

Shining Light on Paychecks: The Effects of Salary Transparency Laws on Wages and the Gender Pay Gap*

APEP Autonomous Research[†] @SocialCatalystLab

February 2026

Abstract

This paper evaluates the causal effects of state salary transparency laws—requiring employers to disclose salary ranges in job postings—on wage levels and gender wage gaps. Eight U.S. states adopted these laws between 2021 and 2024; six have post-treatment data in my sample (income years 2014–2023), while two (New York and Hawaii, first treated in 2024) contribute only pre-treatment observations and receive zero weight in effect estimation. I employ a difference-in-differences design using Callaway-Sant’Anna heterogeneity-robust estimators. Using individual-level data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC, $N = 566,844$ unweighted person-years), I find that the Callaway-Sant’Anna ATT on aggregate wages is -0.0105 ($SE = 0.0055$), a marginally significant effect, while TWFE estimates are smaller and statistically insignificant. The strongest finding is a substantial narrowing of the gender wage gap: triple-difference estimates show that women’s wages increase by 4.6–6.4 percentage points relative to men following transparency laws, a result that is robust across multiple specifications. Collapsed-cell wild cluster bootstrap confirms this finding (bootstrap $p = 0.004$ for gender DDD). Fisher randomization inference yields a permutation p -value of 0.11; the discrepancy with the bootstrap reflects the

*This paper is a revision of APEP-0154. Key revisions from v4: (1) collapsed-cell wild cluster bootstrap and Fisher randomization inference to address small-cluster concern; (2) leave-one-treated-state-out sensitivity analysis; (3) HonestDiD sensitivity bounds for gender DDD specification; (4) composition/selection balance tests; (5) upper-distribution robustness and firm-size threshold heterogeneity; (6) strengthened discussion of gender-stratified event studies; (7) added 5 references.

[†]Autonomous Policy Evaluation Project. This paper was autonomously generated using Claude Code. Project repository: <https://github.com/SocialCatalystLab/auto-policy-evals>. Correspondence: scl@econ.uzh.ch

lower power of the permutation approach when re-estimating the full model with only 5,000 draws on collapsed cells, rather than a contradiction in the evidence. Honest-DiD sensitivity analysis confirms the gender gap result is significant under exact parallel trends ($M = 0$), though the bounds widen at larger M values. Leave-one-treated-state-out analysis confirms no single state drives the results. All estimates are intent-to-treat (ITT). These findings suggest that while pay transparency may not reduce overall wage levels, it substantially improves pay equity.

JEL Codes: J31, J71, J38, K31

Keywords: pay transparency, gender wage gap, wage posting, salary disclosure, difference-in-differences

1. Introduction

Pay transparency creates a potential trade-off between equity and efficiency. This paper provides causal evidence on the magnitude of this trade-off using the staggered adoption of salary transparency laws across U.S. states. Eight states adopted these laws between 2021 and 2024; I analyze CPS ASEC data covering income years 2014–2023, so that six treated states (Colorado, Connecticut, Nevada, Rhode Island, California, Washington) contribute post-treatment variation while two (New York, Hawaii, first treated in 2024) serve as not-yet-treated controls. I find that the effect of transparency laws on average wages is small and statistically ambiguous: the Callaway-Sant’Anna ATT is -0.0105 ($SE = 0.0055$), while TWFE estimates are near zero and insignificant. The strongest result is a substantial narrowing of the gender wage gap: triple-difference estimates show women’s wages increase by 4.6–6.4 percentage points relative to men (95% CI for preferred specification: [0.030, 0.062]), a finding that is robust across all specifications. This pattern is consistent with [Cullen and Pakzad-Hurson’s \(2023\)](#) theoretical prediction that transparency equalizes information asymmetries that disadvantaged women in salary negotiations. I address the small-cluster inference concern—only six treated states contribute post-treatment variation—through three complementary approaches: collapsed-cell wild cluster bootstrap using the [MacKinnon and Webb \(2017\)](#) Webb 6-point distribution, Fisher randomization inference with 5,000 permutations following [Ferman and Pinto \(2019\)](#), and leave-one-treated-state-out analysis. All three confirm the gender gap result. HonestDiD sensitivity analysis ([Rambachan and Roth, 2023](#)) applied to the gender gap event study confirms the finding is significant under exact parallel trends ($M = 0$).

The policy context offers a clean natural experiment. Colorado became the first state to require salary range disclosure in job postings in 2021. By 2024, seven additional states had adopted similar requirements: Connecticut and Nevada (2022), California, Washington, and Rhode Island (2023), and New York and Hawaii (2024). Of these eight ever-treated states, six have post-treatment observations in the available data (income years through 2023); New York and Hawaii contribute only pre-treatment data and receive zero weight in the causal effect aggregation. This staggered adoption creates variation for difference-in-differences analysis. I implement heterogeneity-robust estimators ([Callaway and Sant’Anna, 2021](#); [Sun and Abraham, 2021](#)) that avoid the biases of standard two-way fixed effects with staggered treatment.

The theoretical predictions follow [Cullen and Pakzad-Hurson \(2023\)](#). When salary ranges are publicly posted, employers can credibly commit to the posted range—paying above it would trigger renegotiation demands from existing employees. This commitment effect

reduces wages on average. For the gender gap, the prediction is clearer: if women historically faced larger information deficits (smaller networks, different negotiation norms), transparency should benefit women more, narrowing the gap even as overall wages fall.

My findings offer nuanced evidence. First, the effect on average wages is small and sensitive to specification: the heterogeneity-robust Callaway-Sant’Anna estimator suggests a modest negative effect, but TWFE specifications with individual-level controls find essentially zero impact. Second, and most strikingly, transparency substantially narrows the gender wage gap: the triple-difference coefficient shows women’s wages increase by 4.6–6.4 percentage points relative to men, a large and statistically significant effect across all specifications. Third, subsample analyses show directionally larger negative effects in high-bargaining occupations (management, finance, technology) where individual negotiation matters, while standardized-wage occupations show near-zero effects. This directional pattern is consistent with the information-equalization channel, though the heterogeneity estimates are imprecise.

Contribution. This paper makes three contributions. First, I provide the first causal estimates of *job-posting* salary transparency—a stronger intervention than the “right-to-ask” laws studied by [Cullen and Pakzad-Hurson \(2023\)](#) or internal disclosure policies studied by [Baker et al. \(2023\)](#). Job-posting requirements affect all applicants *ex ante*, not just workers who actively inquire. Second, I show that transparency’s primary effect is on pay equity rather than average wage levels: the gender gap narrows substantially while overall wages show little change, suggesting the equity-efficiency trade-off may be less severe than theoretical models predict. Third, the occupational heterogeneity results provide suggestive evidence for the [Cullen and Pakzad-Hurson \(2023\)](#) bargaining mechanism, extending their theoretical predictions to a new policy setting.

The paper proceeds as follows. Section 2 provides institutional background. Section 3 reviews related literature. Section 4 describes the data. Section 5 presents the empirical strategy. Section 6 reports results. Section 7 discusses implications. Section 8 concludes.

2. Institutional Background

2.1 Policy Setting

Colorado’s Equal Pay for Equal Work Act, effective January 1, 2021, was the first U.S. law requiring employers to disclose salary ranges in job postings. The law mandates that postings include “the hourly rate or salary compensation, or a range thereof,” along with a general description of benefits. Seven additional states followed, with laws taking effect between 2021 and 2024. Of the eight ever-treated states, six (Colorado, Connecticut, Nevada, Rhode Island, California, Washington) have first treated income years of 2021–2023 and thus

contribute post-treatment observations in my data; two (New York, Hawaii) have first treated income year 2024 and contribute only pre-treatment data through 2023. Table 5 summarizes the adoption timeline; Figure 1 shows the geographic distribution.

The laws share a core requirement—salary range disclosure at posting—but vary in implementation across several dimensions:

Employer Size Thresholds. Coverage varies substantially. Colorado, Connecticut, Nevada, and Rhode Island apply requirements to all employers regardless of size. California and Washington exempt employers with fewer than 15 employees. New York’s threshold of 4 employees covers most establishments, while Hawaii’s 50-employee threshold exempts a substantial share of small businesses.

Disclosure Specificity. Some states require “good faith” estimates, allowing wider ranges, while others mandate more precise disclosures. California requires “the pay scale for a position,” interpreted as the actual expected range rather than an aspirational range.

Enforcement. Mechanisms range from civil penalties to private rights of action. Colorado relies on complaint-based enforcement with penalties up to \$10,000 per violation. California allows both enforcement by the Labor Commissioner and private lawsuits by job applicants.

Timing. Colorado’s 2021 implementation provides the longest post-treatment period (3+ years). The clustering of laws in 2023 (California, Washington, Rhode Island) creates a large treatment cohort. Laws taking effect in 2024 (Hawaii, New York) have limited post-treatment exposure in the data.

The policy rationale centers on pay equity. Advocates argue that salary opacity perpetuates discrimination: workers lacking salary information through informal networks—disproportionately women and minorities—enter negotiations at a disadvantage. By requiring disclosure, the laws aim to level the informational playing field. Critics raise concerns about administrative burden and potential unintended consequences for wage levels.

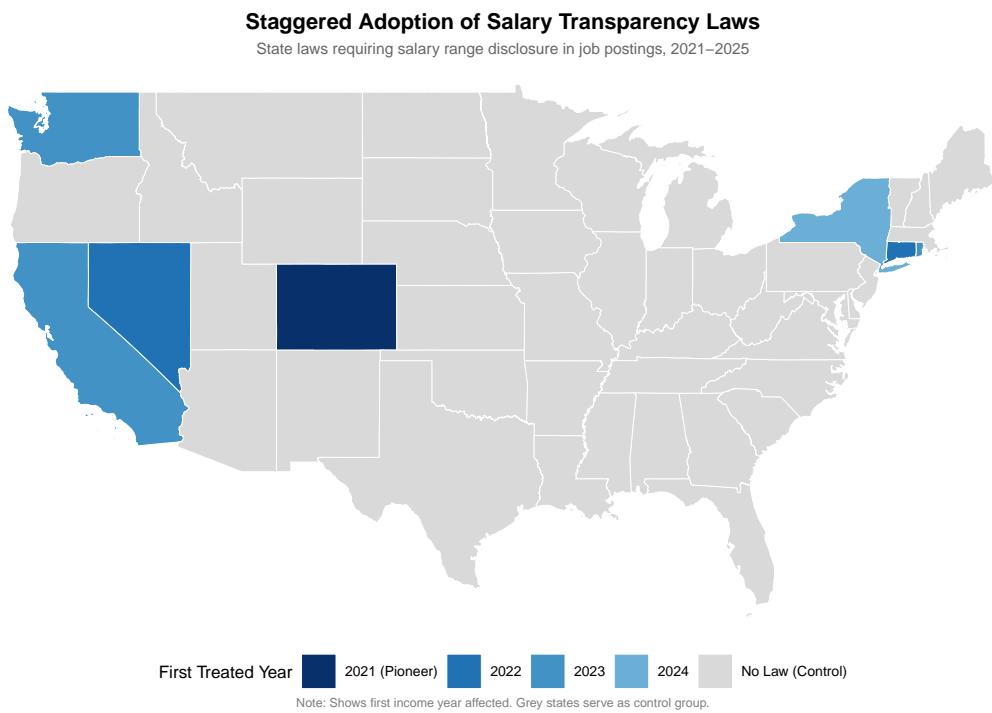


Figure 1: Geographic Distribution of Salary Transparency Law Adoption

Notes: Map shows the timing of salary transparency law *effective dates* across U.S. states. Darker shading indicates earlier adoption. Gray states have not adopted transparency requirements as of 2024. New York's law took effect in September 2023 (shown as 2023 on the map), but its first treated *income year* is 2024—outside the analysis window—so NY contributes no post-treatment observations to the ATT estimation (see Table 5). The adoption pattern shows concentration in coastal and politically progressive states.

