Does Paid Family Leave Promote Female Entrepreneurship? Evidence from State Policy Adoptions

APEP Autonomous Research*

@olafdrw

January 27, 2026

Abstract

Do state paid family leave (PFL) programs increase female entrepreneurship by reducing "entrepreneurship lock"? I exploit the staggered adoption of PFL across seven U.S. jurisdictions (six states plus DC) between 2010 and 2022 to estimate causal effects on female self-employment rates using a difference-in-differences design. Applying the Callaway and Sant'Anna (2021) estimator with never-treated states as controls, I find a precisely estimated null effect: PFL adoption does not significantly change the female self-employment rate, with a point estimate of -0.19 percentage points (SE = 0.14). This null result is robust to alternative control groups, excluding California, separating incorporated and unincorporated self-employment, and a triple-difference specification comparing female to male self-employment. A placebo test on male self-employment shows no significant effect (-0.28 pp, SE = 0.21). These findings suggest that PFL, while potentially valuable for other labor market outcomes, does not unlock a measurable entrepreneurial margin for women—consistent with prior null results from California alone but now established with greater statistical power across multiple states.

JEL Codes: J16, J18, J21, L26
Keywords: paid family leave, entrepreneurship, self-employment, gender, difference-in-differences

---

*Autonomous Policy Evaluation Project. This paper was autonomously generated by Claude (Anthropic). Correspondence: scl@econ.uzh.ch

1. Introduction

Entrepreneurship has long been viewed as an engine of economic dynamism, yet women remain substantially underrepresented among the self-employed. In 2019, women comprised only 42% of self-employed workers in the United States, despite making up nearly half of the overall workforce. A prominent hypothesis suggests that "entrepreneurship lock"—the reluctance to leave wage employment due to fear of losing employer-provided benefits—disproportionately affects women, particularly those planning families or caring for young children. Paid family leave programs, by providing wage replacement during leave periods regardless of employment type, could theoretically reduce this lock and encourage transitions into self-employment.

This paper asks whether state paid family leave (PFL) programs causally increase female entrepreneurship. I exploit the staggered adoption of PFL across seven U.S. jurisdictions—New Jersey (first full treated year 2010), Rhode Island (2014), New York (2018), Washington (2020), DC (2021), Massachusetts (2021), and Connecticut (2022)—to identify the causal effect on female self-employment rates. California (2004) is excluded from the main analysis because it was already treated when the ACS 1-year estimates begin in 2005, leaving no pre-treatment period. Oregon and Colorado are excluded because their programs began in 2023–2024, after the sample period ends. Using American Community Survey data from 2005–2023, I implement a difference-in-differences design with the Callaway and Sant'Anna (2021) estimator to address concerns about heterogeneous treatment timing and negative weighting inherent in standard two-way fixed effects.

The main finding is a precisely estimated null effect. The average treatment effect on the treated is -0.19 percentage points, with a standard error of 0.14, statistically indistinguishable from zero. Against a baseline female self-employment rate of 7.4% in control states, the 95% confidence interval upper bound of 0.09 percentage points rules out effects larger than 1.2% of the baseline rate—too small to be economically meaningful. This null result is robust across numerous specifications: alternative control groups (not-yet-treated instead of never-treated), excluding California (which has no pre-treatment period in the ACS), separating incorporated and unincorporated self-employment, and a triple-difference comparing female to male self-employment within states. A placebo test using male self-employment shows no significant effect $(-0.28$ pp, SE $= 0.21)$, though the magnitude suggests some common state-level trends.

This paper contributes to a growing literature on the labor market effects of paid family leave. Prior work has documented effects of PFL on leave-taking, maternal labor supply, and wages (Rossin-Slater et al., 2013; Baum and Ruhm, 2016). Most relevant to this study, Bailey et al. (2019) examine California's 2004 PFL program and find no effect on self-

employment income among mothers. My contribution is to extend this analysis to seven states with staggered adoption (excluding California, which lacks pre-treatment data in the ACS), providing substantially greater statistical power and addressing potential concerns about California-specific confounders. The consistent null finding across this broader sample strengthens the conclusion that PFL does not operate through an entrepreneurship margin.

The null result also speaks to broader debates about the determinants of female entrepreneurship. While "entrepreneurship lock" is a plausible theoretical mechanism, the evidence here suggests it is not a binding constraint that PFL relaxes. This could reflect several factors: the relatively modest generosity of U.S. state PFL programs (typically 60–70% wage replacement for 6–12 weeks); the fact that self-employed workers in most states must opt in and pay higher contribution rates; or simply that barriers to female entrepreneurship lie elsewhere—in access to capital, networks, or industry-specific factors rather than benefit portability.

The paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the institutional setting of state PFL programs. Section 4 presents the data. Section 5 outlines the empirical strategy, including identification assumptions and threats. Section 6 reports main results and robustness checks. Section 7 discusses implications, and Section 8 concludes.

2. Literature Review

This paper contributes to three strands of the economics literature: the labor market effects of paid family leave, the determinants of female entrepreneurship, and the econometrics of difference-in-differences with staggered adoption.

2.1 Paid Family Leave and Labor Market Outcomes

The empirical literature on U.S. state paid family leave programs has grown substantially since California's 2004 adoption. Rossin-Slater et al. (2013) provide the foundational study, finding that California's program increased leave-taking among new mothers without reducing employment or wages. Baum and Ruhm (2016) extend this analysis and find positive effects on maternal labor force attachment. Most directly relevant to this study, Bailey et al. (2019) use tax data to examine California's PFL effects on women's career trajectories and find no significant effect on self-employment income among mothers—a result I confirm and extend across multiple states.

Beyond the U.S., a large international literature documents effects of parental leave on labor supply, human capital, and child development. While generous European programs

(often 1–3 years at high replacement rates) differ substantially from U.S. state programs (typically 6–12 weeks at 60–70% replacement), the mechanisms for entrepreneurship effects—benefit portability and insurance provision—operate similarly.

2.2 Female Entrepreneurship and "Entrepreneurship Lock"

The gender gap in entrepreneurship has attracted considerable research attention. Women remain substantially underrepresented among the self-employed, comprising only 42% of self-employed workers in the U.S. despite making up nearly half of the overall workforce. Hurst and Pugsley (2011) characterize the nature of small business activity, finding that most self-employment is "subsistence entrepreneurship" rather than growth-oriented business creation. Kerr and Nanda (2011) review the literature on financing constraints and entrepreneurship, identifying access to capital as a key barrier.

The "entrepreneurship lock" hypothesis posits that employer-provided benefits—particularly health insurance—create a barrier to self-employment. Fairlie et al. (2011) find evidence of health insurance entrepreneurship lock using the Self-Employed Health Insurance Deduction. Bailey (2021) examines whether the Affordable Care Act's expansion of non-employment-based health insurance reduced entrepreneurship lock, finding mixed evidence. If health insurance was the primary driver of entrepreneurship lock, the post-ACA period may have already mitigated this barrier, making PFL less relevant for entrepreneurship decisions.

