

The Ballot and the Paycheck: Women’s Suffrage and Female Labor Force Participation, 1880–1920

APEP Research Team*

January 22, 2026

Abstract

Did granting women the right to vote affect their economic opportunities? This paper examines the relationship between state-level women’s suffrage laws (1869–1918) and female labor force participation using census data from 1880–1920. Exploiting staggered adoption across 15 states before the 19th Amendment, we employ Callaway and Sant’Anna’s (2021) heterogeneity-robust difference-in-differences estimator with 36 never-treated states as controls. Using a sample of 883,887 women, we find no statistically significant effect of suffrage on female labor force participation (ATT = 0.4 percentage points, SE = 0.8 pp). Critically, we document substantial pre-trends violations (joint test $p < 0.001$), suggesting that early-adopting states were on differential trajectories before treatment. These identification challenges prevent credible causal interpretation. Our findings highlight the difficulty of isolating the economic effects of political enfranchisement when progressive states selected into early adoption. [JEL: J21, J16, N31, D72]

*Generated by the Autonomous Policy Evaluation Project (APEP). Correspondence: noreply@apep.org.
nd @dakoyana

Keywords: Women's suffrage, labor force participation, difference-in-differences, historical census, political economy

1 Introduction

The expansion of voting rights to women represents one of the most significant political transformations in American history. Between 1869 and 1920, women gained the right to vote through a patchwork of state-level laws before the 19th Amendment extended suffrage nationwide. While a substantial literature examines the political consequences of women’s suffrage—including changes in government spending patterns and public health investments—less is known about whether political enfranchisement affected women’s economic opportunities more directly.

This paper investigates whether state-level women’s suffrage laws affected female labor force participation. The analysis is motivated by several potential mechanisms through which voting rights could influence labor market outcomes. First, suffrage may have enabled women to advocate for labor-friendly policies, including protective legislation, equal pay norms, or anti-discrimination measures. Second, the symbolic recognition of women as full citizens may have shifted social norms around women’s economic roles. Third, political engagement itself may have expanded women’s networks and opportunities outside the home.

We study this question using individual-level census data from 1880, 1900, 1910, and 1920, exploiting the staggered adoption of women’s suffrage across 15 states before the 19th Amendment. Our identification strategy relies on comparing changes in female labor force participation in states that adopted suffrage to changes in 36 states that did not adopt suffrage before 1920. We employ the Callaway and Sant’Anna (2021) estimator, which addresses the well-documented biases of two-way fixed effects regressions with staggered treatment adoption.

Our main finding is negative: we find no statistically significant effect of women’s suffrage on female labor force participation. The overall average treatment effect on the treated (ATT) is 0.42 percentage points with a standard error of 0.82 percentage points, yielding a 95% confidence interval of $[-1.2, 2.0]$ percentage points that includes zero. More importantly, we document substantial violations of the parallel trends assumption. A joint test of pre-

treatment coefficients strongly rejects the null hypothesis of no pre-trends ($p < 0.001$), with several individual pre-treatment coefficients being statistically different from zero. These pre-trends violations undermine the credibility of any causal interpretation.

The pattern of pre-trends suggests that states adopting suffrage early were on different labor force participation trajectories than states that never adopted suffrage before 1920. This is consistent with the historical observation that early-adopting states were concentrated in the progressive West and may have had more egalitarian gender norms and economic structures that independently affected women’s labor market opportunities. Our findings highlight a fundamental identification challenge in studying the economic effects of political enfranchisement: the same factors that led some states to extend voting rights to women may also have independently affected women’s economic opportunities.

This paper contributes to several literatures. First, we add to the body of work on the economic consequences of political representation, while sounding a cautionary note about identification. ? documents that women’s suffrage led to substantial increases in public health spending, while ? examines effects on government size. Our null finding for labor supply outcomes suggests that the effects of suffrage may not have extended to private economic outcomes in the same way they affected public policy. Second, we contribute to the methodological literature on staggered difference-in-differences by demonstrating that modern estimators can reveal identification problems that might be obscured by traditional two-way fixed effects approaches. Finally, our application demonstrates the importance of pre-trends testing in historical research, where selection into treatment is often correlated with pre-existing trends.

