3 to +3	23 and 45,126	1.553	0.037
Married	−6 to +6	23 and 45,126	1.553	0.037
Married	Full 0 to +24	23 and 45,126	1.553	0.037

Note: This table reports Wald tests of the joint significance of the policy interaction coefficients by marital status over specified event-time windows. Each test evaluates the null hypothesis that the policy had no differential effect for the respective group within the window. Standard errors are clustered at the individual level.

As we can see in Figure 8, there is a suggestion that the effect of the policy was different for each subgroup indicating non-married women had a greater positive effect on their LFP from the policy's implementation. The trend for unmarried women is less smooth due to a great difference in the number of observations as married had 64,053 observations and unmarried had 15,730. This important difference likely explains the larger confidence intervals in the unmarried model compared to those for the married model. The unmarried trend exhibited a persistent increase in LFP effect of the policy while the married model exhibited a negligible effect that in the long term slightly decreased. These individual trend results were confirmed in the formal joint tests below, showing there is statistical significance for narrow and larger windows.

To formally confirm the significance of the differential effect between married and unmarried mothers I will implement Model 5 which incorporates a series of three-way interaction terms with a binary dummy control variable , *married*, along with the post-policy binary treatment variable and the month event time dummies B.

Model 5: FEOLS Main interaction term including married dummy variable and baseline interaction terms

$$(5)\ LFP_{its} = \sum_{m=-24}^{+24} \delta_m B_{it}^{m} \times PostPolicy_{ts} married_{it} + \sum_{m=-24}^{+24} \beta_m B_{it}^{m} \times PostPolicy_{ts}$$

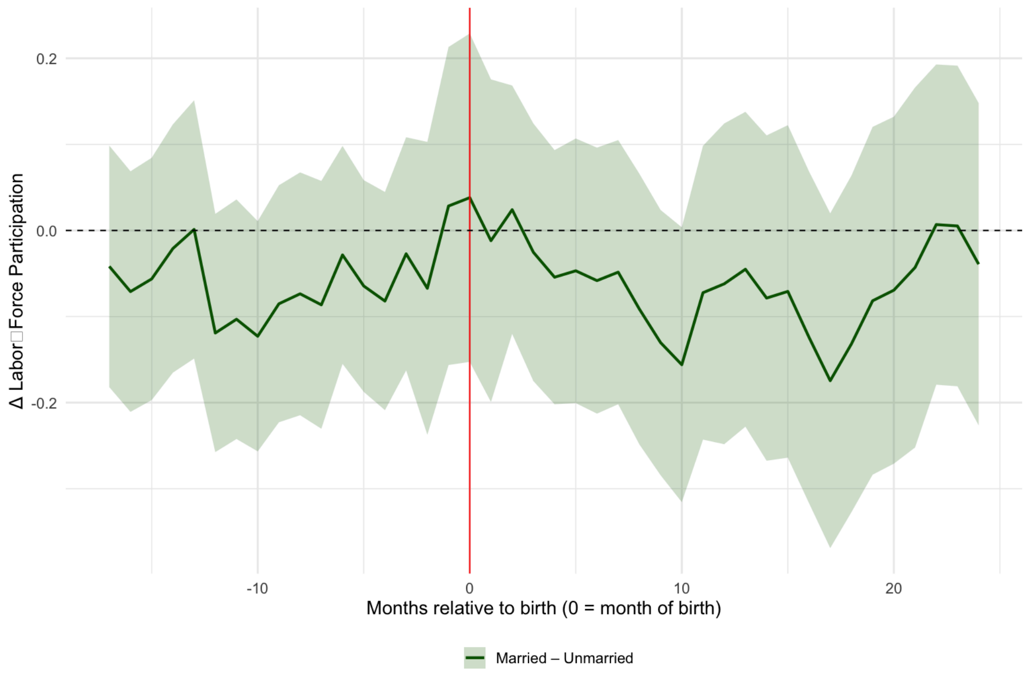
$$+ \sum_{m=-24}^{+24} \lambda_t \times B_{it}^{m} + \sum_{m=-24}^{+24} \theta_s \times B_{it}^{m} + a_i + \lambda_t + \epsilon_{it}$$

In the following Figure 9 coefficients β from Regression (5) are plotted and they show the differential effect from unmarried to married. As expected from the previous subgroup regression, the differential effect is negative, since the baseline of unmarried exhibits a higher effect than the married group. This baseline coefficient are the β coefficients in Regression (5), which explains the unmarried policy effect and the policy effect for married mothers is the addition of β and δ .