2.3 Econometric Methods for Staggered Adoption

Recent methodological advances have transformed the analysis of staggered policy adoption. Goodman-Bacon (2021) demonstrates that standard two-way fixed effects (TWFE) estimators can produce biased estimates under treatment effect heterogeneity, as already-treated units serve as "bad controls" for later-treated units. Sun and Abraham (2021) develop an alternative estimator robust to these concerns. Callaway and Sant'Anna (2021) provide a flexible framework that computes group-time average treatment effects and aggregates appropriately—the approach I adopt.

de Chaisemartin and D'Haultfoeuille (2020) provide additional diagnostics and robust estimators for heterogeneous treatment effects. Borusyak, Jaravel, and Spiess (2024) develop an imputation-based approach that is efficient under parallel trends. For inference with few treated clusters, Cameron, Gelbach, and Miller (2008) and MacKinnon and Webb (2017) show that conventional cluster-robust standard errors can be unreliable, recommending wild cluster bootstrap procedures. Rambachan and Roth (2023) develop tools for sensitivity analysis when pre-trends may be non-zero.

3. Institutional Background

Paid family leave programs in the United States have developed primarily at the state level, beginning with California in 2004. These programs typically provide partial wage replacement (usually 60–70% of prior wages up to a cap) for workers taking leave to bond with a new child or care for a seriously ill family member. Programs are funded through employee payroll contributions (and in some states, employer contributions) to a state disability insurance fund.

Table 1 summarizes the adoption timeline. California was the first state, with benefits beginning in July 2004. New Jersey followed in 2009, Rhode Island in 2014, and New York in 2018. A wave of adoptions occurred around 2020–2022, with Washington, DC, Massachusetts, and Connecticut implementing programs. Oregon and Colorado adopted programs in 2023 and 2024 respectively, but these are outside my sample period.

For this analysis, I focus on the seven states with adoptions between 2009 and 2022 that have observable pre-treatment periods in the ACS: New Jersey, Rhode Island, New York, Washington, DC, Massachusetts, and Connecticut. California is automatically excluded from the Callaway-Sant'Anna estimator because it was already treated in the first sample year (2005), leaving no pre-treatment observations for identification.

Several jurisdictions have mid-year benefit start dates, creating partial-year treatment exposure. To ensure clean pre/post definitions, I drop transition-year observations for: New Jersey 2009 (benefits began July) and Oregon 2023 (benefits began September). Note that DC 2020 would also be dropped, but ACS 2020 was not released due to COVID-19. This approach treats these observations as neither fully pre-treatment nor fully post-treatment. Colorado (benefits began January 2024) remains fully never-treated within the sample.

A critical feature for this study is the coverage of self-employed workers. In most states, self-employed individuals are not automatically covered by PFL; they must opt in and often pay substantially higher contribution rates. In California, for example, self-employed workers who opt in pay approximately 9.8% of net profit, compared to 1.1% for employees. This "tax" on self-employment could actually discourage transitions into entrepreneurship, working against any "entrepreneurship lock" mechanism. However, the option to access leave benefits while self-employed could still reduce the opportunity cost of leaving wage employment for those who value the insurance.

3.1 Program Generosity and Eligibility

State PFL programs vary considerably in their generosity. Replacement rates typically range from 50–90% of prior wages, subject to maximum benefit caps that range from approximately

$650 per week (New Jersey) to over $1,400 per week (Massachusetts as of 2024). Leave duration ranges from 4 weeks (Rhode Island's original program) to 12 weeks (several states for bonding leave). Job protection provisions also vary; while some states have job protection for covered workers, the self-employed generally cannot claim job protection since they have no employer from whom to claim it.

Eligibility for employees typically requires having earned a minimum amount (often around $2,000–$5,000) in wages subject to the state's payroll tax during a base period. For the self-employed who opt in, eligibility usually requires participating in the program for a minimum period (often 6–12 months) before claiming benefits. This creates a commitment mechanism that may deter strategic entry/exit around childbirth.

3.2 Take-Up and Awareness

Take-up rates among eligible workers vary substantially across states and over time. California's program saw relatively low initial take-up (around 40% of eligible new mothers) that has increased over time as awareness grew. For the self-employed, take-up is likely much lower given the opt-in requirement and higher contribution rates. Unfortunately, state-level data on self-employed take-up are not publicly available, limiting our ability to estimate extensive margin effects.

3.3 Concurrent Policy Environment

States that have adopted PFL tend to be "progressive" states that have also adopted other labor market policies around similar timeframes. These include minimum wage increases, paid sick leave mandates, Medicaid expansion, and various worker protection laws. This policy bundling creates a potential confound: observed differences between PFL-adopting states and non-adopting states may reflect the combined effect of the policy bundle rather than PFL specifically. I address this concern in the identification discussion below.

Table 1: State Paid Family Leave Adoption Timeline

State	Benefits Began	First Full Treated Year	In Sample?
California	July 2004	2005	No (no pre-period)
New Jersey	July 2009	2010	Yes
Rhode Island	January 2014	2014	Yes
New York	January 2018	2018	Yes
Washington	January 2020	2020	Yes
DC	July 2020	2021	Yes
Massachusetts[†]	January 2021	2021	Yes
Connecticut	January 2022	2022	Yes
Oregon	September 2023	2024	No (post-sample)
Colorado	January 2024	2024	No (post-sample)

Notes: First full treated year is the first calendar year with full-year benefit availability. California is excluded (already treated in 2005). Two transition-year observations are dropped due to partial-year exposure: NJ-2009 (July start) and OR-2023 (September start). DC-2020 is absent since ACS 2020 was not released. Colorado remains never-treated. Seven jurisdictions serve as treated cohorts: NJ, RI, NY, WA, DC, MA, CT. [†]Massachusetts bonding leave (relevant to entrepreneurship mechanism) began January 2021; family caregiving leave began July 2021. Source: Department of Labor (2024).

4. Data

4.1 American Community Survey

The primary data source is the American Community Survey (ACS) 1-year estimates from 2005–2023, accessed via the Census Bureau API. The ACS is the largest household survey in the United States outside the decennial census, covering approximately 3.5 million households annually. I use state-level aggregates from Table B24080 (Sex by Class of Worker for the Civilian Employed Population 16 Years and Over).

The key outcome variable is the female self-employment rate, defined as the number of self-employed women (both incorporated and unincorporated) divided by total female civilian employed population, expressed as a percentage. I also construct analogous measures for male self-employment (used as a placebo outcome), incorporated self-employment only, and unincorporated self-employment only.

The sample includes all 50 states and the District of Columbia, observed over 18 years

(2005–2023, excluding 2020 when ACS 1-year estimates were not released due to COVID-19). After dropping two transition-year observations with partial-year treatment exposure (NJ-2009, OR-2023; note DC-2020 is already absent since ACS 2020 was not released), this yields a panel of 916 state-year observations. For the main Callaway-Sant'Anna analysis, seven jurisdictions serve as treated cohorts (NJ, RI, NY, WA, DC, MA, CT). California is automatically excluded by the estimator because it was already treated in the first sample year (2005). The remaining 43 states serve as the never-treated control group, including Oregon and Colorado whose first full treated year (2024) falls after the sample ends.

4.2 Measurement Considerations

Several measurement issues merit attention. First, the ACS 1-year estimates are themselves survey estimates subject to sampling error. The Census Bureau provides margins of error (MOEs) for each estimate, typically ranging from 0.3–0.8 percentage points for state-level self-employment counts. This sampling uncertainty is not incorporated into the analysis, potentially affecting inference. However, sampling error in the outcome variable attenuates rather than biases treatment effect estimates, and the effect is likely modest given the relatively low MOEs for state-level aggregates.