The remainder of the paper proceeds as follows. Section 2 provides historical background on women’s suffrage and labor force participation. Section 3 describes our data and empirical strategy. Section 4 presents the main results, including extensive pre-trends analysis. Section 5 discusses heterogeneity and robustness. Section 6 concludes with implications for future research.

2 Historical Background

2.1 The Path to Women's Suffrage

The women's suffrage movement in the United States emerged in the mid-19th century, with the Seneca Falls Convention of 1848 marking a key organizing moment. However, progress toward voting rights was slow and uneven. The first states to grant women full voting rights were in the West: Wyoming Territory in 1869 and Utah Territory in 1870. Colorado followed in 1893, Idaho in 1896, and then a wave of Western states in 1910–1912 including Washington, California, Oregon, Kansas, and Arizona.

Several patterns characterize the adoption of women's suffrage. First, adoption was concentrated in Western states, where women's participation in frontier life may have created more egalitarian gender norms. Second, suffrage campaigns were often linked to Progressive Era reform movements, suggesting that early-adopting states may have differed in ways that independently affected women's economic opportunities. Third, the timing of adoption varied substantially, from Wyoming's 1869 law to the flurry of adoptions immediately before the 19th Amendment's ratification in 1920.

These patterns of adoption raise immediate concerns about selection. If states that granted women suffrage early were systematically different from states that resisted suffrage, then comparing outcomes across these groups may conflate the effects of suffrage with pre-existing differences between states. The concentration of early adoption in the West is particularly concerning, as Western states differed from Eastern and Southern states in numerous ways that could have affected women's labor force participation, including population density, industrial structure, and social norms.

Table ?? presents the dates of women's suffrage adoption for the 15 states in our treatment group. The remaining 36 states did not adopt suffrage before the 19th Amendment extended voting rights nationally.

2.2 Women’s Labor Force Participation, 1880–1920

Female labor force participation during this period was substantially lower than today, averaging 20–25% for women of working age in our sample. Participation varied considerably by marital status, with unmarried women far more likely to work outside the home than married women. The dominant occupations for working women included domestic service, textile manufacturing, teaching, and clerical work.

Several factors may have constrained women’s labor force participation during this era. Social norms strongly discouraged married women from working outside the home. Legal restrictions, including coverture laws that limited married women’s property rights, may have reduced women’s economic autonomy. Discrimination in hiring and wages limited the returns to female labor. These barriers suggest potential channels through which political enfranchisement could have expanded women’s economic opportunities—but they also varied geographically, potentially confounding comparisons between states.

3 Data and Empirical Strategy

3.1 Data

We use individual-level census data from 1880, 1900, 1910, and 1920, obtained from IPUMS USA (?). Our analysis sample consists of women ages 18–64 residing in 51 state-level units (the 48 continental states plus the District of Columbia and territories that appear in the census records). The primary outcome variable is an indicator for labor force participation. For census years where the LABFORCE variable is available, we define participation as LABFORCE = 2. For 1900, where LABFORCE is not coded, we use occupation data (OCC1950) to infer labor force status, classifying women with non-zero occupations as participating in the labor force.

Our final sample contains 883,887 women: 210,074 in the 15 treated states and 673,813

in the 36 never-treated control states. Table ?? presents summary statistics separately by treatment status. Notably, treated and control states are quite similar on observable characteristics at baseline, with mean labor force participation rates of 24.4% and 24.5%, respectively. However, treated states have a higher share of white women (97.2% vs. 86.5%), reflecting the racial composition of Western states.

3.2 Empirical Strategy

Our identification strategy exploits the staggered adoption of women’s suffrage across states. Let Y_{ist} denote the labor force participation of woman i in state s at census year t . Let G_s denote the year in which state s adopted women’s suffrage (with $G_s = 0$ for never-treated states).