From Figure 9 and Table 11 in the appendix, we observe that the majority of the estimated coefficients are negative and after the slight peak aroundchildbirth the rate decreases consistently over the post-natal months and only after month +20 stabilize. This reflects the overall pattern that unmarried mothers experienced a greater increase in LFP following the policy, compared to their married counterparts. The only notable deviation from this trend occurs in the months immediately surrounding childbirth, where the trajectory of LFP increases for married women and dips for unmarried women, producing a short-term reversal in the differential effect.

Figure 9 Event-Study: Post-Policy Effect on LFP

Baseline (Unmarried), Married total, and Married − Unmarried differential



Event-Time Window	DF (Model)	F Statistic	p-value
−3 to 3	7 and 76,830	1.690	0.1006
−6 to 6	13 and 76,830	1.289	0.211
−17 to 24 (full frame)	42 and 76,830	1.871	5.32e−4

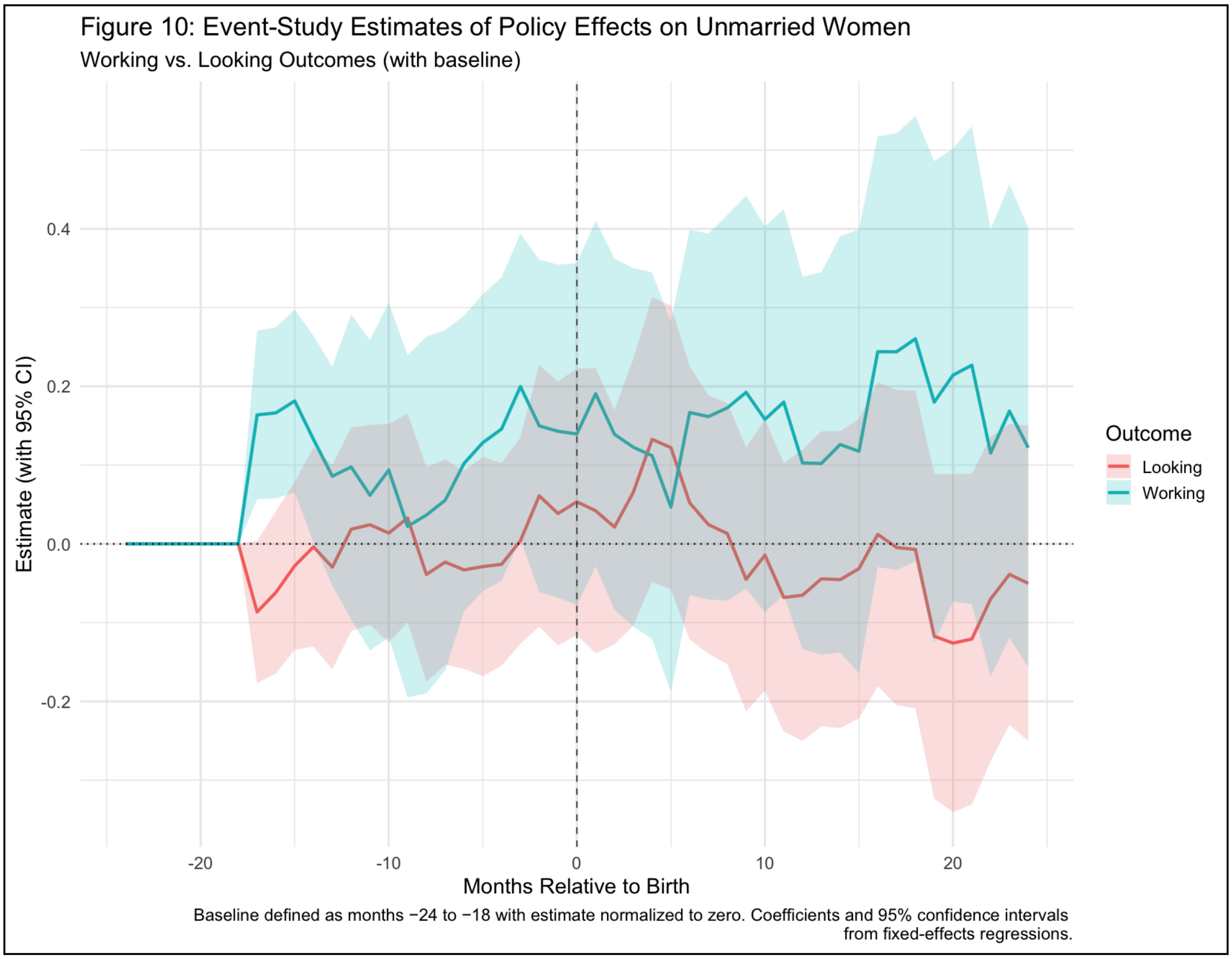
Table 10: Joint Significance of Married × Post-Policy Interaction Terms

