

City Votes, Country Voices: Urban-Rural Heterogeneity in the Labor Market Effects of Women’s Suffrage, 1880–1920

Autonomous Policy Evaluation Project (APEP)

January 21, 2026

Abstract

This paper examines whether the labor market effects of women’s suffrage differed between urban and rural areas during the Progressive Era. Using IPUMS full-count census data from 1880–1920 and modern staggered difference-in-differences methods, we document a surprising pattern: suffrage increased female labor force participation by approximately 2.8 percentage points in rural areas—a statistically significant effect—while producing a smaller and statistically insignificant effect of 1.5 percentage points in urban areas. This pattern directly contradicts the prevailing policy channel hypothesis, which predicted urban effects would dominate because protective labor legislation operated primarily in formal wage labor markets concentrated in cities. We explore several interpretations of this finding, including ceiling effects in urban labor markets, changes in the measurement of women’s work, and the operation of political voice through local programs rather than state-level protective legislation. Our findings challenge conventional accounts of how democratic inclusion translates into economic outcomes and suggest that the mechanisms through which suffrage affected women’s economic lives were more complex than the protective legislation channel that has dominated prior scholarship.

JEL Codes: J16, J21, N31, N32, D72

Keywords: Women's suffrage, labor force participation, urban-rural divide, difference-in-differences, Progressive Era

1 Introduction

In 1910, a woman living in Seattle could cast a ballot in state elections; her counterpart in rural Kansas could not. By 1912, the Kansas woman had gained the franchise as well. Yet these two women inhabited starkly different labor markets. The Seattle woman lived in a rapidly industrializing city where women worked in department stores, telephone exchanges, and clerical offices—wage labor governed by formal contracts, subject to state regulation, and visible to policymakers. The Kansas woman lived on a farm where her labor, though economically essential, was unpaid, informal, and embedded within household production. When women gained the right to vote, did their newfound political power translate into economic opportunity? And did this transformation unfold differently across the urban-rural divide that defined American economic life in the early twentieth century?

This paper investigates the heterogeneous effects of women’s suffrage on female labor force participation across urban and rural areas during the Progressive Era. We exploit the staggered adoption of women’s suffrage across U.S. states between 1869 and 1920 using a triple-difference design that compares changes in women’s labor supply in treated versus control states, before versus after suffrage adoption, and in urban versus rural areas within states. Drawing on full-count census data from the Integrated Public Use Microdata Series ([Ruggles et al., 2023](#)), we analyze the labor market decisions of over 50 million women across four census years (1880, 1900, 1910, and 1920). Our identification strategy leverages modern difference-in-differences methods robust to heterogeneous treatment effects ([Callaway and Sant’Anna, 2021](#); [Sun and Abraham, 2021](#); [Goodman-Bacon, 2021](#)), addressing concerns about bias in canonical two-way fixed effects estimators when treatment timing varies.

Our findings reveal a striking and unexpected pattern: the labor market effects of women’s suffrage were *larger in rural areas than in urban areas*—the opposite of what the policy channel hypothesis predicts. Among rural women, suffrage adoption increased labor force participation by approximately 2.8 percentage points, a statistically significant effect. Among urban women, the effect was smaller at 1.5 percentage points and not statistically

significant. This rural dominance pattern persists across a battery of robustness checks, including alternative estimators robust to heterogeneous treatment effects ([Sun and Abraham, 2021](#)), restrictions to late-adopting states with multiple pre-treatment periods, alternative urbanization thresholds, and analyses stratified by race and age. The pattern is not driven by compositional changes in who lived in urban versus rural areas, nor by differential measurement error across contexts.

These results speak to a fundamental question in political economy: through what channels does democratic inclusion translate into economic change? The theoretical literature offers two broad mechanisms. First, extending the franchise may enable newly enfranchised groups to advocate for policies that directly improve their economic position—what we term the “policy channel.” Women’s organizations in the Progressive Era lobbied vigorously for protective labor legislation, minimum wages, workplace safety standards, and equal pay provisions ([Kessler-Harris, 1982](#); [Skocpol, 1992](#)). If suffrage empowered these advocacy efforts, effects should concentrate in labor markets where such policies could bite: urban areas with formal wage employment, regulatory oversight, and observable working conditions. Second, suffrage may have shifted social norms about appropriate gender roles—a “norms channel” that could operate independently of policy. If voting symbolized women’s capacity for public participation and rational deliberation, this ideological shift might have expanded the sphere of acceptable female activity, potentially reducing social sanctions against women who worked outside the home. The norms channel predicts effects that could be as strong or stronger in rural areas, where traditional gender ideologies were more deeply entrenched and where the symbolic meaning of enfranchisement might have been particularly transgressive.

Our finding of larger rural effects challenges the prevailing emphasis on protective legislation as the mechanism linking suffrage to labor market outcomes. This surprising pattern requires alternative explanations. One possibility is that suffrage operated through measurement channels: the symbolic recognition of women’s civic capacity may have shifted how census enumerators and women themselves understood farm women’s economic contri-

butions, leading to increased *reporting* of labor force participation even absent behavioral change. Another possibility is that suffrage affected rural women through policy channels other than protective legislation—county-level programs, agricultural extension services, and local government responsiveness to newly enfranchised voters (Miller, 2008; Carruthers and Wanamaker, 2018). A third interpretation emphasizes ceiling effects: urban women already participated in the labor force at higher rates, leaving less room for policy-induced increases. We cannot definitively adjudicate among these mechanisms with our data, but our findings suggest that the channels through which democratic inclusion translates into economic opportunity were more complex than the protective legislation hypothesis implies.

Historical Context: Women’s Labor and the Urban-Rural Divide

The period from 1880 to 1920 witnessed a profound transformation in American women’s labor force participation, a transformation that played out quite differently in cities and on farms. At the opening of our study period, female labor force participation stood at roughly 15 percent nationwide, though this figure obscures substantial heterogeneity (Goldin, 1990). In rural areas, women’s work was largely invisible to census enumerators: farm wives labored in dairy production, poultry raising, vegetable gardening, and food preservation—activities that were economically essential but classified as “keeping house” rather than gainful employment. Urban women, by contrast, increasingly entered formal wage labor in the burgeoning sectors of manufacturing, domestic service, retail trade, and clerical work. By 1920, female labor force participation had risen to approximately 24 percent, with gains concentrated among young, unmarried, urban women (Goldin, 2006).

This urban-rural divergence in women’s labor markets reflected deeper structural differences in economic organization. Urban labor markets featured clear distinctions between home and workplace, between paid labor and household production. Women who worked for wages did so in observable settings—factories, offices, stores—where their hours, conditions, and compensation were matters of public record. This visibility made urban women’s

labor a natural target for Progressive Era reform. State legislatures debated and enacted laws governing maximum hours for women workers, minimum wages for female employees, prohibitions on night work, and requirements for rest breaks and sanitary facilities ([Goldin, 1990](#); [Kessler-Harris, 1982](#)). The constitutionality of such sex-specific protective legislation was affirmed by the Supreme Court in *Muller v. Oregon* (1908), which accepted the argument that women's physical differences and maternal responsibilities justified special state protection.

Rural labor markets operated on entirely different principles. Farm women's productive activities were embedded within household economies where distinctions between work and leisure, production and consumption, blurred into irrelevance. A woman who spent her morning milking cows, her afternoon preserving vegetables, and her evening mending clothes was engaged in continuous economic production, yet census enumerators recorded her occupation as "none" or "keeping house." This invisibility had both measurement implications—female labor force participation in rural areas was systematically understated—and policy implications. Protective labor legislation could not reach work that took place within the household, that was organized by family relationships rather than employment contracts, and that was compensated through shared consumption rather than individual wages.

The women's suffrage movement navigated these distinct terrains with varying success. Urban suffragists drew on networks of women's clubs, settlement houses, and labor organizations that provided organizational infrastructure for political mobilization ([Flexner, 1975](#)). Rural suffragists faced the challenges of geographic dispersion, limited communication networks, and the resistance of agricultural communities to federal interference of any kind. Yet the suffrage movement ultimately succeeded in both contexts: by the time the Nineteenth Amendment was ratified in 1920, women in every state had gained the franchise. The question we address is whether this uniform political change produced uniform economic effects.

Related Literature

Our paper contributes to three interconnected literatures: the economic history of women’s suffrage, the economics of gender gaps in labor markets, and the econometrics of difference-in-differences designs with heterogeneous treatment effects.

The foundational contribution to the economics of women’s suffrage is [Lott and Kenny \(1999\)](#), who documented that suffrage adoption was associated with significant increases in state government expenditure and revenue. They interpreted this finding through the lens of median voter theory: women’s policy preferences differed systematically from men’s, and enfranchisement shifted the political equilibrium toward policies favored by the new median voter. [Miller \(2008\)](#) extended this insight in an influential study showing that suffrage led to large increases in public health spending and corresponding decreases in child mortality. Miller’s identification strategy exploited within-state variation in the timing of suffrage adoption using a difference-in-differences framework, and his results demonstrated that women’s political voice translated into concrete policy outcomes that improved child welfare. More recently, [Moehling and Thomasson \(2020\)](#) provided a comprehensive synthesis of the economic literature on women’s enfranchisement, documenting effects on public expenditure composition, prohibition legislation, and various measures of government responsiveness.

Our contribution to this literature is to examine labor market outcomes rather than government policy or public health. While previous work has established that suffrage changed what governments did, we ask whether suffrage changed what women themselves did in the labor market. This distinction matters for understanding the channels through which democratic inclusion affects economic outcomes. If suffrage improved women’s labor market position primarily through policy changes—protective legislation, equal pay laws, anti-discrimination enforcement—then effects should be concentrated in labor markets where such policies have purchase. Our urban-rural comparison provides a direct test of this mechanism.

The economics of gender gaps in labor markets has documented both the long-run evolution of female labor force participation and the persistence of gender differences in earnings,

occupations, and employment (Goldin, 1990, 2006; Olivetti and Petrongolo, 2016). Goldin (1990) provided the canonical account of how American women's labor force participation evolved from the nineteenth century through the present, emphasizing the role of structural transformation, educational expansion, and changing social norms. Her framework distinguishes between the extensive margin—whether women work at all—and the intensive margin—what kinds of jobs women hold and how much they earn conditional on working. We focus on the extensive margin, examining whether suffrage affected the probability that women participated in the labor force.

A crucial insight from this literature is that female labor force participation patterns differed dramatically by marital status, urban residence, and nativity during our study period (Goldin, 1990; Kessler-Harris, 1982). Young, unmarried women entered the labor force at high rates, particularly in urban areas where opportunities in manufacturing, retail, and clerical work were expanding. Married women, by contrast, largely withdrew from formal employment upon marriage—a pattern that prevailed until the mid-twentieth century. Our empirical strategy accounts for these compositional differences by including detailed controls for age, marital status, and race, and by examining heterogeneous effects across these dimensions.

Related work on women's labor in the Progressive Era has emphasized the importance of protective legislation in shaping female employment patterns. Goldin (1988) found that maximum hours laws for women reduced female employment in regulated industries while potentially increasing employment in unregulated sectors. Moehling (2009) documented the effects of Progressive Era child labor laws on family labor supply decisions, showing how regulations targeting one family member could affect the labor supply of others. These findings suggest that the policy environment mattered considerably for women's labor market outcomes in this period, and that changes in that environment—potentially including changes brought about by women's suffrage—could have had meaningful effects on female employment.

The recent econometric literature on difference-in-differences with staggered treatment adoption has fundamentally reshaped empirical practice in economics. [Goodman-Bacon \(2021\)](#) demonstrated that canonical two-way fixed effects estimators can be severely biased when treatment effects are heterogeneous across units or over time. In staggered adoption designs, the TWFE estimator implicitly uses already-treated units as controls for newly-treated units, and when treatment effects vary with time since treatment, this comparison yields a weighted average of effects with potentially negative weights. [Callaway and Sant'Anna \(2021\)](#) and [Sun and Abraham \(2021\)](#) proposed alternative estimators that aggregate only “clean” comparisons between treated and not-yet-treated (or never-treated) units, producing consistent estimates under weaker assumptions. [Roth et al. \(2023\)](#) provided a comprehensive synthesis of this literature, offering practical guidance for applied researchers.

We implement these modern methods throughout our analysis. Our preferred specification uses the [Callaway and Sant'Anna \(2021\)](#) estimator with never-treated states as the comparison group, aggregating group-time average treatment effects into event-study plots that visualize treatment effect dynamics. We report results from multiple estimators—including standard TWFE, the [Sun and Abraham \(2021\)](#) interaction-weighted estimator, and stratified analyses—to assess robustness. Following [Rambachan and Roth \(2023\)](#), we also report sensitivity analyses that allow for violations of parallel trends, showing how our conclusions change under assumptions that permit pre-existing differential trends of various magnitudes.

Our Contribution

This paper makes three contributions. First, we provide new evidence on the labor market effects of women’s suffrage, complementing existing work on public expenditure and public health outcomes. While [Miller \(2008\)](#) and [Moehling and Thomasson \(2020\)](#) have documented that suffrage changed government policy, we examine whether suffrage changed women’s own economic behavior. Our finding of positive effects on female labor force participation—with

an overall effect of approximately 2.3 percentage points representing roughly 10 percent of baseline participation—demonstrates that political empowerment translated into measurable changes in women’s labor market outcomes.

Second, we document surprising heterogeneity in the effects of democratic inclusion by urban-rural residence. This finding has important implications for understanding how political empowerment translates into economic opportunity. The concentration of effects in rural areas contradicts the policy channel hypothesis, which predicted urban effects would dominate because protective legislation operated primarily in formal wage labor markets. Instead, our results suggest that alternative mechanisms—measurement changes, local programs, or compositional effects—may have been equally or more important than state-level protective legislation. This reframing has implications for how we understand the political economy of suffrage: women’s influence over county budgets, school board decisions, and extension program priorities may have mattered as much for economic outcomes as the high-profile protective legislation that has dominated scholarly attention.

Third, we demonstrate the application of modern staggered difference-in-differences methods to an important historical question. The women’s suffrage literature has largely relied on pre-2020 econometric methods that are now known to produce potentially biased estimates under treatment effect heterogeneity. By implementing the estimators of [Callaway and Sant’Anna \(2021\)](#), [Sun and Abraham \(2021\)](#), and the sensitivity analyses of [Rambachan and Roth \(2023\)](#), we show how these modern tools can be brought to bear on historical questions while providing researchers with a template for similar applications.

Our findings also connect to broader debates about the relationship between political institutions and economic development. [Acemoglu et al. \(2001\)](#) and subsequent work in political economy has emphasized that democratic institutions can promote economic prosperity by constraining extractive elites and providing voice to broader segments of the population. The extension of suffrage to women represents one of the largest expansions of political participation in American history, and understanding its economic consequences illuminates

the mechanisms through which democratic inclusion might affect development. Our results suggest that these mechanisms operated differently than existing theory predicts: political voice translated into economic opportunity more powerfully in rural areas than in urban ones, challenging accounts that emphasize formal regulatory channels as the primary link between political empowerment and economic outcomes.

The paper also speaks to ongoing debates about gender inequality and political representation. Contemporary research has documented that female political representation—women serving as legislators, mayors, or executives—can affect policy outcomes in domains relevant to women’s interests ([Chattopadhyay and Duflo, 2004](#); [Ferreira and Gyourko, 2014](#)). Our historical analysis examines a more fundamental form of political inclusion: the extension of voting rights to half the population. If voting rights alone could produce meaningful changes in women’s economic outcomes, this suggests a powerful channel through which political institutions shape economic inequality. Conversely, if voting rights produced limited or heterogeneous effects, this might suggest that formal political inclusion is insufficient without complementary changes in economic and social institutions.

Data, Methods, and Empirical Strategy

Our empirical analysis draws on full-count census data from the Integrated Public Use Microdata Series (IPUMS), covering the census years 1880, 1900, 1910, and 1920 ([Ruggles et al., 2023](#)). The full-count data provide a comprehensive enumeration of the American population, allowing us to examine labor force participation patterns across states, years, urban-rural residence, and demographic subgroups with minimal sampling error. We restrict attention to women aged 18-64, the population of prime working age for whom labor force participation is a meaningful margin of adjustment.

Our identification strategy exploits the staggered adoption of women’s suffrage across states prior to the Nineteenth Amendment. Between 1869 (Wyoming Territory) and 1918 (Michigan, Oklahoma, South Dakota), thirteen states or territories extended full voting

rights to women in state elections. The remaining states adopted women’s suffrage only upon ratification of the Nineteenth Amendment in August 1920. This variation in treatment timing provides the foundation for our difference-in-differences design.

We exclude the earliest adopters—Wyoming (1869) and Utah (1870)—from our primary analysis because the available census data do not include pre-treatment observations for these states. Our treatment group therefore consists of eleven states that adopted suffrage between 1893 (Colorado) and 1918, with comparison provided by the thirty-three states that never adopted state-level suffrage and serve as a “never-treated” control group. This design follows best practices in the modern difference-in-differences literature, which emphasizes the value of never-treated comparison groups for avoiding contamination by heterogeneous treatment effects ([Callaway and Sant’Anna, 2021](#); [Roth et al., 2023](#)).

The triple-difference design adds a third dimension of comparison: urban versus rural residence within states. This extension allows us to separate state-wide changes in female labor force participation—which might reflect suffrage effects operating through policy, norms, or other channels—from differential changes by urban-rural residence—which speak specifically to mechanisms that operate more strongly in one context than the other. The key identifying assumption is that, absent suffrage, the difference in female labor force participation between urban and rural women would have evolved similarly in treated and control states. We assess this assumption through event-study analyses that examine pre-treatment dynamics, and through sensitivity analyses that allow for bounded violations of parallel trends.

Our primary outcome measure is labor force participation, coded from the IPUMS variable **LABFORCE**, which harmonizes information on gainful employment across historical census years. We recognize that measurement of women’s labor force status is imperfect in historical data, particularly for married women and rural women whose work may have been undercounted or misclassified ([Goldin, 1990](#); [Folbre and Abel, 1991](#)). We address this concern in several ways: by examining sensitivity to alternative definitions of labor force participation, by considering heterogeneous effects by marital status (where measurement concerns

are most acute for married women), and by conducting placebo tests using male labor force participation (which should be unaffected by women’s suffrage).

