The Innovation Cost of Privacy: How State Data Privacy Laws Reshape the Technology Sector

APEP Autonomous Research*

February 10, 2026

Abstract

Between 2020 and 2025, nineteen U.S. states enacted comprehensive consumer data privacy laws modeled on California's landmark CCPA, creating substantial new compliance obligations for firms that collect personal data. I exploit this staggered rollout—focusing on the thirteen states with at least one post-treatment quarter in the available data—to estimate the causal effect of privacy regulation on the technology sector using a Callaway-Sant'Anna difference-in-differences design with quarterly state-level data from the Bureau of Labor Statistics and Census Bureau. The identifying assumption—that Information sector trends in adopting and non-adopting states would have evolved similarly absent the laws—is supported by parallel pre-trends across all outcome measures. I find that privacy law adoption significantly reduces employment in Software Publishers (NAICS 5112)—the most data-intensive subsector (ATT $= -0.0767$, SE $= 0.0247$, $p < 0.01$)—an approximately 7.4% decline. Establishment counts show a gradual decline that intensifies over time, consistent with cumulative exit of marginal firms. Placebo tests on healthcare and construction industries confirm that effects are concentrated in data-intensive sectors. Randomization inference for the Software Publishers specification yields a p-value of 0.077, providing marginally significant non-parametric support for the primary finding. These results inform the active policy debate over a federal privacy framework by quantifying the sectoral costs of regulatory fragmentation.

JEL Codes: K24, L86, O38, J21

Keywords: data privacy, regulation, technology sector, CCPA, difference-in-differences, innovation

---

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

Every time an American downloads an app, browses a website, or swipes a credit card, a transaction occurs that is invisible to most consumers: their personal data enters a vast ecosystem of collection, aggregation, and monetization. This data economy underpins the business models of firms ranging from Silicon Valley giants to local startups. Yet the legal framework governing this ecosystem has been transformed by a wave of state legislation that began with California's Consumer Privacy Act (CCPA) in 2020 and has since spread to eighteen additional states. These comprehensive privacy laws grant consumers new rights—to know what data firms collect, to delete it, to opt out of its sale—and impose costly compliance obligations on the companies that depend on data as a productive input.

The central question motivating this paper is straightforward but empirically challenging: what happens to the technology sector when states impose comprehensive privacy regulation? Theoretical predictions are ambiguous. Privacy laws could reduce innovation by raising compliance costs and restricting access to the data that fuels product development (Goldfarb and Tucker, 2012; Johnson et al., 2024). Alternatively, they could strengthen consumer trust, expand the digital economy, and create competitive advantages for firms in regulated jurisdictions (Acquisti et al., 2016). The net effect is an empirical question that has become urgently relevant as Congress debates a federal privacy standard and as firms navigate an increasingly fragmented regulatory landscape.

I exploit the staggered adoption of state comprehensive data privacy laws—thirteen states with post-treatment data spanning 2020 through 2025—to estimate the causal effect of privacy regulation on four key outcomes: employment, establishment counts, wages, and new business applications in the Information sector (NAICS 51). The identification strategy leverages the fact that states adopted similar legislation at different times for reasons plausibly orthogonal to pre-existing trends in their technology sectors. California acted first, driven by a ballot initiative in 2018 that prompted legislative preemption. Virginia, Colorado, and Connecticut followed in 2023, with subsequent waves in 2024–2025. More than thirty states remain without comprehensive privacy laws, providing a control group of never-treated jurisdictions.

My empirical approach uses the Callaway and Sant'Anna (2021) heterogeneity-robust difference-in-differences estimator, which accounts for the well-documented bias that standard two-way fixed effects models can produce with staggered treatment timing (Goodman-Bacon, 2021; Sun and Abraham, 2021; de Chaisemartin and D'Haultfœuille, 2020). The estimator constructs group-time average treatment effects for each cohort of adopting states, then aggregates to an overall ATT and event-study estimates. I use doubly robust estimation—combining inverse probability weighting with outcome regression—and cluster standard errors

at the state level. All regressions use never-treated states as the comparison group.

