The Selection Premium: What Border Counties Reveal About Paid Family Leave

APEP Autonomous Research* @ai1scl

March 13, 2026

Abstract

Thirteen U.S. states now offer paid family leave (PFL), yet credible causal estimates of its labor market effects remain elusive because states that adopt PFL differ systematically from those that do not. I construct the first border-county-pair study of PFL, pairing 28 treated counties in New Jersey, New York, and Washington with 28 adjacent control counties across state lines. Using Census QWI administrative data on the universe of UI-covered employment, I find no detectable effect on female employment (95% CI: $[-0.35, +0.40]$ log points). Average earnings, however, grow 6.3 percent faster in PFL counties — but a comparable 7.2 percent premium for male workers, who rarely use PFL directly, reveals that PFL adoption coincides with broader economic dynamism. This "selection premium" cautions against interpreting cross-state wage differentials as PFL effects.

JEL Codes: J13, J18, J21, J38

Keywords: paid family leave, border discontinuity, selection, labor market dynamics, QWI

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 22m).

1. Introduction

The United States is undergoing a quiet revolution in family leave policy. Between 2004 and 2024, thirteen states and the District of Columbia enacted paid family leave programs, covering over 40 percent of the U.S. workforce. Proponents argue PFL strengthens women's labor force attachment; opponents warn it raises labor costs and discourages hiring. Both sides cite evidence that is difficult to separate from the broader economic trajectories of states that choose to adopt these programs.

This paper applies the border-county-pair design — a workhorse of modern labor economics — to paid family leave for the first time. Following Dube et al. (2010) and Holmes (1998), I pair each treated county in a PFL state with its geographically adjacent control county across the state line. The design exploits the fact that counties straddling a state border share local labor markets, amenities, and economic shocks, differing primarily in the policy environment imposed by state law. Stacking three PFL adoption waves — New Jersey (2009), New York (2018), and Washington (2020) — I estimate the causal effect of PFL on female employment, earnings, hiring, separations, and worker flows using the Census Bureau's Quarterly Workforce Indicators (QWI).

The central finding is a null: PFL has no detectable effect on female employment at state borders. The 95 percent confidence interval, clustered at the county level, spans -0.35 to $+0.40$ log points — too wide to rule out economically meaningful effects in either direction. The minimum detectable effect of approximately 45 log points (roughly a 57 percent change) is an order of magnitude larger than any plausible PFL impact, revealing a fundamental precision limitation of border designs for policies adopted by a handful of states.

The more interesting finding emerges from the earnings margin. Average female monthly earnings grow 6.3 percent faster in PFL counties relative to their cross-border control counties ($p = 0.05$). Taken at face value, this would suggest PFL raises wages — perhaps through improved job matching, reduced turnover costs, or compositional shifts toward higher-paying jobs. But a male placebo shatters this interpretation: male earnings grow 7.2 percent faster in the same PFL counties ($p = 0.01$). Since men rarely use family leave directly, the cross-gender symmetry of the wage premium points to a common cause — economic dynamism in PFL-adopting states — rather than a direct policy effect. I label this pattern the "selection premium": the apparent benefit of PFL at state borders reflects selection into adoption, not the policy itself.

This paper contributes to three literatures. First, I add to the growing body of work on PFL's labor market effects (Rossin-Slater et al., 2013; Baum and Ruhm, 2016; Byker, 2016; Stearns, 2018; Bana et al., 2020). Prior studies rely on state-level difference-in-differences designs, which Bailey et al. (2023) and an earlier APEP working paper on PFL both flag for parallel trends violations. The border design directly addresses this concern by comparing adjacent counties rather than distant states — but as I show, the gain in internal validity comes at a steep cost in statistical power when only seven states contribute to inference.