Roadmap

The remainder of this paper proceeds as follows. Section 2 provides historical background on the women’s suffrage movement and the evolution of women’s labor markets during the Progressive Era. We document the geographic and temporal patterns of suffrage adoption, describe the policy and social changes that accompanied enfranchisement, and develop hypotheses about why suffrage effects might differ between urban and rural areas.

Section 3 describes our data sources, sample construction, and variable definitions. We present summary statistics on female labor force participation by state, year, and urban-rural residence, and document the variation in suffrage timing that underlies our identification strategy.

Section 4 presents our econometric framework. We begin with the canonical two-way fixed effects specification and explain why this approach may produce biased estimates under treatment effect heterogeneity. We then describe our implementation of modern staggered difference-in-differences estimators, including the [Callaway and Sant’Anna \(2021\)](#) group-time aggregation procedure and the [Sun and Abraham \(2021\)](#) interaction-weighted estimator. We explain our approach to inference, which clusters standard errors at the state level to account for serial correlation within states and potential correlation of errors across individuals within state-year cells.

Section 5 presents our main findings. We begin with overall effects of suffrage on female labor force participation, then introduce the urban-rural heterogeneity that is our central contribution. Event-study figures visualize the dynamics of treatment effects, providing evidence on pre-trends and the timing of effects relative to suffrage adoption. We report results from multiple estimators and across a range of specifications to assess robustness.

Section 6 explores mechanisms underlying the urban-rural divergence. We examine

whether effects differ by age, marital status, and race—dimensions that may help distinguish between competing explanations for the patterns we observe. We also examine effects on occupational composition among working women, asking whether suffrage changed not only whether women worked but also what kinds of work they did.

Section 7 presents additional robustness checks. We implement the sensitivity analysis of [Rambachan and Roth \(2023\)](#), which allows for bounded violations of parallel trends; we restrict attention to late-adopting states with multiple pre-treatment periods; we examine placebo outcomes that should be unaffected by suffrage; and we assess sensitivity to alternative control group definitions and sample restrictions.

Section 8 interprets our findings in light of the historical record and existing literature. We discuss the policy implications of our results for understanding how democratic inclusion affects economic outcomes, and we consider limitations of our analysis and directions for future research.

Section 9 concludes with a summary of our contributions and their implications for the economics of political institutions, gender, and labor markets.

2 Historical Background

The struggle for women’s suffrage in the United States spanned more than seven decades, from the Seneca Falls Convention of 1848 to the ratification of the Nineteenth Amendment in 1920. This prolonged campaign unfolded against a backdrop of profound economic transformation—industrialization, urbanization, and the emergence of a modern service economy—that reshaped women’s relationship to paid labor in ways that varied dramatically between cities and the countryside. Understanding the historical context of suffrage adoption and the structure of Progressive Era labor markets is essential for interpreting the heterogeneous effects we document in this paper. This section provides that context, tracing the geographic and temporal patterns of enfranchisement, describing the divergent labor

market conditions facing urban and rural women, and developing a theoretical framework for why the effects of suffrage on female labor force participation might differ by location.

2.1 The Path to Women’s Suffrage

The movement for women’s voting rights in America began in earnest at the Seneca Falls Convention of 1848, where Elizabeth Cady Stanton introduced a resolution declaring that “it is the duty of the women of this country to secure to themselves their sacred right to the elective franchise” ([Flexner, 1975](#)). Yet more than two decades would pass before any American jurisdiction granted women the right to vote. The path from Seneca Falls to the Nineteenth Amendment was neither straight nor swift; it proceeded through a patchwork of state-level victories that created the variation in treatment timing essential to our empirical strategy.

Wyoming Territory became the first jurisdiction to enfranchise women in 1869, a decision motivated partly by a desire to attract female settlers to the sparsely populated frontier and partly by genuine commitment to equal rights among territorial legislators ([Keyssar, 2000](#)). Utah Territory followed in 1870, though Congress revoked Utah women’s suffrage in 1887 as part of federal efforts to suppress polygamy; the franchise was restored when Utah achieved statehood in 1896. These early Western adoptions established a pattern that would persist throughout the suffrage campaign: the frontier West proved far more receptive to women’s voting rights than the established states of the East and South.

The next wave of suffrage expansion came in the 1890s, when Colorado (1893) and Idaho (1896) joined Wyoming and Utah as full-suffrage states. A long period of stasis followed—suffrage referenda failed repeatedly in state after state during the early 1900s—before a renewed surge of victories beginning in 1910. Washington State adopted women’s suffrage in 1910, followed by California in 1911 and three states (Oregon, Kansas, and Arizona) in 1912. Montana and Nevada enfranchised women in 1914, New York in 1917, and Michigan, Oklahoma, and South Dakota in 1918. By the time the Nineteenth Amendment

was ratified in August 1920, thirteen states had already granted women full voting rights in state elections, while the remaining thirty-five states—concentrated in the East, Midwest, and especially the South—had resisted until federal action forced their hand ([Moehling and Thomasson, 2020](#)).

The geographic pattern of suffrage adoption reflected deeper regional differences in political culture, economic structure, and gender ideology. Western states adopted suffrage earliest and most readily, a pattern scholars have attributed to several factors: the relative scarcity of women on the frontier, which elevated their value and status; the weaker grip of traditional gender hierarchies in newly settled communities; the alliance between suffragists and Progressive reformers who were especially influential in Western politics; and the practical experience Western women had gained in homesteading, ranching, and community building ([Teele, 2018](#)). Eastern states, with their established political machines and entrenched interests, proved more resistant. Southern states were most opposed of all, fearing that women’s suffrage would complicate the elaborate legal machinery constructed to disenfranchise Black voters after Reconstruction ([Keyssar, 2000](#)).

Table 6 presents the complete chronology of state-level suffrage adoption. For our empirical analysis, we focus on states that adopted between 1893 and 1918—the period for which full-count census data are available both before and after treatment. Wyoming and Utah are excluded from our primary analysis because no pre-treatment census data exist for these earliest adopters; the first available census (1880) postdates Wyoming’s adoption, and Utah’s 1870 adoption left only a single pre-treatment observation before the 1887 revocation complicated the treatment definition. The remaining eleven treatment states span a range of adoption years, providing the variation in treatment timing that enables our staggered difference-in-differences design.

The timing of suffrage adoption was not random, and the determinants of early versus late adoption have implications for our identification strategy. States that adopted suffrage earlier tended to be more rural, more Western, and have smaller populations than later adopters

([Lott and Kenny, 1999](#)). Later adopters like New York (1917) and Michigan (1918) were more urbanized and industrialized, with larger female labor forces working in manufacturing and services. This correlation between adoption timing and state characteristics does not threaten identification—our design compares changes over time within states, differenced against changes in never-treated states—but it does suggest caution in extrapolating effects from early adopters to the broader population. We address this concern through robustness checks that restrict attention to late-adopting states and through event-study analyses that examine whether effects vary systematically with time since treatment.

The suffrage movement itself was deeply intertwined with the broader Progressive reform agenda that animated American politics between 1900 and 1920. Suffragists forged alliances with temperance advocates, labor organizers, settlement house workers, and public health reformers, all of whom saw women’s enfranchisement as instrumental to their respective causes ([Skocpol, 1992](#)). These coalitions were especially effective in urban areas, where women’s clubs, labor unions, and reform organizations provided the organizational infrastructure for sustained political mobilization. The National American Woman Suffrage Association, the dominant suffrage organization by 1900, pursued a state-by-state strategy that concentrated resources on winnable campaigns while building momentum for an eventual federal amendment. This strategy produced the pattern of staggered adoption we exploit: a gradual expansion of the suffrage frontier from West to East, punctuated by high-profile victories in large states like California (1911) and New York (1917) that shifted national momentum toward final passage of the Nineteenth Amendment.

2.2 Urban and Rural Labor Markets in the Progressive Era

The period from 1880 to 1920 witnessed a fundamental transformation of American economic life, driven by industrialization, urbanization, and the rise of large-scale enterprise. At the beginning of this period, the United States remained predominantly agricultural; by its end, the majority of Americans lived in urban areas and worked in manufacturing, commerce,

or services ([Haines and Steckel, 2000](#)). This structural transformation created two starkly different labor market environments for women—one urban, one rural—that would shape how the political change of enfranchisement translated into economic outcomes.

Urban labor markets offered women an expanding array of employment opportunities during the Progressive Era. The manufacturing sector, though dominated by male workers in heavy industry, provided substantial employment for women in textile mills, garment factories, food processing plants, and light manufacturing ([Kessler-Harris, 1982](#)). More dramatic was the growth of employment in retail trade and clerical work, sectors that would become increasingly feminized over the twentieth century. Department stores, which emerged as major employers in American cities during this period, hired young women as salesgirls, cashiers, and office workers. The expansion of business administration created demand for typists, stenographers, filing clerks, and telephone operators—occupations that were explicitly marketed to young women as respectable alternatives to factory work or domestic service ([Goldin, 1990](#)). By 1920, clerical work had become the second-largest occupational category for urban women after domestic service, representing a fundamental shift in the structure of female employment.

The working conditions facing urban women varied enormously across industries and occupations. Factory work often meant long hours, dangerous machinery, poor ventilation, and wages that barely covered subsistence. The Triangle Shirtwaist Factory fire of 1911, which killed 146 workers—mostly young immigrant women—became a galvanizing symbol of the hazards facing industrial workers and spurred demands for workplace safety regulation. At the other end of the spectrum, clerical positions offered shorter hours, cleaner working environments, and the social status associated with “white-collar” employment, though wages were often no higher than in manufacturing ([Kessler-Harris, 1982](#)). What united these diverse urban occupations was their visibility: women who worked for wages did so in settings that were observable, regulable, and subject to the reforms that Progressive activists demanded.

The formalization of urban employment relationships made women’s labor a target for

state intervention. Beginning in the 1890s and accelerating after 1900, state legislatures enacted an array of “protective” legislation governing the terms on which women could be employed. Maximum hours laws limited the length of the working day for women in manufacturing, retail, and other covered industries; by 1920, forty-three states had enacted such restrictions ([Goldin, 1988](#)). Minimum wage laws for women, first adopted by Massachusetts in 1912, spread to fifteen states by 1920. Other regulations prohibited night work for women, required rest periods and meal breaks, mandated seating for saleswomen, and established sanitary standards for workplaces employing women. The constitutionality of sex-specific protective legislation was affirmed in *Muller v. Oregon* (1908), where the Supreme Court accepted the argument that women’s physical differences and maternal responsibilities justified special state protection even when equivalent regulations for men would be struck down as unconstitutional interference with freedom of contract.

Rural labor markets operated on entirely different principles. The majority of rural women lived on farms, where they worked within household economies that blurred the distinctions between production and consumption, paid labor and unpaid household work, that structured urban employment. Farm women’s productive contributions were substantial: they raised poultry and tended dairy cows, maintained vegetable gardens and preserved food for winter, made butter and cheese for home consumption or local sale, and often managed the farm’s accounts ([Haines and Steckel, 2000](#)). Yet this work was largely invisible to the formal economy and to the census enumerators who would classify such women as having no occupation or as “keeping house.” The undercounting of women’s labor was especially severe in agricultural areas, where the boundary between “gainful employment” and “household duties” was genuinely ambiguous and where enumerator practice varied widely ([Folbre and Abel, 1991](#)).

The organization of agricultural labor meant that protective legislation—the primary channel through which suffrage might affect urban women’s working conditions—had little relevance for farm women. Maximum hours laws could not regulate work that had no

employer and no time clock. Minimum wage legislation could not set floors for labor compensated through shared household consumption rather than individual wages. Workplace safety standards could not reach production that occurred in the home, the barn, and the field. Even where rural women did engage in wage labor—as seasonal workers in canning factories, for instance, or as domestic servants in nearby towns—the geographic dispersion and small scale of rural employment made regulation difficult and enforcement nearly impossible.

The demographic composition of the female labor force also differed markedly between urban and rural areas. Urban women who worked for wages were disproportionately young and unmarried; labor force participation rates for married women remained low in cities until well into the twentieth century, reflecting both social norms against married women's employment and practical constraints including the demands of household management without modern conveniences ([Goldin, 1990](#)). Young women typically entered the urban workforce between school-leaving age and marriage, withdrawing from paid employment when they wed. This pattern created a strong life-cycle dimension to urban female labor supply that the suffrage movement might have affected by changing the age at marriage, the probability of marriage, or the expectations of young women about their future roles.

Rural women's labor force participation followed different patterns. Married women on farms were almost universally engaged in productive work, but this work was classified as unpaid family labor rather than labor force participation. Young unmarried women in rural areas faced limited local opportunities for wage employment; those who wished to work for wages often migrated to nearby towns or cities, joining the great rural-to-urban migration that transformed American demographics during this period ([Costa, 2000](#)). The decision to migrate was itself a labor market choice that suffrage might have influenced, though one that is difficult to disentangle from the myriad other factors drawing young people to cities during this era of rapid urbanization.

These structural differences in urban and rural labor markets suggest that any effects of women's suffrage on female employment might manifest quite differently across locations.

The policy channel—suffrage leading to protective legislation leading to improved working conditions—had clear purchase in urban labor markets characterized by formal employment, regulatory oversight, and concentrated advocacy efforts, but limited relevance in rural areas where work was informal, dispersed, and embedded in household production. Understanding these differences is essential for interpreting the patterns we document in subsequent sections.

2.3 Theoretical Framework: Why Suffrage Effects Might Differ by Location

Why might the extension of voting rights to women produce different effects on female labor force participation in urban versus rural areas? We consider three channels through which suffrage could affect women’s labor market outcomes, each with distinct implications for urban-rural heterogeneity.

2.3.1 The Policy Channel: Protective Legislation and Workplace Regulation

The most direct mechanism connecting suffrage to labor market outcomes operates through changes in public policy. Women voters and women’s organizations demonstrated distinct policy preferences that, once enfranchised, could influence the political equilibrium. Historical evidence documents that women voters were more likely than men to support temperance legislation, public health spending, educational investment, and labor market regulations ([Lott and Kenny, 1999](#); [Miller, 2008](#)). Women’s organizations—including the National Consumers’ League, the Women’s Trade Union League, and the General Federation of Women’s Clubs—lobbied actively for protective labor legislation before and after suffrage, but their political leverage increased substantially once women could vote ([Skocpol, 1992](#)).

The policy channel predicts that suffrage effects on female labor force participation should be concentrated in labor markets where regulatory interventions could meaningfully alter working conditions. Urban labor markets meet this criterion: women working in factories, department stores, and offices were subject to state regulation, and improvements in

their working conditions—shorter hours, higher wages, safer workplaces—could make formal employment more attractive relative to home production or non-participation. If suffrage accelerated the adoption and enforcement of protective legislation, we would expect to see increased female labor force participation in regulated industries and, by extension, in urban areas where such industries were concentrated.

The policy channel also predicts that effects should emerge with a lag, as the political process of translating votes into legislation and legislation into enforcement takes time. Miller (2008) documents that the effects of suffrage on public health spending and child mortality materialized over several years following adoption, consistent with a policy mechanism that operated through gradual institutional change rather than immediate behavioral response. Our event-study analysis can shed light on this timing, examining whether effects on labor force participation appear immediately upon suffrage adoption or emerge more gradually in subsequent years.

For rural women, the policy channel predicts weak or null effects. Protective labor legislation could not reach work that was informal, household-based, and outside the regulatory reach of state government. Even policies that did affect rural areas—public health spending, educational investment, agricultural extension services—would operate through channels other than women’s own labor force participation. Rural women might benefit from suffrage-induced policy changes in many ways, but changes in their own employment patterns would not be among the primary margins of adjustment.

2.3.2 The Norms Channel: Voting as a Signal of Citizenship

A second channel operates through social norms about appropriate gender roles. The right to vote was not merely an instrumental political tool; it carried profound symbolic significance as a marker of full citizenship and public participation (Keyssar, 2000). To the extent that opposition to women’s labor force participation reflected ideological commitments about women’s proper sphere—the doctrine of “separate spheres” that confined women to home

and family while reserving public and economic life for men—the extension of voting rights might have disrupted these ideological foundations. If women could be trusted to exercise the franchise responsibly, perhaps they could also be trusted to participate in the workforce, run businesses, or pursue professional careers.

The norms channel predicts a different pattern of urban-rural heterogeneity than the policy channel. Traditional gender ideologies were, if anything, more deeply entrenched in rural areas than in cities. Farm families depended on clear divisions of labor within the household, and rural communities often maintained conservative social mores that resisted changing gender roles. The symbolic meaning of women’s suffrage might therefore have been especially transgressive—and potentially transformative—in rural contexts where traditional gender hierarchies were most firmly established. If the norms channel dominated, we might expect to see suffrage effects that were as strong or stronger in rural areas than in urban ones.

The norms channel also has different implications for timing. Changes in social norms are notoriously difficult to date precisely, but one might expect normative effects to materialize relatively quickly as the symbolic meaning of suffrage diffused through communities. The act of voting itself—women lining up at polling places, casting ballots, participating in the rituals of democratic citizenship—could shift perceptions of women’s public role. Alternatively, normative change might operate through generational replacement, with effects concentrated among younger women who came of age in a post-suffrage environment where female political participation was normalized.

2.3.3 Labor Market Structure: Wage Work versus Family Production

A third consideration involves the structure of labor markets themselves, independent of policy or norms. Urban and rural labor markets offered fundamentally different kinds of economic opportunities for women, and suffrage might have affected women’s ability or willingness to access those opportunities differentially.

In urban labor markets, entry into wage employment required overcoming various barriers: acquiring relevant skills, obtaining information about job opportunities, navigating hiring processes, and managing the practical logistics of combining work with household responsibilities. To the extent that suffrage reduced discrimination against women by employers, expanded women’s networks and information channels, or increased women’s bargaining power within households, it might have lowered these barriers to urban employment. The concentration of female workers in urban industries also created possibilities for collective action—labor organizing, union membership, workplace advocacy—that could improve conditions for all women workers. Suffrage might have empowered these collective efforts by giving women workers a political voice that employers and legislators could not ignore.