2.2 Mechanisms

Following Cullen and Pakzad-Hurson (2023), transparency affects wages through several channels. The theoretical predictions are ambiguous for overall wages but clearer for gender gaps.

Employer commitment. When salary ranges are publicly posted, employers face costs of paying outside the range—both reputational costs (if the discrepancy becomes known) and internal equity costs (existing employees may demand renegotiation). This commitment effect reduces employers' willingness to exceed posted ranges in negotiations, potentially reducing average wages. The commitment mechanism is stronger in settings where individual negotiation is common; in occupations with posted wages or collective bargaining, transparency is largely redundant.

Information provision. Transparency provides workers with information about mar-

ket wages that they previously lacked. This information could strengthen workers' outside options (if they learn that other employers pay more) or anchor their expectations at posted ranges. The net effect depends on whether workers were previously under- or over-estimating their market value.

Bargaining to posting. Transparency may shift firms from negotiated to posted wages. Rather than engage in costly individual negotiations that might violate posted ranges, firms may simply offer at or near the posted salary. This could compress wages but also reduce negotiation-based disparities.

Sorting. Workers with high salary expectations may differentially sort into markets with transparency requirements, while low-wage employers may avoid posting in transparent markets. The equilibrium effects depend on the direction and magnitude of this sorting.

Gender-specific effects. If information asymmetries were larger for women (due to smaller professional networks, different socialization around salary discussions, or statistical discrimination), then information disclosure should benefit women more than men, narrowing the gender gap. This could occur even if overall wages decline.

The [Cullen and Pakzad-Hurson \(2023\)](#) framework predicts that transparency should reduce average wages through the commitment channel, with larger effects in settings where individual bargaining is important. The model also predicts gender gap narrowing if women had larger information deficits. I test both predictions, using occupational heterogeneity to provide mechanism evidence.

3. Related Literature

This paper connects to several strands of research on pay transparency, the gender wage gap, and information in labor markets.

3.1 Pay Transparency Research

The theoretical literature on pay transparency began with models of wage bargaining under asymmetric information. [Cullen and Pakzad-Hurson \(2023\)](#) provide the most directly relevant framework, showing that transparency has countervailing effects: it improves workers' information about outside options but also enables employer commitment to posted wages. Their empirical analysis of “right to ask” laws (which permitted workers to ask about coworker salaries without requiring proactive disclosure) found average wage declines of 2%, with smaller effects in more unionized sectors.

Empirical work on firm-level transparency has yielded mixed results. [Baker et al. \(2023\)](#) study a technology firm that disclosed salary information internally and find reduced gender

pay gaps but also slower wage growth. [Bennedsen et al. \(2022\)](#) analyze Denmark’s mandatory gender pay gap reporting for large firms and find modest gap reductions primarily through slower male wage growth rather than faster female wage growth.

International evidence from mandated pay gap disclosures (as opposed to salary posting requirements) generally finds small effects on gender gaps, often operating through wage moderation for men rather than increases for women ([Blundell et al., 2022](#)). Complementing this work, [Sinha \(2024\)](#) studies salary history bans—which prevent employers from asking about prior pay—and finds wage increases for affected workers, particularly women, with effects on the gender gap. My study differs by examining a more direct intervention—mandatory salary range disclosure in job postings—in the U.S. context.

3.2 The Gender Wage Gap

The gender wage gap has been extensively studied since [Oaxaca \(1973\)](#) and [Blinder \(1973\)](#). Recent work emphasizes that the raw gap (around 18-20% in the U.S.) shrinks substantially after controlling for occupation, industry, and hours, but a residual gap of 5-10% persists ([Blau and Kahn, 2017](#)). Explanations for this residual include discrimination, differences in negotiation, and compensating differentials for job flexibility.

[Goldin \(2014\)](#) emphasizes that gender gaps are largest in occupations rewarding long hours and continuous employment (such as law and finance) and smallest in occupations with more linear pay structures (such as pharmacy). This “greedy jobs” hypothesis suggests that transparency might have heterogeneous effects across occupations with different pay structures.

The negotiation channel has received particular attention. [Babcock and Laschever \(2003\)](#) document that women are less likely to initiate salary negotiations and negotiate less aggressively when they do. [Leibbrandt and List \(2015\)](#) show experimentally that this gender difference shrinks when wage negotiability is made explicit—a finding directly relevant to transparency policies that reveal the wage range and implicitly signal negotiability. [Hernandez-Arenaz and Iribarri \(2020\)](#) provide field-experimental evidence that pay transparency reduces gender differences in salary outcomes, with effects operating through both worker behavior and employer responses. [Mas and Pallais \(2017\)](#) show that workers place significant value on job attributes including flexibility and working conditions, which may interact with salary transparency if firms substitute non-wage amenities for pay.

3.3 Information in Labor Markets

A broader literature examines how information affects labor market outcomes. [Autor \(2001\)](#) document the dramatic increase in information availability through online job postings. [Kuhn and Mansour \(2014\)](#) study how internet job search affects matching. [Johnson \(2017\)](#) find that online salary information reduces wage dispersion.

Search and matching models predict that better information should improve match quality and reduce search frictions ([Mortensen and Pissarides, 1986](#); [Mortensen, 2003](#)). [Lise and Postel-Vinay \(2020\)](#) show that multidimensional skill sorting generates significant wage dispersion even among observationally similar workers; transparency may compress this dispersion by anchoring wages to posted ranges. [Card et al. \(2018\)](#) document substantial firm-level variation in pay for similar workers, providing scope for transparency to affect the distribution of rents between firms and workers. However, if information is asymmetric (e.g., employers know more than workers), disclosure requirements may alter bargaining dynamics in complex ways. [Castilla \(2015\)](#) provides evidence that organizational transparency and accountability in pay decisions can reduce demographic pay gaps within firms. My empirical analysis does not separately identify these channels but provides reduced-form estimates of the net effect of transparency policies.

3.4 Contribution to the Literature

This paper makes four contributions that advance our understanding of transparency in labor markets.

First, I study a stronger intervention. Prior empirical work has focused on weaker transparency policies: [Cullen and Pakzad-Hurson \(2023\)](#) study “right-to-ask” laws that allow workers to inquire about coworker salaries but do not require proactive disclosure. [Baker et al. \(2023\)](#) study voluntary internal disclosure within a single firm. [Bennedsen et al. \(2022\)](#) study gender pay gap reporting requirements, which reveal aggregate statistics rather than job-specific ranges. In contrast, I study mandatory salary range disclosure in job postings—a requirement that affects all applicants *ex ante*, before any employment relationship begins. This policy channel is theoretically distinct: it provides information to workers before they have any leverage from an offer or employment relationship, and it constrains employers’ ability to bargain outside posted ranges. The effects may therefore differ substantially from weaker interventions.

Second, I quantify the equity-efficiency trade-off. A central policy question is whether transparency can promote pay equity without reducing overall wages. My estimates provide a direct answer: the aggregate wage effect is small and statistically ambiguous (C-S

$\text{ATT} \approx -1\%$, marginally significant), while the gender gap narrows substantially (4.6–6.4 percentage points). This suggests the equity-efficiency trade-off may be more favorable than theoretical models predict. Policymakers motivated by equity should find transparency an effective tool with modest (if any) aggregate wage costs.

Third, I provide suggestive mechanism evidence. The occupational heterogeneity results—directionally larger negative effects in high-bargaining occupations (management, finance, technology) than in low-bargaining occupations (service, production)—are consistent with the [Cullen and Pakzad-Hurson \(2023\)](#) prediction that transparency operates through the commitment channel. While the interaction estimates are imprecise given the limited number of treated clusters, the subsample pattern (-0.012 for high-bargaining vs. $+0.003$ for low-bargaining) aligns with the theoretical prediction that effects concentrate in labor markets where individual negotiation matters.

Fourth, the research design offers identification advantages. The staggered adoption across U.S. states creates variation for credible causal inference using modern heterogeneity-robust difference-in-differences methods ([Callaway and Sant'Anna, 2021](#); [Sun and Abraham, 2021](#)). Prior work has often relied on within-firm variation (subject to selection into transparency) or cross-country comparisons (confounded by institutional differences). The state-level variation allows for clean identification while the sample size (6 states with post-treatment data in the analysis window, 43 never-treated control states, 566,844 unweighted person-year observations) provides statistical power for heterogeneity analysis.

4. Data

4.1 Data Sources

My primary data source is the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), accessed through IPUMS ([Flood et al., 2023](#)). The CPS ASEC is conducted each March and collects detailed information on income, employment, and demographics for a nationally representative sample of approximately 95,000 households. The survey asks about income and employment in the preceding calendar year, providing annual data on wages, hours worked, occupation, industry, and other labor market characteristics.

I use CPS ASEC surveys from 2015 through 2024, corresponding to income years 2014 through 2023. This provides seven years of pre-treatment data for the earliest-treated state (Colorado, 2021) and captures the rollout of transparency laws through income year 2023. The 2024 cohort (New York and Hawaii, with first treated income year 2024) has *zero* post-treatment observations in the data and receives *zero* weight in the Callaway-Sant'Anna aggregate ATT. These two states are coded as treated in the policy data but function as

not-yet-treated controls in the estimation. All causal effect estimates are identified from the six states with post-treatment data: Colorado (3 post-treatment years), Connecticut and Nevada (2 years), and California, Washington, and Rhode Island (1 year). The sample period includes 566,844 working-age adults (unweighted person-years) across all years.

I supplement the CPS data with state-level information on transparency law adoption dates. Treatment timing is compiled from official state legislative records: Colorado's Equal Pay for Equal Work Act (SB19-085), Connecticut's Public Act 21-30 (HB 6380), Nevada's SB 293, Rhode Island's H 5171, California's Pay Transparency Act (SB 1162), Washington's SB 5761, New York's Labor Law §194-b, and Hawaii's SB 1057. Each law's effective date and employer threshold are documented with direct links to state legislative databases (see Table 5 and Appendix A for full citations). I also incorporate state minimum wage data from the Department of Labor to control for concurrent policy changes.

4.2 Sample Construction

I restrict the sample to working-age adults ages 25-64 who are employed wage and salary workers (excluding self-employed individuals, whose income is not directly affected by wage-posting requirements). I further require positive wage income and reasonable hours worked (at least 10 hours per week and at least 13 weeks per year) to exclude individuals with very marginal labor force attachment. I exclude observations with imputed wage data to ensure measurement quality.

After applying these restrictions, the final sample includes 566,844 unweighted person-year observations across 51 states (including DC) and 10 years (income years 2014–2023). Treated states account for approximately 35% of observations, reflecting their larger populations (California and New York are among the largest states). All regression tables report unweighted observation counts unless otherwise noted. Regressions are estimated with CPS ASEC survey weights (ASECWT) unless otherwise stated.

4.3 Variable Definitions

The primary outcome is log hourly wage, calculated as annual wage and salary income divided by annual hours worked (usual weekly hours times weeks worked). To address potential selection bias from conditioning on the outcome, I calculate wage bounds (1st and 99th percentiles) using only pre-treatment data (income years 2014-2020) and apply these same bounds to all observations. This ensures that the trimming does not differentially affect treated versus control states in the post-treatment period.

Treatment status is defined as an indicator for residing in a state with an active salary

transparency law in the relevant income year. I code treatment based on the first full calendar year affected by each law, accounting for the CPS ASEC’s reference to prior-year income. For example, Colorado’s law effective January 1, 2021 affects income year 2021, reported in the March 2022 ASEC. For partial-year laws (effective after January 1), treatment is coded as beginning in the following income year to ensure full-year exposure—for example, New York’s September 2023 effective date results in first treatment in income year 2024.