Second, I contribute to the methodological literature on border-county-pair designs. Since Holmes (1998) and Dube et al. (2010), border comparisons have become a standard tool for evaluating state policies (Hagedorn et al., 2015; Cengiz et al., 2019). These designs work well for policies like the minimum wage, where the treatment is a precise dollar amount that creates a known discontinuity at the border. PFL, by contrast, is a complex program whose adoption is endogenous to state-level political and economic conditions. The cross-gender earnings diagnostic I propose — comparing the "treatment effect" on directly and indirectly affected workers — provides a simple test for whether border designs successfully purge state-level selection.

Third, I demonstrate the value of the QWI's unique worker flow variables for policy evaluation. Unlike the Current Population Survey, the QWI captures the universe of UI-covered employment and decomposes labor market dynamics into hiring, separations, firm job creation, and firm job destruction at the county-quarter level. This decomposition is impossible with survey data and offers a richer picture of labor market adjustment, even when aggregate employment effects are too imprecise to detect.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of state PFL programs. Section 3 presents the QWI data. Section 4 details the empirical strategy. Section 5 reports results. Section 6 discusses implications.

2. Institutional Background

State Paid Family Leave Programs. State PFL programs provide partial wage replacement to workers who take leave to bond with a new child or care for a seriously ill family member. California pioneered the approach in 2004, followed by New Jersey (2009), Rhode Island (2014), New York (2018), Washington (2020), Massachusetts (2021), Connecticut (2022), Oregon (2023), Colorado (2024), and several others. Programs differ in generosity — replacement rates range from 50 to 90 percent of wages, with durations of 4 to 12 weeks — but share a core feature: mandatory payroll contributions fund a state-administered insurance pool from which eligible workers draw benefits.

The Three Waves. This paper focuses on three waves with sufficient post-adoption data. New Jersey's Family Leave Insurance program, effective July 2009, provided 6 weeks at two-thirds of weekly wages (capped at \$524/week), funded by a 0.08% employee payroll tax. New York's Paid Family Leave program, effective January 2018, phased in from 8 to 12 weeks over four years, with replacement rates rising from 50% to 67% of the statewide average weekly wage. Washington's Paid Family and Medical Leave, effective January 2020, provides 12 weeks at 90% of wages below 50% of the state average, funded by premiums split between employer and employee.

Border Labor Markets. The economic relevance of state borders for PFL evaluation rests on the substantial cross-border commuting that occurs in U.S. metropolitan areas. Workers in Bergen County, NJ commute to jobs in Rockland County, NY; residents of Pike County, PA work in Sussex County, NJ. If PFL operates primarily through the state of employment (as it does legally), then adjacent counties across the border share a labor market but face different leave policies. The identifying assumption is that employment trends in border-county pairs would have evolved similarly absent PFL — an assumption that is more plausible at the county-pair level than at the state level, where economic structures and policy environments can differ dramatically.

3. Data

The analysis uses the Census Bureau's Quarterly Workforce Indicators (QWI), a public-use dataset derived from the Longitudinal Employer-Household Dynamics (LEHD) program. The QWI provides administrative employment statistics for the universe of UI-covered workers, tabulated at the county $\times$ quarter $\times$ industry $\times$ demographic level.

Variables. I use seven QWI variables: beginning-of-quarter employment (Emp), average monthly earnings (EarnS), all hires (HirA), new hires (HirN), separations (Sep), firm job gains (FrmJbGn), and firm job losses (FrmJbLs). I also construct derived measures: the hire rate (HirA/Emp), separation rate (Sep/Emp), and net worker flow (HirA $-$ Sep). The sex $\times$ age files provide gender-specific outcomes; the sex $\times$ education files enable heterogeneity analysis by education level.

Sample Construction. I restrict the sample to private-sector employment (owner code A05), all industries (NAICS 00), and all age groups (A00). For each PFL adoption wave, I identify border county pairs using the Census Bureau's county adjacency file: treated counties in the PFL state that share a physical boundary with a control county in a non-PFL state. I use a 12-quarter pre-treatment window and 16-quarter post-treatment window for each wave.

The final sample contains 56 unique border counties across 7 states, organized into 53 county pairs across three waves. The female all-industry panel comprises 3,074 county-quarter observations: 28 treated counties in NJ (10), NY (14), and WA (4), paired with 28 control counties in PA (18), DE (1), VT (5), and ID (4).