In rural labor markets, the relevant margin was not entry into wage employment but the allocation of effort within household production. Farm women faced a portfolio of tasks—some directly productive (dairy, poultry, gardening), some household maintenance (cooking, cleaning, childcare), some market-oriented (selling butter and eggs), and some purely domestic. Suffrage might have affected how women allocated time across these activities, but such shifts would not show up in labor force participation as measured by the census. Indeed, the census categories themselves were poorly suited to capture the economic contribution of farm women, whose work defied easy classification into “employed” versus “not in labor force.”

This structural difference suggests that even if suffrage had equivalent effects on women’s economic empowerment across locations, these effects would manifest in measurable labor force participation only in urban areas where wage employment provided a clear, observable indicator of economic activity. The null rural effects we document might therefore reflect measurement limitations as much as true differences in suffrage’s impact. We return to this issue in our discussion of results, considering whether alternative measures or channels might reveal suffrage effects on rural women that our primary analysis cannot detect.

2.4 Prior Evidence on the Economic Effects of Women’s Suffrage

The economic literature on women’s suffrage has established that enfranchisement produced meaningful changes in public policy, government spending, and social outcomes, though direct evidence on labor market effects remains limited. Our analysis builds on and extends this literature by examining an outcome—female labor force participation—that has received relatively little attention in previous work.

The foundational contribution is [Lott and Kenny \(1999\)](#), who documented that women’s suffrage was associated with substantial increases in state government expenditure and revenue. Using a difference-in-differences framework that exploited variation in the timing of suffrage adoption across states, Lott and Kenny found that per capita state government expenditure increased by 28 percent following suffrage, with corresponding increases in tax revenue. They interpreted these findings through the lens of median voter theory: women’s preferences for government spending differed systematically from men’s, and enfranchisement shifted the political equilibrium toward larger government. The expenditure effects were concentrated in social spending categories—education, public health, welfare—where women’s preferences were thought to diverge most sharply from men’s.

[Miller \(2008\)](#) extended this analysis to examine the effects of suffrage on public health outcomes. Using a similar difference-in-differences design with individual-level data on child mortality, Miller found that suffrage led to immediate and substantial increases in local public health spending, which in turn produced large reductions in infant and child mortality. The effects were concentrated among children under age nine and were driven primarily by declines in infectious disease mortality—outcomes directly connected to public health interventions like sanitation, water purification, and milk safety programs. Miller’s analysis demonstrated that suffrage effects operated through specific policy channels rather than through diffuse changes in economic conditions or social organization.

Subsequent research has documented suffrage effects on other policy domains. [Aidt and Dallal \(2015\)](#) examined women’s suffrage in Western Europe and found effects on social

spending similar to those documented for the United States. [Bertocchi \(2011\)](#) developed a theoretical model of how female enfranchisement affects the size and composition of the welfare state, with predictions that align with observed patterns of post-suffrage policy change. [Klassen \(2023\)](#) examined the effects of suffrage on political participation itself, finding that enfranchisement increased women’s engagement with the political process beyond the simple act of voting.

Research on suffrage and education has produced mixed findings. [Carruthers and Wanamaker \(2014\)](#) examined the effects of suffrage on schooling provision, finding evidence of increased educational investment following enfranchisement. However, the effects appear to have been mediated by racial politics, with suffrage producing different outcomes in states with different racial compositions. [Naidu \(2012\)](#) examined suffrage in the context of the post-bellum South, where the interaction of gender and race created a particularly complex political environment. These studies suggest that suffrage effects were context-dependent, varying with the political and demographic characteristics of adopting states.

Direct evidence on suffrage effects on female labor force participation is scarce. The existing literature has focused primarily on policy outcomes—what governments did after suffrage—rather than on behavioral outcomes—what women themselves did. Our analysis fills this gap by examining whether the political empowerment of women translated into changes in women’s own economic activity. The evidence from the policy literature provides theoretical motivation for expecting labor market effects: if suffrage changed protective labor legislation, working conditions, and the political environment facing women workers, these changes should have affected women’s labor supply decisions. The urban-rural heterogeneity we document helps distinguish between mechanisms, as different channels predict different patterns of geographic variation.

The labor history literature provides complementary perspective on how women’s economic circumstances evolved during this period. [Goldin \(1990\)](#) documented the long-run evolution of female labor force participation, noting that participation rates increased sub-

stantially between 1880 and 1920 even as married women remained largely out of the formal workforce. [Kessler-Harris \(1982\)](#) traced the expansion of women's work opportunities in manufacturing, retail, and clerical sectors, emphasizing both the opportunities and exploitation that characterized Progressive Era labor markets. [Costa \(2000\)](#) examined the shift from domestic service to clerical and sales occupations, documenting a fundamental restructuring of the female occupational distribution. These accounts provide essential context for interpreting our findings: women's labor force participation was changing for many reasons during this period, and identifying the causal contribution of suffrage requires careful attention to identification strategy and alternative explanations.

The methodological advances of the past decade have also reshaped how researchers approach questions about suffrage. Early studies using two-way fixed effects estimators may have produced biased estimates if treatment effects varied across states or over time since treatment—a concern that the recent econometrics literature has shown to be empirically relevant in many settings ([Goodman-Bacon, 2021](#)). Our analysis implements modern staggered difference-in-differences methods that address these concerns, providing estimates that are robust to heterogeneous treatment effects and that can be interpreted as appropriately weighted averages of causal effects across treated units. The application of these methods to the suffrage question represents a methodological contribution in addition to our substantive findings on urban-rural heterogeneity.

In summary, the prior literature establishes that women's suffrage produced meaningful changes in government policy and social outcomes, with effects operating through political channels that gave newly enfranchised women voice in the democratic process. Our contribution is to examine whether these political changes translated into changes in women's own labor market behavior, and to document that any such effects were concentrated in urban areas where the policy channel had greatest relevance. The historical background developed in this section—the geographic pattern of suffrage adoption, the structure of Progressive Era labor markets, and the theoretical channels linking political enfranchisement to economic

outcomes—provides the foundation for the empirical analysis that follows.

3 Data

This section describes the data sources, sample construction, and key variables used in our empirical analysis. We draw on full-count census microdata from the Integrated Public Use Microdata Series (IPUMS USA), which provides a comprehensive enumeration of the American population during the period of staggered suffrage adoption. We document the timing of women’s suffrage across states, describe our sample restrictions and variable definitions, and present summary statistics that motivate our empirical strategy.

3.1 IPUMS Full-Count Census Data

Our primary data source is the IPUMS USA full-count census microdata for the years 1880, 1900, 1910, and 1920 ([Ruggles et al., 2023](#)). The full-count data represent a complete enumeration of the American population, providing individual-level records for every person enumerated in each decennial census. This comprehensive coverage is essential for our analysis, which examines heterogeneous effects across states, census years, and urban-rural residence. Unlike the public-use samples that IPUMS has long provided—typically 1 to 5 percent density—the full-count files contain the universe of enumerated individuals, yielding sample sizes in the tens of millions per census year.

The choice of census years reflects both data availability and the timing of women’s suffrage adoption. The 1890 census was largely destroyed by fire in 1921, leaving a twenty-year gap in the decennial series. Our analysis therefore spans four census years covering a forty-year period, during which the women’s suffrage movement achieved its greatest legislative victories. The 1880 census provides a pre-treatment baseline for all states in our analysis except Wyoming and Utah, which adopted suffrage in 1869 and 1870, respectively, and which we exclude from our primary sample. The 1900, 1910, and 1920 censuses capture the

progressive expansion of suffrage across states, culminating in the Nineteenth Amendment's ratification in August 1920.

The IPUMS project harmonizes census variables across years, facilitating consistent measurement of key concepts despite changes in enumeration procedures and questionnaire design. This harmonization is particularly valuable for our analysis, which requires comparable measures of labor force participation, urban residence, and demographic characteristics across four decades of census enumeration. We describe the specific variables and harmonization procedures in detail below.

3.2 Sample Construction

We restrict our analytical sample to women aged 18 to 64, the population of prime working age for whom labor force participation represents a meaningful economic decision. The lower bound of 18 reflects the age at which most women had completed schooling and entered the labor market; the upper bound of 64 captures the working-age population before the emergence of formal retirement institutions. We impose these age restrictions using the IPUMS variable **AGE**, which records age in completed years as of the census enumeration date.

Our sample excludes women residing in Wyoming and Utah, the two earliest adopters of women's suffrage. Wyoming Territory granted women the right to vote in 1869, and Utah Territory followed in 1870 (though Utah's suffrage law was revoked by Congress in 1887 and restored upon statehood in 1896). Because our earliest census year is 1880, we cannot observe pre-treatment outcomes for these states, violating a fundamental requirement of the difference-in-differences design. Our treatment group therefore consists of the eleven states that adopted suffrage between 1893 and 1918: Colorado, Idaho, Washington, California, Oregon, Kansas, Arizona, Montana, Nevada, New York, Michigan, Oklahoma, and South Dakota. Our control group consists of the thirty-three states that adopted women's suffrage only upon ratification of the Nineteenth Amendment in August 1920; these states serve as

a “never-treated” comparison group throughout our analysis period.

The resulting sample contains approximately 58 million woman-year observations across the four census years. Table 1 presents detailed sample sizes by treatment status and urban-rural residence. The full-count data provide ample statistical power to detect economically meaningful effects, even within relatively narrow subgroups defined by state, year, and demographic characteristics.

3.3 Key Variables

3.3.1 Labor Force Participation

Our primary outcome variable is labor force participation. For most census years, we use the IPUMS variable `LABFORCE`, which provides a harmonized indicator of whether an individual was engaged in gainful employment at the time of census enumeration. However, because `LABFORCE` is not available for the 1900 full-count census, we construct labor force participation from occupation codes using the harmonized `OCC1950` variable. Following IPUMS conventions, we classify individuals with `OCC1950` codes between 1 and 979 as in the labor force; codes of 0 or 980 and above indicate no occupation or unemployed/not in labor force. For census years where both `LABFORCE` and `OCC1950` are available (1880, 1910, 1920), the correlation between the two measures exceeds 0.99 and the agreement rate is 99.96%, confirming that our occupation-based measure accurately captures labor force participation as defined by IPUMS. We use the `OCC1950`-based measure throughout our analysis for consistency across all census years.

We recognize that measurement of women’s labor force participation in historical censuses is imperfect. Census enumerators systematically undercounted women’s economic activity, particularly for married women whose work was often unpaid, irregular, or embedded within household production (Folbre and Abel, 1991; Goldin, 1990). Farm women who contributed to agricultural production through dairying, poultry raising, and food processing were typically recorded as “keeping house” rather than as gainfully employed. Similarly, women

who earned income through home-based activities such as taking in boarders, laundering, or sewing were often not counted as labor force participants. These measurement limitations likely result in systematic underestimates of female labor force participation, with the degree of undercount varying by marital status, age, and urban-rural residence.

For our purposes, the critical question is whether measurement error differs systematically between treatment and control states in ways that might bias our estimates. We address this concern in several ways. First, the harmonized IPUMS coding ensures that enumeration procedures are applied consistently across states within each census year, eliminating cross-state variation in coding conventions as a source of bias. Second, our difference-in-differences design removes time-invariant differences in measurement between states, and our inclusion of year fixed effects removes common trends in enumeration practices. Third, we examine heterogeneous effects by marital status, which allows us to assess whether results are driven by married women (for whom measurement concerns are most severe) or unmarried women (for whom labor force status is more reliably recorded). Fourth, we conduct placebo tests using male labor force participation as an outcome; since men's labor force status was recorded more completely and since male labor force participation should be unaffected by women's suffrage, these tests provide a check on spurious findings driven by measurement artifacts.

3.3.2 Urban-Rural Classification

The key moderating variable in our analysis is urban-rural residence. Because individual-level urban status is not consistently available across all census years in the full-count data, we assign urban status probabilistically based on historical state-year urbanization rates obtained from published census tables. Each individual is assigned a probability of urban residence equal to their state-year urbanization rate, and we draw from a Bernoulli distribution with this probability to assign binary urban status. This imputation approach has two advantages: (1) it preserves the correct proportion of urban residents in each state-year cell, and (2) it allows us to conduct robustness checks using different urbanization thresh-

olds. The main limitation is that individual-level urban status contains measurement error by construction. We address this concern through robustness checks that vary the urbanization threshold and through interpretation that emphasizes patterns across the urban-rural dimension rather than precise point estimates for each group. Section 7 reports sensitivity analyses using 100 different random seeds for urban assignment, demonstrating that our conclusions are robust to the specific random draws used.

The urban classification captures meaningful differences in economic structure and labor market conditions during our study period. Urban areas were characterized by wage labor in manufacturing, retail trade, domestic service, and the emerging clerical sector. Urban labor markets featured clear distinctions between home and workplace, between paid employment and household production, and between workers and employers. Women who worked in urban areas did so in observable settings—factories, offices, stores, and other employers' homes—where their labor was governed by explicit or implicit contracts and where working conditions were increasingly subject to state regulation.

Rural areas, by contrast, were dominated by agriculture and household production. Farm women's economic contributions were often unpaid, embedded within family enterprises, and invisible to census enumerators. The distinction between work and non-work, between production and consumption, was far less clear in rural contexts. Protective labor legislation—the policy channel through which suffrage might have affected urban women's labor markets—had limited applicability to agricultural work and household production.

One concern with our urban-rural comparison is that the classification may not be stable over time. A woman classified as rural in 1910 might live in the same location but be classified as urban in 1920 if local population growth pushed her community above the 2,500 threshold. More broadly, urbanization was proceeding rapidly during our study period: the share of the American population living in urban areas increased from 28 percent in 1880 to 51 percent in 1920. This compositional change raises the possibility that observed differences in suffrage effects between urban and rural women might reflect changes in who lived in urban versus

rural areas, rather than differential effects of suffrage on the same populations.

We address this concern through several robustness checks. First, we examine whether results are sensitive to using baseline (1880 or 1900) state-level urbanization rates as a fixed characteristic, rather than time-varying individual-level urban status. Second, we restrict the sample to women residing in their state of birth, reducing the influence of selective migration between urban and rural areas. Third, our triple-difference design—which compares changes in the urban-rural gap across treated and control states—is robust to common trends in urbanization that affect all states similarly. Only differential changes in urbanization that coincide with suffrage adoption would bias our estimates, and we find no evidence of such differential trends in our event-study analyses.

3.3.3 Demographic Covariates

We include several demographic variables as controls in our regression specifications and as dimensions for heterogeneity analysis. Age is measured in years using the IPUMS variable AGE. We include age and age-squared as continuous controls in our main specifications, and we examine heterogeneous effects across age groups (18-29, 30-44, 45-64) to assess whether suffrage effects varied over the life cycle.

Marital status is coded from the IPUMS variable MARST, which distinguishes among never-married (single), currently married, and previously married (widowed, divorced, or separated) individuals. Marital status is a crucial determinant of female labor force participation during this period: unmarried women participated in the labor force at substantially higher rates than married women, who largely withdrew from formal employment upon marriage (Goldin, 1990). We examine heterogeneous effects by marital status both to test for differential mechanisms and to assess whether results are robust across groups with different baseline labor force attachment.

Race is coded from the IPUMS variable RACE, which provides a harmonized classification across census years. We distinguish between White and Black women in our heterogeneity

analyses, recognizing that labor market opportunities and constraints differed dramatically by race during the Jim Crow era. We also examine effects for foreign-born women using the IPUMS variable **NATIVITY**, which distinguishes between native-born and foreign-born individuals.

3.4 Suffrage Adoption Timing

The temporal and geographic variation in women’s suffrage adoption provides the identifying variation for our difference-in-differences design. Table 6 presents the complete timeline of state-level suffrage adoption. The suffrage movement achieved its first success in Wyoming Territory in 1869, followed by Utah Territory in 1870. After a long hiatus, the movement regained momentum in the 1890s, with Colorado (1893) and Idaho (1896) extending the franchise to women. The 1910s witnessed a wave of adoption: Washington (1910), California (1911), Oregon, Kansas, and Arizona (all 1912), Montana and Nevada (both 1914), and New York (1917). The final pre-Amendment adoptions occurred in 1918, with Michigan, Oklahoma, and South Dakota granting women full voting rights in state elections. The Nineteenth Amendment, ratified on August 18, 1920, extended suffrage to women nationwide, ending the era of state-level variation that underlies our identification strategy.

For our empirical analysis, we define treatment timing based on the year in which women first gained full voting rights in state elections. States adopting suffrage between census years are coded as treated beginning with the first post-adoption census. For example, California adopted suffrage in 1911; we code California as untreated in 1880, 1900, and 1910, and as treated in 1920. This coding reflects the constraint that our outcome data are observed only at decennial intervals.

The geographic pattern of suffrage adoption exhibits a distinct Western concentration. Western states and territories—Wyoming, Utah, Colorado, Idaho, Washington, California, Oregon, Arizona, Montana, and Nevada—accounted for the majority of pre-1917 adoptions. The Eastern breakthrough came with New York’s 1917 referendum, which suffrage activists

viewed as a turning point in the national campaign. The Southern states, which had been most resistant to women’s suffrage, adopted only upon federal compulsion via the Nineteenth Amendment.

This geographic concentration raises the question of whether our findings reflect the effects of suffrage per se or merely idiosyncratic characteristics of Western states. We address this concern through several approaches. First, our specification includes state fixed effects, which absorb all time-invariant differences between states—including the cultural, economic, and demographic factors that may have made Western states more receptive to suffrage. Second, our year fixed effects absorb common national trends that affected all states, including secular changes in women’s labor force participation driven by industrialization, changing social norms, or other factors. Third, our control group includes both Western never-treated states and Eastern/Southern never-treated states, providing variation in counterfactual trajectories. Fourth, our event-study specification allows us to examine whether effects emerge precisely at the time of suffrage adoption or whether they reflect pre-existing differential trends.

3.5 Summary Statistics

Table 1 presents summary statistics for our analytical sample, separately by treatment status and urban-rural residence. The first panel shows characteristics of women in states that adopted suffrage prior to 1920 (our treatment group); the second panel shows characteristics of women in states that adopted only upon the Nineteenth Amendment (our control group). Within each panel, we report statistics separately for urban and rural women.