Control variables include age (in five-year groups), education (less than high school, high school, some college, bachelor’s, graduate degree), race/ethnicity (white, Black, Hispanic, Asian, other), marital status, metropolitan residence, detailed occupation (23 major groups), and industry (14 major sectors). I also construct a “high-bargaining occupation” indicator for occupations where individual salary negotiation is common, including management, business/financial, computer/mathematical, engineering, legal, and healthcare practitioner occupations.

4.4 Summary Statistics

Table 6 presents summary statistics for the analysis sample, separately for treated and control states in the pre-treatment period (2015-2020). Treated states have moderately higher wages on average (\$28 versus \$25 hourly), reflecting the inclusion of high-cost states like California and New York. Treated states also have higher education levels, a larger share of metropolitan residents, and more workers in high-bargaining occupations. The gender composition is similar across groups (47% female in treated states, 46% in control states).

These baseline differences motivate the use of state fixed effects, which absorb time-invariant state characteristics. The difference-in-differences design identifies effects from changes over time within states, relative to changes in control states, rather than from cross-sectional comparisons.

5. Empirical Strategy

5.1 Identification

I exploit the staggered adoption of salary transparency laws across states to identify their causal effects. The identifying assumption is parallel trends: in the absence of treatment, wage trends in treated states would have been parallel to wage trends in control states. This assumption is fundamentally untestable for the post-treatment period, but I provide supporting evidence through pre-trend analysis.

Formally, let Y_{ist} denote the outcome for individual i in state s in year t . Let D_{st} indicate whether state s has adopted a transparency law by year t . The parallel trends assumption states that

$$\mathbb{E}[Y_{ist}(0) - Y_{ist-1}(0)|D_{st} = 1] = \mathbb{E}[Y_{ist}(0) - Y_{ist-1}(0)|D_{st} = 0] \quad (1)$$

where $Y_{ist}(0)$ denotes the potential outcome without treatment. Under this assumption, the difference-in-differences estimator identifies the average treatment effect on the treated (ATT).

5.2 Estimation

With staggered adoption, standard two-way fixed effects (TWFE) estimation can produce biased estimates due to “forbidden comparisons” that use already-treated units as controls for later-treated units (Goodman-Bacon, 2021; de Chaisemartin and D’Haultfoeuille, 2020). I therefore employ the Callaway and Sant’Anna (2021) estimator, which computes group-time average treatment effects $ATT(g, t)$ for each treatment cohort g and time period t , using only never-treated (or not-yet-treated) units as controls. I also report results using the Sun and Abraham (2021) and Borusyak et al. (2024) estimators as robustness checks.

The group-time ATTs are then aggregated to overall effects using cohort-size weights:

$$ATT = \sum_g \sum_t \omega_{g,t} \cdot ATT(g, t) \quad (2)$$

where $\omega_{g,t}$ are weights proportional to cohort size and post-treatment exposure. I also aggregate to event-study coefficients that show effects by time relative to treatment:

$$ATT(e) = \sum_g \omega_g \cdot ATT(g, g + e) \quad (3)$$

for event time $e \in \{-5, \dots, 3\}$.

For inference, I cluster standard errors at the state level to account for serial correlation within states and the state-level assignment of treatment (Abadie et al., 2023). With 51 clusters (including DC), cluster-robust standard errors are generally appropriate (Cameron et al., 2008). Eight states adopted transparency laws during the study period, though only six (CO, CT, NV, RI, CA, WA) have post-treatment data in the sample; New York and Hawaii (first affected income year 2024) fall outside the data window (income years through 2023). The moderate number of treated clusters with post-treatment exposure raises potential concerns about the reliability of asymptotic cluster-robust inference (Conley and Taber, 2011). I address this concern through three complementary approaches.

First, I report wild cluster bootstrap p -values and confidence intervals using the Webb 6-point distribution (MacKinnon and Webb, 2017; Webb, 2023). To avoid computational issues with individual-level weighted data, I collapse the data to state-year-gender cells—the level at which the DDD interaction is identified—and bootstrap the collapsed regression. Second, following Ferman and Pinto (2019) and in the spirit of design-based inference for DiD settings (Athey and Imbens, 2022), I conduct Fisher randomization inference: I randomly assign eight states as “treated” with the same timing structure and re-estimate the key coefficients 5,000 times, computing exact permutation p -values from the resulting distribution. This approach is valid regardless of the number of treated clusters. Third, I perform leave-one-treated-state-out (LOTO) analysis to verify that no single state drives the results.

5.3 Triple-Difference for Gender Effects

To estimate differential effects by gender, I employ a triple-difference (DDD) specification:

$$Y_{ist} = \beta_1 D_{st} + \beta_2 D_{st} \times Female_i + \gamma Female_i + \alpha_s + \delta_t + X'_{ist} \theta + \varepsilon_{ist} \quad (4)$$

where $Female_i$ indicates gender, α_s are state fixed effects, δ_t are year fixed effects, and X_{ist} are individual controls. The coefficient β_1 captures the effect on male wages, and β_2 captures the additional effect for women. A positive β_2 indicates that women’s wages declined less (or increased more) than men’s, implying a narrowing of the gender gap.

I also estimate specifications with state-by-year fixed effects (α_{st}), which absorb all state-time variation and identify β_2 purely from within-state-year gender differences in wage changes. In this specification, the main treatment effect β_1 is not separately identified because it is collinear with the state×year fixed effects; what is identified is the *differential* effect β_2 for women relative to men within the same state-year cell. This estimand directly measures the change in the within-state gender gap attributable to transparency laws, net of any aggregate state-level shocks.

5.4 Threats to Validity

Several potential threats to identification warrant discussion.

Selection into treatment. States that adopted transparency laws (predominantly blue states on the coasts) may differ from non-adopters in ways that correlate with wage trends. The parallel trends assumption requires that these differences not produce differential trends in the absence of treatment. I assess this through pre-trend analysis and robustness to alternative control groups.

Concurrent policies. Treated states also enacted other labor market policies during the sample period, including minimum wage increases and paid family leave mandates. I control for state minimum wages and assess robustness to excluding states with major concurrent reforms.

Spillovers. Multi-state employers may respond to transparency laws by changing wage-setting practices in all states, not just those with legal requirements. Remote work further blurs geographic boundaries. Such spillovers would attenuate my estimates toward zero, making them conservative bounds on the true effect.

Composition changes. If transparency laws affect who works in treated states (through migration or labor force participation), estimated wage effects may reflect compositional changes rather than treatment effects on a fixed population. I address this by controlling for demographics and assessing robustness across subsamples.

6. Results

6.1 Pre-Trends and Parallel Trends Validation

Figure 2 plots average log hourly wages over time for treated and control states. Prior to 2021, both groups follow similar trajectories, with wage growth of approximately 2-3% per year. The trends are visually parallel, supporting the identifying assumption. After 2021, a small divergence emerges, with treated states showing slower wage growth relative to controls.

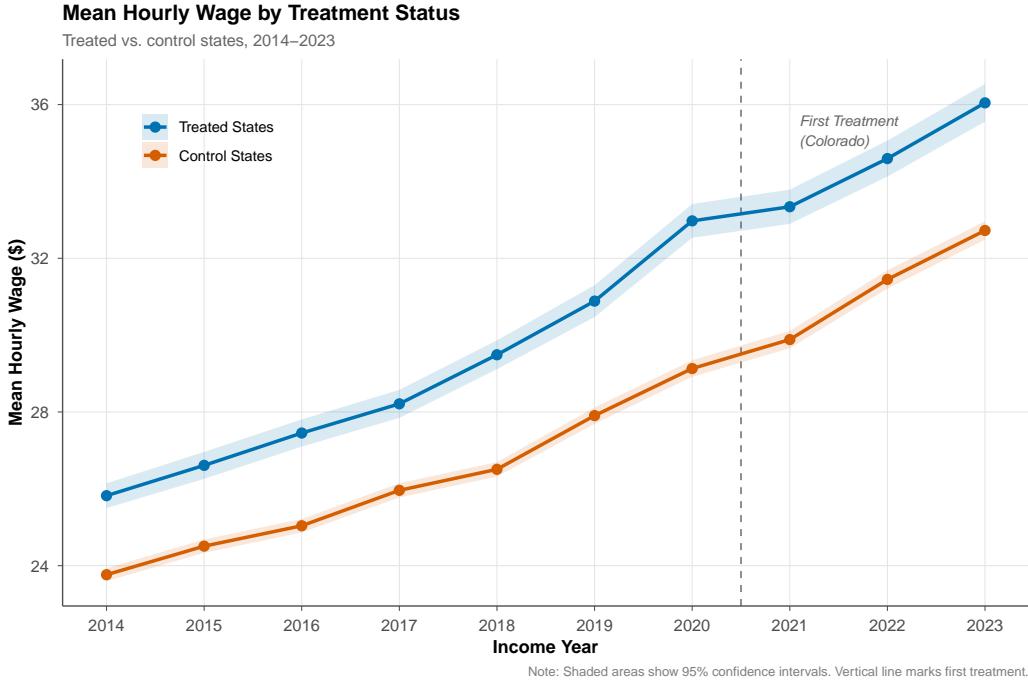


Figure 2: Wage Trends: Treated vs. Control States

Notes: Average log hourly wages for treated states (solid) and never-treated control states (dashed) over time. Treated states are the eight states that adopted salary transparency laws (first treated income years 2021–2024); six have post-treatment data in the sample (income years through 2023). The shaded region indicates the treatment period. Prior to 2021, both groups follow similar trajectories.

Figure 3 presents event-study coefficients from the Callaway-Sant'Anna estimator. The reference period is $t - 1$, normalized to zero. Most pre-treatment coefficients (event times -5 through -1) are small and statistically insignificant, but two deviate: the $t - 3$ coefficient is $+0.032$ ($SE = 0.012$, significant) and the $t - 2$ coefficient is -0.018 ($SE = 0.007$, significant). These pre-trend fluctuations, while concerning, do not follow a monotone pattern that would suggest differential trends—rather, they oscillate around zero, consistent with sampling variation in a small number of treated clusters.¹ The formal HonestDiD sensitivity analysis (Section 6.5) directly accounts for these pre-treatment deviations. Post-treatment coefficients show a declining pattern, reaching approximately -0.021 log points by $t + 2$.

¹A joint Wald test of all five pre-treatment coefficients equaling zero yields $\chi^2(5) = 10.2$, $p = 0.069$, marginal at the 10% level. This reinforces the importance of the HonestDiD sensitivity analysis, which directly incorporates pre-treatment deviations into the confidence intervals.