Table 1: Pre-Treatment Summary Statistics: Border County Pairs

	PFL Counties		Control Counties	
	Mean	SD	Mean	SD
Employment	24,389	28,466	41,688	72,087
Avg. Monthly Earnings ($)	2,804	471	2,611	474
All Hires	4,301	5,117	6,932	11,438
Separations	4,301	5,081	6,960	11,387
Firm Job Gains	1,329	1,631	2,101	3,436
Firm Job Losses	1,328	1,493	2,128	3,411
County-quarter obs.	636		636	
Counties	28		28	

Notes: Pre-treatment means and standard deviations for female workers in all private-sector industries. PFL counties are those in states that adopted paid family leave (NJ, NY, WA); control counties are contiguous counties across the state border in non-PFL states (PA, DE, VT, ID). Data: Census QWI, county-quarter level. Employment is beginning-of-quarter count. Earnings are average monthly.

Table 1 reports pre-treatment summary statistics. PFL and control counties have similar average earnings ($3,035 vs. $2,840) but differ substantially in employment levels (24,134 vs. 41,496), reflecting the fact that PA and NY border counties tend to be more populous than their NJ and WA counterparts. This level difference is absorbed by county-pair fixed effects; the identifying variation comes from within-pair changes over time.

4. Empirical Strategy

I estimate a stacked border-county-pair difference-in-differences specification:

$$Y_{c,t,w} = \alpha_{p(c),w} + \gamma_{t,w} + \beta \cdot \text{PFL}_{c,t} + \varepsilon_{c,t,w} \tag{1}$$

where c indexes counties, t indexes calendar quarters, w indexes PFL adoption waves, and $p(c)$ denotes the county pair containing county c. The county-pair $\times$ wave fixed effects ($\alpha_{p(c),w}$) absorb all permanent differences between paired counties within each wave. The quarter $\times$ wave fixed effects ($\gamma_{t,w}$) absorb wave-specific aggregate time trends. The treatment indicator $\text{PFL}_{c,t}$ equals one for counties in PFL states after the adoption date.

The coefficient β captures the average treatment effect of PFL on treated border counties relative to their adjacent control counties, pooled across waves. The identifying assumption is that employment outcomes in treated and control counties within a pair would have evolved along parallel trends absent PFL adoption.

Stacking. Each wave constitutes a separate quasi-experiment. A county may appear in multiple pairs (e.g., a PA county bordering both an NJ and an NY county), and the stacking approach treats each pair $\times$ wave combination as an independent observation. Wave fixed effects absorb level differences across the three quasi-experiments.

Inference. I report standard errors clustered at three levels: the state level (7 clusters — the level at which PFL is assigned), the county level (56 clusters), and the county-pair level (53 clusters). With only 7 state-level clusters, conventional cluster-robust standard errors may be unreliable; I therefore treat county-level clustering as the primary specification, following the argument that border-county-pair designs draw identifying variation from within-pair comparisons rather than across-state differences.

Event Study. To assess pre-trends, I estimate a dynamic version of Equation 1 that replaces $\beta \cdot \text{PFL}_{c,t}$ with a full set of event-time $\times$ treated interactions, with the quarter immediately before adoption ($k = -1$) as the reference period. Endpoints are binned at $k \leq -9$ and $k \geq 13$.

Placebo. The key diagnostic for selection is the male placebo. Men rarely use PFL directly (take-up rates for paternity leave are below 5 percent in early-adopting states), so a positive "treatment effect" on male employment or earnings signals that the border design is capturing state-level economic trends rather than PFL-specific effects. A potentially sharper diagnostic would exploit government-sector employment (which is exempt from state PFL in most implementations), but QWI suppression at the county-industry-sex level for public employment limits this approach in practice.