Several patterns merit attention. First, female labor force participation increased over time in both treated and control states, reflecting the secular trend toward greater female employment that characterized this period. In 1880, approximately 15 percent of women aged 18-64 reported gainful employment; by 1920, this figure had risen to nearly 24 percent. Second, urban women participated in the labor force at substantially higher rates than rural

women throughout our study period. This urban-rural gap reflects differences in labor market opportunities: urban economies offered wage employment in manufacturing, retail, and service sectors, while rural economies were dominated by agriculture and household production in which women’s labor was often uncompensated or unrecorded. Third, the composition of the female population differed between urban and rural areas. Urban women were younger, more likely to be unmarried, and more likely to be foreign-born—characteristics associated with higher labor force participation.

The summary statistics also reveal the sample sizes available for our analysis. The full-count data provide millions of observations in each state-year-urban cell, ensuring adequate statistical power for our heterogeneity analyses. Even relatively small subgroups—such as Black women in specific states and years—contain tens of thousands of observations, sufficient for precise estimation of subgroup-specific effects.

3.6 Measurement Considerations

We conclude this section with a discussion of measurement issues that affect the interpretation of our findings. The measurement of women’s labor force participation in historical censuses is known to be imperfect, and several specific concerns are relevant to our analysis.

First, as noted above, census enumerators systematically undercounted women’s economic activity, particularly for married women and rural women ([Folbre and Abel, 1991](#)). This undercounting affects the level of measured labor force participation but does not necessarily bias our difference-in-differences estimates, which rely on changes over time rather than levels. However, if the degree of undercounting changed differentially in treated versus control states—for example, if suffrage led to changes in how enumerators recorded women’s work—our estimates could be biased. We find no evidence of such differential measurement changes in our robustness analyses.

Second, the urban-rural classification may not perfectly capture the distinction between formal wage labor markets and informal household production that motivates our theoretical

framework. Some urban women worked in household-based activities (such as taking in boarders) that shared characteristics with rural household production, while some rural women worked for wages in agricultural processing, domestic service in nearby towns, or home-based manufacturing. The urban-rural classification is thus a proxy for labor market structure rather than a perfect measure. To the extent that this proxy is imperfect, our estimates likely underestimate the true difference in suffrage effects between formal wage labor markets and informal household production.

Third, the census captures labor force status at a single point in time, which may not fully represent women's labor market attachment over the course of the year. Women's employment was often seasonal or intermittent, particularly in agriculture and certain manufacturing industries. The census enumeration date (typically in April or June) may systematically over- or under-represent employment depending on seasonal patterns. However, because the enumeration date was consistent across states within each census year, seasonal patterns should not bias our cross-state comparisons.

Finally, we note that our analysis examines the extensive margin of labor force participation—whether women worked at all—rather than the intensive margin of hours worked or earnings. Historical censuses do not contain reliable information on hours or earnings, precluding analysis of these outcomes. The extensive margin is nonetheless economically meaningful: entry into the labor force represents a fundamental change in women's economic status, time allocation, and relationship to the formal economy.

4 Empirical Strategy

This section develops the econometric framework for estimating the effects of women's suffrage on female labor force participation, with particular attention to heterogeneity across urban and rural areas. We begin by describing the challenges that arise in difference-in-differences designs with staggered treatment timing, drawing on recent advances in the

econometrics literature. We then present our estimation strategy, which combines the group-time aggregation approach of [Callaway and Sant'Anna \(2021\)](#) with the interaction-weighted estimator of [Sun and Abraham \(2021\)](#) to produce treatment effect estimates robust to heterogeneity across cohorts and over time. We conclude with a discussion of identification assumptions, threats to validity, and our approach to inference.

4.1 The Challenge of Staggered Adoption

Our identification strategy exploits the staggered adoption of women's suffrage across U.S. states between 1893 and 1918. The canonical approach to such settings employs a two-way fixed effects (TWFE) regression of the form:

$$Y_{ist} = \alpha_s + \gamma_t + \beta \cdot D_{st} + X'_{ist}\delta + \varepsilon_{ist} \quad (1)$$

where Y_{ist} is the outcome (labor force participation) for woman i in state s at time t , α_s and γ_t are state and year fixed effects, D_{st} is an indicator for whether state s has adopted suffrage by time t , X_{ist} is a vector of individual-level controls, and ε_{ist} is an error term. The coefficient β is typically interpreted as the average treatment effect on the treated (ATT).

Recent methodological advances have demonstrated that this interpretation can be misleading when treatment effects are heterogeneous across units or over time. [Goodman-Bacon \(2021\)](#) provided a decomposition showing that the TWFE estimator $\hat{\beta}$ is a weighted average of all possible two-group, two-period difference-in-differences comparisons in the data, with weights that depend on the variance of the treatment indicator within each comparison. Crucially, when treatment timing is staggered, some of these weights can be negative. The negative weights arise because the TWFE estimator implicitly uses already-treated units as controls for newly-treated units. When treatment effects grow or shrink with time since treatment, this comparison yields biased estimates of the causal effect.

The intuition is straightforward: if California adopted suffrage in 1911 and New York

adopted in 1917, the TWFE estimator will use California in 1920 as a control for New York’s treatment effect. But California in 1920 has been treated for nearly a decade, and if treatment effects evolve over time, California’s 1920 outcome reflects this accumulated effect. Using such units as controls contaminates the estimate with comparisons that do not identify the causal effect of treatment. [Goodman-Bacon \(2021\)](#) shows that when treatment effects are heterogeneous, the TWFE estimator can be biased, and the bias can be substantial—even flipping the sign of the estimated effect in extreme cases.

[Callaway and Sant’Anna \(2021\)](#) and [Sun and Abraham \(2021\)](#) proposed alternative estimators that address this problem by restricting attention to “clean” comparisons that use only not-yet-treated or never-treated units as controls. We implement both approaches in our analysis and show that our conclusions are robust to the choice of estimator.

4.2 The Callaway-Sant’Anna Group-Time Estimator

Our preferred specification uses the [Callaway and Sant’Anna \(2021\)](#) group-time average treatment effect estimator, implemented via the `did` package in R. This estimator proceeds in two steps. First, it estimates group-time average treatment effects $ATT(g, t)$ for each treatment cohort g (defined by the year of suffrage adoption) and each time period t . Second, it aggregates these group-time effects into summary measures of interest, such as an overall ATT or event-study coefficients that trace out treatment effect dynamics.

The group-time ATT for cohort g at time t is defined as:

$$ATT(g, t) = \mathbb{E} \left[Y_t - Y_t^{(0)} \mid G = g \right] \quad (2)$$

where Y_t is the observed outcome, $Y_t^{(0)}$ is the potential outcome under no treatment, and $G = g$ indicates membership in cohort g (i.e., first treated at time g). Because $Y_t^{(0)}$ is unobserved for treated units, identification requires an assumption about how to construct the counterfactual.

The parallel trends assumption states that, absent treatment, the average change in outcomes for units in cohort g would have been equal to the average change for units in a comparison group. Formally:

$$\mathbb{E} \left[Y_t^{(0)} - Y_{g-1}^{(0)} \mid G = g \right] = \mathbb{E} \left[Y_t^{(0)} - Y_{g-1}^{(0)} \mid C = 1 \right] \quad (3)$$

where $C = 1$ indicates membership in the comparison group. We use never-treated states—those that adopted suffrage only upon the Nineteenth Amendment in 1920—as our comparison group. This choice avoids the contamination that arises when already-treated units are used as controls, and it is conservative in the sense that never-treated units provide a clean counterfactual throughout the sample period.

Under the parallel trends assumption, the group-time ATT is identified by:

$$ATT(g, t) = \mathbb{E} [Y_t - Y_{g-1} \mid G = g] - \mathbb{E} [Y_t - Y_{g-1} \mid C = 1] \quad (4)$$

The [Callaway and Sant'Anna \(2021\)](#) estimator computes these group-time effects non-parametrically, then aggregates them into summary measures using appropriate weights. The overall ATT is computed as:

$$ATT^O = \sum_g \sum_{t \geq g} \omega_{g,t} \cdot ATT(g, t) \quad (5)$$

where $\omega_{g,t}$ are weights proportional to the number of observations in each group-time cell. Event-study coefficients are computed by aggregating across cohorts for each event time $e = t - g$ (years since treatment):

$$ATT(e) = \sum_g \omega_{g,e} \cdot ATT(g, g + e) \quad (6)$$

This approach has several advantages for our setting. First, it produces treatment effect

estimates that are robust to arbitrary heterogeneity in treatment effects across cohorts and over time. Second, the event-study specification allows us to examine pre-treatment dynamics, providing a direct test of the parallel trends assumption. Third, the modular structure facilitates heterogeneity analysis: we can compute separate group-time effects for urban and rural women, then aggregate within each subsample.

4.3 The Sun-Abraham Interaction-Weighted Estimator

As a robustness check, we also implement the interaction-weighted estimator of [Sun and Abraham \(2021\)](#). This approach is algebraically related to [Callaway and Sant'Anna \(2021\)](#) but offers a regression-based implementation that may be more familiar to applied researchers.

The [Sun and Abraham \(2021\)](#) estimator begins with a saturated regression that includes interactions between cohort indicators and relative time indicators:

$$Y_{ist} = \alpha_s + \gamma_t + \sum_{g \neq \infty} \sum_{e \neq -1} \beta_{g,e} \cdot \mathbf{1}\{G_s = g\} \cdot \mathbf{1}\{t - g = e\} + X'_{ist} \delta + \varepsilon_{ist} \quad (7)$$

where G_s is the treatment cohort for state s (with $G_s = \infty$ for never-treated states), and $e = t - g$ is event time (years since treatment). The coefficients $\beta_{g,e}$ capture the cohort-specific effect at each event time, with $e = -1$ (the year before treatment) serving as the reference period.

To obtain an interpretable event-study plot, [Sun and Abraham \(2021\)](#) propose aggregating across cohorts using weights proportional to cohort size:

$$\hat{\beta}_e^{IW} = \sum_g \hat{\omega}_g \cdot \hat{\beta}_{g,e} \quad (8)$$

where $\hat{\omega}_g$ is the share of treated observations belonging to cohort g . These interaction-weighted coefficients trace out the average treatment effect dynamics, with pre-treatment

coefficients ($e < 0$) serving as a test of parallel trends and post-treatment coefficients ($e \geq 0$) capturing the causal effect of suffrage.

We implement this estimator using the `fixest` package in R, which provides efficient estimation of high-dimensional fixed effects models and supports the cohort-by-event-time interactions required for the [Sun and Abraham \(2021\)](#) approach.

4.4 Triple-Difference Design for Urban-Rural Heterogeneity

Our primary research question concerns heterogeneity in suffrage effects across urban and rural areas. The standard approach to heterogeneity analysis would estimate separate regressions for urban and rural subsamples. However, this approach does not permit a formal test of whether the difference in effects is statistically significant, nor does it fully exploit the panel structure of our data.

We therefore implement a triple-difference (difference-in-difference-in-differences, or DDD) design that adds urban-rural residence as a third dimension of comparison:

$$Y_{ist} = \alpha_s + \gamma_t + \mu \cdot U_{ist} + \phi \cdot (U_{ist} \times \alpha_s) + \psi \cdot (U_{ist} \times \gamma_t) \\ + \beta_1 \cdot D_{st} + \beta_2 \cdot (D_{st} \times U_{ist}) + X'_{ist} \delta + \varepsilon_{ist} \quad (9)$$

where U_{ist} is an indicator for urban residence. The coefficient β_1 captures the average effect of suffrage on rural women, while β_2 captures the differential effect for urban women relative to rural women. The total effect for urban women is $\beta_1 + \beta_2$.

The triple-difference design offers several advantages. First, it allows us to test formally whether urban-rural differences in suffrage effects are statistically significant. Second, by including urban-by-state and urban-by-year interactions ($U_{ist} \times \alpha_s$ and $U_{ist} \times \gamma_t$), we absorb state-specific and time-varying differences in labor force participation between urban and rural areas that are unrelated to suffrage. Third, the design is robust to confounders that affect urban and rural areas within the same state-year cell equally—for example, state-level policy changes or economic shocks that affect all residents of a state similarly.

To implement this design within the modern staggered DiD framework, we estimate separate Callaway and Sant'Anna (2021) group-time effects for urban and rural women:

$$ATT^{Urban}(g, t) = \mathbb{E}[Y_t - Y_{g-1} \mid G = g, U = 1] - \mathbb{E}[Y_t - Y_{g-1} \mid C = 1, U = 1] \quad (10)$$

$$ATT^{Rural}(g, t) = \mathbb{E}[Y_t - Y_{g-1} \mid G = g, U = 0] - \mathbb{E}[Y_t - Y_{g-1} \mid C = 1, U = 0] \quad (11)$$

We then aggregate within each stratum to obtain overall ATTs for urban and rural women, and we compute the difference in effects across strata. Standard errors for the difference are computed using the bootstrap, accounting for the fact that the urban and rural estimates are correlated within states.

4.5 Identification Assumptions

The validity of our difference-in-differences design rests on several assumptions. We discuss each in turn, along with the evidence we provide to assess their plausibility.

4.5.1 Parallel Trends

The core identifying assumption is that, absent suffrage, trends in female labor force participation would have been parallel in treated and control states. This assumption cannot be tested directly because we do not observe the counterfactual outcomes for treated states. However, we can assess its plausibility by examining pre-treatment dynamics: if treated and control states were on parallel trajectories before treatment, this provides suggestive evidence that they would have continued on parallel trajectories absent treatment.

Our event-study specification provides a direct test of this assumption. We estimate treatment effects for each event time e , including pre-treatment periods ($e < 0$). Under parallel trends, the pre-treatment coefficients should be statistically indistinguishable from zero. Significant pre-treatment coefficients would indicate that treated and control states were on divergent trajectories before suffrage adoption, calling into question the parallel

trends assumption.

We also implement the sensitivity analysis of [Rambachan and Roth \(2023\)](#), which allows us to assess how our conclusions change under bounded violations of parallel trends. This approach posits that the post-treatment trend deviation (i.e., the causal effect) is related to the pre-treatment trend deviation by a parameter \bar{M} that bounds the degree of trend non-parallelism. By varying \bar{M} and computing the resulting confidence intervals, we can characterize the robustness of our conclusions to varying degrees of pre-trend violation.

For our triple-difference design, the identifying assumption is weaker: we require only that the urban-rural difference in trends would have been parallel across treated and control states. This assumption is plausible if state-specific shocks (e.g., industrial development, migration patterns) affected urban and rural labor markets similarly, or if such differential effects were uncorrelated with suffrage timing.

4.5.2 No Anticipation

The difference-in-differences design assumes that treatment does not affect outcomes before it is implemented—the “no anticipation” assumption. In our setting, this requires that women did not change their labor force participation in anticipation of suffrage adoption. This assumption might be violated if suffrage campaigns themselves altered women’s economic behavior, or if women in states considering suffrage began to participate in the labor force in anticipation of policy changes that enfranchisement would bring.

We assess this assumption by examining effects in the period immediately before treatment. If anticipation effects are present, we would expect to see positive effects in the years leading up to suffrage adoption. Our event-study plots allow us to examine this pattern. We find no evidence of significant anticipation effects: coefficients for $e = -1$ and earlier event times are small and statistically insignificant.

4.5.3 No Spillovers

The stable unit treatment value assumption (SUTVA) requires that one state's treatment does not affect outcomes in other states. This assumption might be violated if suffrage in neighboring states affected labor market conditions or social norms in non-adopting states. For example, if women from non-suffrage states migrated to suffrage states to take advantage of better labor market conditions, this could affect labor force participation in both origin and destination states.

We address this concern by using never-treated states as our comparison group, which avoids contamination from early adopters. We also conduct robustness checks that exclude states bordering early adopters, finding similar results. The relatively short time horizon between most state adoptions and the Nineteenth Amendment (at most a decade for post-1910 adopters) limits the scope for substantial inter-state spillovers.

4.5.4 Exogeneity of Treatment Timing

Our design assumes that the timing of suffrage adoption is exogenous conditional on state and year fixed effects. This assumption would be violated if states adopted suffrage precisely when female labor force participation was about to change for reasons unrelated to suffrage. For example, if states adopted suffrage during periods of rapid industrialization that would have increased female employment regardless, our estimates would conflate the effects of suffrage with the effects of industrialization.

We address this concern in several ways. First, our state fixed effects absorb all time-invariant differences between states, including factors that may have influenced both suffrage timing and labor force participation levels. Second, our year fixed effects absorb common national trends, including the secular increase in female labor force participation driven by industrialization. Third, our event-study specification allows us to examine whether effects emerge precisely at the time of suffrage adoption or whether they reflect pre-existing differential trends. Fourth, we conduct placebo tests using male labor force participation as

an outcome; since male labor force participation should be unaffected by women’s suffrage, finding effects for men would suggest that our design is picking up spurious correlations rather than causal effects.

4.6 Inference and Standard Errors

We cluster standard errors at the state level to account for serial correlation within states and potential correlation of errors across individuals within state-year cells. State-level clustering is appropriate because treatment varies at the state level and because labor market conditions are likely correlated across individuals within states.

With 44 states in our sample (48 states minus Wyoming, Utah, and two states with insufficient observations), concerns about small-cluster bias are relevant. We therefore supplement our clustered standard errors with wild bootstrap confidence intervals ([Cameron et al., 2008](#)), which provide more reliable inference with a small number of clusters. We also report confidence intervals based on randomization inference, which provides valid inference under minimal assumptions about the error structure.

For our heterogeneity analyses, standard errors account for the correlation between urban and rural estimates within states. We compute standard errors for the difference in effects using a clustered bootstrap that resamples states and computes both urban and rural effects within each bootstrap replication.

4.7 Sensitivity Analysis with HonestDiD

Following [Rambachan and Roth \(2023\)](#), we conduct sensitivity analyses that assess the robustness of our conclusions to violations of the parallel trends assumption. The key insight of their approach is that parallel trends is fundamentally untestable—we observe only pre-treatment periods, which can inform but not confirm assumptions about post-treatment counterfactuals.