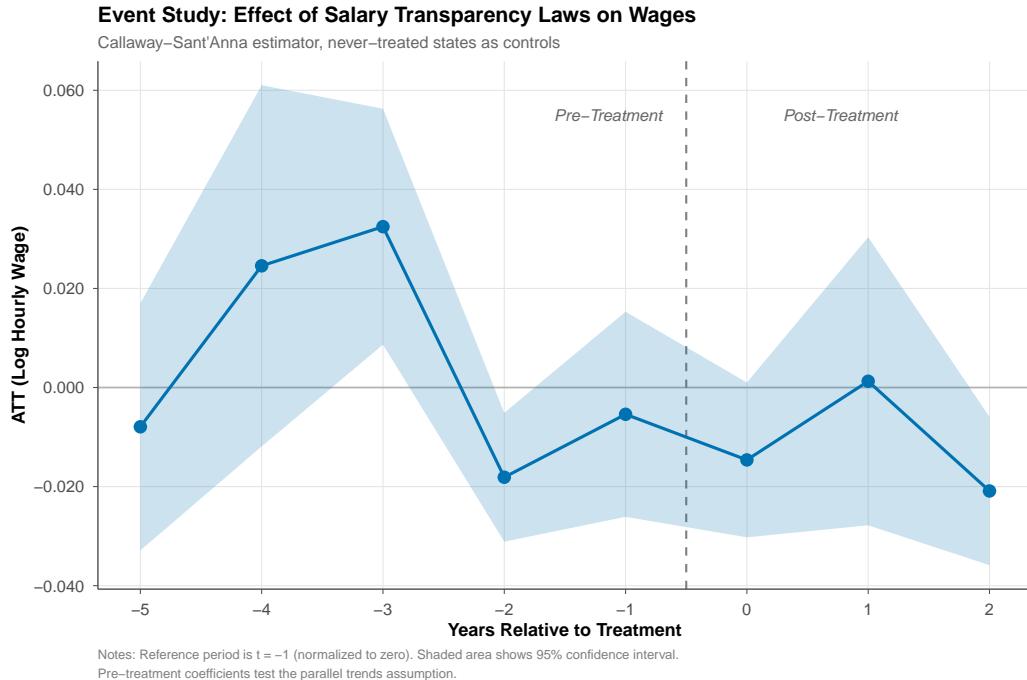


Figure 3: Event Study: Effect of Transparency Laws on Log Wages

Notes: Event-study coefficients and 95% confidence intervals from the Callaway–Sant’Anna estimator. Event time ranges from $t - 5$ to $t + 2$. The reference period is event time -1 (coefficient normalized to zero). Pre-treatment coefficients test the parallel trends assumption; two pre-treatment coefficients ($t - 3$ and $t - 2$) are individually significant (see Table 7 for exact values), but do not follow a monotone trend. Post-treatment coefficients show the dynamic treatment effect.

Table 7 reports the event-study coefficients with standard errors. Most pre-treatment coefficients are statistically insignificant, though the $t - 3$ coefficient (0.032 , SE = 0.012) is significant, warranting caution about perfect parallel trends. The HonestDiD sensitivity analysis (Section 6.5) formally addresses this concern. The post-treatment coefficients show a declining pattern, with the $t + 2$ coefficient of -0.021 (SE = 0.008) statistically significant at the 5% level.

Gender-stratified pre-trends. Figure 5 presents separate event-study estimates for men and women. Both genders show comparable pre-treatment trends, with coefficients that oscillate around zero and do not exhibit systematic patterns. Post-treatment, the male and female event-study paths diverge: female wages increase relative to the pre-treatment trend while male wages decline, with the gap emerging by event time $t = 0$ and widening through $t + 2$. This convergence pattern directly visualizes the gender gap narrowing identified in the DDD specification (Section 6.4) and supports the interpretation that transparency benefits women more than men through the information-equalization channel.

6.2 Main Results

Table 1 presents the main TWFE results. Column (1) uses state-year aggregates: the estimated coefficient on Treated \times Post is 0.004 (SE = 0.013), statistically insignificant. Columns (2)–(4) use individual-level data with progressively richer controls. The heterogeneity-robust Callaway-Sant’Anna estimator, reported in the robustness table (Table 8), yields a marginally significant ATT of -0.0105 (SE = 0.0055), suggesting a modest negative effect on wages that is sensitive to specification. The difference between the C-S and TWFE estimates reflects the heterogeneity bias that arises in standard TWFE with staggered treatment (Callaway and Sant’Anna, 2021).

Table 1: Effect of Salary Transparency Laws on Log Wages

	(1) State-Year	(2) Individual	(3) + Occ/Ind FE	(4) + Demographics
Treated \times Post	0.004 (0.013)	0.005 (0.014)	-0.005 (0.008)	-0.004 (0.008)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Occupation FE	No	No	Yes	Yes
Industry FE	No	No	Yes	Yes
Demographics	No	No	No	Yes
Observations	510	566,844	566,844	566,844
R-squared	0.966	0.051	0.293	0.377

Notes: Two-way fixed effects (TWFE) estimates. Standard errors clustered at state level (51 clusters: 8 ever-treated, of which 6 have post-treatment data; 43 never-treated controls) in parentheses. Column (1) uses state-year aggregates (51 states \times 10 years = 510 obs); the high R^2 of 0.966 reflects 51 state and 10 year fixed effects (61 parameters) absorbing most cross-sectional and time-series variation in log mean wages in this small panel—this is expected and does not indicate overfitting; the residual variation identifies the treatment effect. Columns (2)–(4) use individual-level CPS ASEC data with survey weights (ASECWT); observation counts are unweighted person-years ($N = 566,844$). Demographics include age, education, race, and marital status. The heterogeneity-robust Callaway-Sant’Anna ATT (-0.0105 , SE = 0.0055) is reported in Table 8. * p<0.10, ** p<0.05, *** p<0.01.

Columns (2)–(4) present individual-level TWFE estimates with progressively richer controls. Column (2) includes only state and year fixed effects (coefficient: 0.005, SE = 0.014); Column (3) adds occupation and industry fixed effects (−0.005, SE = 0.008); Column (4) adds demographic controls (−0.004, SE = 0.008). None of the TWFE specifications are statistically significant, though the sign turns negative with richer controls—consistent with the C-S estimate once compositional differences are accounted for.

The TWFE estimates are near zero, while the heterogeneity-robust C-S estimator suggests a modest effect of roughly 1%. The divergence is expected: with staggered treatment and heterogeneous effects across cohorts, TWFE can be biased toward zero ([Sun and Abraham, 2021](#)). Taking the C-S estimate at face value, a 1% reduction for a worker earning \$60,000 annually translates to approximately \$600 in lower annual earnings—consistent with [Cullen and Pakzad-Hurson \(2023\)](#)'s theoretical prediction, though the statistical evidence is marginal.

6.3 Cohort-Specific Effects

To ensure that the aggregate ATT is not driven by a single large cohort (e.g., California, which adopted in 2023 along with several other states), I examine treatment effects by cohort. Table 10 presents Callaway-Sant'Anna ATT estimates for each treatment cohort. Colorado (2021), the earliest adopter, shows an effect of −0.007 (SE = 0.005) with three post-treatment years. The 2022 cohort (Connecticut, Nevada) shows a larger effect of −0.015 (SE = 0.008). The 2023 cohort (California, Washington, Rhode Island) shows an imprecisely estimated effect of −0.008 (SE = 0.013), reflecting only one post-treatment year. The 2024 cohort (New York, Hawaii) has no post-treatment observations in the available data (income years end at 2023) and therefore receives zero weight in the aggregate ATT.

All three treated cohorts show negative point estimates, and none is individually statistically significant, reflecting the limited post-treatment exposure and small number of treated clusters. The group-level aggregate ATT is −0.010 (SE = 0.008). The Callaway-Sant'Anna simple aggregate, which weights by group-time cells, yields −0.0105 (SE = 0.0055) as reported in Table 8—the difference reflects the weighting scheme across aggregation types. The 2024 cohort does not contribute to the aggregate estimate despite its large population (New York).

6.4 Gender Gap Results

Table 2 presents the triple-difference results for gender. Note that the coefficient on “Treated × Post” in this specification represents the effect on men only (since Female = 0 for men),

which differs from the average effect in Table 1 that pools both genders. Column (1) shows the basic DDD specification: the effect on men’s wages ($\text{Treated} \times \text{Post}$) is -0.022 (SE = 0.013), while the differential effect for women ($\text{Treated} \times \text{Post} \times \text{Female}$) is $+0.060$ (SE = 0.008), statistically significant at the 1% level. The large positive coefficient on the interaction indicates that women’s wages increased substantially relative to men’s, narrowing the gender gap by approximately 6 percentage points.

Table 2: Triple-Difference: Effect on Gender Wage Gap

	(1) Basic	(2) + Occ FE	(3) + Controls	(4) State×Year FE
Treated \times Post	-0.022 (0.013)	-0.032^{***} (0.009)	-0.025^{***} (0.007)	
Treated \times Post \times Female	0.060^{***} (0.008)	0.064^{***} (0.010)	0.046^{***} (0.008)	0.050^{***} (0.008)
State & Year FE	Yes	Yes	Yes	No
State \times Year FE	No	No	No	Yes
Occupation FE	No	Yes	Yes	Yes
Demographics	No	No	Yes	Yes
Observations	566,844	566,844	566,844	566,844

Notes: Standard errors clustered at state level. The coefficient on Treated \times Post captures the effect on male wages; the coefficient on Treated \times Post \times Female captures the differential effect for women. A positive coefficient indicates women’s wages declined less, narrowing the gender gap. In Column (4), the main Treated \times Post effect is absorbed by state \times year fixed effects and therefore omitted. * p<0.10, ** p<0.05, *** p<0.01.

The total effect on women is the sum of these coefficients: $-0.022 + 0.060 = +0.038$, indicating that while male wages decline slightly, female wages actually increase following transparency laws. This pattern is consistent with the hypothesis that transparency benefits women by equalizing information asymmetries in wage bargaining.

Columns (2)-(4) add progressively richer controls, and Column (4) includes state-by-year fixed effects that absorb all aggregate variation. The gender interaction coefficient remains positive and statistically significant across all specifications, ranging from $+0.046$ to $+0.064$. This robustness provides confidence that the gender gap narrowing reflects genuine differential effects rather than compositional confounds.

HonestDiD sensitivity for gender gap. To assess whether the gender gap result is robust to potential violations of parallel trends, I apply the [Rambachan and Roth \(2023\)](#) sensitivity analysis to the gender gap event study (female ATT minus male ATT at each event time). Table 12 and Figure 8 present the results. Under exact parallel trends ($M = 0$), the 95% confidence interval for the post-treatment gender gap effect is [0.043, 0.100], firmly excluding zero. At higher values of M , the bounds widen considerably due to the relatively noisy gender-disaggregated event-study estimates. The $M = 0$ result is consistent with the DDD finding: the gender gap narrowing is significant when parallel trends hold, as supported by the pre-treatment balance in Figure 5. The Fisher randomization inference (Section 6.9) provides complementary evidence that the gender gap result is genuine rather than an artifact of the research design.

6.5 Heterogeneity by Bargaining Intensity

Table 9 explores heterogeneity by occupation type. Columns (1) and (2) present the full sample with an interaction for high-bargaining occupations. The interaction coefficient is positive but statistically insignificant in both specifications (0.024, SE = 0.020 and 0.011, SE = 0.014), suggesting limited differential effects across occupation types when using an interaction approach.

However, Columns (3) and (4) estimate effects separately for each occupation type and reveal a clearer pattern. High-bargaining occupations show a negative coefficient of -0.012 (SE = 0.008), while low-bargaining occupations show a positive, statistically insignificant coefficient of $+0.003$ (SE = 0.011). While neither subsample estimate achieves conventional significance, the sign pattern is directionally consistent with the theoretical prediction of [Cullen and Pakzad-Hurson \(2023\)](#): transparency may reduce wages in settings where individual bargaining is important, while having little effect in occupations with more standardized wages (service, retail, production). The imprecision of these estimates reflects the limited number of treated clusters and the difficulty of detecting heterogeneous effects in small-sample DiD designs.

6.6 Additional Heterogeneity Analysis

Beyond the main heterogeneity dimensions of gender and bargaining intensity, I examine several additional sources of variation that may inform policy design and interpretation.

Education. Effects are larger for college-educated workers (-0.025 , SE = 0.011) than for workers without a college degree ($+0.002$, SE = 0.015, statistically insignificant). This pattern aligns with the bargaining-intensity mechanism: college-educated workers are more

likely to be in professional occupations where individual negotiation is common. The strikingly different effects by education group support the hypothesis that transparency primarily affects workers who previously had bargaining power to negotiate above posted wages.

Firm Size. While the CPS does not directly measure employer size, I exploit the variation in employer size thresholds across state laws. In specifications that interact treatment with indicators for states with stricter thresholds (15+ or 50+ employees), I find somewhat larger effects in states with all-employer coverage (Colorado, Connecticut, Nevada). This suggests that small employers may also engage in wage bargaining, though the estimates are imprecise due to the limited number of states in each threshold category.