Note: This table reports Wald tests of the joint significance of the Married × post-policy interaction coefficients across specified event-time windows. Each test evaluates the null hypothesis that the policy had no differential effect for married individuals within the given window. Standard errors are clustered at the individual level.

Despite these patterns, most of the individual interaction term coefficients are not statistically significant on their own. However, Table 10 presents results from Wald tests of joint significance across time windows. These tests show that the differential effect is significant over the narrow window (−3 to +3 months) at the 0.1 alpha level, and highly significant across the full window (−17 to +24 months), as indicated by a very low p-value. Together, these findings suggest a short-term gain in LFP for married mothers around childbirth, but a larger, more sustained positive effect for unmarried mothers over the full post-policy horizon.

To further understand the drivers of the positive effect among unmarried women, I decomposed the labor force participation outcome into two components: actively looking for work and being employed. This allows for a more nuanced examination of whether the observed gains reflect increased job attainment, job-seeking behavior, or both. To do so I will use the previously used Model 3 and 4 and I will run the regression using data for unmarried women.

Figure 10 shows the decomposition of LFP effects for unmarried women into two components, the probability of being employed (working) and the probability of being unemployed but actively looking for work. Overall the trend for working is higher than looking for work in the majority of the graph, except for 5 months after birth where there is a sharp increase in looking for work which rises to 12.2% and working dips to 4.7% , as seen in Table 13 in the appendix. In the months prior to childbirth, both trends seem to increase respectively as a whole with some dips and after the intersection at 5 months after we see a decreasing trend in looking for work while working continues to increase. Since the working trajectory remains elevated across most of the post-birth window this suggests a sustained employment response to the policy. Meanwhile, the “looking for work” component is positive and negative at times but as it is more volatile and its slope change indicates a decrease in looking for work preceding the slight increase building up to the +5 month .



These patterns suggest that paid leave policies may have helped facilitate labor market attachment among unmarried mothers, by enabling returns to employment maintaining a persistently high level of working with an upward trend. The policy also appears to have increased the rate of looking punctually, however there is a persistent decrease in the postpartum window, indicating the policy favored attachment to employment. However the confidence intervals are noticeably very large and include zero in most event times suggesting the data might not be strong enough for statistical significance.

5. Limitations

This study is subject to two main limitations: (1) reduced precision when evaluating subgroup effects due to imbalanced sample sizes, and (2) a loss of statistical power associated with the inclusion of high-dimensional control variables interacting with dynamic event-study terms.

The first limitation became evident during the heterogeneous effects analysis. Because the research design required splitting the sample by race and marital status, the number of observations across subgroups became highly unequal. For example, the group of people of color, POC, contained approximately 17,646 observations, whereas the non-POC group contained over 73,000. Similarly, there were around 15,730 observations for unmarried individuals, compared to more than 64,000 for married individuals. These imbalances significantly affect the precision of the estimated effects. Larger subsample sizes typically result in smoother, more stable estimates, as seen in the non-POC model , while smaller groups like the POC and unmarried samples yielded noisier trends and wider confidence intervals. This asymmetry limits the reliability of inference for underrepresented groups. Future studies would benefit from datasets with more balanced representation across key demographic categories, either through oversampling or by pooling multiple survey panels.