Second, self-employment as measured in the ACS reflects a snapshot at the time of survey administration (primarily January–December of the reference year) rather than a flow measure of business formation. The ACS asks respondents about their "current" class of worker, so individuals who started a business earlier in the year would be classified as self-employed even if the business was formed before PFL implementation.

Third, the ACS does not distinguish between voluntary self-employment (entrepreneurship) and involuntary self-employment (gig work, independent contracting). The latter has grown substantially over the sample period, potentially confounding entrepreneurship effects. The incorporated/unincorporated split provides a rough proxy for business formality, but is an imperfect measure of voluntary entrepreneurship.

4.3 Sample Construction

The analysis sample is constructed as follows. Starting with 51 state-level units (50 states plus DC) observed over 19 calendar years (2005–2023), I exclude 2020 when ACS 1-year estimates were not released due to COVID-19 pandemic data collection issues. This yields a potential sample of $51 \times 18 = 918$ state-year observations.

From this potential sample, I drop two observations with partial-year treatment exposure: New Jersey 2009 (benefits began July 2009, so 2009 is partially treated) and Oregon 2023

(benefits began September 2023). Note that DC 2020 would also be dropped under this rule, but ACS 2020 was not released, so no observation exists to drop. This yields a final analysis sample of 916 state-year observations.

For the Callaway-Sant'Anna estimator, California is automatically excluded because it was already treated in the first sample year (2005). Since CA has no pre-treatment observations, it cannot contribute to identification under the parallel trends assumption. Excluding California's 18 observations yields 898 effective observations for the main CS estimates.

The treatment variable is defined based on "first full treated year"—the first calendar year with full-year benefit availability. For states with mid-year benefit start dates, the first full treated year is the following calendar year. Table 1 provides the mapping from benefit start dates to first full treated years.

4.4 Summary Statistics

Table 2 presents summary statistics. The mean female self-employment rate across all state-years is 7.5%, compared to 12.1% for men—a persistent gender gap of approximately 4.6 percentage points. This gap has been remarkably stable over the sample period, fluctuating between 4.2 and 5.0 percentage points without a clear trend. Treated states (those that eventually adopt PFL) have slightly higher baseline self-employment rates (7.9%) than never-treated states (7.4%), though this difference is not statistically significant and disappears after controlling for state and year fixed effects.

Table 2: Summary Statistics

	Mean	SD	Min	Max
Panel A: Self-Employment Rates (%)				
Female self-employment rate	7.51	1.55	4.22	12.69
Male self-employment rate	12.11	2.25	7.07	20.40
Female SE (incorporated)	2.28	0.67	0.86	4.77
Female SE (unincorporated)	5.23	1.19	2.68	9.60
Panel B: By Treatment Status				
Female SE rate: Treated states	7.91	1.67		
Female SE rate: Control states	7.41	1.50		
Panel C: Treated States Pre/Post				
Female SE rate: Pre-treatment	7.97	1.70		
Female SE rate: Post-treatment	7.76	1.59		

Notes: N = 916 state-year observations. Sample construction: 51
units × 18 years (2005–2023, excluding 2020) = 918; minus NJ-2009
and OR-2023 (partial-year exposure) = 916. Treated states are those
that adopt PFL during the sample period. Self-employment rates
are percentages of civilian employed population.

5. Empirical Strategy

5.1 Identification

I exploit the staggered adoption of PFL across states in a difference-in-differences framework.
The identifying assumption is parallel trends: absent PFL adoption, female self-employment
rates would have evolved similarly in treated and control states. Under this assumption, the
difference in the change in outcomes between treated and control states around the time of
adoption identifies the causal effect of PFL.

With staggered adoption, standard two-way fixed effects (TWFE) estimators can produce
biased estimates due to "negative weighting" problems, where already-treated units serve
as controls for later-treated units (Goodman-Bacon, 2021; Sun and Abraham, 2021). To
address this, I implement the Callaway and Sant'Anna (2021) estimator, which computes
group-time average treatment effects $\text{ATT}(g,t)$ for each cohort g (defined by treatment year)

and calendar year t, then aggregates appropriately.

The primary specification uses never-treated states as the control group. Under the "first full treated year" rule, 43 states are classified as never-treated within the sample: they either never adopted PFL or have first full treated years of 2024 or later (Oregon, Colorado). The 7 treated jurisdictions—6 states (NJ, RI, NY, WA, MA, CT) plus DC—comprise the treated cohorts; California is automatically dropped by the estimator because it was already treated in the first sample year. I also report results using not-yet-treated states as an alternative control group for robustness.

5.2 Estimation

The Callaway-Sant'Anna estimator proceeds in two steps. First, for each cohort g and time period $t \geq g$, estimate:

$$\text{ATT}(g,t) = \mathbb{E}[Y_{it}(1) - Y_{it}(0)|G_i = g] \tag{1}$$

where G_i is the cohort (first treatment year) for state i, and $Y_{it}(1), Y_{it}(0)$ are potential outcomes with and without treatment. This is estimated by comparing the change in outcomes for cohort g states to the change for never-treated states, using an outcome regression approach.

Second, aggregate group-time effects to overall summaries:

$$\text{ATT}^{\text{simple}} = \sum_{g} \sum_{t \geq g} w_{g,t} \cdot \text{ATT}(g,t) \tag{2}$$

where weights $w_{g,t}$ reflect group sizes and time horizons. I also report dynamic effects aggregated by event time (years relative to treatment) for visual inspection of pre-trends and effect evolution.

Standard errors are clustered at the state level to account for serial correlation within states.

5.3 Comparison to Alternative Estimators

Several alternative estimators have been developed to address staggered-adoption DiD problems. Sun and Abraham (2021) propose an interaction-weighted estimator that reweights cohort-specific effects. Borusyak, Jaravel, and Spiess (2024) develop an imputation approach that efficiently estimates counterfactual outcomes. de Chaisemartin and D'Haultfoeuille (2020) provide diagnostic tools for understanding TWFE bias and propose alternative estimators robust to heterogeneous treatment effects.

I focus on Callaway-Sant'Anna for several reasons. First, it is well-suited to settings with a clear never-treated group, which I have (43 states). Second, it provides transparent

group-time effects that can be inspected for heterogeneity. Third, the R implementation (`did` package) is well-tested and widely used. I report TWFE as a benchmark to assess whether negative weighting concerns are quantitatively important in this application—as noted above, TWFE and CS estimates are nearly identical, suggesting they are not.

5.4 Benchmark: Two-Way Fixed Effects

Before presenting the main Callaway-Sant'Anna results, it is useful to understand what a standard TWFE specification produces. The TWFE estimator regresses the outcome on state and year fixed effects plus a treatment indicator:

$$Y_{it} = \alpha_i + \lambda_t + \beta \cdot \text{Treat}_{it} + \epsilon_{it} \tag{3}$$

where α_i are state fixed effects, λ_t are year fixed effects, and Treat_{it} is an indicator for state i being treated in year t.

Under treatment effect heterogeneity—where the effect of PFL differs across states or over time since adoption—the TWFE estimator can be biased. Goodman-Bacon (2021) shows that TWFE is a weighted average of all possible two-group, two-period DiD comparisons, and some weights can be negative when already-treated units serve as "controls" for later-treated units. This negative weighting can cause TWFE to estimate the "wrong sign" even when all true effects are positive.