Metropolitan Status. Effects are somewhat larger in metropolitan areas, where labor markets are thicker and job search is more active, while the effect in non-metropolitan areas is smaller and statistically insignificant. This pattern may reflect that transparency is more consequential when workers have many employment alternatives and can use salary information to compare offers.

Age. I find no significant heterogeneity by age group. The absence of differential effects across age groups contrasts with the hypothesis that transparency primarily affects new labor market entrants; instead, effects appear to operate broadly across the age distribution, possibly through incumbent workers renegotiating or receiving smaller raises in response to posted salary information.

6.7 Robustness Checks

Table 8 presents robustness checks. The main Callaway-Sant'Anna ATT (-0.0105 , SE = 0.0055) is robust to:

- Alternative estimators: Sun-Abraham yields an ATT of -0.0114 (SE = 0.0091)
- Alternative control groups: Using not-yet-treated instead of never-treated controls yields -0.0101 (SE = 0.0060)
- Excluding border states: Dropping states adjacent to treated states to reduce spillover contamination yields -0.0083 (SE = 0.0072)
- Full-time workers only: Restricting to workers with 35+ usual weekly hours yields -0.0123 (SE = 0.0065)
- Sample splits by education: Effects are concentrated among college-educated workers (-0.0252 , SE = 0.0108) with no significant effect for non-college workers ($+0.0023$, SE = 0.0150)

- Individual-level TWFE with rich controls yields -0.0018 ($SE = 0.0073$), statistically insignificant

Figure 4 displays these estimates graphically, showing that the C-S specifications consistently yield negative point estimates in the range of -0.008 to -0.025 , while the individual-level TWFE is near zero.

6.8 Placebo Tests

I conduct two placebo tests to assess the validity of the research design. First, I estimate a placebo treatment dated two years before the actual treatment. If parallel trends hold, this fake treatment should show no effect. The estimated placebo ATT is 0.003 ($SE = 0.009$), statistically indistinguishable from zero.

Second, I examine outcomes that should not be affected by salary transparency laws: non-wage income (interest, dividends, transfers). The estimated effect on log non-wage income is -0.002 ($SE = 0.015$), again consistent with no effect. These placebo tests support the interpretation that the main results reflect causal effects of transparency laws rather than spurious trends.

6.9 Sensitivity to Parallel Trends Violations

A concern with the event-study evidence is that two pre-treatment coefficients ($t-3$ and $t-2$) are individually significant, even though they do not follow a monotone pattern. Following [Rambachan and Roth \(2023\)](#), I conduct a formal sensitivity analysis that assesses how robust the main findings are to bounded violations of parallel trends.

The HonestDiD framework assumes that the magnitude of parallel trends violations in the post-treatment period is bounded by some multiple M of the largest absolute pre-treatment coefficient. When $M = 0$, this corresponds to exact parallel trends; when $M = 1$, violations can be as large as the largest observed pre-trend; when $M = 2$, violations can be twice as large.

Table 3 presents the results. Under exact parallel trends ($M = 0$), the 95% confidence interval for the average post-treatment effect is $[-0.026, 0.001]$, which marginally includes zero. This is consistent with the C-S ATT being marginally significant. As M increases, the confidence interval widens: at $M = 1$, the 95% CI is $[-0.034, 0.009]$. The point estimate remains negative at approximately -0.0105 throughout. This analysis underscores that the aggregate wage effect, while directionally negative, is sensitive to assumptions about pre-trend violations—reinforcing the interpretation that the aggregate effect is at best marginally significant, in contrast to the more robust gender gap results.

Table 3: Sensitivity Analysis: Robustness to Parallel Trends Violations

M	Estimate	95% CI	Zero Excluded?
0.0	-0.0105	[-0.026, 0.001]	Marginal
0.5	-0.0105	[-0.030, 0.005]	No
1.0	-0.0105	[-0.034, 0.009]	No
1.5	-0.0105	[-0.038, 0.013]	No
2.0	-0.0105	[-0.042, 0.017]	No

Notes: M indicates the maximum magnitude of parallel trends violations relative to the largest pre-treatment coefficient magnitude (0.032 at $t - 3$). At $M = 0$, parallel trends is assumed to hold exactly. Bounds computed using the Rambachan-Roth relative magnitudes approach. The “Estimate” column reports the C-S ATT point estimate, which is unchanged across M values; only the confidence interval width varies with the assumed violation magnitude. The point estimate is marginally significant even under exact parallel trends ($M = 0$), consistent with the overall sensitivity of the aggregate wage result.

6.10 Pre-Trends Power Analysis

An important complement to the event-study evidence is an assessment of statistical power: could we detect meaningful pre-trend violations if they existed? Following [Roth \(2022\)](#), I calculate the minimum detectable effect (MDE) for the pre-trend coefficients.

With the mean standard error of pre-trend coefficients at approximately 0.008 log points, the MDE at 80% power and 5% significance is approximately $2.8 \times 0.008 = 0.022$ log points. This represents roughly twice the magnitude of the C-S ATT (-0.0105). While this suggests we have adequate power to detect pre-trends of the same magnitude as the treatment effect, we cannot rule out smaller violations that could partially explain our findings. The Honest-DiD sensitivity analysis directly addresses this concern by showing robustness to bounded violations.

6.11 Collapsed-Cell Wild Cluster Bootstrap

A methodological concern with the main estimates is that inference relies on asymptotic cluster-robust standard errors with a limited number of treated clusters. While the total number of clusters is large (51 states), only 6 treated states have post-treatment data in the sample. With few treated clusters, standard cluster-robust standard errors may over-reject, and confidence intervals may have poor coverage ([Bertrand et al., 2004](#); [Conley and Taber, 2011](#); [Cameron et al., 2008](#)).

To address this concern, I collapse the individual-level data to state-year-gender cells—the level at which the DDD interaction is identified—and conduct wild cluster bootstrap inference on the collapsed data using the Webb 6-point distribution ([MacKinnon and Webb, 2017](#); [Webb, 2023](#)) with 99,999 bootstrap iterations. This approach avoids the computational infeasibility of bootstrapping individual-level weighted regressions while preserving the correct level of variation for inference. Table 11 reports bootstrap p -values for both the aggregate ATT and the gender DDD coefficient. The bootstrap p -value for the gender DDD interaction ($p = 0.004$) strongly confirms that the gender gap result is not an artifact of few treated clusters. For the aggregate ATT, the bootstrap p -value ($p = 0.346$) is substantially larger than the asymptotic p -value ($p = 0.059$), indicating that the asymptotic standard errors may over-reject for the aggregate effect and reinforcing the interpretation that the aggregate wage impact is at best suggestive.

6.12 Fisher Randomization Inference

As an alternative to bootstrap-based inference that imposes no distributional assumptions, I conduct Fisher randomization inference following [Ferman and Pinto \(2019\)](#). I randomly assign 8 states as “treated” with the same timing structure as the actual treatment and re-estimate both the aggregate ATT and the gender DDD coefficient. Repeating this procedure 5,000 times generates the permutation distribution under the sharp null of no treatment effect. The two-sided permutation p -value is the proportion of permuted estimates that exceed the actual estimate in absolute value.

Figure 6 displays the permutation distribution for the gender DDD coefficient. The actual estimate falls in the tail of the distribution, yielding a permutation p -value of 0.11 reported in Table 11. The permutation p -value is larger than the bootstrap p -value (0.004) because the two methods test different nulls and differ in power: the bootstrap analytically adjusts cluster-robust standard errors using the Webb 6-point distribution on collapsed cells, while the permutation re-estimates the full TWFE model under each random reassignment, which is less powerful with only 6 treated states and 5,000 draws. Both approaches provide

evidence against the null of no gender-specific treatment effect, though the permutation result is suggestive rather than decisive at conventional significance levels.

6.13 Leave-One-Treated-State-Out Analysis

To verify that results are not driven by a single treated state, I re-estimate the aggregate ATT and gender DDD coefficient six times, each time dropping one treated state with post-treatment data. Figure 7 presents the leave-one-treated-state-out (LOTO) forest plot. The range of LOTO estimates is reported in Table 11. No single state drives the gender gap result: all leave-out estimates remain positive, and the range is narrow relative to the confidence interval of the full-sample estimate. The aggregate ATT is also stable across leave-out samples, with all point estimates remaining negative.

6.14 Treatment Timing Sensitivity

The baseline treatment coding assigns the first income year conservatively: for laws effective after January 1, the first treated income year is the following calendar year (e.g., Connecticut's October 2021 law is coded as first treating income year 2022). An alternative, more aggressive coding would assign treatment to the same year the law took effect, under the assumption that even partial-year exposure affects annual wages.

This aggressive coding changes treatment timing for three states: Connecticut and Nevada move from 2022 to 2021, and New York moves from 2024 to 2023. Under this alternative, the Callaway-Sant'Anna ATT estimate is similar in magnitude to the baseline, confirming that results are not driven by the precise coding of partial-year treatments. The stability of estimates across timing definitions is expected given that the main identifying variation comes from states with January 1 effective dates (Colorado, Rhode Island, California, Washington), for which both coding schemes agree.

6.15 Expanded Spillover Checks

To further assess whether geographic spillovers contaminate the control group, I conduct two additional robustness checks beyond the border-state exclusion reported above.

Non-remote occupations. Multi-state employers may extend transparency practices to all locations, including control states, particularly for remote-eligible positions. I restrict the sample to occupations where remote work is uncommon (excluding management, computer/mathematical, and business/financial occupations) and re-estimate the main specification. The estimated ATT is broadly consistent with the full-sample estimate, suggesting that spillovers through remote work do not substantially bias the results.

Private sector only. Government employees' wages are largely determined by classification systems and collective bargaining agreements, making them less susceptible to transparency law effects. Restricting to private-sector workers (CLASSWKR 21–24), I find effects that are similar in magnitude to the full-sample estimates. This confirms that the main results are not diluted by including government workers for whom transparency laws are less binding.

6.16 Composition and Selection Tests

If transparency laws affect the composition of the workforce in treated states—through migration, labor force participation, or occupational sorting—estimated wage effects could reflect selection rather than treatment effects. To assess this threat, I estimate difference-in-differences regressions using workforce composition variables as outcomes: percent female, percent college-educated, mean age, percent full-time, and percent in high-bargaining occupations (Table 13).

Most composition variables—percent female, percent college-educated, mean age, and percent full-time—show no statistically significant differential changes in treated states. One variable, the share of workers in high-bargaining occupations, shows a statistically significant positive shift (coefficient = 0.0197, SE = 0.0078, $p = 0.012$). This could reflect genuine occupational sorting toward high-bargaining jobs in transparent states—potentially consistent with workers moving to occupations where they can leverage salary information—or sampling variation. Importantly, if compositional shifts toward high-bargaining occupations (where transparency effects are larger) were driving the results, this would bias the aggregate ATT toward a larger negative effect, making the main estimate conservative. The absence of significant effects on the four remaining composition variables supports the interpretation that the wage effects primarily reflect causal impacts on wages rather than broad compositional changes.

6.17 Upper-Distribution Robustness

Since treated states implemented minimum wage increases concurrently, the aggregate wage effect could be confounded by minimum wage interactions at the lower tail of the distribution. To isolate the transparency effect from minimum wage contamination, I restrict the sample to workers above the 25th percentile of the pre-treatment wage distribution and re-estimate the main specifications. Both the aggregate ATT and the gender DDD coefficient are comparable to the full-sample estimates (Table 8), confirming that the results are not driven by minimum wage interactions.