The second major limitation concerns the loss of statistical power when introducing additional control variables that interacted with the dynamic event-study indicators. In moving from the baseline model Model 0 to the fully adjusted specification Model 1, the number of parameters increased substantially due to a proliferation of interaction terms. This rise in model complexity inflated standard errors, making it more difficult to reject the null hypothesis. As a result, both joint significance tests and individual coefficient tests produced higher p-values, reducing the ability to detect meaningful effects. To mitigate this issue, future research could consider simplifying the model, for example, by reducing the number of event-time dummies, or increasing the sample size to compensate for the added dimensionality.

6. Conclusions

This study examined the impact of a paid parental leave policy on labor force participation (LFP) among women, with a focus on dynamic effects around childbirth and heterogeneity by race and marital status. When accounting for state-specific time-varying controls and evaluating all women in the sample, the policy appeared to generate a short-term increase in labor force participation. While the average monthly effect followed the expected pattern, wide confidence intervals limited statistical power: joint hypothesis tests could not confirm that the observed short-term rise in LFP was significantly different from zero. Nevertheless, both Model 0 (baseline) and Model 1 (fully adjusted) support the conclusion that the policy had a significant overall effect. Importantly, both models reveal a decline in LFP in the longer run, indicating that while the policy may have encouraged short-term retention or return to work, these gains did not persist over time.

The analysis of heterogeneous effects by race revealed stronger policy impacts for mothers of color (POC) relative to non-POC mothers. As shown in Figure 5, the LFP trajectory for POC mothers diverged notably from the flat pattern seen among non-POC women, reaching peaks of up to 20% above baseline in the months surrounding childbirth. Joint significance tests over the −3 to +3 month window confirmed that the increase in labor force participation around childbirth was statistically meaningful for both groups. In Model 2, which included three-way interactions to directly estimate differential effects by race, the results closely mirrored the pattern observed in the figure. Coefficients around childbirth were statistically significant and indicated that POC mothers experienced an increase in LFP that was 16.1 percentage points higher than their non-POC counterparts. Further decomposition of the POC response showed that the effect was driven by increased employment (working) and a decline in job-seeking (looking), suggesting that the policy promoted job attachment and reduced the risk of skill depreciation during leave. During the statistically significant −6 to +6 month window, the probability of working for POC mothers fluctuated between +10% and +20%, indicating a substantial and sustained short-term employment response.

The analysis by marital status also uncovered significant heterogeneity. Unmarried mothers experienced a larger and more sustained increase in labor force participation compared to married mothers, who showed only a brief and modest rise (peaking at +4.7%) around the month of childbirth. While the short-term gain for married women was visible, it was relatively small and not sustained. In contrast, unmarried women exhibited a consistently elevated LFP trajectory throughout the post-policy horizon. When decomposed into employment and job search components, the results showed that the positive effect for unmarried women was largely driven by increased employment. Their working rates were persistently high and upward trending over the entire observation window, peaking at a 20% increase shortly after childbirth. The probability of looking for work showed a short-term rise followed by a decline after five months, reinforcing the interpretation that the policy strengthened employment attachment rather than simply increasing job search.

In sum, this study provides strong evidence that the paid parental leave policy improved labor force outcomes for women overall, while having disproportionately positive effects for historically disadvantaged groups, namely mothers of color and unmarried women. These findings suggest that targeted leave policies may help reduce labor market inequality by supporting those most vulnerable to employment disruptions around childbirth.

8. Bibliography

Anderson, Deborah J., et al. "The Motherhood Wage Penalty: Which Mothers Pay It and Why?" *American Economic Review*, vol. 92, no. 2, Apr. 2002, pp. 354–58. *DOI.org (Crossref)*, https://doi.org/10.1257/000282802320191606.

Bailey, Martha Jane and Byker, Tanya and Patel, Elena and Ramnath, Shanthi, The Long-Term Effects of California's 2004 Paid Family Leave Act on Women's Careers: Evidence from U.S. Tax Data (December 2019). CEPR Discussion Paper No. DP14217, Available at SSRN: https://ssrn.com/abstract=3518555

Byker, Tanya S. "Paid Parental Leave Laws in the United States: Does Short-Duration Leave Affect Women's Labor-Force Attachment?" *American Economic Review*, vol. 106, no. 5, May 2016, pp. 242–46. *DOI.org (Crossref)*, https://doi.org/10.1257/aer.p20161118.