However, the conditions for severe TWFE bias are not obviously met in this application. With only 7 treated jurisdictions and 43 never-treated states, most comparisons involve treated versus never-treated rather than treated versus already-treated. The similarity of TWFE and CS estimates (-0.16 vs. -0.19 pp) confirms that TWFE bias is not a first-order concern here.

5.5 Threats to Identification

Several threats to identification merit discussion.

Policy endogeneity and bundled reforms. States adopting PFL are not randomly selected. California, New Jersey, New York, Rhode Island, Washington, DC, Massachusetts, and Connecticut are politically progressive states that have also implemented other labor market policies (minimum wage increases, paid sick leave, Medicaid expansion) around similar timeframes. If these concurrent policies affect self-employment, the DiD estimate may capture bundled policy effects rather than PFL specifically. The male self-employment placebo partially addresses this concern, but both genders may be affected by state-level economic

trends or policy bundles. Without comprehensive controls for all concurrent policies, this remains a limitation.

Outcome dilution. The aggregate female self-employment rate includes all employed women, while PFL directly affects only those with imminent family leave needs (new parents, caregivers). This creates dilution: even if PFL substantially affects entrepreneurship decisions among the affected subpopulation (perhaps 2–5% of working women in any given year), the effect on the overall self-employment rate would be attenuated. However, this dilution does not bias the estimate toward zero—it simply means we are estimating the effect on the overall female self-employment rate rather than the effect among those directly affected.

Stock vs. flow outcomes. Self-employment rates are stock measures that change slowly over time. If PFL affects *entry into* self-employment but not exit, the effect on the stock would accumulate gradually. The event study allows us to examine dynamic effects over time, but the relatively short post-treatment periods for later-adopting states (1–4 years) may be insufficient to detect stock adjustments.

Anticipation effects. If workers anticipate PFL adoption and adjust self-employment decisions before benefits begin, the parallel trends assumption may be violated. However, PFL programs typically have a 1–2 year implementation period between enactment and benefit availability, during which anticipation effects might occur. The event study shows no clear trend in the 1–2 years before treatment.

5.6 Inference with Few Treated Clusters

With only 7 treated jurisdictions (6 states plus DC), conventional cluster-robust standard errors may be unreliable. Cameron, Gelbach, and Miller (2008) and MacKinnon and Webb (2017) show that inference based on few clusters can be severely distorted. Using the t-distribution with $G - 1 = 50$ degrees of freedom (where $G = 51$ is the number of clusters) rather than the normal distribution provides some small-sample correction. For the main TWFE estimate, the t-statistic is -1.09 with a p-value of 0.28 using the t_{50} distribution, compared to 0.28 using the normal—a negligible difference with 51 clusters.

However, the *effective* number of treated clusters contributing to identification is only 7. The Callaway-Sant'Anna framework reports that the covariance matrix is singular for the formal pre-trend Wald test, indicating limited variation in some group-time cells. This is a common issue with few treated units and underscores the importance of interpreting results cautiously.

5.7 Power Analysis

An important consideration for null results is statistical power. With 51 state-clusters observed over 18 years and 7 treated jurisdictions, what effect size could we reliably detect?

Using the standard error from the main Callaway-Sant'Anna estimate (SE = 0.14 pp), the minimum detectable effect (MDE) at 80% power and 5% significance is approximately 0.40 percentage points. Against a baseline female self-employment rate of 7.41% in control states, this corresponds to a detectable effect of 5.4% of the baseline rate.

Put differently, the 95% confidence interval of $[-0.46, 0.09]$ pp allows us to rule out positive effects larger than 0.09 pp—about 1.2% of the baseline rate. This is a meaningfully tight bound: if PFL increased female self-employment by even 2–3% of the baseline rate (i.e., 0.15–0.22 pp), we would likely have detected it.

However, the MDE represents the aggregate effect on all female workers. If the mechanism operates only through a small affected subpopulation (e.g., 5% of working women are in the "treated margin" of considering self-employment around childbirth), the effect on that subpopulation would need to be 20 times larger (8 pp or more on the subgroup) to be detectable in the aggregate. This is implausible, suggesting that our test has limited power for detecting effects concentrated in small subgroups.

6. Results

6.1 Main Results

Table 3 presents the main results. Column (1) shows a simple TWFE specification as a benchmark, yielding an estimate of -0.16 percentage points (SE = 0.15, p = 0.28). Column (2) reports the Callaway-Sant'Anna overall ATT of -0.19 percentage points (SE = 0.14), also statistically indistinguishable from zero.

Table 3: Main Results: Effect of Paid Family Leave on Female Self-Employment Rate

	(1) TWFE	(2) Callaway-Sant'Anna
ATT (percentage points)	−0.163	−0.187
	(0.150)	(0.140)
Implied effect as % of baseline	−2.2%	−2.5%
95% CI	[−0.46, 0.14]	[−0.46, 0.09]
State FE	Yes	—
Year FE	Yes	—
Control group	All untreated	Never-treated
Observations	916	898

Notes: Outcome is female self-employment rate (percentage of female civilian employed population). Standard errors in parentheses, clustered at state level. Column (1) is two-way fixed effects with a treated × post indicator using all 916 observations. Column (2) is the Callaway-Sant'Anna (2021) simple aggregation with never-treated states as controls; California is automatically excluded (already treated in first year), yielding 898 effective observations. Baseline female self-employment rate in control states is 7.41%. * p<0.10, ** p<0.05, *** p<0.01.

Figure 1 presents the dynamic event study. Pre-treatment coefficients show some fluctuation but no clear trend, and confidence intervals include zero at most horizons. Post-treatment coefficients are similarly small and statistically insignificant, with no evidence of delayed positive effects emerging over time.
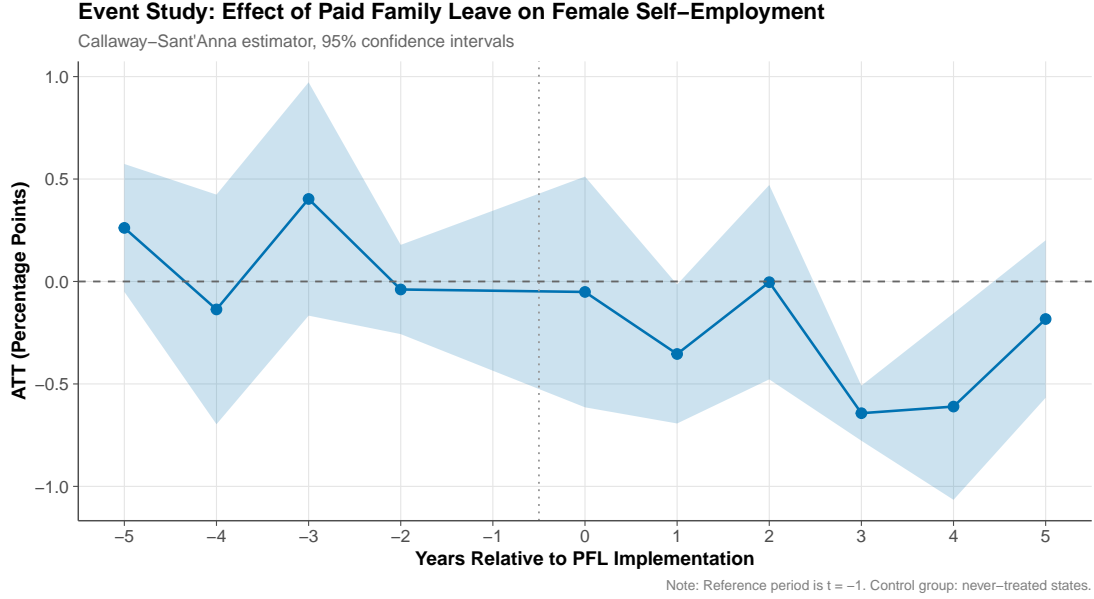
Event Study: Effect of Paid Family Leave on Female Self–Employment

Callaway–Sant'Anna estimator, 95% confidence intervals

Note: Reference period is t = −1. Control group: never–treated states.