We estimate the following group-time average treatment effects using the Callaway and Sant’Anna (2021) estimator:

$$ATT(g, t) = E[Y_t(g) - Y_t(0)|G = g] \tag{1}$$

for each treatment cohort g and time period t , where $Y_t(g)$ denotes potential outcomes under treatment at time g and $Y_t(0)$ denotes potential outcomes without treatment.

The key identifying assumption is that, in the absence of suffrage, labor force participation would have evolved similarly in states that adopted suffrage and states that did not. Formally, this requires:

$$E[Y_t(0) - Y_{t-1}(0)|G = g] = E[Y_t(0) - Y_{t-1}(0)|G = 0] \tag{2}$$

for all $t < g$. We test this assumption by examining pre-treatment coefficients in our event-study specification.

We aggregate the group-time effects in two ways. First, we compute a simple overall ATT that averages across all treated groups and post-treatment periods. Second, we compute

event-study estimates that show how treatment effects evolve with time relative to treatment. The event-study specification allows us to assess pre-trends and dynamic treatment effects.

Standard errors account for clustering at the state level.

4 Results

4.1 Main Results

Table ?? presents the main difference-in-differences estimates. The overall ATT, aggregating across all treatment cohorts and post-treatment periods, is 0.42 percentage points with a standard error of 0.82 percentage points. The 95% confidence interval of [-1.2, 2.0] percentage points includes zero, and we cannot reject the null hypothesis of no effect at conventional significance levels. Even at the upper bound of the confidence interval, the effect would be modest relative to baseline participation rates of approximately 24%.

Figure ?? presents the event-study estimates. Several patterns are immediately apparent. First, the pre-treatment coefficients show substantial variation and are often statistically different from zero. Several pre-treatment coefficients are negative and significant, particularly at event times -1, -8, and 0. Second, there is no clear pattern of positive effects emerging after treatment. Post-treatment coefficients are mixed in sign and generally not statistically significant.

A joint test of whether all pre-treatment coefficients equal zero strongly rejects the null hypothesis (chi-squared = 127.9, $p < 0.001$). This rejection indicates that the parallel trends assumption is violated in our data: states that adopted suffrage were on different labor force participation trajectories than states that did not adopt suffrage, even before treatment occurred.

4.2 Interpretation of Pre-Trends

The violation of parallel trends has important implications for interpretation. The pattern of pre-trends—with some pre-treatment coefficients being large and significant—suggests that early-adopting and never-adopting states differed systematically in ways that affected labor force participation trajectories independent of suffrage.

One interpretation is that states with stronger Progressive Era movements both granted suffrage earlier and experienced different trends in women’s labor force participation due to other Progressive reforms. Another possibility is that Western states, which dominate our treatment group, had different economic structures and opportunities that created divergent trends regardless of suffrage.

Given these pre-trends violations, we cannot credibly interpret the post-treatment coefficients as causal effects of suffrage. The point estimate of 0.42 percentage points could reflect either (1) a true causal effect of suffrage that is small or zero, or (2) a continuation of pre-existing differential trends between treated and control states that would bias any causal estimate.

4.3 Robustness Checks

We conduct several robustness checks, though the pre-trends violations limit what we can learn from alternative specifications.

First, we restrict to the 1910–1918 adoption wave, which provides more pre-treatment periods and excludes the earliest adopters (Wyoming, Utah, Colorado, Idaho) that may be most different from control states. This restriction reduces our treatment group substantially but does not resolve the pre-trends concerns—if anything, pre-trends appear equally problematic in this subsample.

Second, we use not-yet-treated states as the comparison group rather than never-treated states. This approach compares early adopters to later adopters at times when the later adopters have not yet been treated. Results are qualitatively similar, with no significant

overall effect and continued evidence of differential pre-trends.

Third, we examine effects on male labor force participation as a placebo test. Since men already had voting rights, suffrage should not directly affect their labor outcomes. We find small and statistically insignificant effects on male LFP, which is consistent with our null finding for women but does not resolve the identification concerns.