"Family and Medical Leave Act of 1993." *United States Code*, vol. 29, 103–3, 5 Feb. 1993, pp. 6–35.

Goldin, Claudia. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, vol. 104, no. 4, Apr. 2014, pp. 1091–119. *DOI.org (Crossref)*, https://doi.org/10.1257/aer.104.4.1091.

Kleven, Henrik, et al. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, vol. 11, no. 4, Oct. 2019, pp. 181–209. *DOI.org (Crossref)*, https://doi.org/10.1257/app.20180010.

Rose, Julie, et al. *The Ethnicity Motherhood Pay Penalty*.

7. Appendix

Table 11:

Table Married Individuals

Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	0.004	0.032	0.911		[-0.058, 0.066]
-16	-0.016	0.029	0.584		[-0.073, 0.041]
-15	0.005	0.034	0.887		[-0.061, 0.071]
-14	0.005	0.035	0.88		[-0.064, 0.075]
-13	-0.007	0.036	0.84		[-0.077, 0.062]
-12	-0.028	0.035	0.434		[-0.097, 0.042]
-11	-0.011	0.038	0.771		[-0.086, 0.063]
-10	-0.024	0.038	0.521		[-0.098, 0.05]
-9	-0.036	0.04	0.369		[-0.114, 0.042]
-8	-0.041	0.04	0.306		[-0.119, 0.037]
-7	-0.045	0.04	0.254		[-0.124, 0.033]
-6	-0.028	0.042	0.507		[-0.11, 0.054]
-5	-0.019	0.042	0.652		[-0.101, 0.063]
-4	-0.03	0.043	0.479		[-0.114, 0.053]
-3	0.005	0.043	0.903		[-0.08, 0.091]
-2	0.007	0.043	0.869		[-0.078, 0.092]
-1	0.022	0.044	0.615		[-0.064, 0.109]
0	0.028	0.046	0.544		[-0.062, 0.117]
1	0.023	0.047	0.631		[-0.07, 0.116]
2	-0.005	0.048	0.91		[-0.099, 0.088]
3	-0.023	0.048	0.629		[-0.118, 0.072]
4	-0.011	0.05	0.822		[-0.109, 0.087]
5	-0.012	0.05	0.81		[-0.111, 0.086]
6	-0.018	0.053	0.732		[-0.121, 0.085]
7	-0.012	0.054	0.818		[-0.118, 0.093]
8	-0.041	0.056	0.462		[-0.151, 0.069]
9	-0.035	0.057	0.542		[-0.146, 0.077]
10	-0.037	0.056	0.511		[-0.148, 0.074]
11	-0.046	0.057	0.425		[-0.158, 0.066]
12	-0.035	0.058	0.547		[-0.15, 0.079]
13	-0.024	0.06	0.687		[-0.141, 0.093]
14	-0.033	0.062	0.589		[-0.154, 0.087]
15	-0.019	0.063	0.764		[-0.142, 0.104]
16	-0.069	0.065	0.287		[-0.197, 0.058]
17	-0.092	0.067	0.171		[-0.223, 0.04]
18	-0.087	0.068	0.2		[-0.219, 0.046]
19	-0.122	0.068	0.074	•	[-0.255, 0.012]
20	-0.087	0.064	0.178		[-0.213, 0.039]
21	-0.108	0.065	0.095	•	[-0.235, 0.019]
22	-0.093	0.066	0.164		[-0.223, 0.038]
23	-0.067	0.068	0.327		[-0.201, 0.067]
24	-0.094	0.07	0.18		[-0.231, 0.043]

Table 12:

Table Unmarried Individuals					
Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	0.099	0.043	0.022	**	[0.015, 0.183]
-16	0.088	0.051	0.086	•	[-0.012, 0.188]
-15	0.114	0.055	0.039	**	[0.006, 0.221]
-14	0.082	0.053	0.124		[-0.022, 0.186]
-13	0.044	0.062	0.476		[-0.077, 0.165]
-12	0.119	0.092	0.197		[-0.062, 0.3]
-11	0.124	0.098	0.204		[-0.067, 0.315]
-10	0.131	0.099	0.186		[-0.063, 0.324]
-9	0.099	0.106	0.349		[-0.109, 0.307]
-8	0.036	0.108	0.739		[-0.175, 0.247]
-7	0.029	0.109	0.792		[-0.184, 0.242]
-6	0.01	0.089	0.912		[-0.165, 0.185]
-5	0.077	0.089	0.387		[-0.097, 0.25]
-4	0.14	0.089	0.117		[-0.035, 0.315]
-3	0.146	0.095	0.125		[-0.04, 0.332]
-2	0.198	0.11	0.071	•	[-0.017, 0.413]
-1	0.112	0.117	0.34		[-0.118, 0.342]
0	0.141	0.121	0.247		[-0.097, 0.379]
1	0.151	0.119	0.204		[-0.082, 0.384]
2	0.117	0.103	0.254		[-0.084, 0.319]
3	0.145	0.105	0.169		[-0.061, 0.351]
4	0.168	0.104	0.107		[-0.036, 0.372]
5	0.162	0.106	0.127		[-0.046, 0.371]
6	0.14	0.106	0.186		[-0.067, 0.348]
7	0.151	0.107	0.158		[-0.058, 0.36]
8	0.152	0.111	0.173		[-0.066, 0.37]
9	0.153	0.111	0.166		[-0.063, 0.37]
10	0.158	0.114	0.167		[-0.066, 0.382]
11	0.079	0.119	0.511		[-0.156, 0.313]
12	0.053	0.123	0.666		[-0.188, 0.294]
13	0.085	0.124	0.496		[-0.159, 0.329]
14	0.12	0.124	0.332		[-0.122, 0.363]
15	0.132	0.127	0.298		[-0.116, 0.38]
16	0.174	0.124	0.16		[-0.069, 0.416]
17	0.215	0.124	0.084	•	[-0.029, 0.458]
18	0.178	0.127	0.162		[-0.071, 0.427]
19	0.071	0.13	0.583		[-0.183, 0.326]
20	0.058	0.135	0.665		[-0.205, 0.322]
21	0.012	0.138	0.93		[-0.258, 0.282]
22	0.005	0.126	0.969		[-0.241, 0.251]
23	-0.02	0.129	0.879		[-0.273, 0.234]
24	0.049	0.132	0.708		[-0.209, 0.308]

Table 13

Estimates for Married Working vs Looking Models						
Event Time	Estimate (Working)	SE (Working)	p-value (Working)	Estimate (Looking)	SE (Looking)	p-value (Looking)
−17.000	0.164	0.054	0.003	−0.086	0.046	0.061
−16.000	0.166	0.055	0.003	−0.062	0.052	0.241
−15.000	0.181	0.059	0.002	−0.028	0.054	0.608
−14.000	0.132	0.067	0.048	−0.004	0.065	0.953
−13.000	0.086	0.071	0.226	−0.030	0.066	0.655
−12.000	0.098	0.099	0.322	0.019	0.066	0.778
−11.000	0.062	0.100	0.539	0.024	0.065	0.707
−10.000	0.094	0.108	0.387	0.014	0.071	0.846
−9.000	0.022	0.111	0.841	0.033	0.067	0.627
−8.000	0.037	0.115	0.751	−0.039	0.069	0.576
−7.000	0.055	0.110	0.615	−0.023	0.066	0.728
−6.000	0.102	0.096	0.287	−0.033	0.064	0.606
−5.000	0.129	0.096	0.181	−0.029	0.071	0.685
−4.000	0.146	0.098	0.138	−0.026	0.066	0.692
−3.000	0.200	0.099	0.044	0.004	0.067	0.952
−2.000	0.150	0.108	0.164	0.061	0.085	0.472
−1.000	0.143	0.108	0.185	0.039	0.085	0.652
0.000	0.140	0.111	0.208	0.053	0.086	0.536
1.000	0.191	0.112	0.089	0.042	0.092	0.649
2.000	0.139	0.114	0.221	0.021	0.076	0.779
3.000	0.123	0.116	0.291	0.065	0.086	0.449
4.000	0.112	0.119	0.346	0.133	0.092	0.150
5.000	0.047	0.120	0.696	0.122	0.092	0.184
6.000	0.167	0.118	0.159	0.052	0.088	0.556
7.000	0.162	0.119	0.173	0.024	0.084	0.770
8.000	0.173	0.125	0.167	0.013	0.085	0.877
9.000	0.192	0.127	0.132	−0.045	0.085	0.599
10.000	0.158	0.125	0.206	−0.014	0.088	0.870
11.000	0.180	0.125	0.150	−0.068	0.087	0.434
12.000	0.103	0.121	0.395	−0.065	0.094	0.490
13.000	0.102	0.124	0.410	−0.044	0.095	0.643
14.000	0.126	0.135	0.350	−0.045	0.096	0.637
15.000	0.118	0.144	0.414	−0.031	0.097	0.746
16.000	0.244	0.139	0.081	0.012	0.098	0.902
17.000	0.244	0.141	0.085	−0.005	0.102	0.964
18.000	0.260	0.144	0.071	−0.007	0.103	0.945
19.000	0.180	0.156	0.249	−0.117	0.105	0.265
20.000	0.214	0.147	0.145	−0.126	0.110	0.250
21.000	0.227	0.155	0.143	−0.121	0.107	0.259
22.000	0.115	0.145	0.426	−0.070	0.105	0.505
23.000	0.169	0.147	0.251	−0.039	0.097	0.691
24.000	0.122	0.142	0.391	−0.050	0.102	0.624