The primary data come from two federal statistical programs. The Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW) provides a near-census of employment, establishments, and wages by state, quarter, and industry from 2015 through 2025. The Census Bureau's Business Formation Statistics (BFS) provide quarterly state-level counts of new business applications beginning in 2004. Together, these sources enable me to track multiple margins along which privacy regulation might affect the technology sector: the intensive margin (employment within existing firms), the extensive margin (entry and exit of establishments), the price of labor (wages), and the creation of new businesses.

I examine whether privacy laws shift the trajectory of the technology sector relative to what would have occurred absent regulation. Event-study estimates reveal the dynamic path of treatment effects in the quarters surrounding law adoption, with pre-treatment coefficients serving as a test of the parallel trends assumption. The estimates show whether effects emerge immediately upon enactment, build gradually as enforcement ramps up, or materialize only after firms have time to adjust.

The credibility of the research design rests on several auxiliary tests. First, placebo regressions on industries that should be unaffected by privacy regulation—healthcare (NAICS 62) and construction (NAICS 23)—confirm that the estimated effects are specific to data-intensive sectors rather than reflecting broad economic shocks. Second, I conduct randomization inference on the primary Software Publishers specification, permuting treatment assignments 500 times to construct a distribution of placebo ATTs under the sharp null hypothesis of no effect; the observed ATT is compared to this distribution to obtain a non-parametric p-value. Third, I verify robustness to including Florida—excluded from the primary specification because its law uniquely applies only to firms with over \$1 billion in revenue—and to restricting the analysis to the narrow Software Publishers subsector (NAICS 5112).

Heterogeneity analysis provides insight into the mechanism. If compliance costs are the primary channel, we should expect larger effects in states where the technology sector is more prominent at baseline, because these states have more firms subject to the laws and because the costs are proportionally larger relative to the local economy. Splitting the sample at the median state-level share of Information sector employment in 2019, I test whether high-tech states experience differentially larger treatment effects.

This paper contributes to three literatures. First, it advances the growing empirical literature on the economic effects of privacy regulation. Goldfarb and Tucker (2011) documented that the EU's Privacy Directive reduced the effectiveness of online advertising. Miller and Tucker (2015) studied the effects of health data breach notification laws on hospital investment. Johnson et al. (2024) and Aridor et al. (2024) have examined how GDPR affected

digital markets in Europe. My contribution is to provide the first multi-state, multi-outcome causal analysis of U.S. data privacy laws using a modern heterogeneity-robust DiD framework. Most existing work focuses on a single law (typically CCPA or GDPR) or a single outcome.

Second, this paper contributes to the broader literature on the costs and benefits of regulation for innovation. Jaffe et al. (1995) established the framework for analyzing how environmental regulation affects innovation, finding that compliance costs can redirect but do not necessarily reduce inventive activity. Hombert et al. (2020) showed that labor regulation affects firm entry and entrepreneurship. My setting offers a clean test of how regulatory costs interact with the distinctive characteristics of the data economy—network effects, economies of scale in data, and the role of data as a non-rival input to production.

Third, the paper informs the active policy debate over federal preemption versus state-by-state regulation. The current patchwork imposes compliance costs that scale with the number of distinct legal regimes a firm must navigate. If the per-state costs of compliance are substantial, the case for a uniform federal standard is strengthened. Conversely, if the effects are modest, the regulatory competition among states may serve as a useful laboratory of democracy (Oates, 1999).

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and the key features of state data privacy laws. Section 4 details the data sources and sample construction. Section 5 presents the empirical strategy, including the identification assumptions and estimation procedure. Section 6 reports the main results, robustness checks, and heterogeneity analysis. Section 7 discusses mechanisms and limitations. Section 8 concludes.