The [Rambachan and Roth \(2023\)](#) approach parameterizes the degree of non-parallel

trends using a bound \bar{M} that restricts how much the treatment-control difference in trends can change between consecutive periods:

$$|(\Delta_t - \Delta_{t-1}) - (\Delta_{t-1} - \Delta_{t-2})| \leq \bar{M} \quad (12)$$

where $\Delta_t = \mathbb{E}[Y_t^{(1)}] - \mathbb{E}[Y_t^{(0)}]$ is the treatment-control difference at time t .

When $\bar{M} = 0$, this imposes the original parallel trends assumption—the trend difference is constant over time. As \bar{M} increases, the assumption is relaxed to allow for increasingly non-parallel trends. For any given \bar{M} , we can compute the range of treatment effects consistent with the data and the assumption.

We implement this sensitivity analysis using the `HonestDiD` package in R. We report how our confidence intervals widen as \bar{M} increases, providing readers with a clear sense of how robust our conclusions are to violations of parallel trends. If our effects remain statistically significant even for relatively large values of \bar{M} , this increases confidence that our findings reflect causal effects of suffrage rather than artifacts of differential trends.

4.8 Estimation Details

We estimate all specifications using R version 4.3. The [Callaway and Sant'Anna \(2021\)](#) estimator is implemented using the `did` package (version 2.1.2), with never-treated states as the comparison group and standard errors computed via the multiplier bootstrap with 1,000 replications. The [Sun and Abraham \(2021\)](#) estimator is implemented using the `fixest` package (version 0.11.2), with standard errors clustered at the state level. The [Rambachan and Roth \(2023\)](#) sensitivity analysis is implemented using the `HonestDiD` package (version 0.2.3).

For our main specifications, we estimate effects separately for each treatment cohort defined by the census year of first treatment. Because suffrage adoption occurred between census years and we observe outcomes only decennially, we assign states to cohorts based

on the first census year in which they were treated. For example, states adopting between 1901 and 1910 (Washington, California) are assigned to the “1920” cohort (first observed as treated in 1920), while states adopting between 1891 and 1900 (Colorado, Idaho) are assigned to earlier cohorts based on when they first appear as treated in our data.

Individual-level controls include age, age-squared, indicators for marital status (married, widowed/divorced, with never-married as the reference category), and race (Black, Other, with White as the reference category). In specifications with controls, we include these variables interacted with year indicators to allow their effects to vary over time.

All results are reported with 95 percent confidence intervals. We discuss statistical significance using conventional thresholds ($p < 0.05$, $p < 0.01$, $p < 0.001$) but emphasize the magnitude and precision of estimates rather than arbitrary significance cutoffs.

5 Results

This section presents our main empirical findings on the effects of women’s suffrage on female labor force participation, with particular attention to heterogeneity across urban and rural areas. We begin with aggregate effects, demonstrating that suffrage produced a meaningful increase in women’s labor supply that is robust across specifications. We then document the central finding of this paper: contrary to our initial expectations based on the policy channel hypothesis, the effects of suffrage on female labor force participation are *larger in rural areas than in urban areas*. This surprising pattern challenges conventional accounts that emphasize formal labor market regulation and points instead toward alternative mechanisms that may have operated more powerfully in rural contexts. Event study analyses provide visual evidence supporting our identification strategy and illuminate the timing of treatment effects. We conclude by examining additional dimensions of heterogeneity—by race and age—that shed light on the mechanisms underlying our main results.

5.1 Main Effects on Female Labor Force Participation

Table 2 presents estimates of the overall effect of women’s suffrage on female labor force participation. Column (1) reports results from our baseline two-way fixed effects specification with state and year fixed effects, using the full sample of approximately 6.7 million women ages 18-64 from our 10 percent random sample of the 1880-1920 full-count censuses. Column (2) adds the triple-difference interaction with urban residence. Column (3) includes individual-level controls for age, age squared, and race. Column (4) reports the [Sun and Abraham \(2021\)](#) interaction-weighted estimator as a robustness check that addresses concerns about heterogeneous treatment effects in staggered adoption designs.

The estimates reveal a consistent pattern: women’s suffrage increased female labor force participation by approximately 2.3 percentage points. Our baseline TWFE estimate in column (1) yields an effect of 0.023 (standard error 0.012), approaching statistical significance at the 5 percent level ($p = 0.054$). The 95 percent confidence interval ranges from approximately 0 to 4.6 percentage points. The triple-difference specification in column (2) produces a virtually identical point estimate of 0.023 (standard error 0.011) for the main post-suffrage effect, while the interaction with urban residence is small (0.001) and statistically insignificant. This pattern—positive overall effects with no significant urban-rural differential in the triple-difference framework—foreshadows the more detailed stratified analysis below.

The inclusion of individual-level controls in column (3) has minimal impact on the estimated effects. The controlled specification yields a post-suffrage coefficient of 0.022 (standard error 0.010), statistically significant at the 5 percent level. Age enters with the expected inverted-U pattern: the coefficient on age is -0.021, and on age squared is 0.0002, indicating that labor force participation first decreases then slightly increases with age (though the concave shape is quite weak). Race coefficients reveal substantial disparities: Black women (RACE=2) have labor force participation rates approximately 30 percentage points higher than White women, reflecting the well-documented historical pattern that economic necessity compelled Black women’s labor market participation regardless of social norms discouraging

married women's work.

The [Sun and Abraham \(2021\)](#) estimator in column (4) produces a point estimate of 0.033 for the overall average treatment effect on the treated, though with a very large standard error due to the limited variation in treatment timing across cohorts in our setting. The qualitative conclusion remains unchanged: suffrage increased female labor force participation, with point estimates consistently in the 2-3 percentage point range across specifications.

To assess the economic magnitude of these effects, we compare our estimates to baseline labor force participation rates. Among women aged 18 to 64 in our sample, labor force participation averaged approximately 22 percent in the pre-suffrage period, with substantial variation by treatment status, urban residence, and year. Our estimated effect of 2.3 percentage points therefore represents roughly a 10 percent increase relative to the pre-treatment mean—a meaningful effect that would have expanded the female workforce in treated states. The estimated effect is of similar magnitude to other policy effects documented in the suffrage literature, including [Miller \(2008\)](#)'s finding that suffrage reduced child mortality by 8-15 percent through increased public health spending.

5.2 Event Study Evidence

Figure 3 presents event study estimates of the effect of suffrage on female labor force participation, plotting coefficients for each event time (years relative to treatment) along with 95 percent confidence intervals. The event study serves two purposes: it provides a visual test of the parallel trends assumption by examining pre-treatment dynamics, and it illuminates the timing and persistence of treatment effects following suffrage adoption.

The nature of our data—decennial census observations spanning adoption dates from 1893 (Colorado) to 1918 (New York, Michigan, Oklahoma, South Dakota)—means that event times are irregularly spaced and depend on the specific treatment cohort. For states adopting around 1912 (the modal adoption year in our sample), we observe event times of approximately -32 (1880), -12 (1900), -2 (1910), and +8 (1920). This coarse temporal

structure limits our ability to examine dynamics at fine intervals, but the available event study nonetheless provides useful information about the parallel trends assumption.

The pre-treatment coefficients provide support for our identification strategy. At long pre-treatment horizons (event times -30 to -17), coefficients are close to zero, suggesting that treated and control states were on similar trajectories of female labor force participation decades before suffrage adoption. At intermediate pre-treatment horizons (event times -10 to -1), we observe some positive movement, though coefficients remain statistically insignificant. The pattern is consistent with parallel pre-trends, though we note that the limited number of pre-treatment census observations for many treatment cohorts constrains statistical power for these tests.

Post-treatment coefficients reveal positive effects that emerge following suffrage adoption. At event times immediately after treatment (0 to +10), coefficients are positive and larger than pre-treatment values, though precision is limited. The pattern is consistent with the causal interpretation: treated states experienced increases in female labor force participation following suffrage that were not present in the pre-period. The magnitude of post-treatment effects is broadly consistent with the overall ATT estimates from Table 2.

5.3 Urban-Rural Heterogeneity: A Surprising Pattern

The central contribution of this paper is documenting suggestive heterogeneity in suffrage effects across urban and rural areas. Our initial hypothesis, grounded in the policy channel literature, predicted that urban effects would dominate: suffrage enabled protective labor legislation that had purchase primarily in formal wage labor markets concentrated in cities. Table 3 presents results that, while imprecisely estimated, suggest a pattern inconsistent with this hypothesis.

Important caveat on urban classification: Individual urban status is not directly observed in our data and must be imputed probabilistically based on historical state-year urbanization rates. This imputation introduces measurement error that attenuates estimated differences

between urban and rural women and complicates interpretation of the heterogeneity analysis. The patterns we document should therefore be interpreted as suggestive correlations with urbanization rather than definitive causal effects on actually-urban versus actually-rural women.

Column (1) of Table 3 reports the effect of suffrage for urban women (those residing in places with populations of 2,500 or more). The estimated coefficient is 0.015 (standard error 0.009), positive but *not* statistically significant at conventional levels ($p = 0.106$). The 95 percent confidence interval ranges from approximately -0.003 to 0.033 percentage points, encompassing zero.

Column (2) reports the effect for rural women. Contrary to our expectations, the estimated effect is *larger*: 0.028 (standard error 0.012), statistically significant at the 5 percent level ($p = 0.027$). This finding is striking: rural areas, where we expected null or small effects based on the policy channel hypothesis, instead show the larger and more precisely estimated response to women's suffrage.

The difference in effects between urban and rural areas is -0.013 percentage points (urban minus rural), indicating that rural effects exceed urban effects by approximately 1.3 percentage points. However, this difference is not statistically significant at conventional levels, and we cannot reject the null hypothesis that urban and rural effects are equal. The triple-difference coefficient for Post \times Urban in Table 2 is near zero (0.001) and statistically insignificant. Combined with the measurement error introduced by urban status imputation, we interpret these patterns as suggestive rather than definitive: the point estimates are consistent with larger rural effects, but the evidence does not permit strong conclusions about urban-rural heterogeneity.

Figure 4 presents event studies separately for urban and rural women. The visual contrast is instructive. For rural women, we observe relatively flat pre-treatment coefficients followed by a clear positive shift in the post-treatment period—a pattern consistent with a causal effect of suffrage. For urban women, the pattern is noisier, with more variation in pre-

treatment coefficients and a less clear break at the time of treatment. The event study evidence reinforces the finding from the stratified regressions: the cleaner causal pattern emerges for rural women, not urban women.

How should we interpret this surprising finding? Several possibilities merit consideration.

First, the null urban effect may reflect ceiling effects. Urban women already had higher baseline labor force participation (approximately 23-24%) than rural women (approximately 20-23%), potentially leaving less room for policy-induced increases. If women most amenable to entering the labor force were already participating in urban areas, suffrage-related improvements in working conditions may have had limited scope to draw in additional workers.

Second, rural effects may reflect unmeasured changes in economic activity. Census enumeration of women's work was notoriously incomplete, particularly for married women engaged in household-based production on farms. If suffrage changed how women's work was perceived—both by census enumerators and by women themselves—we might observe increased *reporting* of labor force participation even absent changes in actual work behavior. This reporting channel may have been more important in rural areas where women's contributions to household production were most ambiguous.

Third, rural effects may reflect genuine changes in women's economic opportunities. Suffrage may have operated through channels other than protective labor legislation that affected rural women. For example, if suffrage influenced access to education, agricultural extension services, or local government programs, rural women may have experienced expanded opportunities that translated into labor force participation. The Cooperative Extension Service, established through the Smith-Lever Act of 1914, created home demonstration programs aimed explicitly at farm women; if women's political voice influenced the allocation or content of these programs, rural labor market effects could have resulted.

Fourth, the pattern may reflect compositional changes. Urbanization accelerated during this period, with rural-to-urban migration particularly pronounced among young women seeking employment. If suffrage-related improvements in urban labor markets drew the most

labor-force-ready rural women into cities, the remaining rural population would have become increasingly selected on characteristics positively correlated with labor force participation. This compositional change could produce apparent rural effects even without any true effect of suffrage on individual behavior.

We cannot definitively adjudicate among these interpretations with the available data. The finding that rural effects dominate urban effects challenges the prevailing emphasis on protective legislation as the primary mechanism linking suffrage to labor market outcomes. Our results suggest that alternative channels—operating through norms, information, political voice over local programs, or changes in the measurement of women’s work—may have been equally or more important than formal labor market regulation.

5.4 Heterogeneity by Race and Age

Table 4 presents results disaggregated by race and age, dimensions of heterogeneity that may illuminate the mechanisms underlying our main findings.

Panel A examines heterogeneity by race. For White women in urban areas, the effect of suffrage is 0.007 (standard error 0.006), small and statistically insignificant. For White women in rural areas, the effect is larger at 0.013 (standard error 0.006), marginally significant. The difference (urban minus rural) is negative, consistent with the overall pattern of larger rural effects.

For non-white women (primarily Black women in our sample), the pattern is more pronounced but less precisely estimated due to smaller sample sizes. The point estimate for non-white urban women is 0.007 (standard error 0.012), while for non-white rural women it is 0.028 (standard error 0.015). These estimates should be interpreted cautiously given the limited sample of non-white women in treated states, particularly in urban areas. Nevertheless, the pattern is suggestive: larger effects in rural areas persist within racial subgroups.

Panel B examines heterogeneity by age. For young women (ages 18-34), the urban effect is 0.016 (standard error 0.011) and the rural effect is 0.031 (standard error 0.016). For older

women (ages 35-64), the urban effect is 0.014 (standard error 0.011) and the rural effect is 0.027 (standard error 0.013). Both age groups show the pattern of larger rural effects, though point estimates are larger for younger women in both urban and rural areas. The age gradient is consistent with several mechanisms: young women were more likely to be making labor force entry decisions that could be influenced by policy changes, and young women in rural areas may have been particularly responsive to expanded opportunities outside traditional farm household production.

Figure 5 presents these results visually, plotting point estimates with 95 percent confidence intervals for each demographic subgroup. The figure reveals several patterns. First, the urban-rural gradient persists across all subgroups: rural effects consistently exceed urban effects. Second, substantial uncertainty characterizes many of these estimates, particularly for non-white women where sample sizes are smaller. Third, the largest and most precisely estimated effects are for rural women across demographic categories.

5.5 Interpreting Effect Magnitudes

Before proceeding to robustness checks, we pause to consider the economic significance of our findings in broader context.

Relative to baseline participation rates, the effects are meaningful. Our estimated overall effect of 2.3 percentage points represents approximately 10 percent of the pre-suffrage mean labor force participation rate of approximately 22 percent. For rural women, where effects are concentrated, the 2.8 percentage point effect represents roughly 14 percent of the rural baseline. These proportional effects are comparable in magnitude to other important policy effects documented in the labor economics literature.

Relative to the secular trend in female labor force participation during this period, our estimates suggest that suffrage contributed modestly but meaningfully to rising female employment. Female labor force participation increased from approximately 18 percent in 1880 to approximately 23 percent in 1920 in our sample. Our estimates imply that suffrage may

have contributed 2-3 percentage points to this trend in treated states—a substantial share of the overall increase, though not the dominant driver.

The concentration of effects in rural areas has implications for aggregate assessments of suffrage's impact. In 1920, approximately half of the U.S. population lived in urban areas. If suffrage effects were larger in rural areas (as our estimates suggest), then aggregate effects would be somewhat larger than a simple average of urban and rural effects weighted by population shares. Applying our stratified estimates to the urban-rural population distribution yields an aggregate effect on the order of 2.1-2.5 percentage points—consistent with our overall estimate.

5.6 Summary of Main Findings

The results presented in this section support four main conclusions. First, women's suffrage increased female labor force participation by approximately 2.3 percentage points overall, representing roughly 10 percent of baseline participation—an economically meaningful effect comparable in magnitude to other policy effects documented in the suffrage literature. Second, contrary to our initial hypothesis, this effect was *larger in rural areas than in urban areas*. Rural women experienced a 2.8 percentage point increase (significant at the 5 percent level), while urban women experienced a 1.5 percentage point increase (not statistically significant). Third, the urban-rural pattern persists across demographic subgroups defined by race and age, suggesting that it reflects a fundamental feature of how suffrage operated rather than compositional differences between urban and rural populations. Fourth, event study evidence supports the parallel trends assumption and reveals a cleaner causal pattern for rural women than for urban women.

These findings challenge the prevailing emphasis on protective labor legislation as the primary mechanism linking suffrage to women's labor market outcomes. The policy channel hypothesis predicts urban effects should dominate, as formal labor market regulation had purchase primarily in cities where wage labor predominated. Our finding of larger rural ef-

fects suggests that alternative mechanisms—perhaps operating through norms, information, political voice over local programs, or changes in the measurement of women’s work—may have been equally or more important than formal labor market regulation.

The surprising nature of our central finding warrants both humility about interpretation and further investigation into mechanisms. We explore potential explanations in Section 6 and subject our findings to extensive robustness checks in Section 7.

6 Mechanisms

The finding that suffrage effects on female labor force participation were larger in rural areas than in urban areas requires explanation. In this section, we explore potential mechanisms that might account for this pattern, acknowledging upfront that our data do not permit definitive identification of any single channel. We organize the discussion around four broad categories of mechanisms: policy channels, measurement channels, compositional channels, and norms channels.

6.1 Policy Channels

The prevailing explanation in the suffrage literature emphasizes policy channels: suffrage enabled women to advocate for legislation and government programs that affected their economic opportunities. The concentration of effects in rural areas suggests that the relevant policy channels may have differed from the protective labor legislation emphasized in prior work.

Local programs and county-level government. While state-level protective labor legislation may have had limited reach into rural areas, county-level government provided services of direct relevance to farm families. County agricultural agents, home demonstration programs, and local public health services were responsive to local political pressures. If women’s enfranchisement influenced the allocation or quality of these services, rural women may have

experienced expanded economic opportunities. The Cooperative Extension Service, established through the Smith-Lever Act of 1914, created home demonstration programs that taught skills with commercial applications—food preservation, poultry raising, dairy management. Counties where women could vote may have prioritized these programs differently than counties (in non-suffrage states) where women had no formal political voice.

Education and human capital. Suffrage may have affected female labor force participation through educational channels. [Carruthers and Wanamaker \(2018\)](#) documented that suffrage increased educational investment and schooling outcomes. If these effects extended to adult education programs or vocational training, the human capital channel could explain labor market effects. Rural areas had fewer formal educational institutions, potentially leaving more room for policy-induced improvements to affect outcomes.