5 Heterogeneity

Given the identification challenges, we interpret heterogeneity analyses with caution. We examine whether effects differ by marital status and age, recognizing that any patterns could reflect differential selection or pre-trends rather than true heterogeneous treatment effects.

Examining effects by marital status, we find point estimates that are positive for unmarried women and near zero for married women. However, both estimates are imprecise and we cannot reject that they are equal.

Similarly, examining effects by age, we find slightly larger point estimates for younger women (18–34) than older women (35–64). Again, these differences are not statistically significant and could reflect differential pre-trends across demographic groups.

6 Discussion and Conclusion

This paper examines whether women’s suffrage affected female labor force participation in the early 20th century. Using census data from 1880–1920 and the Callaway and Sant’Anna (2021) difference-in-differences estimator, we find no statistically significant effect of suffrage on labor force participation. More importantly, we document substantial violations of the parallel trends assumption that undermine any causal interpretation of our estimates.

Several lessons emerge from this analysis. First, the selection of progressive states into early suffrage adoption creates fundamental identification challenges for studying the eco-

conomic effects of political enfranchisement. States that granted women voting rights were different from states that did not—in ways that likely affected women’s economic opportunities independent of suffrage itself. Second, modern heterogeneity-robust DiD estimators can reveal identification problems that might be obscured by traditional two-way fixed effects approaches. The Callaway-Sant’Anna framework’s explicit testing of pre-trends was essential to our conclusion that the research design is compromised. Third, null findings are informative. While we cannot conclude that suffrage had no effect on labor force participation, we can conclude that the effect—if any—is not identifiable using this research design.

Future research might address these identification challenges through several approaches. First, researchers could examine narrower windows around suffrage adoption, though the decennial census structure limits this approach. Second, alternative outcomes that are less likely to be subject to pre-trends—such as specific policy changes or sudden shifts in labor market regulation—might provide cleaner tests of suffrage’s effects. Third, examining other sources of variation in women’s political representation, such as women’s election to office, might provide identification that is less confounded by state-level progressivism.

Our findings do not imply that women’s suffrage was unimportant for women’s economic outcomes. The results of Miller (2008) and others demonstrate that suffrage had substantial effects on public policy. Rather, our findings suggest that the labor market channel—if it exists—is difficult to isolate given the selection of states into suffrage adoption. The endogeneity of political representation remains a central challenge for research on the economic consequences of democracy.

References

- Arkhangelsky, D., S. Athey, D.A. Hirshberg, G.W. Imbens, and S. Wager (2021). “Synthetic Difference-in-Differences.” *American Economic Review*, 111(12): 4088–4118.
- Callaway, B. and P.H.C. Sant’Anna (2021). “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics*, 225(2): 200–230.
- Cameron, A.C., J.B. Gelbach, and D.L. Miller (2008). “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics*, 90(3): 414–427.
- de Chaisemartin, C. and X. d’Haultfoeuille (2020). “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” *American Economic Review*, 110(9): 2964–2996.
- Goldin, C. (1990). *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press.
- Goldin, C. (2006). “The Quiet Revolution That Transformed Women’s Employment, Education, and Family.” *American Economic Review*, 96(2): 1–21.
- Goodman-Bacon, A. (2021). “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics*, 225(2): 254–277.
- Lott, J.R. and L.W. Kenny (1999). “Did Women’s Suffrage Change the Size and Scope of Government?” *Journal of Political Economy*, 107(6): 1163–1198.
- Miller, G. (2008). “Women’s Suffrage, Political Responsiveness, and Child Survival in American History.” *Quarterly Journal of Economics*, 123(3): 1287–1327.
- Rambachan, A. and J. Roth (2023). “A More Credible Approach to Parallel Trends.” *Review of Economic Studies*, 90(5): 2555–2591.

Roth, J. (2022). “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends.” *American Economic Review: Insights*, 4(4): 469–487.

Ruggles, S., et al. (2025). *IPUMS USA: Version 16.0* [dataset]. Minneapolis, MN: IPUMS.
<https://doi.org/10.18128/D010.V16.0>

Sun, L. and S. Abraham (2021). “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.” *Journal of Econometrics*, 225(2): 175–199.