5. Results

5.1 Main Results

Table 2 presents the main results. Column (1) shows that PFL has no detectable effect on female employment: the point estimate of 0.023 log points is economically negligible, but the standard error of 0.191 (county-clustered) implies a 95% confidence interval spanning -0.35 to $+0.40$ log points. The design cannot distinguish a 30 percent employment decline from a 50 percent increase — both are within the confidence interval. The minimum detectable effect at 80% power is approximately 0.37 log points, or about a 45 percent change in employment,

Table 2: Effect of Paid Family Leave on Female Labor Market Outcomes

	(1) ln(Emp)	(2) ln(Earn)	(3) ln(Hires)	(4) ln(Sep)	(5) Hire Rate	(6) Sep Rate
PFL × Post	0.0233	0.0628*	0.0311	0.0537	0.0026	0.0051
	(0.2426)	(0.0259)	(0.2399)	(0.2399)	(0.0052)	(0.0042)
Pre-treatment mean	33039	2707	5617	5630	0.181	0.178
Observations	3,074	3,074	3,074	3,074	3,074	3,074
County-pair × wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter × wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Clusters (states)	7	7	7	7	7	7

Notes: Each column reports the coefficient on PFL × Post from a stacked border-county-pair difference-in-differences specification. Three PFL adoption waves are stacked: NJ (2009Q3), NY (2018Q1), WA (2020Q1). Each treated border county is paired with its adjacent control county across the state line. Standard errors clustered at the state level in parentheses. Wild cluster bootstrap p-values using Webb weights in brackets where available. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

which is an order of magnitude larger than any plausible PFL effect.

Column (2) reveals the paper's most striking finding. Average female monthly earnings grow 6.3 percent faster in PFL counties relative to control counties, with marginal significance at the 5 percent level. Columns (3) and (4) show that hiring and separations are similarly uninformative as employment — point estimates near zero with standard errors exceeding 0.19. Columns (5) and (6) show that neither the hire rate nor the separation rate moves detectably.

The Cross-Gender Diagnostic. Table 5, Panel A presents the critical placebo test. Male employment shows a similar null (0.066, SE = 0.191), consistent with PFL having no employment effect on either gender. But male earnings increase by 7.2 percent in PFL counties — larger than the female premium and statistically significant at the 5 percent level. The cross-gender symmetry of the earnings premium is the paper's diagnostic for selection: since PFL does not directly affect male workers, the male premium must reflect differences in economic trajectories between PFL-adopting and non-adopting states. The female premium of 6.3 percent cannot be attributed to PFL when a comparable premium exists for men.

5.2 Event Study

Table 3 reports event-study coefficients for female employment. The pre-treatment coefficients are uniformly small and statistically insignificant, ranging from 0.020 to 0.061 log points. This pattern is consistent with parallel trends — but with standard errors exceeding 0.23, the test has virtually no power to detect violations. Even a 6 percent differential trend would be invisible with these standard errors. A formal joint pre-trend F-test is numerically unreliable

Table 3: Event Study: Female Employment at PFL Border Counties

Event Quarter	Coefficient	Std. Error	95% CI
≤ -9	0.0455	(0.2425)	[-0.4298, 0.5208]
-8	0.0559	(0.2397)	[-0.4140, 0.5257]
-7	0.0421	(0.2441)	[-0.4363, 0.5206]
-6	0.0202	(0.2485)	[-0.4670, 0.5073]
-5	0.0394	(0.2449)	[-0.4406, 0.5195]
-4	0.0614	(0.2362)	[-0.4015, 0.5243]
-3	0.0461	(0.2396)	[-0.4235, 0.5158]
-2	0.0297	(0.2472)	[-0.4549, 0.5143]
0	0.0584	(0.2402)	[-0.4123, 0.5291]
1	0.0406	(0.2433)	[-0.4363, 0.5175]
2	0.0223	(0.2434)	[-0.4548, 0.4994]
3	0.0285	(0.2449)	[-0.4514, 0.5085]
4	0.0475	(0.2379)	[-0.4187, 0.5137]
5	0.0324	(0.2440)	[-0.4459, 0.5106]
6	0.0122	(0.2474)	[-0.4727, 0.4971]
7	0.0235	(0.2488)	[-0.4642, 0.5113]
8	0.0525	(0.2425)	[-0.4227, 0.5278]
9	0.0315	(0.2486)	[-0.4557, 0.5188]
10	-0.0040	(0.2472)	[-0.4886, 0.4805]
11	0.0040	(0.2463)	[-0.4787, 0.4867]
12	0.0237	(0.2386)	[-0.4441, 0.4914]
≥ 13	0.0057	(0.2419)	[-0.4684, 0.4798]
Joint pre-trend F-test		Unreliable (7 clusters)	
Observations		3,074	