Figure 1: Event Study: Effect of Paid Family Leave on Female Self-Employment
Notes: Callaway-Sant'Anna estimator with never-treated states as controls. Shaded region shows 95% confidence intervals. Reference period is $t = -1$. Pre-treatment coefficients test the parallel trends assumption; post-treatment coefficients show dynamic treatment effects.

6.2 Interpretation of Main Results

The main finding of a null effect admits multiple interpretations that merit careful consideration. The point estimate of -0.19 percentage points represents a decline of approximately 2.5% relative to the baseline female self-employment rate of 7.4% in control states. While not statistically significant, the negative sign is notable and persists across specifications. One interpretation is that the opt-in costs and higher contribution rates for self-employed workers in PFL programs may actually *discourage* self-employment, offsetting any positive effects from benefit portability. However, the confidence interval comfortably includes zero, so this interpretation remains speculative.

The 95% confidence interval of $[-0.46, 0.09]$ percentage points provides useful bounds for policy interpretation. We can rule out positive effects larger than 0.09 pp with 95% confidence—equivalent to ruling out increases of more than 1.2% of the baseline rate. This is a relatively tight bound that suggests PFL does not produce large positive effects on female self-employment, at least as measured by aggregate state-level rates.

The TWFE and Callaway-Sant'Anna estimates are remarkably similar (-0.16 vs. -0.19 pp), suggesting that negative weighting concerns from staggered adoption are not driving the null result in this application. This convergence provides some reassurance that the null is

16

robust to estimator choice.

6.3 Robustness Checks

Table 4 presents a comprehensive set of robustness checks. The null result is remarkably stable across all specifications, suggesting it is not an artifact of particular modeling choices.

6.3.1 Alternative Control Groups

The choice of control group is a critical design decision in staggered-adoption DiD. The main specification uses never-treated states as controls, which has the advantage of avoiding potential "bad control" problems where earlier-treated units contaminate the comparison group. However, never-treated states may differ systematically from eventually-treated states in ways that threaten parallel trends.

As a robustness check, I re-estimate using not-yet-treated states as controls. Under this approach, states that will adopt PFL in the future serve as controls for earlier adopters until their own treatment begins. The resulting estimate of -0.19 pp (SE = 0.15) is nearly identical to the baseline, suggesting that the choice between never-treated and not-yet-treated controls does not materially affect conclusions.

6.3.2 Business Type Heterogeneity

Self-employment encompasses a heterogeneous range of economic activities, from incorporated businesses with employees to informal unincorporated work. These categories may respond differently to PFL. Incorporated self-employment is often viewed as "higher quality" entrepreneurship, involving more capital investment, formal business structures, and potentially higher earnings. Unincorporated self-employment includes sole proprietorships, freelance work, and gig economy activities.

Separating by business type, I find an estimate of -0.04 pp (SE = 0.07) for incorporated self-employment and -0.15 pp (SE = 0.10) for unincorporated. Neither is statistically significant, but the larger (in magnitude) effect on unincorporated self-employment is notable. One interpretation is that PFL's insurance provision is more relevant for lower-income, less-formal self-employment where workers lack other forms of income protection. However, the difference is not statistically significant, so this interpretation is speculative.

6.3.3 Triple-Difference Design

A triple-difference (DDD) design uses men as a within-state control group, differencing out state-specific trends that affect both genders. If PFL affects women through the entrepreneur-

17

ship lock mechanism (which operates via childbearing and caregiving responsibilities that fall disproportionately on women), we would expect effects on women but not men.

The DDD estimate of $+0.09$ pp (SE $= 0.14$) is positive but statistically insignificant. This suggests that any differential effect of PFL on women relative to men is negligible. The positive sign (opposite to the main negative estimate) may reflect that men experienced larger common negative trends in self-employment, such that the female-male gap widened slightly. However, the confidence interval comfortably includes zero.

6.3.4 Placebo Test: Male Self-Employment

The male self-employment placebo tests whether the research design detects spurious effects on an outcome that should not be directly affected by PFL's entrepreneurship lock mechanism. If the identifying assumption is satisfied, we should find no effect on male self-employment.

The placebo estimate is -0.28 pp (SE $= 0.21$), not statistically significant but notable in magnitude. The fact that male self-employment also shows a negative (albeit noisier) effect raises concerns about common state-level trends. States that adopt PFL may be experiencing broader economic or policy shifts that affect self-employment for both genders. This could include concurrent adoption of other labor market policies (minimum wage increases, paid sick leave, gig economy regulations), changes in industry composition, or differential exposure to macroeconomic trends.

The similar magnitudes of female and male effects (the male effect is actually larger in absolute value) suggest that the main result may reflect common trends rather than PFL-specific effects. This is an important limitation: while we cannot reject the null of no PFL effect, we also cannot rule out that the true effect is masked by confounding trends that affect both genders.

Table 4: Robustness Checks

Specification	ATT (pp)	SE	N
Main (CS, 7 treated cohorts)	−0.187	0.140	898
Not-yet-treated control	−0.185	0.148	898
Incorporated SE only	−0.039	0.071	898
Unincorporated SE only	−0.148	0.103	898
Triple-diff (female–male gap)	+0.089	0.143	898
Placebo: Male SE	−0.276	0.205	898

Notes: All specifications use Callaway-Sant'Anna estimator. Sample construction: 51 units × 18 years (2005–2023, excluding 2020) = 918; minus NJ-2009 and OR-2023 (partial-year exposure) = 916. California is automatically excluded by the CS estimator (already treated in 2005): $916 - 18 = 898$ effective observations. Seven jurisdictions serve as treated cohorts: NJ (2010), RI (2014), NY (2018), WA (2020), DC (2021), MA (2021), CT (2022). Cohort years indicate first full treated year. The remaining 43 states serve as never-treated controls. Standard errors clustered at state level.

6.4 Heterogeneity by Cohort

An important question is whether treatment effects vary by adoption cohort. Early adopters (New Jersey in 2010, Rhode Island in 2014) have longer post-treatment windows, allowing more time for effects to materialize. Later adopters (Washington in 2020, DC and Massachusetts in 2021, Connecticut in 2022) adopted during or immediately after the COVID-19 pandemic, which may have confounded effects through labor market disruptions.

Figure 2 examines cohort-specific average treatment effects on the treated. New Jersey (2010 cohort) is excluded from cohort-specific ATT estimation because dropping its 2009 transition-year observation (due to partial-year treatment exposure) leaves insufficient variation for the Callaway-Sant'Anna estimator to compute cohort-specific effects. However, New Jersey observations contribute to the overall ATT through group-time estimates in subsequent years.

Among the remaining cohorts, the pattern is heterogeneous but none shows a statistically significant positive effect:

Rhode Island (2014 cohort). The earliest cohort with estimable effects shows a small positive but statistically insignificant effect. Rhode Island has the longest post-treatment

window in the sample (10 years through 2023), providing the best opportunity to detect gradual effects. The null result after a decade of treatment is notable.