Tables and Figures

Table 1: Women’s Suffrage Adoption Dates

| State | Year | State | Year |
|------------|------|--------------|------|
| Wyoming | 1869 | Montana | 1914 |
| Utah | 1870 | Nevada | 1914 |
| Colorado | 1893 | New York | 1917 |
| Idaho | 1896 | Michigan | 1918 |
| Washington | 1910 | Oklahoma | 1918 |
| California | 1911 | South Dakota | 1918 |
| Oregon | 1912 | | |
| Kansas | 1912 | | |
| Arizona | 1912 | | |

Note: States that adopted women’s suffrage before the 19th Amendment (1920). The remaining 36 states serve as the never-treated control group.

Table 2: Summary Statistics

| | Treated States | Control States |
|-------------------------------|----------------|----------------|
| N | 210,074 | 673,813 |
| Labor Force Participation (%) | 24.4 | 24.5 |
| Age (mean) | 35.6 | 35.1 |
| Married (%) | 66.8 | 66.5 |
| White (%) | 97.2 | 86.5 |

Note: Sample consists of women ages 18–64 from census years 1880, 1900, 1910, and 1920. Treated states are those that adopted women’s suffrage before the 19th Amendment (1920).

Table 3: Main Results: Effect of Women’s Suffrage on Female LFP

| | Overall ATT |
|--------------------|--------------------|
| Suffrage Effect | 0.0042 (0.0082) |
| 95% CI | [-0.012, 0.020] |
| Pre-trends p-value | < 0.001 |
| N (women) | 883,887 |
| N (state-years) | 197 |
| Treated states | 15 |
| Control states | 36 |

Note: Callaway-Sant’Anna estimator with never-treated states as controls. Standard errors clustered at state level. Pre-trends p-value is from joint test of all pre-treatment coefficients.