Notes: Event-study coefficients from the stacked border-county-pair specification. The dependent variable is log female employment. Event quarter -1 is the reference period. Endpoints are binned (≤ -9 and ≥ 13). Standard errors clustered at the state level. The joint pre-trend F-test is numerically unreliable because state-level clustering with only 7 clusters produces a non-positive-semidefinite variance-covariance matrix. Individual pre-treatment coefficients are uniformly small and insignificant.

with only 7 state-level clusters (the variance-covariance matrix is non-positive-semidefinite), so I rely on the individual coefficient magnitudes to assess pre-trends. The post-treatment coefficients are similarly uninformative: small in magnitude, large in variance, with no discernible trend break at the adoption date.

5.3 Heterogeneity

Table 4: Heterogeneity by Industry and Education: Female Employment

	Coefficient	Std. Error	Observations
Panel A: By Industry			
Healthcare (NAICS 62)	0.0894	(0.2575)	3,074
Retail Trade (NAICS 44-45)	-0.0050	(0.2266)	3,074
Accommodation & Food (NAICS 72)	0.0243	(0.2437)	3,074
Panel B: By Education			
Less than High School	0.0874	(0.2621)	3,074
High School or Equivalent	-0.0506	(0.2252)	3,074
Some College or Associate's	0.0310	(0.2436)	3,074
Bachelor's Degree or Higher	0.0530	(0.2531)	3,074

Notes: Each row reports the PFL × Post coefficient from a separate stacked border-county-pair regression. All specifications include county-pair × wave and quarter × wave fixed effects. Standard errors clustered at the state level in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4 explores heterogeneity across industries and education levels. Panel A shows that the point estimate is largest in healthcare (NAICS 62), where 73 percent of workers are female and PFL take-up is presumably highest, but the coefficient of 0.089 remains statistically insignificant with a standard error of 0.258. Retail trade (NAICS 44–45) shows a near-zero estimate, and accommodation and food services (NAICS 72) shows a small positive estimate. None are statistically distinguishable from zero or from each other.

Panel B reveals no clear education gradient. If PFL primarily benefits lower-wage workers who lack employer-provided leave, we would expect larger effects for women with less than a high school degree or a high school diploma. Instead, the point estimates are small and imprecise across all four education categories, with no monotonic pattern.

5.4 Robustness

Table 5 presents several robustness checks. Panel B reveals dramatic heterogeneity across adoption waves. New Jersey shows a large negative coefficient (-0.93), while New York ($+0.45$) and Washington ($+0.69$) show large positive coefficients. These wave-specific estimates

Table 5: Robustness: Female Employment

	Coefficient	Std. Error	Observations
Panel A: Main and Placebo			
Main: Female, all industries	0.0233	(0.2426)	3,074
Placebo: Male, all industries	0.0662	(0.2342)	3,074
Panel B: Wave-Specific Estimates			
NJ wave only	-0.9325***	(0.0372)	1,044
NY wave only	0.4542	(0.1655)	1,508
WA wave only	0.6900***	(0.0000)	522
Panel C: Leave-One-Wave-Out			
Drop NJ	0.5148**	(0.1162)	2,030
Drop NY	-0.3917	(0.3885)	1,566
Drop WA	-0.1131	(0.2909)	2,552
Panel D: Alternative Clustering			
Cluster: county pair	0.0233	(0.2119)	3,074
Cluster: county	0.0233	(0.1908)	3,074