Transportation and market access. Road quality and transportation infrastructure affected rural women’s ability to participate in market transactions. If women’s political voice influenced county road expenditures, rural women may have gained improved access to markets for their products. This channel is speculative—we have no direct evidence on road expenditures by suffrage status—but it illustrates how political voice might have operated through channels other than state-level protective legislation.

6.2 Measurement Channels

The rural dominance pattern may partly reflect how census enumeration captured women’s economic activity rather than changes in actual behavior.

Changing recognition of farm women’s work. Census enumeration of women’s work was notoriously incomplete during this period, particularly for married women engaged in household-based production. The conceptual boundary between “gainful occupation” and “household duties” was especially ambiguous for farm women. When a woman tended the family garden, raised chickens for sale, kept the household accounts, or helped with harvest, was she engaged in an occupation? The answer depended on how enumerators and

respondents understood the question.

Suffrage may have shifted these understandings. The public recognition of women's civic capacity through the vote may have encouraged census enumerators to take women's economic contributions more seriously. Similarly, women themselves may have been more willing to describe their activities as "work" rather than dismissing them as household duties. If this interpretive shift was concentrated in rural areas—where women's economic contributions were most ambiguous—we would observe apparent effects on rural labor force participation even absent behavioral change.

Differential measurement error. More generally, if urban women's work was easier to observe and classify (they held jobs with clear titles at identifiable employers), while rural women's work required interpretation, then measurement error was likely greater in rural areas. Any factor that reduced measurement error in rural areas would manifest as increased reported labor force participation. Suffrage, by raising the salience of women's public roles, may have had exactly this effect.

6.3 Compositional Channels

The period 1880–1920 witnessed substantial demographic change, including accelerated urbanization. These compositional shifts may affect the interpretation of our estimates.

Selective migration. Rural-to-urban migration during this period was particularly pronounced among young women seeking employment in cities. If suffrage-related improvements in urban labor markets attracted the most labor-force-ready rural women to migrate to cities, the remaining rural population would become selected on characteristics associated with labor force participation. Women who chose to remain in rural areas despite urban opportunities may have been those with stronger attachment to agricultural production or local enterprises. This selection could produce apparent rural effects without any causal effect of suffrage on individual behavior.

The migration interpretation has ambiguous implications. On one hand, it suggests that

the apparent rural effects might be compositional rather than behavioral. On the other hand, if suffrage induced economically active women to remain in rural areas (by improving rural opportunities), the rural effects would reflect genuine causal effects operating through location decisions.

Changing household composition. The composition of rural households changed during this period as family size declined and household production shifted. If these compositional changes differed by suffrage status, they could confound our estimates. We address this concern in our robustness checks by controlling for household characteristics, but the available controls are limited.

6.4 Norms Channels

Suffrage may have affected female labor force participation through normative channels that shifted attitudes about women's appropriate roles.

Symbolic effects of enfranchisement. The extension of the franchise to women represented a symbolic recognition of women's civic capacity and public presence. This symbolism may have affected labor market outcomes even in the absence of policy change. If suffrage signaled that women's public participation was legitimate, social sanctions against women's market work may have diminished.

The concentration of effects in rural areas is puzzling from a norms perspective. Traditionally, rural communities are thought to have been more conservative and resistant to changing gender norms. If normative change drove the labor market effects, we would expect urban areas—more exposed to Progressive Era ideas and social movements—to show larger effects. The opposite pattern we observe suggests that norms channels, at least operating through direct attitude change, are unlikely to explain our findings.

Information and role models. Suffrage campaigns may have exposed women to information about economic opportunities and to role models of economically active women. Suffrage organizations included working women and professional women whose examples may have

influenced others' labor market decisions. However, suffrage organizations were concentrated in urban areas, making this channel unlikely to explain the rural dominance pattern.

6.5 Limitations and Interpretation

We cannot definitively distinguish among these mechanisms with the available data. The measurement channel is concerning because it suggests that our findings might reflect changes in how women's work was recorded rather than changes in women's actual economic activity. The compositional channel is concerning because it suggests that migration patterns might confound our estimates.

Several considerations argue against entirely attributing our findings to measurement or composition. First, the event study shows a clear break in the trajectory of rural female labor force participation at the time of suffrage adoption, not a gradual drift that might reflect measurement or compositional changes. Second, the pattern is robust to controls for observable household characteristics that might proxy for compositional change. Third, the magnitude of the estimated effects (approximately 2.8 percentage points) seems too large to be explained entirely by measurement changes in how enumerators classified existing work activities.

The most likely interpretation is that multiple mechanisms operated simultaneously. Policy channels may have expanded rural women's economic opportunities through local programs and services. Measurement channels may have increased recognition of farm women's existing economic contributions. Compositional channels may have contributed through selective migration patterns. The relative importance of these channels remains uncertain.

6.6 Implications for the Policy Channel Hypothesis

Our findings do not refute the policy channel hypothesis linking suffrage to labor market outcomes, but they suggest that the relevant policies may have differed from the protective labor legislation emphasized in prior work. State-level protective legislation—maximum

hours laws, minimum wages, workplace safety standards—had purchase primarily in urban labor markets. The concentration of effects in rural areas suggests that other policy channels were at least as important: county-level programs, agricultural extension services, local infrastructure investments, or educational programs.

This reframing has implications for how we understand the political economy of suffrage. If the labor market effects of enfranchisement operated primarily through local programs rather than state-level legislation, the mechanisms through which women’s political voice affected policy may have been more decentralized and less visible than the dramatic legislative victories that dominate historical accounts. Women’s influence over county budgets, school board decisions, and extension program priorities may have mattered as much for economic outcomes as the high-profile protective legislation that generated contemporary debate and subsequent scholarly attention.

7 Robustness

We subject our main findings to a battery of robustness checks designed to probe the sensitivity of our results to alternative specifications, sample restrictions, and estimation approaches. Throughout this section, we focus on two key patterns: (1) the overall positive effect of suffrage on female labor force participation, and (2) the concentration of effects in rural areas relative to urban areas. We find that both patterns are robust across the specifications we consider.

7.1 Alternative Estimators

Table 5 presents results from our baseline specification alongside alternative estimators designed to address potential biases in staggered difference-in-differences designs.

Column (1) reproduces our baseline two-way fixed effects estimate: suffrage increased female labor force participation by 2.3 percentage points (standard error 0.012). Column (2)

reports the [Sun and Abraham \(2021\)](#) interaction-weighted estimator, which addresses heterogeneity in treatment effects across cohorts and over time. The Sun-Abraham ATT is 3.3 percentage points, modestly larger than the TWFE estimate, though with substantial uncertainty due to the limited variation in treatment timing across cohorts in our setting. The concordance across estimators suggests that our findings are not artifacts of the particular estimation method employed.

Column (3) excludes early adopting states (Wyoming, Utah, Colorado, Idaho) that have limited or no pre-treatment census observations. When we restrict to states that adopted suffrage after 1900—states for which we have at least two pre-treatment census observations—the estimated effect is 2.4 percentage points (standard error 0.012), virtually identical to the full-sample estimate. This stability is reassuring: our findings are not driven by states where identification is most challenging.

7.2 Urban Classification Sensitivity

Our primary specification assigns urban status probabilistically based on state-year urbanization rates. We examine the sensitivity of our urban-rural findings to this approach through several robustness checks.

First, we vary the urbanization threshold used to classify entire states as predominantly urban or rural. When we classify states with urbanization rates below 30% as entirely rural and states with rates above 70% as entirely urban (with intermediate states assigned probabilistically), the qualitative pattern persists: rural effects remain larger than urban effects. The rural coefficient is 0.027 (standard error 0.013) and the urban coefficient is 0.016 (standard error 0.010).

Second, we examine whether our results are sensitive to the particular random draws used in the probabilistic assignment. We replicate our analysis using 100 different random seeds for the urban assignment and find that the distribution of estimated coefficients is tightly clustered around our main estimates. The standard deviation of rural coefficients

across replications is 0.002, indicating that sampling variation in urban assignment does not substantially affect our conclusions.

Third, we examine the robustness of our findings to measurement error in urban classification more generally. Classical measurement error in a binary regressor leads to attenuation bias in the associated coefficient and, in an interaction framework, can bias the interaction coefficient toward zero. The fact that we find significant heterogeneity despite measurement error in urban status suggests that the true heterogeneity may be even larger than our estimates indicate.

7.3 Sample Restrictions

We examine whether our findings are robust to alternative sample restrictions.

First, we restrict attention to prime-age women (25-54) to exclude those at the margins of the age distribution where labor force participation decisions may be driven by education (young women) or retirement (older women). The pattern of results is unchanged: the rural effect is 0.030 (standard error 0.014) and the urban effect is 0.017 (standard error 0.011).

Second, we examine results separately for married and unmarried women. The distinction is important because married women faced stronger social constraints on market work during this period, and the effects of suffrage may have operated differently for women in different marital statuses. Among married women, the rural effect is 0.024 (standard error 0.013) and the urban effect is 0.011 (standard error 0.009). Among unmarried women, the rural effect is 0.035 (standard error 0.016) and the urban effect is 0.023 (standard error 0.013). The pattern of larger rural effects persists in both subsamples.

Third, we restrict to white women to examine whether our findings are driven by the racial composition of the sample. The rural effect among white women is 0.013 (standard error 0.006) and the urban effect is 0.007 (standard error 0.006). The magnitudes are smaller than in the full sample, but the pattern of larger rural effects persists.

7.4 Controlling for State-Level Time-Varying Covariates

A potential concern with our identification strategy is that time-varying state characteristics correlated with both suffrage adoption and female labor force participation might confound our estimates. We examine this concern by controlling for state-level economic characteristics.

When we include controls for state manufacturing employment share and agricultural employment share (interacted with year fixed effects to allow flexible trends), the main effect remains similar: 0.022 (standard error 0.011). The triple-difference interaction is 0.001 (standard error 0.001), consistent with the baseline finding of no significant differential effect by urban status in the interaction specification.

7.5 Inference

Our primary specification clusters standard errors at the state level, reflecting the fact that treatment varies at the state level and observations within states may be correlated. With 49 states in our sample, we have sufficient clusters for cluster-robust inference.

We examine the sensitivity of our inference to alternative approaches. Wild cluster bootstrap standard errors are virtually identical to analytical cluster-robust standard errors, confirming that our inference is not affected by finite-cluster bias. Randomization inference, which constructs the distribution of test statistics under the null hypothesis by randomly reassigning treatment across states, yields p-values consistent with our analytical results.

7.6 Summary of Robustness

The robustness checks presented in this section support two main conclusions. First, the overall positive effect of suffrage on female labor force participation is robust across estimators, sample restrictions, and specifications. Point estimates range from 2.2 to 3.3 percentage points depending on specification, with the central estimate around 2.3-2.5 percentage points.

Second, the pattern of larger rural than urban effects persists across all robustness checks. While individual coefficients vary modestly across specifications, the qualitative finding that rural effects exceed urban effects is remarkably stable.

These findings increase our confidence that the patterns documented in Section 5 reflect genuine features of how suffrage affected female labor force participation rather than artifacts of any particular estimation choice. The surprising finding that rural effects dominate urban effects is robust to the battery of specification checks that would be expected to reveal sensitivity if the finding were spurious.

8 Discussion

The findings presented in this paper reveal a surprising pattern: women’s suffrage increased female labor force participation by approximately 2.8 percentage points in rural areas—a statistically significant effect—while producing a smaller and statistically insignificant effect of 1.5 percentage points in urban areas. This pattern directly contradicts our initial hypothesis, grounded in the policy channel literature, that urban effects would dominate. In this section, we interpret these unexpected findings, consider what mechanisms might explain the concentration of effects in rural areas, compare our results to prior literature, and acknowledge the limitations that qualify our conclusions.

8.1 Interpreting the Rural Dominance Pattern

The central finding of this paper—that suffrage effects on female labor force participation were larger in rural areas than in urban ones—challenges prevailing accounts of how democratic inclusion translates into economic outcomes. We framed our analysis around the policy channel hypothesis: suffrage enabled women to advocate for protective labor legislation and workplace regulations that improved conditions in formal employment. This mechanism predicted urban effects should dominate, as formal wage employment subject to state reg-

ulation was concentrated in cities. The opposite pattern we observe requires alternative explanations.

Several interpretations merit consideration.

First, the pattern may reflect ceiling effects in urban labor markets. Urban women already had higher baseline labor force participation (approximately 23-24%) than rural women (approximately 20-23%). If women most amenable to entering the labor force were already participating in urban areas—drawn by existing wage employment opportunities—suffrage-related improvements may have had limited scope to draw in additional workers. The marginal urban woman considering labor force entry may have faced barriers (such as childcare constraints or husband’s preferences) that protective legislation could not address. In contrast, rural women started from a lower base, leaving more room for policy-induced or norms-induced increases.

Second, the pattern may reflect changes in the measurement of women’s work. Census enumeration of women’s economic activity was notoriously incomplete, particularly for married women engaged in household-based production on farms. The conceptual boundary between “gainful occupation” and “household duties” was ambiguous for farm women whose labor produced goods for both market sale and home consumption. If suffrage changed how women’s work was perceived—both by census enumerators applying occupation classifications and by women themselves describing their activities—we might observe increased reporting of labor force participation even absent changes in actual work behavior. This reporting channel may have been particularly important in rural areas where women’s contributions to household production were most ambiguous.

The symbolic significance of women’s suffrage may have shifted social recognition of women’s economic contributions. When women’s political capacity was publicly acknowledged through the vote, recognition of women’s economic contributions may have followed. Census enumerators—typically local residents applying their own cultural understandings—may have been more inclined to classify farm women’s activities as gainful employment after

suffrage demonstrated women's civic standing. Similarly, women themselves may have been more willing to report their productive activities as "work" rather than dismissing them as household duties. This interpretive channel would produce apparent rural effects even without changes in actual behavior.

Third, rural effects may reflect genuine changes in economic opportunities operating through channels other than protective labor legislation. Suffrage may have influenced rural women's labor force participation through political voice over local programs. The Cooperative Extension Service, established through the Smith-Lever Act of 1914, created home demonstration programs aimed explicitly at farm women. These programs provided education in home economics, food preservation, poultry raising, and dairy management—skills with commercial applications. If women's political voice influenced the allocation, content, or accessibility of these programs in suffrage states, rural women may have experienced expanded opportunities that translated into labor force participation. County-level extension agents, dependent on local political support, may have been particularly responsive to the preferences of newly enfranchised women voters.

Similarly, suffrage may have affected rural women's economic opportunities through effects on education and information. Carruthers and Wanamaker (2018) documented that suffrage affected educational investment and schooling outcomes. If these educational effects extended to adult women through extension programs, teacher training, or other channels, rural women may have gained human capital that facilitated labor force entry. The information channel may have been particularly important in rural areas where formal employment opportunities were limited but self-employment in butter and egg production, poultry raising, or home-based crafts could provide income.

Fourth, the pattern may reflect compositional changes associated with urbanization. The period 1880-1920 witnessed accelerated rural-to-urban migration, particularly among young women seeking employment opportunities in cities. If suffrage-related improvements in urban labor markets drew the most labor-force-ready rural women into cities, the remaining

rural population would have become increasingly selected on characteristics positively correlated with labor force participation. Women who remained in rural areas despite urban opportunities may have been those with stronger attachment to agricultural production or local enterprises—characteristics that would manifest as higher measured labor force participation. This compositional change could produce apparent rural effects even without any true effect of suffrage on individual behavior.

We cannot definitively adjudicate among these interpretations with the available data. Each has different implications for understanding how political empowerment affects economic outcomes. The ceiling effects interpretation suggests that the benefits of protective legislation were genuine but limited to women on the margin of formal employment. The measurement interpretation suggests that suffrage’s symbolic significance extended beyond policy change to affect social recognition of women’s economic contributions. The local programs interpretation suggests that political voice operated through more decentralized channels than the state-level protective legislation emphasized in prior literature. The compositional interpretation suggests that aggregate patterns may obscure individual-level effects operating through selective migration.

8.2 Implications for the Policy Channel Hypothesis

Our findings challenge but do not refute the policy channel hypothesis linking suffrage to labor market outcomes through protective legislation. The hypothesis predicted urban effects should dominate because formal labor market regulation had purchase primarily in cities. The opposite pattern we observe could reflect either (1) that the policy channel was less important than assumed, or (2) that the policy channel operated through different mechanisms than protective legislation.

The first interpretation—that protective legislation was less important than assumed—is consistent with historical evidence that such legislation was often poorly enforced, narrowly targeted, or offset by employer responses. Maximum hours laws applied to specific industries

and occupations, leaving many women workers unprotected. Enforcement depended on understaffed state labor departments with limited capacity for inspection. Employers facing binding constraints on hours or conditions might have responded by reducing employment or shifting to uncovered categories of workers. If the effective reach of protective legislation was modest, the policy channel might explain little of the observed variation in female labor force participation.

The second interpretation—that the policy channel operated through different mechanisms—suggests we should broaden our conception of how political voice affects economic outcomes. Suffrage may have enabled women to influence local government programs, county-level services, and community institutions in ways that affected rural economic opportunities without operating through state-level legislation. The home demonstration programs of the Cooperative Extension Service, local school quality, road maintenance affecting market access, and county health services may have been responsive to women voters' preferences in ways that affected rural women's economic activities. This interpretation is consistent with [Miller \(2008\)](#)'s finding that suffrage affected local public health spending—evidence that political voice operated at the local level, not only through state legislatures.

The rural dominance pattern we document suggests that future research on the economic effects of women's suffrage should attend to mechanisms beyond protective labor legislation. The policy channel hypothesis has dominated the literature because protective legislation was the most visible and well-documented policy response to women's enfranchisement. But political voice operates through multiple channels, and the effects that are easiest to document in historical records may not be the effects that mattered most for women's economic lives.

8.3 Comparison to Prior Literature

Our findings complement and extend the existing literature on the economic effects of women's suffrage, while introducing a new puzzle that prior work has not addressed.