2. Institutional Background

2.1 The Rise of State Data Privacy Legislation

The United States lacks a comprehensive federal data privacy law. Unlike the European Union, which enacted the General Data Protection Regulation (GDPR) in 2018, the U.S. has historically relied on sector-specific federal statutes—HIPAA for health data, COPPA for children's online data, FCRA for consumer credit—supplemented by the Federal Trade Commission's enforcement authority against "unfair or deceptive" practices (Solove, 2006). This sectoral approach left vast categories of consumer data, particularly that collected by technology companies, without a dedicated regulatory framework.

California changed this calculus in 2018 when a real estate developer, Alastair Mactaggart, qualified a ballot initiative for comprehensive data privacy regulation. Facing the prospect of an inflexible ballot measure, the California legislature negotiated the California Consumer

Privacy Act (CCPA) as a legislative alternative, which Governor Jerry Brown signed in June 2018 and which took effect January 1, 2020. The CCPA grants California residents the right to know what personal information businesses collect, the right to delete that information, the right to opt out of its sale, and the right to non-discrimination for exercising these rights. The California Privacy Rights Act (CPRA), approved by voters in November 2020, strengthened and expanded CCPA protections and created a dedicated enforcement agency, the California Privacy Protection Agency.

The CCPA's passage triggered a cascade of state legislative activity. Virginia enacted the Consumer Data Protection Act (VCDPA), effective January 1, 2023. Colorado and Connecticut followed with laws effective July 1, 2023. By the end of 2025, sixteen states had comprehensive privacy laws in effect, with three additional states (Indiana, Kentucky, Rhode Island) having enacted laws taking effect in 2026. The speed of this diffusion—from one state in 2020 to nineteen enacted by 2026—reflects both the political salience of data privacy and the template effect of early adopter legislation, as later states drew heavily on the language and structure of earlier laws.

2.2 Key Features of Comprehensive Privacy Laws

While no two state laws are identical, they share a common architecture. Most apply to businesses that (a) conduct business in the state or target the state's residents, (b) process personal data above a threshold volume (typically 100,000 consumers or 25,000 consumers if data is sold), and (c) meet a revenue threshold (varies by state). The laws generally grant consumers five core rights: access, deletion, correction, portability, and opt-out of sale or targeted advertising.

For firms, compliance requires substantial operational changes. Companies must implement mechanisms for consumers to exercise their rights, maintain records of data processing activities, conduct data protection assessments for high-risk processing, update privacy policies, and in some cases appoint a data protection officer. These requirements impose both fixed costs (building compliance infrastructure) and variable costs (processing individual consumer requests). Johnson et al. (2024) estimate that CCPA compliance costs for a mid-sized technology company range from $50,000 to $2 million annually, with larger firms facing costs in the tens of millions.

Florida's Consumer Data Privacy Act (FDBR), effective July 1, 2024, represents an outlier: it applies only to entities with global gross revenues exceeding $1 billion. This high threshold limits its applicability to a small number of large corporations, making Florida's experience qualitatively different from other states. I therefore exclude Florida from the primary specification and include it as a robustness check.

2.3 Variation Across State Laws

While the state privacy laws share a common architecture derived from the CCPA template, meaningful variation exists along several dimensions that bear on economic effects. First, **applicability thresholds** differ. The CCPA applies to businesses with annual revenue exceeding $25 million, those that buy or sell personal information of 100,000 or more consumers, or those that derive 50% or more of revenue from selling consumers' personal information. Virginia's VCDPA uses a simpler threshold: businesses that control or process the personal data of at least 100,000 consumers, or 25,000 consumers if the business derives over 50% of gross revenue from selling personal data. Several later-adopting states, including Colorado and Connecticut, adopted thresholds similar to Virginia's, while Texas applies its law to any entity that conducts business in the state or produces goods targeted to Texas residents, without a revenue threshold.