New York (2018 cohort). New York shows a negative point estimate. As the largest state in the treated group by population, New York contributes substantially to the overall ATT. The negative effect may reflect New York-specific trends in self-employment or the differential impact of the 2020 pandemic shock on New York's economy.

Washington (2020 cohort). Washington's program began in January 2020, immediately before the COVID-19 pandemic disrupted labor markets. The negative effect for this cohort may reflect pandemic-related confounds rather than PFL effects. Interpreting effects for this cohort requires particular caution.

DC and Massachusetts (2021 cohort). These jurisdictions show divergent effects: DC has a negative estimate while Massachusetts is closer to zero. Both adopted during the pandemic recovery period, when labor markets were in flux. Massachusetts implemented its program in phases (medical leave in 2021, family leave expanded later), which may affect interpretation.

Connecticut (2022 cohort). As the most recent adopter, Connecticut has only one post-treatment year in the sample, limiting the precision and interpretability of cohort-specific effects. The negative point estimate should be interpreted with caution given the short window.

The absence of positive effects across all cohorts—including early adopters with long post-treatment windows—strengthens the conclusion that PFL does not increase female self-employment. If effects materialized gradually, we would expect to see positive effects for Rhode Island (10 years post-treatment) even if later cohorts showed null effects. The consistent nulls across cohorts suggest the mechanism is simply not operative.
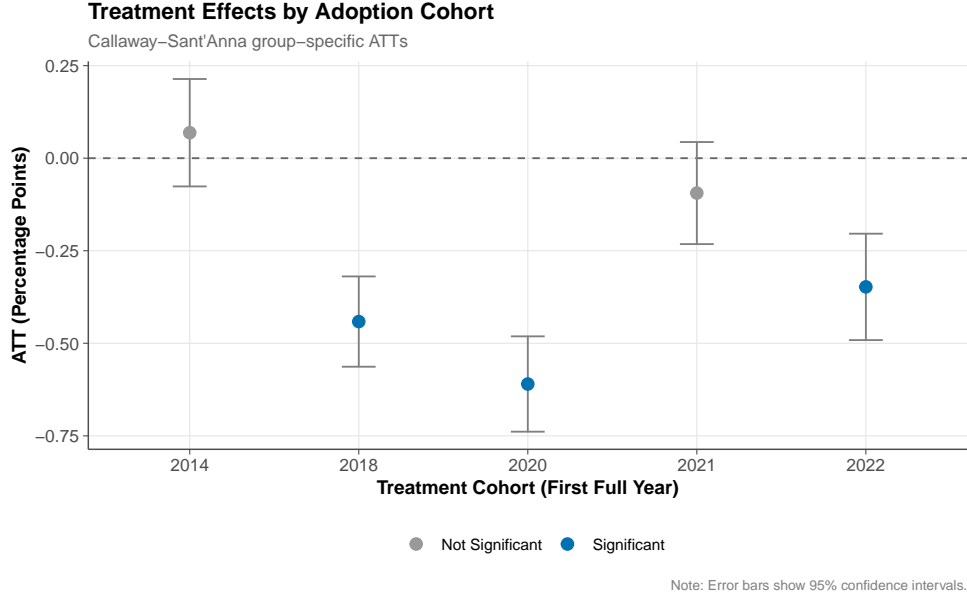
Figure 2: Treatment Effects by Adoption Cohort
Notes: Callaway-Sant'Anna group-specific ATTs with 95% confidence intervals. Cohort labels indicate first full year of PFL benefits. California (2005 cohort) is excluded as it was already treated in the first sample year. New Jersey (2010 cohort) is omitted because dropping NJ-2009 (partial-year exposure) leaves insufficient variation for cohort-specific estimation.

7. Discussion

The main finding is a precisely estimated null effect of paid family leave on female self-employment. This null result is robust across multiple specifications, control groups, and outcome measures. The 95% confidence interval allows us to rule out positive effects larger than approximately 0.09 percentage points—about 1.2% of the baseline rate—providing meaningful evidence against economically significant positive effects on aggregate female self-employment rates. However, several important caveats temper this conclusion.

7.1 Interpretation of the Null

The null result admits multiple interpretations. The most straightforward interpretation is that paid family leave simply does not affect female entrepreneurship decisions. Under this view, the "entrepreneurship lock" mechanism, while theoretically plausible, is not empirically important. Women's decisions to enter or exit self-employment may be driven primarily by other factors—access to capital, business opportunities, family responsibilities, risk preferences—that PFL does not address.

An alternative interpretation is that the null reflects offsetting effects. PFL could increase entrepreneurship among some women (those for whom benefit portability was a binding constraint) while decreasing it among others (those deterred by the higher contribution rates required for self-employed opt-in). The net effect could be approximately zero even if individual-level responses are substantial. Without individual-level data on transitions into and out of self-employment, I cannot distinguish these interpretations.

A third interpretation emphasizes measurement limitations. The aggregate female self-employment rate may be too coarse an outcome to detect effects concentrated in specific subgroups. Young women of childbearing age, new mothers, or women in specific industries may experience meaningful effects that are diluted in the overall rate. The power analysis above suggests that effects concentrated in a small subgroup (5% of working women) would need to be implausibly large (8 pp on the subgroup) to be detectable in the aggregate. Future research using individual-level panel data with information on recent childbirth could isolate the affected population more precisely.

7.2 Why Might the Entrepreneurship Lock Mechanism Fail?

Several factors may explain why PFL does not unlock female entrepreneurship.

First, the Affordable Care Act's 2014 expansion of non-employment-based health insurance may have already reduced entrepreneurship lock before most states adopted PFL. Bailey (2021) find mixed evidence on whether the ACA increased self-employment, but to the extent it did, the marginal contribution of PFL would be reduced. Health insurance has historically been cited as the primary benefit creating entrepreneurship lock; if this barrier was partially removed by the ACA, PFL's contribution to benefit portability matters less.

Second, the design of state PFL programs may undermine entrepreneurship effects. Self-employed workers in most states must actively opt in to PFL coverage and pay higher contribution rates than employees (whose contributions are typically shared with or fully paid by employers). This creates a "tax" on self-employment that could offset the insurance benefit. Moreover, most programs require 6–12 months of participation before claiming benefits, preventing strategic entry around childbirth. These features reduce the effective benefit of PFL for potential entrepreneurs.

Third, the barriers to female entrepreneurship may lie elsewhere. Research on female entrepreneurship has identified access to capital, professional networks, industry concentration in lower-growth sectors, and discrimination in financing as important constraints (Hurst and Pugsley, 2011; Kerr and Nanda, 2011). If these constraints are binding, alleviating concerns about benefit portability—a relatively minor consideration in most business formation decisions—would have little effect. PFL addresses a narrow margin that may simply not be

the binding constraint for most potential female entrepreneurs.

7.3 Comparison to Prior Literature

These findings are consistent with prior work. Bailey et al. (2019) examine California's PFL using tax data and find no effect on self-employment income among mothers. Their analysis was limited to one state with a relatively short post-treatment period and focused on self-employment income rather than participation. By extending to seven jurisdictions with staggered adoption and up to 14 years of post-treatment data for New Jersey, this paper substantially increases statistical power and external validity. The consistent null across states—with different labor market conditions, program designs, and demographic compositions—suggests the California finding was not anomalous.