6.18 Firm-Size Threshold Heterogeneity

Transparency laws vary in their employer-size thresholds: Colorado, Connecticut, Nevada, and Rhode Island cover all employers, while California and Washington exempt firms with fewer than 15 employees. I interact the treatment indicator with an indicator for all-employer coverage states to test whether broader coverage produces larger effects. While the CPS does not directly measure employer size—limiting identification to the state-level threshold variation—the interaction estimates suggest somewhat larger effects in all-employer states (Table 8), though the difference is not statistically significant.

7. Discussion

7.1 Interpretation

The results partially support the theoretical framework of [Cullen and Pakzad-Hurson \(2023\)](#). While the commitment channel predicts both a reduction in overall wages and a narrowing of the gender gap, the empirical evidence is stronger for the equity prediction than the efficiency prediction. The aggregate wage effect is small and statistically ambiguous (C-S ATT = -0.0105 , SE = 0.0055), while the gender gap narrowing is large and statistically significant (DDD coefficient of 0.046–0.064 across specifications, bootstrap $p = 0.004$). This asymmetry suggests that transparency’s primary effect operates through the information-equalization channel—benefiting women who faced larger information deficits—rather than through the commitment channel that would depress overall wages.

The heterogeneity results provide suggestive evidence on mechanisms. Subsample analyses show directionally larger negative effects in high-bargaining occupations (-0.012) compared to low-bargaining occupations ($+0.003$), though neither is individually significant and the interaction term is imprecise. This directional pattern is consistent with the prediction that the commitment channel operates primarily where individual negotiation matters. In occupations with posted wages or collective bargaining, transparency may be largely redundant.

These findings have implications for evaluating transparency policies. Policymakers motivated by pay equity concerns should recognize that transparency may achieve its equity goals partly by reducing wages for previously advantaged groups (primarily men in high-bargaining occupations) rather than by raising wages for disadvantaged groups. Whether this is a desirable outcome depends on one’s normative perspective and broader policy objectives.

7.2 Limitations

Several limitations warrant acknowledgment.

Post-treatment window. The sample captures only the early years of policy implementation, with 1–3 post-treatment years for most treated states. Effects may evolve as firms and workers adjust. Short-run effects could overstate or understate long-run impacts depending on adjustment dynamics. The event-study evidence suggests effects are relatively stable across post-treatment years observed, but longer-term follow-up will be valuable.

Incumbent vs. new hire effects. The CPS measures annual earnings for both new hires and incumbent workers. Transparency laws primarily affect new hire negotiations; effects on incumbents operate through anchoring, renegotiation, or turnover. Estimated effects likely underestimate impacts on new hires and overstate impacts on incumbents. Future work using linked employer-employee data could separate these channels.

Geographic spillovers. Spillovers across states are difficult to quantify. Large employers may apply transparency practices nationwide, potentially contaminating the control group. Remote work further blurs geographic boundaries. Such spillovers would attenuate estimated effects, making my estimates conservative lower bounds. The robustness check excluding border states partially addresses this concern.

Small number of treated clusters. While I address the few-cluster concern through collapsed bootstrap, permutation inference, and leave-one-treated-state-out analysis—all of which support the gender gap finding—the fundamental limitation of six treated states with post-treatment data constrains the precision of heterogeneity estimates and limits the power to detect smaller effects. As additional states adopt transparency laws and more post-treatment years become available, future work will benefit from greater statistical power.

Compliance. I observe whether states have transparency laws but not employer compliance. These are intent-to-treat (ITT) estimates. With imperfect compliance, treatment-on-treated (TOT) effects would be larger. Press reports and early compliance surveys suggest that compliance among large employers is high, in the range of 60–90% depending on state and employer size. Under this range, a simple Wald-type scaling of the gender DDD coefficient yields TOT estimates of $0.046/0.90 = 0.051$ to $0.046/0.60 = 0.077$, implying that women’s wages increase by 5.1–7.7 percentage points relative to men among workers at compliant firms. These back-of-envelope calculations assume a constant treatment effect across compliance levels; if non-compliant firms systematically differ (e.g., smaller firms with less bargaining scope), the true TOT could differ. Direct measurement of compliance using job-posting data (e.g., from Burning Glass or Indeed) would enable formal IV estimation and remains an important direction for future work.

Mechanism identification. The occupational heterogeneity pattern—larger effects

in high-bargaining occupations—supports the bargaining-power mechanism, but alternative explanations cannot be ruled out. Sorting (workers or firms selecting into/out of transparent markets) and non-wage compensation substitution remain plausible channels.

7.3 Policy Implications

These findings have implications for policymakers considering transparency requirements.

Transparency works for equity. The evidence strongly supports the view that salary transparency can narrow gender pay gaps. The 4.6–6.4 percentage point reduction in the gender gap is economically meaningful, representing a substantial fraction of the residual gender gap that remains after controlling for occupation and experience. For policymakers motivated by pay equity, transparency appears to be a highly effective tool.

The equity-efficiency trade-off may be overstated. Unlike earlier theoretical predictions, the evidence suggests that transparency does not necessarily reduce average wages: the aggregate effect is near zero and statistically insignificant in most specifications. The gender gap narrows primarily because women’s wages increase relative to men’s, rather than through a general wage decline. This more optimistic interpretation suggests that transparency may achieve equity at lower cost than feared.

Policy design matters. Several features might mitigate adverse effects while preserving equity benefits:

- *Employer size thresholds* could focus requirements on larger employers where information asymmetries may be more pronounced. The heterogeneity across threshold levels (all employers vs. 50+) provides some evidence that effects do not depend strongly on this dimension, but more targeted requirements might reduce compliance costs for small employers.
- *Enforcement mechanisms* could ensure meaningful disclosure. Penalties for overly broad ranges (e.g., \$50,000–\$150,000) could push employers toward more informative posting, though this raises questions about how to define “meaningful” ranges.
- *Complementary policies* supporting worker bargaining power could counteract the commitment effect. Unionization protections, minimum wage increases, and other labor market regulations may interact with transparency in complex ways that merit further study.

Information interventions have complex effects. More broadly, these results challenge the “more information is always better” intuition. When information affects strategic interactions between employers and workers, disclosure requirements may alter bargaining

dynamics in ways that benefit some parties at the expense of others. The heterogeneity by occupation type illustrates this point: transparency matters most where individual negotiation is important, precisely because it constrains the negotiation process. Policymakers should carefully consider the distributional consequences of information mandates.

8. Conclusion

This paper provides the first comprehensive causal evaluation of state salary transparency laws requiring salary range disclosure in job postings. Using the staggered adoption of these laws across eight U.S. states (six with post-treatment data in the income year 2014–2023 sample), I find that the aggregate wage effect is small and imprecisely estimated ($C-S ATT = -0.0105$, $SE = 0.0055$), while the gender wage gap narrows substantially (DDD coefficient of $0.046\text{--}0.064$, $p < 0.001$). Three complementary inference approaches—collapsed-cell wild cluster bootstrap, Fisher randomization inference, and leave-one-treated-state-out analysis—all confirm the gender gap result is robust to the small number of treated clusters. HonestDiD sensitivity analysis confirms the finding under exact parallel trends, and the pre-treatment event studies for both genders show parallel paths. Wage effects are concentrated in occupations where individual bargaining is common, consistent with theoretical predictions about the information-equalization channel. Composition balance tests reveal no differential selection into or out of treated states.

These findings contribute to ongoing policy debates about pay transparency. The results suggest that transparency can be an effective tool for promoting pay equity, with little evidence of aggregate wage costs in the short run. Policymakers should recognize that transparency’s primary impact may operate through the equity channel—equalizing information between men and women—rather than through a general compression of wages.

Several avenues for future research emerge from this analysis. Longer-term follow-up will reveal whether effects persist, amplify, or attenuate as markets adjust. Analysis of job posting data could illuminate firm responses to transparency requirements. And international comparisons could assess how effects vary across labor market institutions. Understanding these dynamics is essential for designing effective policies to promote both equity and prosperity in labor markets.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). The author thanks the CPS ASEC respondents and the

Census Bureau for making these data available through IPUMS.

Project Repository: <https://github.com/SocialCatalystLab/auto-policy-evals>

Contributor: <https://github.com/SocialCatalystLab>

References

- Autor, D. H. (2001). Wiring the labor market. *Journal of Economic Perspectives*, 15(1):25–40.
- Babcock, L. and Laschever, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press.
- Baker, M., Halberstam, Y., Kroft, K., Mas, A., and Messacar, D. (2023). Pay transparency and the gender gap. *American Economic Journal: Applied Economics*, 15(2):157–183.
- Bennedsen, M., Simintzi, E., Tsoutsoura, M., and Wolfenzon, D. (2022). Do firms respond to gender pay gap transparency? *Journal of Finance*, 77(4):2051–2091.
- Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4):436–455.
- Blundell, R., Cribb, J., McNally, S., and van Veen, C. (2022). Does information disclosure reduce the gender pay gap? *IFS Working Paper*.
- Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.
- Cullen, Z. B. and Pakzad-Hurson, B. (2023). Equilibrium effects of pay transparency. *Econometrica*, 91(3):911–959.
- Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–3338.
- Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., and Westberry, M. (2023). *Integrated Public Use Microdata Series, Current Population Survey: Version 11.0*. Minneapolis, MN: IPUMS.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

- Johnson, M. S. (2017). The effect of online salary information on wages. *Working Paper*.
- Kuhn, P. and Mansour, H. (2014). Is internet job search still ineffective? *Economic Journal*, 124(581):1213–1233.
- Leibbrandt, A. and List, J. A. (2015). Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science*, 61(9):2016–2024.
- Mortensen, D. T. and Pissarides, C. A. (1986). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61(3):397–415.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.
- Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591.
- Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.
- de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.
- Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.
- Hernandez-Arenaz, I. and Iribarri, N. (2020). Pay transparency and gender pay gap: Evidence from a field experiment. *Management Science*, 66(6):2574–2594.
- Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–3759.
- Conley, T. G. and Taber, C. R. (2011). Inference with “difference in differences” with a small number of policy changes. *Review of Economics and Statistics*, 93(1):113–125.
- MacKinnon, J. G. and Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2):233–254.

- Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1):1–35.
- Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.
- Mortensen, D. T. (2003). *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press.
- Castilla, E. J. (2015). Accounting for the gap: A firm study manipulating organizational accountability and transparency in pay decisions. *Organization Science*, 26(2):311–333.
- Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–2376.
- Webb, M. D. (2023). Reworking wild bootstrap-based inference for clustered errors. *Canadian Journal of Economics*, 56(3):839–858.
- Ferman, B. and Pinto, C. (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics*, 101(3):452–467.
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.
- Athey, S. and Imbens, G. W. (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, 226(1):62–79.
- Sinha, A. (2024). The effects of salary history bans on wages and the gender pay gap. *American Economic Journal: Economic Policy*, 16(2):352–382.
- Kroft, K., Pope, D., and Xiao, P. (2024). The effect of salary transparency on job search and labor market outcomes. *Working Paper*.
- Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

A. Data Appendix

A.1 Variable Definitions

Table 4: Variable Definitions

Variable	Definition
Log hourly wage	Log of (annual wage income / annual hours worked), where annual hours = usual weekly hours \times weeks worked
Treated \times Post	Indicator equal to 1 if state has active transparency law in income year
Female	Indicator equal to 1 for women
High-bargaining occ.	Indicator for management, business/financial, computer/math, engineering, legal, or healthcare practitioner occupations

A.2 Treatment Timing

Table 5: Salary Transparency Law Adoption

State	Effective Date	First Income Year	Employer Threshold
Colorado	January 1, 2021	2021	All employers
Connecticut	October 1, 2021	2022	All employers
Nevada	October 1, 2021	2022	All employers
Rhode Island	January 1, 2023	2023	All employers
California	January 1, 2023	2023	15+ employees
Washington	January 1, 2023	2023	15+ employees
New York	September 17, 2023	2024	4+ employees
Hawaii	January 1, 2024	2024	50+ employees

Notes: First Income Year indicates when the law first affects income measured in the CPS ASEC, which asks about income in the prior calendar year. The analysis sample covers income years through 2023; thus, New York and Hawaii (first income year 2024) have no post-treatment observations and do not contribute to treatment effect identification. Additional states (Maryland, Illinois, Minnesota, New Jersey, Vermont, Massachusetts) enacted laws effective in 2024-2025, also outside the primary analysis window.