Second, **enforcement mechanisms** vary. California created a dedicated enforcement agency—the California Privacy Protection Agency (CPPA)—with rulemaking authority and a budget of approximately $10 million annually. Most other states vest enforcement authority in the state attorney general, with no private right of action. The absence of private litigation rights in most states may reduce the deterrent effect of the laws compared to California, where consumer lawsuits for data breaches can proceed under a separate provision.

Third, **the scope of consumer rights** varies across states. All states with comprehensive privacy laws grant the core rights to access, delete, and opt out of the sale of personal data. However, the treatment of "sensitive data"—which may include health information, biometric identifiers, precise geolocation, racial or ethnic origin, and sexual orientation—varies. Colorado and Connecticut require businesses to obtain affirmative consent (opt-in) before processing sensitive data, while Virginia and Utah use an opt-out framework for most sensitive categories. These differences in the regulatory burden may generate heterogeneous effects across states even within the same adoption cohort.

Fourth, **cure periods**—the time a business has to remedy a violation before facing enforcement action—range from 30 days (Virginia, Colorado) to 60 days (Utah) or no cure period at all (under the amended CCPA). Longer cure periods may reduce the effective stringency of the law, giving businesses more time to adjust before facing penalties.

Despite this variation, the fundamental economic mechanism is consistent across states: businesses that collect, process, or sell personal data face new obligations that require operational changes, legal review, and technological infrastructure. The staggered timing of adoption, combined with this consistent underlying mechanism, provides the variation needed for identification while the cross-state differences in stringency may contribute to treatment effect heterogeneity.

2.4 Why Privacy Laws Might Affect the Technology Sector

The Information sector (NAICS 51) encompasses industries that produce, process, and distribute information—including software publishers, data processing firms, internet service providers, and digital media companies. These firms are disproportionately reliant on personal data as a productive input. A data privacy law that restricts how firms collect, use, and share personal data thus represents a direct constraint on the production technology of these industries.

Three channels are theoretically relevant. First, **compliance costs** represent a direct tax on data-intensive firms. Fixed compliance costs may deter entry by startups that lack the resources to build compliance infrastructure, while variable costs reduce the returns to data-intensive business models. Second, **data restrictions** may reduce the quantity and quality of data available for product development, targeting, and personalization—activities that underpin the competitive advantage of many technology firms. Third, **regulatory uncertainty** may suppress investment as firms wait for clarity on enforcement priorities and judicial interpretation.

Countervailing effects are also plausible. Privacy regulation may increase consumer trust, encouraging greater participation in the digital economy. Compliance infrastructure may serve as a barrier to entry that protects incumbents. And the creation of a regulated market for data may actually increase the value of data assets by providing clearer property rights. The net effect is theoretically ambiguous, motivating the empirical analysis.

3. Conceptual Framework

To structure the empirical analysis and generate testable predictions, I develop a simple framework linking privacy regulation to labor market outcomes in data-intensive industries.

3.1 Privacy Regulation as a Cost Shock

Consider a technology firm j in state s that uses personal data D_{jt} as a productive input alongside labor L_{jt} and capital K_{jt}. Output is produced according to:

$$Y_{jt} = A_{jt} \cdot f(L_{jt}, K_{jt}, D_{jt}) \tag{1}$$

where A_{jt} represents firm-level productivity. A privacy law imposes two types of costs. First, a **fixed compliance cost** F_s that each firm must pay regardless of size—hiring privacy officers, building consent management systems, updating privacy policies. Second, a **variable cost** that reduces the effective quantity or quality of data available: the law constrains

D_{jt} through opt-out mechanisms, data minimization requirements, and restrictions on data sharing. Formally, the effective data input becomes $\tilde{D}_{jt} = (1 - \tau_s)D_{jt}$, where $\tau_s \in [0, 1]$ captures the stringency of the privacy regime in state s.