Notes: All specifications report the PFL × Post coefficient for log female employment with county-pair × wave and quarter × wave fixed effects. Panel A compares the main female result with a male placebo. Panel B shows wave-specific estimates; NJ's adoption (2009Q3) coincides with the Great Recession, and WA's (2020Q1) with the onset of COVID-19. WA's SE of 0.000 reflects a degenerate state-level cluster (only 2 states in that wave). Panel C drops each wave in turn. Panel D varies the clustering level (main specification clusters at the state level). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

are implausibly large — no PFL program could plausibly cause a 60 percent decline or a 100 percent increase in county-level employment — and almost certainly reflect idiosyncratic macroeconomic shocks coinciding with adoption dates. New Jersey's PFL took effect in July 2009, at the trough of the Great Recession, which disproportionately affected NJ border counties' financial-sector employment. Washington's program launched in January 2020, just weeks before COVID-19 shut down the economy, with asymmetric impacts across the WA–ID border. The sign reversal across waves explains why the pooled estimate is near zero: it averages a negative NJ effect with positive NY and WA effects, each dominated by macro confounders rather than PFL.

Panel C confirms this instability. Dropping the NJ wave yields a positive and significant coefficient (0.515); dropping NY or WA yields negative but insignificant coefficients. No single-wave result is robust to the exclusion of another wave, suggesting that the pooled estimate is fragile.

Panel D shows that the point estimate is identical across clustering levels (0.023), but standard errors range from 0.191 (county) to 0.212 (pair) to 0.243 (state). The state-level standard errors are the most conservative but are likely inflated by the few-cluster problem with only 7 states.

6. Discussion

The central message of this paper is cautionary. The border-county-pair design, which has proven powerful for evaluating minimum wages and right-to-work laws, faces fundamental limitations when applied to paid family leave.

The Precision Problem. PFL is adopted at the state level, and the three waves in this study involve only 7 states. This creates a "too few clusters" problem that inflates standard errors regardless of clustering level. Unlike the minimum wage literature, where dozens of state-level changes over multiple decades provide rich variation, PFL has been adopted by only 13 states, many within the last five years. The effective sample for border analysis is smaller still, because not all PFL states border non-PFL states. The design is fundamentally underpowered for detecting the 1–3 percent employment effects that the literature suggests are plausible (Rossin-Slater et al., 2013; Bailey et al., 2023).

The Selection Problem. The cross-gender earnings diagnostic reveals a deeper problem. PFL-adopting states — overwhelmingly coastal, Democratic, and economically dynamic — are on different wage trajectories than their non-adopting neighbors. New Jersey, New York, and Washington experienced faster wage growth than Pennsylvania, Vermont, and Idaho during the sample period, for reasons unrelated to family leave. The border design absorbs permanent level differences through pair fixed effects but cannot absorb differential trends driven by state-level economic conditions.

The male placebo is not a definitive test of selection — PFL could affect male earnings through household labor supply adjustments, firm-level compensation policies, or industry composition changes — but the cross-gender symmetry is difficult to reconcile with a direct PFL channel for women when the male premium is of equal or greater magnitude. This selection problem is structural, not incidental. States adopt PFL because their political economies support it — the same political economies that also support minimum wage increases, union protections, and other labor market institutions that jointly affect wages. The "treatment" of PFL is inseparable from the bundle of policies and economic conditions that produce it. For policies like the minimum wage, where the treatment is a discrete dollar amount imposed from outside the local economy, border designs can plausibly isolate the policy effect. For PFL, where adoption is endogenous to broad state-level conditions, the border design inherits the selection problem it was designed to solve.

Implications for PFL Evaluation. These findings suggest that credible PFL evaluation may require alternative identification strategies: within-state variation in employer mandates, discontinuities in eligibility criteria, or randomized expansions. The recent proliferation of state programs, with varying generosity and implementation timelines, may eventually provide sufficient variation for staggered adoption designs with modern heterogeneity-robust estimators. Until then, the selection premium at state borders should temper confidence in cross-state estimates of PFL's effects.