Table 14:

Table 4: Married × Post-Policy × Event Time Interaction Coefficients

Event Time	Estimate	Std. Error	p-value	Significance	95% Confidence Interval
-17	-0.042	0.072	0.56		[-0.182, 0.098]
-16	-0.071	0.071	0.319		[-0.211, 0.069]
-15	-0.056	0.072	0.433		[-0.197, 0.084]
-14	-0.021	0.074	0.776		[-0.165, 0.123]
-13	0.001	0.077	0.988		[-0.149, 0.151]
-12	-0.119	0.071	0.092	•	[-0.257, 0.019]
-11	-0.103	0.071	0.146		[-0.242, 0.036]
-10	-0.123	0.068	0.072	•	[-0.257, 0.011]
-9	-0.085	0.07	0.225		[-0.223, 0.052]
-8	-0.074	0.072	0.306		[-0.215, 0.067]
-7	-0.086	0.073	0.239		[-0.23, 0.058]
-6	-0.028	0.065	0.66		[-0.155, 0.098]
-5	-0.065	0.063	0.304		[-0.187, 0.058]
-4	-0.082	0.065	0.205		[-0.209, 0.045]
-3	-0.027	0.069	0.694		[-0.163, 0.108]
-2	-0.067	0.087	0.439		[-0.237, 0.103]
-1	0.028	0.094	0.763		[-0.156, 0.213]
0	0.038	0.097	0.696		[-0.153, 0.229]
1	-0.012	0.096	0.901		[-0.199, 0.175]
2	0.024	0.074	0.743		[-0.12, 0.168]
3	-0.025	0.076	0.74		[-0.175, 0.124]
4	-0.054	0.075	0.471		[-0.202, 0.093]
5	-0.047	0.078	0.55		[-0.201, 0.107]
6	-0.058	0.079	0.46		[-0.213, 0.096]
7	-0.048	0.078	0.536		[-0.202, 0.105]
8	-0.091	0.08	0.255		[-0.249, 0.066]
9	-0.13	0.079	0.097	•	[-0.284, 0.024]
10	-0.156	0.081	0.056	•	[-0.316, 0.004]
11	-0.072	0.087	0.406		[-0.243, 0.098]
12	-0.062	0.095	0.514		[-0.248, 0.124]
13	-0.045	0.093	0.629		[-0.228, 0.138]
14	-0.079	0.096	0.415		[-0.267, 0.11]
15	-0.071	0.099	0.472		[-0.264, 0.122]
16	-0.124	0.098	0.207		[-0.317, 0.068]
17	-0.175	0.099	0.079	•	[-0.369, 0.02]
18	-0.132	0.1	0.186		[-0.327, 0.064]
19	-0.082	0.103	0.428		[-0.284, 0.12]
20	-0.069	0.103	0.5		[-0.271, 0.132]
21	-0.043	0.107	0.687		[-0.252, 0.166]
22	0.007	0.095	0.943		[-0.179, 0.193]
23	0.005	0.095	0.956		[-0.181, 0.191]
24	-0.039	0.095	0.681		[-0.226, 0.148]