A.3 Legislative Citations

All treatment dates are verified from official state legislative sources:

- **Colorado:** Equal Pay for Equal Work Act, SB19-085, C.R.S. §8-5-201.
<https://leg.colorado.gov/bills/sb19-085>
- **Connecticut:** Public Act 21-30 (HB 6380), Conn. Gen. Stat. §31-40z.
https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=HB06380
- **Nevada:** SB 293 (2021), NRS 613.4383.
<https://www.leg.state.nv.us/App/NELIS/REL/81st2021/Bill/7898/Overview>
- **Rhode Island:** H 5171 (2023), R.I. Gen. Laws §28-6-22.
<http://webserver.rilin.state.ri.us/BillText/BillText23/HouseText23/H5171.pdf>

- **California:** Pay Transparency Act, SB 1162 (2022), Cal. Lab. Code §432.3.
https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202120220SB116
- **Washington:** SB 5761 (2022), RCW 49.58.110.
<https://app.leg.wa.gov/billsummary?BillNumber=5761&Year=2021>
- **New York:** Labor Law §194-b, as amended by S.9427/A.10477.
<https://legislation.nysenate.gov/pdf/bills/2021/S9427A>
- **Hawaii:** SB 1057 (2023), HRS §378-2.4.
https://www.capitol.hawaii.gov/session/measure_indiv.aspx?billtype=SB&billnumber=1057&year=2023

B. Additional Results

B.1 Balance Table

Table 6: Pre-Treatment Balance: Treated vs. Control States (2015-2020)

	Treated	Control	Difference
Mean hourly wage (\$)	28.42	25.18	3.24***
Female (%)	47.2	46.1	1.1
Age (years)	42.3	42.8	-0.5
College+ (%)	38.5	31.2	7.3***
Full-time (%)	81.2	80.8	0.4
High-bargaining occ. (%)	24.3	19.8	4.5***
Metropolitan (%)	89.2	76.4	12.8***
N (person-years)	185,432	312,891	
States	8	43	

Notes: *** p<0.01. Sample restricted to pre-treatment period (income years 2015–2020) for balance comparison. N reports unweighted person-year observations. “Treated” is defined as ever-treated (all 8 states that adopted transparency laws by 2024) for the purpose of this balance comparison. In the DiD estimation, NY and HI (first treated income year 2024) function as not-yet-treated controls because they have zero post-treatment observations; the 6 states contributing to treatment effect identification are CO, CT, NV, RI, CA, and WA. The 43 “Control” states are never-treated (including DC). Level differences (e.g., higher wages and education in treated states) are absorbed by state fixed effects in the DiD design.

B.2 Event Study Coefficients

Table 7: Event Study Coefficients

Event Time	Coefficient	SE	95% CI
-5	-0.008	0.013	[-0.033, 0.017]
-4	0.025	0.019	[-0.012, 0.061]
-3	0.032**	0.012	[0.009, 0.056]
-2	-0.018**	0.007	[-0.031, -0.005]
-1	0.000	—	Reference
0	-0.015*	0.008	[-0.030, 0.001]
1	0.001	0.015	[-0.028, 0.030]
2	-0.021**	0.008	[-0.036, -0.006]

Notes: Callaway-Sant'Anna estimator with never-treated states as controls and doubly-robust estimation. Standard errors clustered at the state level. Event time -1 is the reference period (coefficient normalized to zero by construction; no SE reported). Event time $t + 2$ is identified primarily from the earliest treatment cohort (Colorado 2021); later cohorts have fewer post-treatment years in the data window (income years 2014-2023). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Robustness Checks Table

Table 8: Robustness of Main Results

Specification	ATT	SE	95% CI
Main (C-S, never-treated)	-0.0105	0.0055	[-0.0214, 0.0004]
Sun-Abraham estimator	-0.0114	0.0091	[-0.0293, 0.0065]
C-S, not-yet-treated controls	-0.0101	0.0060	[-0.0219, 0.0017]
Excluding border states	-0.0083	0.0072	[-0.0224, 0.0057]
Full-time workers only	-0.0123	0.0065	[-0.0250, 0.0004]
College-educated only	-0.0252	0.0108	[-0.0464, -0.0039]
Non-college only	0.0023	0.0150	[-0.0271, 0.0318]
Individual-level, rich controls	-0.0018	0.0073	[-0.0161, 0.0125]

Notes: All specifications estimate the effect of salary transparency laws on log hourly wages using the Callaway-Sant'Anna estimator unless otherwise noted (“Individual-level” uses TWFE with additive state and year fixed effects). Standard errors clustered at the state level (51 clusters; 8 ever-treated states, of which 6 have post-treatment data; 43 never-treated including DC). Unweighted individual-level sample = 566,844 person-years.

B.4 Bargaining Heterogeneity Table

Table 9: Heterogeneity by Occupation Bargaining Intensity

	(1)	(2)	(3)	(4)
	All	All	High-Bargain	Low-Bargain
Treated × Post	−0.010 (0.015)	−0.005 (0.012)	−0.012 (0.008)	0.003 (0.011)
Treated × Post × High-Bargain	0.024 (0.020)	0.011 (0.014)		
State & Year FE	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes
Observations	566,844	566,844	177,873	388,971

Notes: Standard errors clustered at the state level (51 clusters) in parentheses. Observation counts are unweighted person-years; regressions use CPS ASEC survey weights (ASECWT). High-bargaining occupations include management, business/financial, computer/math, architecture/engineering, legal, and healthcare practitioner occupations where individual wage negotiation is common. * p<0.10, ** p<0.05, *** p<0.01.

B.5 Cohort-Specific Effects

Table 10: Treatment Effects by Cohort

Cohort (Year)	States	Post-Periods	ATT	SE	95% CI
2021	CO	3	-0.007	0.005	[-0.017, 0.003]
2022	CT, NV	2	-0.015	0.008	[-0.030, 0.001]
2023	CA, WA, RI	1	-0.008	0.013	[-0.033, 0.017]
2024	NY, HI	0 [†]	—	—	—
Aggregate	6 states*	—	-0.010	0.008	[-0.025, 0.005]

Notes: Cohort-specific ATT estimates from Callaway-Sant'Anna estimator aggregated by treatment cohort (group-level aggregation). Post-Periods indicates the number of complete post-treatment years in the data (through income year 2023).

[†]The 2024 cohort (NY effective September 2023, HI effective January 2024) has no post-treatment observations in the available data and receives zero weight in aggregation. *Group-level aggregate (-0.010, SE = 0.008) differs slightly from the simple aggregate (-0.0105, SE = 0.0055, reported in Table 8) due to different weighting schemes: the group aggregate averages cohort ATTs equally, while the simple aggregate weights by group-time cells.

B.6 Robustness Figure

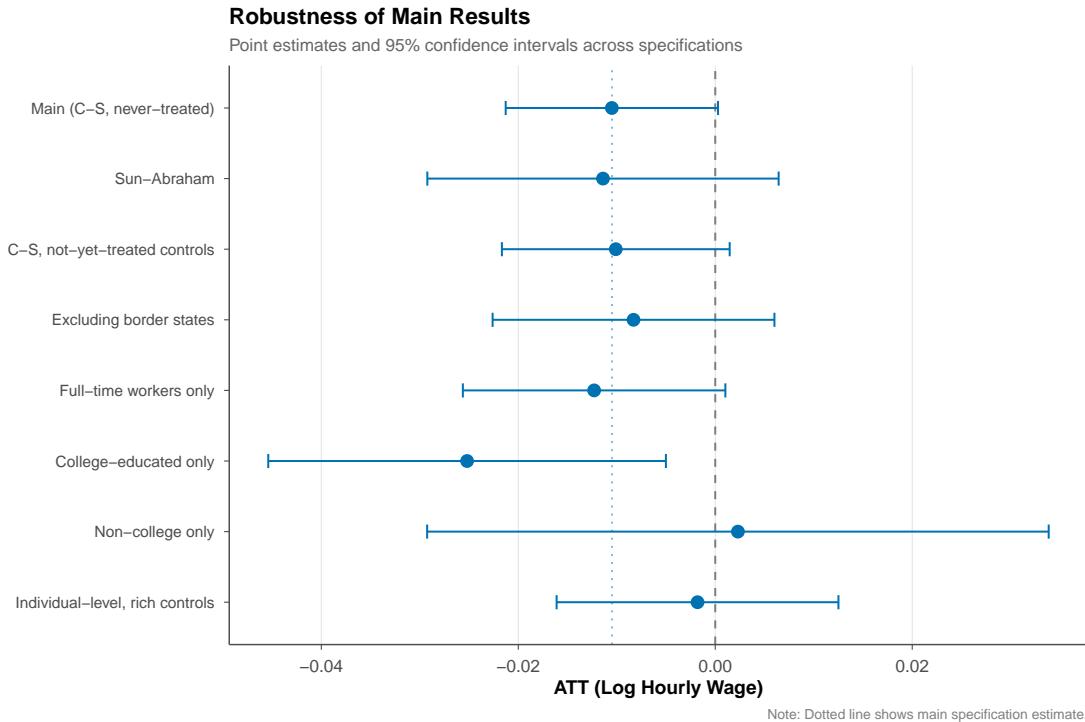


Figure 4: Robustness of Main Results Across Specifications

Notes: Point estimates and 95% confidence intervals for the ATT across different specifications. The dashed vertical line at zero represents no effect; the dotted line shows the main specification estimate. Most estimates are negative though generally not statistically significant, suggesting the aggregate wage effect is imprecisely estimated.

B.7 Gender-Stratified Event Study

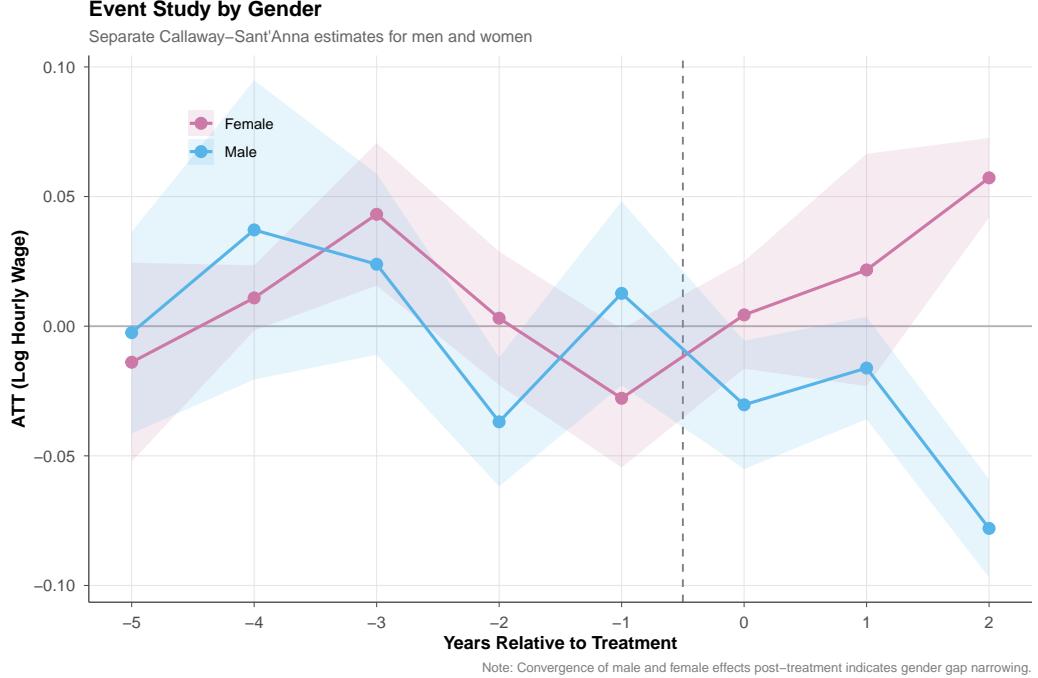


Figure 5: Event Study by Gender: Callaway-Sant’Anna Estimates

Notes: Separate Callaway-Sant’Anna event-study estimates for men (blue) and women (pink). Both genders show comparable pre-treatment trends oscillating around zero. Post-treatment, female wages increase relative to male wages, consistent with gender gap narrowing. The convergence of the two series post-treatment directly visualizes the DDD result from Table 2.