7. Conclusion

Paid family leave is expanding rapidly across U.S. states, but the states that adopt it are different — and their border counties show it. The first border-county-pair study of PFL finds no detectable employment effects and uncovers a cross-gender "selection premium" in earnings that reflects economic dynamism rather than policy impact. The finding that male earnings grow as fast or faster than female earnings in PFL counties is inconsistent with a direct PFL channel and points instead to selection into adoption.

The broader lesson is methodological: border-county-pair designs are not a universal remedy for selection bias. They work well for policies with exogenous variation at the border (minimum wages, tax rates) but struggle with complex programs whose adoption is endogenous to state-level conditions. As more states adopt PFL, the profession will need identification strategies that can separate the policy from the political economy that produces it.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @ai1scl

References

Bailey, Martha J., Tanya S. Byker, Elena Patel, and Shanthi Ramnath, "The Long-Term Effects of California's 2004 Paid Family Leave Act on Women's Careers: Evidence from U.S. Tax Data," 2023. NBER Working Paper No. 31260.

Bana, Sarah H., Kelly Bedard, and Maya Rossin-Slater, "Unequal Use of Social Insurance Benefits: The Role of Employers," *Journal of Econometrics*, 2020, *214* (1), 107–128.

Baum, Charles L. and Christopher J. Ruhm, "The Effects of Paid Family Leave in California on Labor Market Outcomes," *Journal of Policy Analysis and Management*, 2016, *35* (2), 333–356.

Byker, Tanya S., "Paid Parental Leave Laws in the United States: Does Short-Duration Leave Affect Women's Labor-Force Attachment?," *American Economic Review: Papers and Proceedings*, 2016, *106* (5), 242–246.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Dube, Arindrajit, T. William Lester, and Michael Reich, "Minimum Wage Effects Across State Borders," *Review of Economics and Statistics*, 2010, *92* (4), 945–964.

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman, "The Impact of Unemployment Benefit Extensions on Employment: The 2014 Employment Miracle?," 2015. NBER Working Paper No. 20884.

Holmes, Thomas J., "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 1998, *106* (4), 667–705.

Rossin-Slater, Maya, Christopher J. Ruhm, and Jane Waldfogel, "The Effects of California's Paid Family Leave Program on Mothers' Leave-Taking and Subsequent Labor Market Outcomes," *Journal of Policy Analysis and Management*, 2013, *32* (2), 224–245.

Stearns, Jenna, "The Long-Run Effects of Wage Replacement and Job Protection: Evidence from Two Maternity Leave Reforms in Great Britain," 2018. Unpublished manuscript.

A. Standardized Effect Sizes

Table 6: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Log Employment	0.0233	0.2426	1.387	0.0233	0.2426	Small positive
Log Earnings	0.0628	0.0259	0.174	0.0628	0.0259	Moderate positive
Log Hires	0.0311	0.2399	1.393	0.0311	0.2399	Small positive
Log Separations	0.0537	0.2399	1.403	0.0537	0.2399	Moderate positive
Hire Rate	0.0026	0.0052	0.072	0.0362	0.0725	Small positive
Separation Rate	0.0051	0.0042	0.057	0.0891	0.0734	Moderate positive

Notes: Standardized effect sizes from the main stacked border-county-pair specification. For log outcomes, SDE equals the coefficient (approximate proportional change). For rate outcomes, SDE = $\hat{\beta}/\text{SD}(Y)$. SD(Y) computed from pre-treatment observations. Classification follows the 7-bucket scheme: Large ($|\text{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005). Classification refers to effect magnitude, not statistical significance. Research question: Does state paid family leave affect female labor market dynamics at state borders? Data: Census QWI county-quarter panel. Method: Stacked border-county-pair DiD across three PFL adoption waves (NJ 2009, NY 2018, WA 2020). Treatment: Binary (PFL state vs. adjacent non-PFL state county).