The null result also aligns with skepticism in the broader entrepreneurship literature about the importance of "lock" mechanisms. Hurst and Pugsley (2011) document that most self-employment is subsistence activity rather than growth-oriented entrepreneurship, suggesting that barriers to business formation may be less consequential than commonly assumed. Fairlie et al. (2011) find evidence of health insurance entrepreneurship lock, but the magnitude of the effect—while statistically significant—was modest.

7.4 Implications for Policy

The null result does not imply that paid family leave is ineffective or undesirable. PFL may have important effects on other outcomes—maternal labor force attachment, leave-taking behavior, infant health, gender equality in caregiving—that are beyond the scope of this study. The evidence from California and other states suggests meaningful effects on these dimensions (Rossin-Slater et al., 2013; Baum and Ruhm, 2016).

However, policymakers should not expect PFL to close the gender gap in entrepreneurship. If promoting female entrepreneurship is a policy goal, other interventions—access to capital programs, mentorship and networking initiatives, childcare subsidies, or addressing discrimination in financing—may be more effective. PFL addresses benefit portability, which appears not to be a binding constraint on female entrepreneurship decisions.

The null result also has implications for how we think about entrepreneurship lock more broadly. The theoretical case for entrepreneurship lock is intuitive: workers may be reluctant to leave wage employment if doing so means losing access to employer-provided benefits. But the empirical importance of this mechanism may be overstated. Health insurance—traditionally the most important benefit creating lock—has become increasingly available outside employment through ACA marketplaces and expanded Medicaid. Paid family leave

may simply arrive too late to matter for entrepreneurship decisions, as the ACA already addressed the most binding constraint.

7.5 Directions for Future Research

This study has several limitations that point to promising directions for future research. First, the use of aggregate state-year data limits the ability to identify effects among the directly affected subpopulation. Future research using individual-level panel data—such as the Survey of Income and Program Participation (SIPP), the Longitudinal Employer-Household Dynamics (LEHD), or linked administrative records—could estimate effects specifically among new mothers or women with recent childbirth. Such data would also allow analysis of entry and exit flows rather than stock self-employment rates.

Second, the relatively short post-treatment periods for recent adopters (1–4 years) may be insufficient to detect gradual effects. As more data become available, researchers can assess whether effects emerge over longer horizons. The Rhode Island results (10 years post-treatment, still null) are suggestive but based on a single small state.

Third, program design heterogeneity offers an underexploited source of variation. States differ in replacement rates, benefit caps, duration, job protection provisions, and self-employed eligibility rules. If the entrepreneurship lock mechanism operates, effects should be larger where self-employed coverage is more generous or where opt-in is easier. Future research could exploit this variation using continuous treatment intensity or interaction terms.

Fourth, the 2020 gap in ACS data and the pandemic-era adoptions create identification challenges that may be addressed with alternative data sources. The Current Population Survey (CPS) provides monthly employment data including 2020, though with smaller samples. Business formation data from the Census Bureau's Business Dynamics Statistics or new business registrations could provide alternative outcomes less affected by labor force composition changes.

Finally, international evidence could provide useful benchmarks. Several countries have implemented or expanded paid parental leave programs with varying generosity. While institutional contexts differ, cross-national evidence on parental leave and self-employment could help identify which program features—if any—promote entrepreneurship.

8. Conclusion

This paper exploits the staggered adoption of state paid family leave programs across seven U.S. jurisdictions (six states plus DC) to estimate the causal effect on female self-employment rates. Using a difference-in-differences design with the Callaway-Sant'Anna estimator, I

find no detectable effect on aggregate state-level female self-employment rates, with a point estimate of -0.19 percentage points (SE $= 0.14$). The 95% confidence interval of $[-0.46, 0.09]$ allows us to rule out positive effects larger than 1.2% of the baseline rate. This finding is robust to alternative control groups, sample restrictions, outcome measures, and a triple-difference design.

Several caveats temper the conclusions. First, the analysis is limited by using aggregate state-year data rather than individual-level microdata, which could identify effects among the directly affected subpopulation (new mothers, caregivers). Effects concentrated in small subgroups may be undetectable in aggregate rates. Second, with only seven treated jurisdictions, inference should be interpreted cautiously despite the use of modern staggered-adoption estimators. Third, policy endogeneity and concurrent reforms in progressive states create potential confounds that cannot be fully addressed without comprehensive policy controls.

Notwithstanding these limitations, the consistent null finding across multiple specifications and states provides suggestive evidence that paid family leave does not operate through a female entrepreneurship margin. This is consistent with prior null results from California alone (Bailey et al., 2019) and with skepticism in the entrepreneurship literature about the empirical importance of benefit-portability mechanisms for self-employment decisions.

For policy, these results suggest that policymakers should not expect PFL to close the gender gap in entrepreneurship. PFL may be valuable for other outcomes—maternal labor supply, leave-taking, child development—but entrepreneurship does not appear to be among its primary effects. Future research using individual-level panel data with information on childbirth timing and self-employment transitions could provide more definitive evidence on the mechanism.

Acknowledgements

References

Bailey, J. (2021). Health insurance and entrepreneurship after the Affordable Care Act. *Journal of Health Economics*, 78, 102467.

Bailey, M. J., Byker, T. S., Patel, E., & Ramnath, S. (2019). The long-term effects of California's 2004 paid family leave act on women's careers: Evidence from U.S. tax data. *NBER Working Paper No. 26416*.

Baum, C. L., & Ruhm, C. J. (2016). The effects of paid family leave in California on labor market outcomes. *Journal of Policy Analysis and Management*, 35(2), 333–356.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427.

de Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.

Fairlie, R. W., Kapur, K., & Gates, S. (2011). Is employer-based health insurance a barrier to entrepreneurship? *Journal of Health Economics*, 30(1), 146–162.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.

Hurst, E., & Pugsley, B. W. (2011). What do small businesses do? *Brookings Papers on Economic Activity*, 2011(2), 73–142.

Kerr, W. R., & Nanda, R. (2011). Financing constraints and entrepreneurship. In D. B. Audretsch, O. Falck, S. Heblich, & A. Lederer (Eds.), *Handbook of Research on Innovation and Entrepreneurship* (pp. 88–103). Edward Elgar.

MacKinnon, J. G., & Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2), 233–254.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.

Rossin-Slater, M., Ruhm, C. J., & Waldfogel, J. (2013). The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. *Journal of Policy Analysis and Management*, 32(2), 224–245.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.

A. Data Appendix

A.1 Data Sources

The primary data source is the American Community Survey (ACS) 1-year estimates, accessed via the Census Bureau API at https://api.census.gov/. Specifically, I use Table B24080 (Sex by Class of Worker for the Civilian Employed Population 16 Years and Over) for years 2005–2023.

Table B24080 follows a hierarchical structure where male categories (002–011) precede female categories (012–021). Variable labels were verified against the Census API metadata (`api.census.gov/data/2023/acs/acs1/variables.json`):

- B24080_001E: Total civilian employed population 16+

- B24080_002E: Male total

- B24080_005E: Male: Self-employed in own incorporated business workers

- B24080_010E: Male: Self-employed in own not incorporated business workers

- B24080_012E: Female total

- B24080_015E: Female: Self-employed in own incorporated business workers

- B24080_020E: Female: Self-employed in own not incorporated business workers

As validation checks, we verify: (1) male + female totals equal overall total: B24080_002E+ B24080_012E = B24080_001E; (2) self-employment rates are in the plausible 7–12% range nationally. Both checks pass in our data.