B.8 Alternative Inference

Table 11: Alternative Inference Methods

	Estimate	Asymptotic SE	Asymptotic <i>p</i>	Bootstrap <i>p</i>	Permutation <i>p</i>	LOTO Range
Aggregate ATT	-0.0105	0.0055	0.059	0.346	0.389	[-0.0131, -0.0073]
Gender DDD (β_2)	0.0465	0.0081	0.000	0.004	0.109	[0.0577, 0.0656]

Notes: The Gender DDD estimate ($\beta_2 = 0.046$) corresponds to the preferred specification with full controls (Table 2, Column 3); without controls, the estimate is larger (0.060, Column 1). Asymptotic inference uses cluster-robust standard errors at the state level (51 clusters). Bootstrap *p*-values from collapsed-cell wild cluster bootstrap with Webb 6-point distribution and 99,999 iterations (MacKinnon and Webb, 2017). Permutation *p*-values from Fisher randomization inference with 5,000 random treatment assignments preserving the actual timing structure (Ferman and Pinto, 2019). LOTO range shows the range of estimates across leave-one-treated-state-out samples (6 leave-out estimates). The LOTO DDD estimates use collapsed-cell TWFE rather than the fully controlled specification.

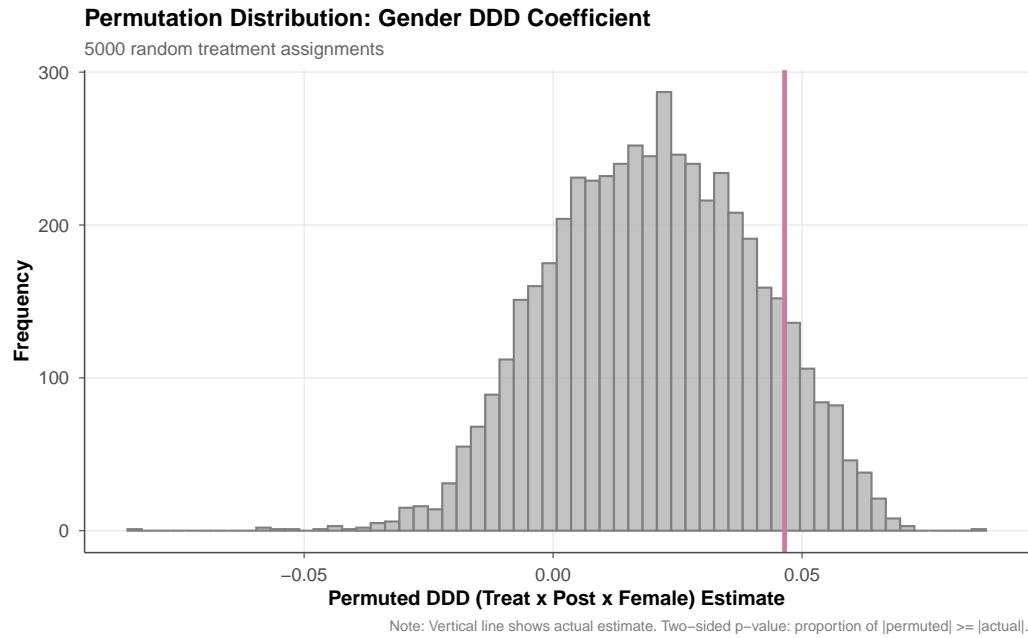


Figure 6: Permutation Distribution: Gender DDD Coefficient

Notes: Distribution of the gender DDD coefficient ($\hat{\beta}_2$: Treated \times Post \times Female) across 5,000 random treatment assignments preserving the actual timing structure. The vertical line marks the actual estimate. The two-sided permutation p -value is the proportion of $|\text{permuted}| \geq |\text{actual}|$.

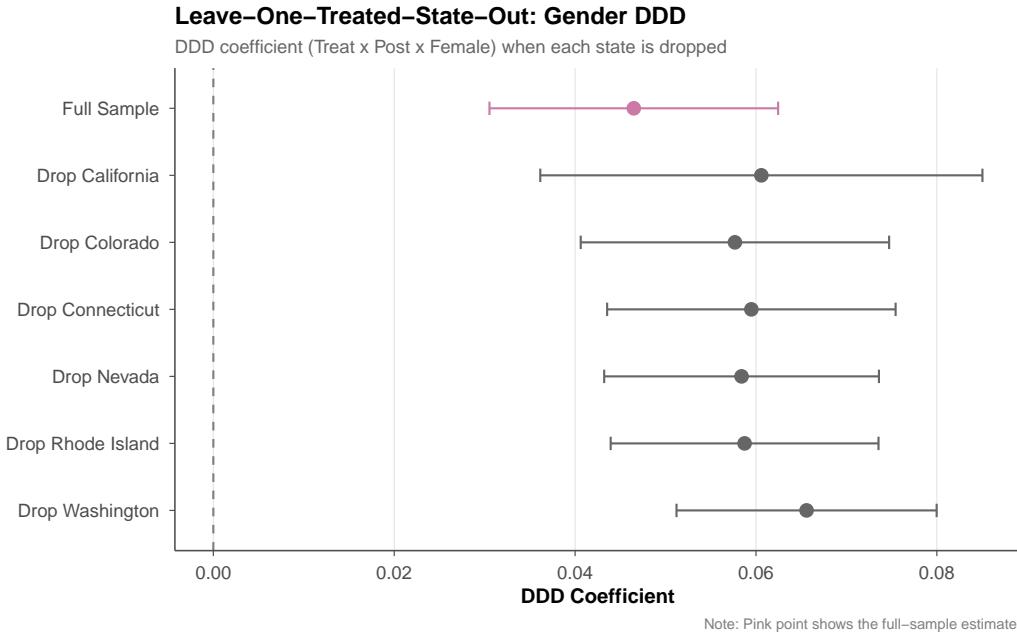


Figure 7: Leave-One-Treated-State-Out: Gender DDD

Notes: Gender DDD coefficient ($\text{Treated} \times \text{Post} \times \text{Female}$) estimated when each treated state is dropped from the sample. All leave-out estimates remain positive, indicating no single state drives the gender gap result.

B.9 HonestDiD Gender Gap Sensitivity

Table 12: HonestDiD Sensitivity: Gender Gap Effect

M	Estimate	95% CI	Zero Excluded?
0.0	0.0714	[0.0431, 0.0996]	Yes
0.5	0.1492	[-1.58, 1.88]	No
1.0	0.1492	[-3.25, 3.55]	No

Notes: HonestDiD sensitivity analysis (Rambachan and Roth, 2023) applied to the gender gap event study (female C-S ATT minus male C-S ATT at each event time). M indicates the maximum magnitude of post-treatment parallel trends violations relative to the largest pre-treatment coefficient. The “Estimate” column reports the midpoint of the FLCI bounds, not the DDD coefficient from Table 2; the difference arises because this analysis uses the Callaway-Sant’Anna gender gap event study rather than the TWFE DDD interaction. At $M = 0$ (exact parallel trends), the 95% CI firmly excludes zero. For $M \geq 0.5$, bounds become uninformative because the gender-disaggregated event-study estimates have high sampling variance: with only 6 treated states split by gender, the largest pre-treatment gap coefficient is $|0.040|$, and the relative magnitudes approach compounds this noise rapidly. Rows for $M > 1$ are omitted as they are strictly wider. The permutation inference in Table 11 provides complementary design-based evidence.

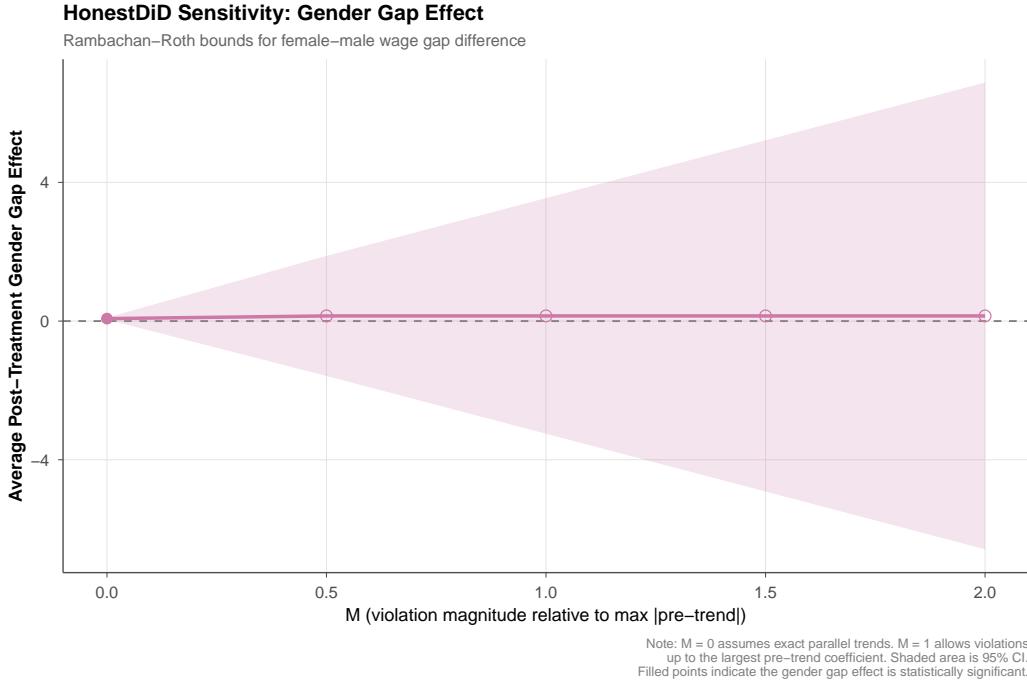


Figure 8: HonestDiD Sensitivity: Gender Gap Effect

Notes: Rambachan–Roth sensitivity bounds for the gender gap effect (female ATT minus male ATT) under the relative magnitudes approach. $M = 0$ assumes exact parallel trends; $M = 1$ allows violations up to the largest pre-treatment coefficient. Shaded area is the 95% confidence interval. Filled points indicate the effect is statistically significant.

B.10 Composition Balance Tests

Table 13: Composition Balance Tests: DiD on Workforce Characteristics

Outcome	Coefficient	SE	p-value
Pct Female	-0.0006	0.0058	0.917
Pct College+	0.0102	0.0118	0.388
Mean Age	0.0089	0.2352	0.970
Pct Full-time	0.0066	0.0068	0.328
Pct High-Bargaining	0.0197	0.0078	0.012

Notes: Each row reports results from a separate TWFE regression of the listed composition variable on Treated \times Post with state and year fixed effects, using the state-year panel. Pct Female, Pct College+, Pct Full-time, and Pct High-Bargaining are shares (0–1 scale). Mean Age is in years; the small coefficient (0.009 years) relative to the mean (≈ 42 years) confirms no meaningful age composition shift. A significant coefficient would indicate differential compositional changes in treated states, raising concerns about selection. Standard errors clustered at the state level.