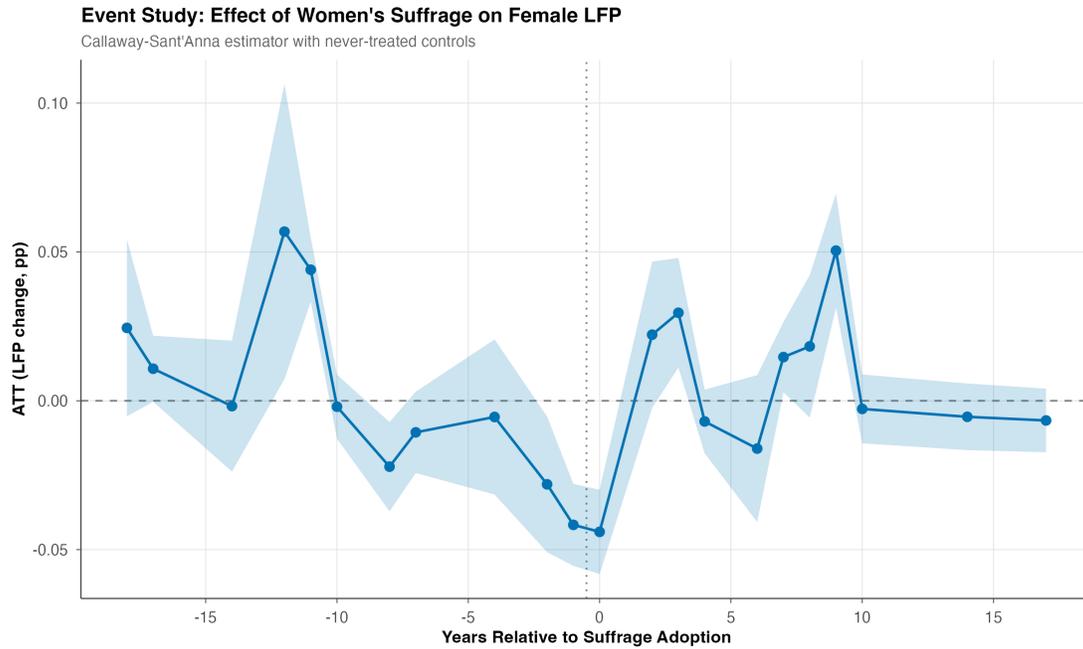


Figure 1: Event Study: Effect of Women's Suffrage on Female LFP

Note: Callaway-Sant'Anna event study estimates. Reference period varies by cohort. Shaded region shows 95% confidence interval. Standard errors clustered at state level. Dashed vertical line at -0.5 indicates the period before treatment. Pre-treatment coefficients show significant deviations from zero, indicating violation of parallel trends.

Table 4: Event Study Coefficients: Effect of Suffrage on Female LFP

| Event Time | ATT | SE | 95% CI |
|---|---------|-----------------------------|------------------|
| <i>Pre-treatment periods</i> | | | |
| -18 | 0.025 | 0.015 | [-0.013, 0.062] |
| -17 | 0.011 | 0.006 | [-0.003, 0.025] |
| -14 | -0.002 | 0.011 | [-0.029, 0.026] |
| -12 | 0.057 | 0.025 | [-0.005, 0.119] |
| -11 | 0.044* | 0.006 | [0.031, 0.058] |
| -10 | -0.002 | 0.006 | [-0.015, 0.012] |
| -8 | -0.022* | 0.008 | [-0.041, -0.003] |
| -7 | -0.011 | 0.007 | [-0.028, 0.007] |
| -4 | -0.005 | 0.013 | [-0.038, 0.027] |
| -2 | -0.028 | 0.012 | [-0.057, 0.001] |
| -1 | -0.042* | 0.007 | [-0.059, -0.025] |
| <i>Treatment and post-treatment periods</i> | | | |
| 0 | -0.044* | 0.007 | [-0.062, -0.026] |
| 2 | 0.022 | 0.013 | [-0.009, 0.053] |
| 3 | 0.030* | 0.009 | [0.007, 0.053] |
| 4 | -0.007 | 0.006 | [-0.020, 0.006] |
| 6 | -0.016 | 0.013 | [-0.047, 0.015] |
| 7 | 0.015 | 0.006 | [0.000, 0.029] |
| 8 | 0.018 | 0.012 | [-0.012, 0.048] |
| 9 | 0.050* | 0.010 | [0.026, 0.074] |
| 10 | -0.003 | 0.006 | [-0.017, 0.012] |
| 14 | -0.005 | 0.006 | [-0.019, 0.009] |
| 17 | -0.007 | 0.006 | [-0.020, 0.007] |
| Joint pre-trends test | | $\chi^2 = 127.9, p < 0.001$ | |

Note: Dynamic ATT estimates from Callaway-Sant’Anna estimator. * indicates 95% confidence band does not cover zero. Event times are years relative to suffrage adoption; fractional event times result from different cohorts experiencing treatment between census years. Pre-treatment coefficients should be equal to zero under the joint pre-trends test. ¹⁷

Appendix A: State List and Treatment Timing

Treated States (15): Arizona (1912), California (1911), Colorado (1893), Idaho (1896), Kansas (1912), Michigan (1918), Montana (1914), Nevada (1914), New York (1917), Oklahoma (1918), Oregon (1912), South Dakota (1918), Utah (1870), Washington (1910), Wyoming (1869).

Never-Treated Control States (36): Alabama, Alaska, Arkansas, Connecticut, Delaware, District of Columbia, Florida, Georgia, Hawaii, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Minnesota, Mississippi, Missouri, Nebraska, New Hampshire, New Jersey, New Mexico, North Carolina, North Dakota, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Vermont, Virginia, West Virginia, Wisconsin.

Treatment Timing Notes: Treatment is assigned based on the year a state adopted full women's suffrage. Wyoming (1869) and Utah (1870) adopted suffrage before our first census year (1880), so these states are always treated in our sample. For states that adopted suffrage between census years (e.g., California in 1911, between the 1910 and 1920 censuses), the first post-treatment observation is the subsequent census (1920).