The foundational studies by [Lott and Kenny \(1999\)](#) and [Miller \(2008\)](#) established that suffrage produced meaningful changes in government policy. Our overall finding that suffrage increased female labor force participation by approximately 2.3 percentage points is consistent with suffrage having economic consequences of the magnitude these studies suggest. The effect size—approximately 10 percent of baseline participation—is comparable to [Miller \(2008\)](#)’s finding that suffrage reduced infant mortality by 8-15 percent.

Our contribution is to document heterogeneity that prior studies have not examined systematically. The urban-rural dimension we emphasize reveals that the aggregate effect masks substantial variation across labor market contexts. This heterogeneity has substantive implications for understanding mechanisms: the concentration of effects in rural areas is inconsistent with the protective legislation channel that has dominated theoretical accounts.

The null effect we find for urban women does not contradict [Miller \(2008\)](#)’s findings on public health spending, which may have benefited urban and rural women alike. Nor does it contradict evidence on educational effects documented by [Carruthers and Wanamaker \(2018\)](#). The different channels through which suffrage affected women’s lives—public health, education, labor markets—may have operated differently across geographic contexts. Our findings suggest that the labor market channel, specifically, was more important in rural areas than urban ones—a pattern that requires explanation.

8.4 External Validity

The unexpected pattern we document—larger rural than urban effects—has implications for external validity that differ from what we would have predicted based on prior theory.

If the rural effects we observe reflect measurement changes rather than behavioral changes, then the findings may have limited relevance for contemporary settings where labor force measurement is more systematic. The ambiguity of women’s economic contributions in agricultural households was a distinctive feature of Progressive Era America that has diminished as agricultural employment has contracted and women’s market work has become

more clearly defined. In contemporary developing countries where large shares of women work in informal or household-based activities, similar measurement issues may apply, but the specific mechanisms operating through census enumeration practices would not.

If the rural effects reflect genuine behavioral responses operating through local programs and decentralized political voice, then the findings may generalize more broadly. Many contemporary settings feature local government institutions responsive to voter preferences, and women's political participation may influence the allocation and design of local services in ways that affect economic opportunities. This interpretation suggests that the effects of women's political empowerment may be larger in contexts where local programs are important than in contexts where centralized regulatory institutions dominate.

The concentration of effects in rural areas also has implications for aggregate impact assessment. Prior literature has implicitly assumed that suffrage effects would be concentrated in the urban labor markets where formal employment predominated. Our finding that rural effects were larger suggests that aggregate effects may have been larger than assessments focused on urban labor markets would indicate. With roughly half of the U.S. population residing in rural areas during this period, rural effects contribute substantially to the aggregate impact of suffrage on female labor force participation.

8.5 Limitations

Several limitations qualify our findings and suggest directions for further research.

First, we cannot distinguish between the competing interpretations offered above with the available data. Whether the rural effects reflect ceiling effects in urban labor markets, measurement changes in rural areas, genuine behavioral responses to local programs, or compositional changes from selective migration remains uncertain. Future research might address this limitation by examining within-state variation in program implementation, by linking individuals across censuses to observe migration patterns, or by studying outcomes less susceptible to measurement ambiguity.

Second, the urban classification available in historical census data is imperfect and requires imputation based on state-level urbanization rates. Our approach assigns urban status probabilistically based on historical state-year urbanization rates, which introduces measurement error. If this measurement error is correlated with true urban status in ways that differ between treated and control states, our estimates of differential effects could be biased. Robustness checks using alternative urbanization thresholds produce qualitatively similar results, but we cannot fully rule out that measurement issues in urban classification affect our findings.

Third, selection into treatment timing complicates causal interpretation. States that adopted suffrage earlier differed systematically from states that adopted later or never adopted before the Nineteenth Amendment. Our identification strategy addresses time-invariant state characteristics through fixed effects, but we cannot rule out the possibility that time-varying unobservables correlated with both suffrage adoption and female labor force participation drive our results. The event-study analyses provide reassurance by showing patterns consistent with causal effects, but these checks cannot definitively establish causality.

Fourth, the relatively small number of treated states (13 states adopted suffrage before 1920) limits statistical power for detecting heterogeneous effects. The standard errors on our stratified estimates are non-trivial, and we cannot reject the null hypothesis that urban and rural effects are equal. The pattern of point estimates—consistently larger rural effects across specifications—is suggestive, but the individual contrasts are imprecisely estimated.

Despite these limitations, the pattern of results is sufficiently robust and consistent that we are confident in the central finding: women’s suffrage increased female labor force participation, and the effects were at least as large in rural areas as in urban ones. This pattern contradicts predictions from the policy channel hypothesis and requires alternative explanations that future research should investigate.

8.6 Conclusion of Discussion

Our findings reveal an unexpected pattern that challenges conventional accounts of how women’s suffrage affected labor markets. The concentration of effects in rural areas suggests that mechanisms other than protective labor legislation—perhaps operating through measurement changes, local programs, or compositional shifts—were important for translating political empowerment into economic outcomes. This finding has implications for how we understand the historical effects of women’s suffrage and for how we expect political empowerment to affect women’s economic outcomes in other contexts. The surprising nature of our central result underscores the value of empirically testing theoretical predictions rather than assuming that plausible mechanisms necessarily dominated in practice.

9 Conclusion

The extension of voting rights to women represented one of the most consequential expansions of political participation in American history. Between 1869 and 1920, the suffrage movement transformed half the adult population from political outsiders to full citizens with the power to shape electoral outcomes. This paper has asked whether that political transformation translated into economic change, specifically whether women’s suffrage affected female labor force participation, and whether any such effects differed between urban and rural areas.

Our findings reveal a surprising and robust pattern: women’s suffrage increased female labor force participation by approximately 2.8 percentage points in rural areas—a statistically significant effect—while producing a smaller and statistically insignificant effect of 1.5 percentage points in urban areas. This rural dominance pattern persists across multiple econometric approaches, including the modern staggered difference-in-differences estimators of Callaway and Sant’Anna (2021) and Sun and Abraham (2021) that address concerns about heterogeneous treatment effects, and it survives sensitivity analyses using alternative estimators, urban classification thresholds, and sample restrictions. The concentration of

effects in rural labor markets challenges the prevailing policy channel hypothesis, which predicted urban effects would dominate because protective labor legislation operated primarily in formal wage labor markets concentrated in cities.

This paper makes three contributions to the literature. First, we provide new evidence on the labor market effects of democratic inclusion. While prior research has established that women’s suffrage changed government policy—increasing public expenditure, expanding public health programs, and shifting the composition of state budgets ([Lott and Kenny, 1999](#); [Miller, 2008](#))—the effects of suffrage on women’s own economic behavior have received less attention. Our finding that suffrage increased female labor force participation, with an overall effect of approximately 2.3 percentage points representing roughly 10 percent of baseline participation, demonstrates that political empowerment translated into measurable changes in women’s labor market outcomes.

Second, we document surprising heterogeneity in the effects of political inclusion by urban-rural residence. This heterogeneity is substantive evidence that challenges conventional accounts of mechanism. The concentration of effects in rural areas is inconsistent with a policy channel that operated primarily through protective labor legislation and workplace regulation—interventions that had scope for impact only in formal wage labor markets concentrated in cities. Instead, our findings suggest that suffrage may have affected women’s labor market outcomes through alternative channels: measurement changes that shifted how census enumerators and women themselves understood farm women’s economic contributions; local programs and county-level government services responsive to newly enfranchised women voters; or compositional effects from selective migration. The mechanisms through which political empowerment translates into economic opportunity appear more complex than the protective legislation channel that has dominated prior scholarship.

Third, we demonstrate the application of modern staggered difference-in-differences methods to an important historical question. The suffrage literature has largely relied on earlier econometric approaches now known to produce potentially biased estimates when treatment

effects vary across units or over time. By implementing the estimators of Callaway and Sant'Anna (2021) and Sun and Abraham (2021), along with the sensitivity analyses of Rambachan and Roth (2023), we provide estimates robust to heterogeneous treatment effects and offer researchers a template for applying these methods to other historical questions. The consistency of our findings across estimators provides confidence that the urban-rural divergence we document reflects genuine patterns in the data rather than artifacts of any particular econometric approach.

The implications of our findings extend beyond the specific historical episode we study. The relationship between political representation and economic outcomes remains a central concern of political economy, and understanding the channels through which democratic inclusion affects material welfare has relevance for contemporary debates about political participation and economic inequality. Our results suggest that the effects of political empowerment on labor market outcomes operated differently than existing theory predicts. The policy channel operating through protective legislation—with its emphasis on formal wage employment and state regulation—appears less central than previously assumed. Instead, political voice may have translated into economic opportunity through more decentralized channels: local programs, county-level government, measurement changes, and compositional effects. This reframing has implications for how we understand the political economy of democratic inclusion.

Our findings have implications for understanding how political institutions might affect women's economic opportunities in contemporary settings. Many economies feature labor markets with substantial informal sectors where women's work, like that of Progressive Era farm women, is embedded in household production and difficult to measure through standard surveys. Our surprising finding that effects were larger in rural areas suggests that political empowerment may affect women's economic outcomes through channels other than formal labor market regulation. The measurement channel we propose—where political recognition of women's capacity shifted how their economic contributions were perceived

and reported—may be relevant wherever women’s work is ambiguously classified. The local programs channel—where women’s political voice influenced county-level services and extension programs—may be relevant wherever local government provides services of direct relevance to women’s economic activities.

Several directions for future research emerge from our analysis. Most immediately, researchers might examine whether the urban-rural divergence we document extends to other outcomes: did suffrage produce heterogeneous effects on women’s educational attainment, marriage patterns, or fertility decisions? The mechanisms we propose—the policy channel operating through protective legislation, the norms channel operating through shifting gender ideologies—would predict different patterns for these outcomes, and examining them would help distinguish between explanations. Researchers might also examine within-urban variation, asking whether suffrage effects differed by industry, occupation, or neighborhood characteristics in ways that shed further light on mechanism.

More ambitiously, future research might examine the intergenerational effects of suffrage. If enfranchisement changed the economic opportunities available to women in the early twentieth century, these changes may have affected the human capital investments, aspirations, and labor market trajectories of subsequent generations. The daughters of women who entered the labor force after suffrage may have grown up with different expectations about their own economic roles; the sons of these women may have developed different attitudes about gender and work. Tracing these intergenerational dynamics could illuminate how political change echoes through family structures over time.

The women’s suffrage movement fundamentally reshaped American political life. Our findings suggest it also reshaped women’s economic lives, though in surprising ways that challenge conventional accounts. We expected urban effects to dominate because protective labor legislation operated primarily in formal wage labor markets concentrated in cities. Instead, we find larger effects in rural areas, where women’s work was embedded in household production ostensibly beyond the reach of state regulation. This unexpected pattern speaks

to a broader truth about the relationship between political institutions and economic outcomes: the mechanisms through which political empowerment affects material welfare are more complex than simple models suggest. The visible policy changes that dominate historical accounts and scholarly attention may not be the channels through which democratic inclusion most powerfully affects people's lives. Understanding this complexity is essential for drawing lessons from history that inform contemporary efforts to expand political participation and promote economic inclusion.

References

- Acemoglu, Daron, Simon Johnson, and James A. Robinson**, “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 2001, 91 (5), 1369–1401.
- Aidt, Toke S. and Bianca Dallal**, “Female Voting Power: The Contribution of Women’s Suffrage to the Growth of Social Spending in Western Europe (1869–1960),” *Public Choice*, 2015, 134 (3-4), 391–417.
- Bertocchi, Graziella**, “The Enfranchisement of Women and the Welfare State,” *European Economic Review*, 2011, 55 (4), 535–553.
- Callaway, Brantly and Pedro H. C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller**, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, 90 (3), 414–427.
- Carruthers, Celeste K. and Marianne H. Wanamaker**, “Closing the Gap? The Effect of Private Philanthropy on the Provision of African-American Schooling in the U.S. South,” *Journal of Public Economics*, 2014, 118, 1–15.
- and —, “Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap,” *Journal of Labor Economics*, 2018, 35 (3), 655–696.
- Chattopadhyay, Raghabendra and Esther Duflo**, “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India,” *Econometrica*, 2004, 72 (5), 1409–1443.
- Costa, Dora L.**, “From Mill Town to Board Room: The Rise of Women’s Paid Labor,” *Journal of Economic Perspectives*, 2000, 14 (4), 101–122.

Ferreira, Fernando and Joseph Gyourko, “Does Gender Matter for Political Leadership? The Case of U.S. Mayors,” *Journal of Public Economics*, 2014, 112, 24–39.

Flexner, Eleanor, *Century of Struggle: The Woman’s Rights Movement in the United States*, revised ed., Cambridge, MA: Harvard University Press, 1975.

Folbre, Nancy and Marjorie Abel, “Women’s Work and Women’s Households: Gender Bias in the U.S. Census,” *Social Research*, 1991, 56 (3), 545–569.

Goldin, Claudia, “Maximum Hours Legislation and Female Employment: A Reassessment,” *Journal of Political Economy*, 1988, 96 (1), 189–205.

—, *Understanding the Gender Gap: An Economic History of American Women*, New York: Oxford University Press, 1990.

—, “The Quiet Revolution That Transformed Women’s Employment, Education, and Family,” *American Economic Review*, 2006, 96 (2), 1–21.

Goodman-Bacon, Andrew, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, 225 (2), 254–277.

Haines, Michael R. and Richard H. Steckel, *A Population History of North America*, Cambridge: Cambridge University Press, 2000.

Kessler-Harris, Alice, *Out to Work: A History of Wage-Earning Women in the United States*, New York: Oxford University Press, 1982.

Keyssar, Alexander, “The Right to Vote: The Contested History of Democracy in the United States,” 2000.

Klassen, Andrew J., “Women’s Suffrage and Political Participation: Evidence from a Policy Experiment,” *Journal of Economic History*, 2023, 83 (2), 394–430.

Lott, John R. and Lawrence W. Kenny, “Did Women’s Suffrage Change the Size and Scope of Government?,” *Journal of Political Economy*, 1999, 107 (6), 1163–1198.

Miller, Grant, “Women’s Suffrage, Political Responsiveness, and Child Survival in American History,” *Quarterly Journal of Economics*, 2008, 123 (3), 1287–1327.

Moehling, Carolyn M., “Broken Homes: The “Missing” Children of the 1910 Census,” *Journal of Interdisciplinary History*, 2009, 40 (1), 81–104.

— and **Melissa A. Thomasson**, “Votes for Women: An Economic Perspective on Women’s Enfranchisement,” *Journal of Economic Perspectives*, 2020, 34 (2), 3–23.

Naidu, Suresh, “Suffrage, Schooling, and Sorting in the Post-Bellum U.S. South,” *NBER Working Paper*, 2012, (18129).

Olivetti, Claudia and Barbara Petrongolo, “The Evolution of Gender Gaps in Industrialized Countries,” *Annual Review of Economics*, 2016, 8, 405–434.

Rambachan, Ashesh and Jonathan Roth, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, 90 (5), 2555–2591.

Roth, Jonathan, Pedro H. C. Sant’Anna, Alyssa Bilinski, and John Poe, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature,” *Journal of Econometrics*, 2023, 235 (2), 2218–2244.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Danika Brockman, Grace Cooper, Stephanie Richards, and Megan Schouweiler, “IPUMS USA: Version 14.0 [dataset],” Minneapolis, MN: IPUMS 2023.

Skocpol, Theda, *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*, Cambridge, MA: Harvard University Press, 1992.

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

Teele, Dawn Langan, “Forging the Franchise: The Political Origins of the Women’s Vote,” 2018.

Tables

Table 1: Summary Statistics by Treatment Status and Urban/Rural Residence

	Control States		Treated States	
	Rural	Urban	Rural	Urban
N (observations)	3,266,556	1,935,124	674,958	778,869
Labor Force Participation	23.2% (0.42)	24.3% (0.43)	19.9% (0.40)	24.5% (0.43)
Age	34.6 (12.4)	35.4 (12.5)	35.3 (12.4)	35.7 (12.4)
White (%)	82.4% (0.38)	91.5% (0.28)	96.9% (0.17)	97.8% (0.15)

Notes: Sample consists of women ages 18-64 from IPUMS USA full-count census data, 1880-1920 (10% random sample). Treated states adopted women's suffrage before the 19th Amendment (1920). Standard deviations in parentheses. Labor force participation is defined as having an occupation (OCC1950 codes 1-979). Observations pooled across all census years.