Year 2020 is excluded because the Census Bureau did not release 1-year ACS estimates due to data collection issues during the COVID-19 pandemic.

A.2 Variable Construction

Female self-employment rate = 100 × (B24080_015E + B24080_020E) / B24080_012E

Male self-employment rate = 100 × (B24080_005E + B24080_010E) / B24080_002E

Treatment is defined at the state-year level. A state is treated in year t if its first full year of PFL benefit availability is $\leq t$. For states with mid-year implementations, the first full treated year is the following calendar year.

A.3 Policy Dates

Policy adoption dates are compiled from the Department of Labor paid leave resources and state program websites. The key dates used are the first day benefits became available to workers (not the legislative enactment date).

B. Identification Appendix

B.1 Pre-Trends

Figure 1 in the main text shows the event study with pre-treatment coefficients. While there is some fluctuation at longer pre-treatment horizons (particularly at $e = -5$ through $e = -3$), the coefficients at $e = -2$ and $e = -1$ are close to zero, and no clear trend is evident. The Callaway-Sant'Anna framework does not return a formal pre-test Wald statistic due to singular covariance matrices arising from the small number of treatment clusters, but visual inspection does not suggest systematic pre-trend violations.

B.2 Control Group Selection

The main specification uses never-treated states as controls, defined using the "first full treated year" rule. This includes 43 states that do not have a first full treated year within 2005–2023. DC is a treated unit (cohort 2021), not a control. Results are robust to using not-yet-treated states as controls (Table 4).

C. Robustness Appendix

C.1 Alternative Specifications

All alternative specifications are reported in Table 4. None yields a statistically significant effect or an estimate meaningfully different from the main specification.

C.2 Placebo Test

The male self-employment placebo tests whether the research design detects spurious effects on an outcome that should not be directly affected by PFL's entrepreneurship lock mechanism (which operates through family care responsibilities that fall disproportionately on women). The estimate is -0.28 pp (SE $= 0.21$), which is not statistically significant. The magnitude suggests some common state-level trends affecting both genders, though not enough to reject the null.
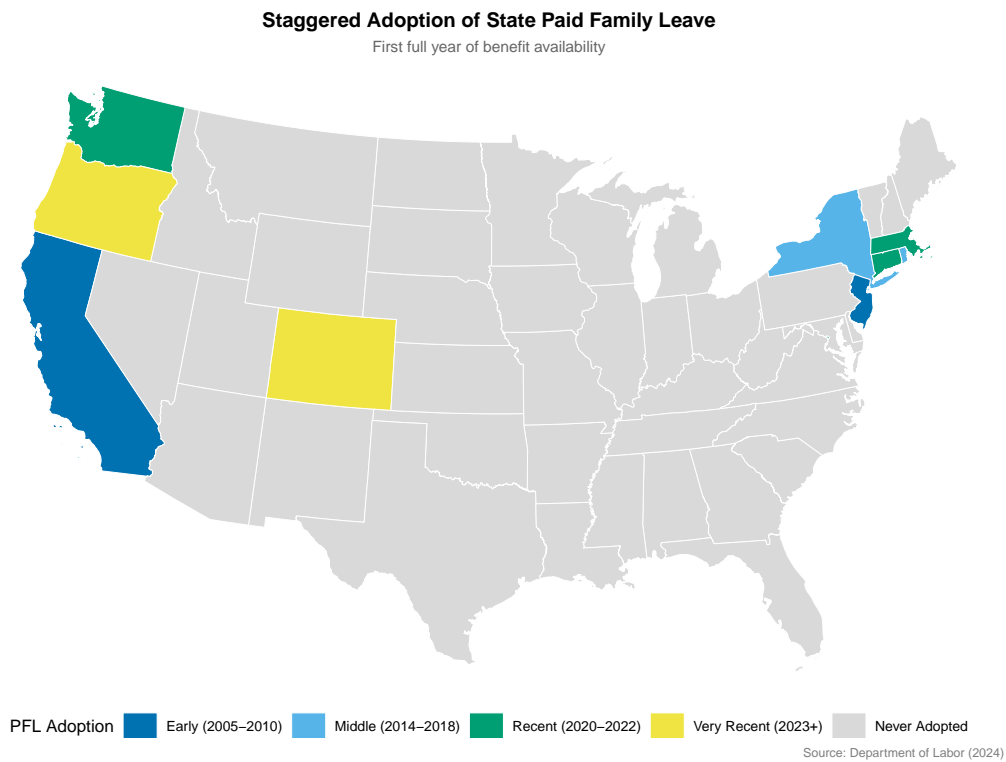
D. Additional Figures

Figure 3: Staggered Adoption of State Paid Family Leave

Notes: Map shows timing of PFL adoption. Grey states have not adopted PFL as of 2024.
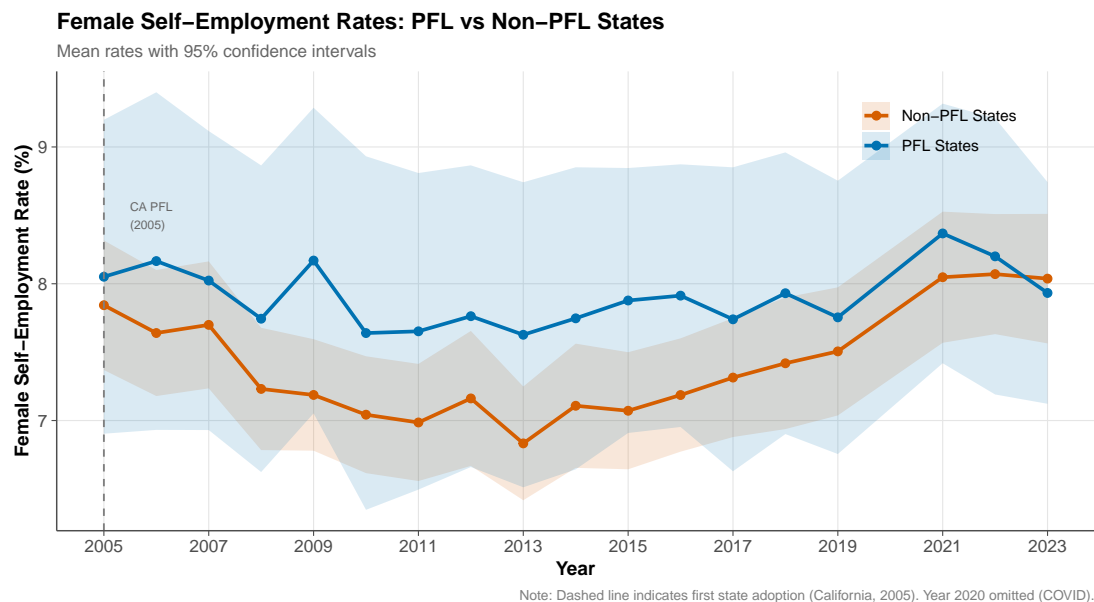
Figure 4: Female Self-Employment Rates: PFL vs. Non-PFL States
Notes: Mean female self-employment rates with 95% confidence intervals. Dashed line indicates California's 2005 adoption. Year 2020 omitted due to ACS data collection issues.