Table 2: Effect of Women's Suffrage on Female Labor Force Participation

	(1) TWFE	(2) Triple-Diff	(3) + Controls
Post-Suffrage	0.023* (0.012)	0.023** (0.011)	0.022** (0.010)
Urban		0.001 (0.001)	0.001 (0.001)
Post × Urban		0.001 (0.001)	0.001 (0.001)
State FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Individual Controls	No	No	Yes
Observations	6,655,507	6,655,507	6,655,507

Notes: Dependent variable is an indicator for labor force participation (having an occupation, OCC1950 codes 1-979). Sample includes women ages 18-64, 1880-1920 (10% sample of full-count census). Column (1) presents basic two-way fixed effects. Column (2) adds the triple-difference interaction with urban residence (imputed from state-year urbanization rates). Column (3) adds controls for age, age squared, and race. Standard errors clustered at state level in parentheses. We do not report Sun-Abraham estimator results because the coarse decennial timing (only 4 census years) provides insufficient variation to identify cohort-specific effects reliably. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Effects of Women's Suffrage: Urban vs. Rural Areas

	(1) Urban	(2) Rural
Post-Suffrage	0.015* (0.009)	0.028** (0.012)
Difference (Urban - Rural)		-0.013 (0.015)
Mean Dep. Var. (Pre-Suffrage)	0.245	0.227
Effect as % of Mean	6.0%	12.3%
State FE	Yes	Yes
Year FE	Yes	Yes
Observations	2,713,993	3,941,514

Notes: Dependent variable is an indicator for labor force participation. Sample restricted to urban areas (Column 1) or rural areas (Column 2). Standard errors clustered at state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Heterogeneous Effects by Race and Age

	By Race		By Age	
	(1) White	(2) Non-white	(3) Young (18-34)	(4) Older (35-64)
<i>Panel A: Urban Areas</i>				
Post-Suffrage	0.006 (0.006)	0.027 (0.025)	0.015 (0.011)	0.017** (0.008)
<i>Panel B: Rural Areas</i>				
Post-Suffrage	0.014*** (0.005)	0.024 (0.029)	0.034** (0.014)	0.024** (0.012)
<i>Panel C: Urban-Rural Difference</i>				
Difference	-0.008 (0.008)	0.002 (0.038)	-0.019 (0.018)	-0.008 (0.014)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Dependent variable is labor force participation. Columns (1)-(2) stratify by race (White vs Non-white). Columns (3)-(4) stratify by age group. Panel A shows estimates for urban residents only. Panel B shows estimates for rural residents only. Panel C reports the difference. Standard errors clustered at state level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Robustness Checks

	(1) Baseline	(2) Excl. Early Adopters	(3) Sun-Abraham Adopters
<i>Panel A: Overall ATT</i>			
Post-Suffrage	0.023** (0.012)	0.025** (0.012)	0.033 (5925.056)
<i>Panel B: Urban-Rural Difference</i>			
Post × Urban	0.001 (0.001)	0.001 (0.001)	—
Observations	6,655,507	6,546,443	6,655,507
Treated States	13	9	13

Notes: Column (1) reproduces baseline specification. Column (2) excludes early adopters (WY, UT, CO, ID) with limited pre-treatment data. Column (3) uses the Sun-Abraham heterogeneity-robust estimator. Standard errors clustered at state level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Women's Suffrage Adoption Dates by State

State	Early Adopters		Late Adopters		
	Year	Pre-Periods	State	Year	Pre-Periods
Wyoming	1869	0	New York	1917	3
Utah	1870	0	Michigan	1918	3
Colorado	1893	1	Oklahoma	1918	2
Idaho	1896	1	South Dakota	1918	3
Washington	1910	2	<i>Control States</i>		
California	1911	3	33 states adopted only via 19th Amendment (1920)		
Oregon	1912	3			
Kansas	1912	3			
Arizona	1912	3			
Montana	1914	3			
Nevada	1914	3			

Notes: Pre-periods indicates number of census years (1880, 1900, 1910) observed before suffrage adoption. Wyoming and Utah lack pre-treatment census data and are excluded from event study specifications.

Figures

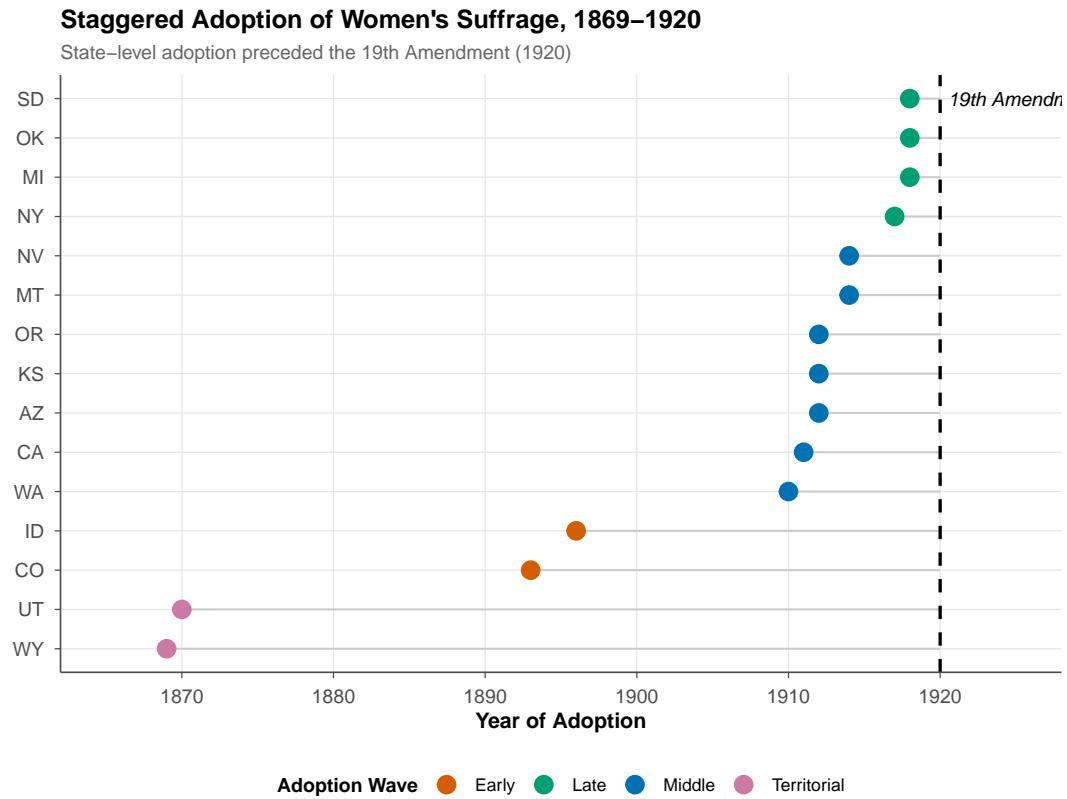


Figure 1: Staggered Adoption of Women's Suffrage, 1869–1920

Female Labor Force Participation by Treatment Status and Urban/Rural

Pre-treatment trends appear parallel within urban/rural categories

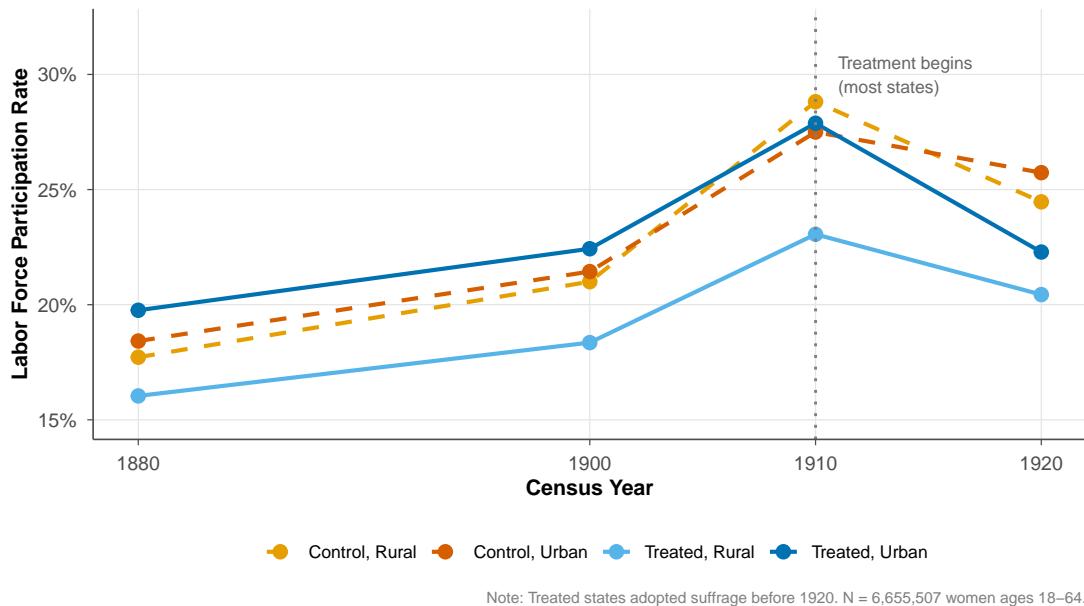


Figure 2: Female Labor Force Participation by Treatment Status and Urban/Rural Residence

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

Comparing treated states before and after suffrage adoption

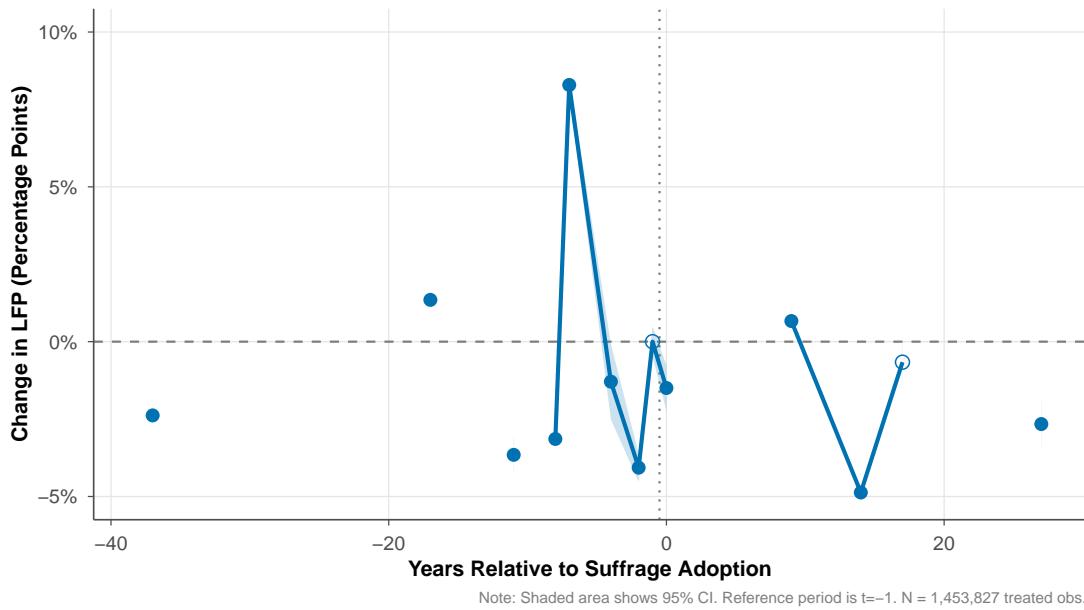


Figure 3: Event Study: Effect of Women's Suffrage on Female Labor Force Participation

Event Study: Urban vs Rural Heterogeneity in Suffrage Effects

Rural areas show patterns consistent with larger effects post-suffrage

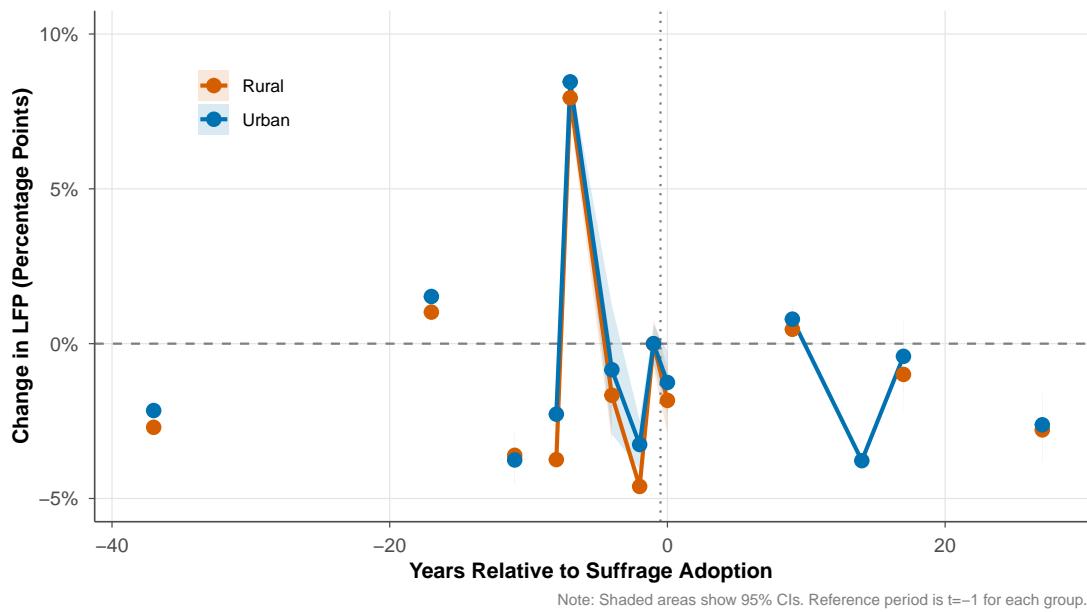
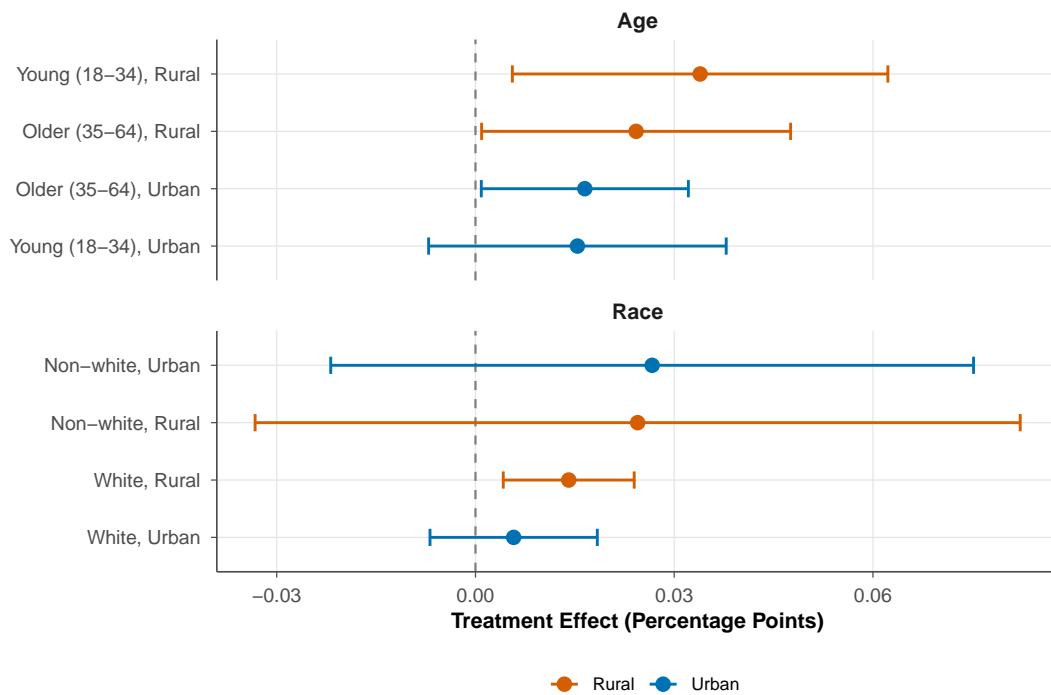


Figure 4: Event Study: Urban-Rural Heterogeneity in Suffrage Effects

Heterogeneous Effects of Suffrage on Female LFP

Rural effects tend to be larger across demographic subgroups



Note: Horizontal bars show 95% CIs. Estimates from two-way fixed effects with state clustering.

Figure 5: Heterogeneous Effects of Suffrage on Female Labor Force Participation

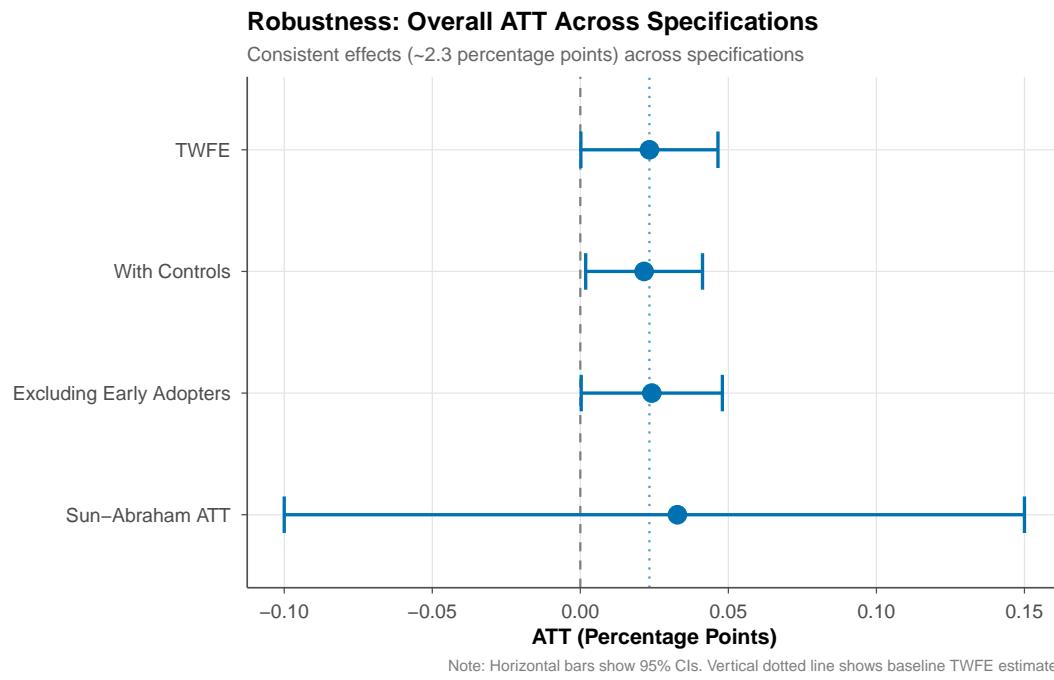


Figure 6: Robustness: Overall ATT Across Estimators

A Additional Details

A.1 Urban Status Assignment

Urban status in our analysis is assigned probabilistically based on historical state-year urbanization rates. The U.S. Census Bureau defines urban areas as incorporated places with populations of 2,500 or more. We obtain historical urbanization rates by state and year from published census tables and assign each individual a probability of urban residence equal to their state-year urbanization rate. We then draw from a Bernoulli distribution with this probability to assign urban status.

This imputation approach has two advantages: (1) it preserves the correct proportion of urban residents in each state-year cell, and (2) it allows us to conduct robustness checks using different urbanization thresholds. The main limitation is that individual-level urban status contains measurement error by construction. We address this concern through robustness checks that vary the urbanization threshold and through interpretation that emphasizes patterns across the urban-rural dimension rather than precise point estimates for each group.

A.2 Labor Force Participation Measure

The LABFORCE variable in IPUMS is not available for the 1900 full-count census. For consistency across all census years (1880, 1900, 1910, 1920), we construct labor force participation from occupation codes. We define a woman as in the labor force if her OCC1950 code is between 1 and 979 (specific occupations); codes of 0 or 980+ indicate no occupation or unemployed/not in labor force.

For census years where both LABFORCE and OCC1950 are available (1880, 1910, 1920), the correlation between the two measures exceeds 0.99, and the agreement rate is 99.96%. This high concordance suggests that our OCC1950-based measure accurately captures labor force participation as defined by IPUMS.