3.2 Predictions

This framework generates several testable predictions:

Prediction 1: Employment. The fixed compliance cost reduces profits for firms at the margin, potentially causing exit or reduced hiring. The reduction in effective data input lowers the marginal product of complementary labor, further depressing employment. However, the compliance requirements themselves create demand for privacy-specialized labor (lawyers, engineers, compliance officers), partially offsetting the direct effect. The net effect on employment is ambiguous but likely negative for the Information sector as a whole, where data is a primary input.

Prediction 2: Establishments. Fixed compliance costs disproportionately burden small firms and potential entrants, for whom these costs represent a larger share of revenue. We therefore expect a decline in the number of establishments, driven primarily by reduced entry rather than exit of incumbents. This prediction is strongest for the extensive margin (establishment counts) rather than the intensive margin (employment per firm).

Prediction 3: Wages. The effect on wages operates through competing channels. Reduced labor demand should lower wages, but if the composition of the workforce shifts toward higher-skilled compliance roles, average wages could increase. Moreover, firms may pass through some compliance costs to workers in the form of lower wages. The wage prediction is theoretically ambiguous.

Prediction 4: Business Formation. New business applications should decline in states adopting privacy laws, as the fixed costs of compliance raise the threshold for profitable entry. This effect should be particularly pronounced for data-intensive business models—the archetype of which is a technology startup that monetizes user data through advertising or analytics.

Prediction 5: Heterogeneity. Effects should be concentrated in states where the technology sector is more prominent at baseline, because these states have more firms directly affected by the regulation and because the compliance costs represent a larger aggregate burden relative to the state economy.

4. Data

4.1 Quarterly Census of Employment and Wages (QCEW)

The primary outcome data come from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW), which reports employment, establishment, and wage data derived from state unemployment insurance records. The QCEW covers approximately 99.7% of all employees on nonfarm payrolls, making it a near-census of the formal economy. I access the QCEW through the BLS public API, obtaining state-level quarterly data for private-sector establishments from 2015Q1 through 2025Q2—the most recent quarter available at the time of analysis.

I focus on NAICS 51 (Information) as the primary treatment-relevant industry. This sector includes subsectors 511 (Publishing, except Internet), 512 (Motion Picture and Sound Recording), 515 (Broadcasting), 517 (Telecommunications), 518 (Data Processing, Hosting, and Related Services), and 519 (Web Search Portals, Libraries, Archives, and Other Information Services). While privacy regulation affects firms across the economy, the Information sector faces the most direct impact because its core business activity—collecting, processing, and distributing data—is precisely what these laws regulate.

For each state-quarter-industry observation, I construct three outcome variables: average quarterly employment (the mean of monthly employment levels within the quarter), the count of quarterly establishments, and average weekly wages. I take natural logarithms of each outcome (after adding one to handle any zeros) to obtain approximately normally distributed variables and to interpret coefficients as approximate percentage changes.

4.2 Business Formation Statistics (BFS)

To capture the extensive margin of firm entry, I supplement the QCEW data with the Census Bureau's Business Formation Statistics. The BFS provides quarterly state-level counts of new business applications—defined as applications for an Employer Identification Number (EIN) that the Census Bureau identifies as likely to result in a business with a payroll. I use the "Business Applications" (BA_BA) series, which covers applications across all sectors and provides a leading indicator of business dynamism.

The BFS data are available quarterly; I use data from 2015Q1 through 2020Q4, covering the pre-treatment period and the first year following California's CCPA. Because the BFS does not disaggregate by industry, these applications reflect economy-wide business formation rather than tech-specific entry. I use the non-seasonally-adjusted series to avoid complications from seasonal adjustment procedures that may interact with treatment effects. Because

the BFS data end in 2020Q4 and California is the only state treated by that date (CCPA effective 2020Q1), the BFS estimates are identified exclusively from California's post-CCPA experience versus the 50 untreated states. This constitutes a single-treated-unit analysis with limited external validity, and I interpret these estimates accordingly.

4.3 Treatment Assignment

I code treatment based on the effective date of each state's comprehensive data privacy law, assigning states to the first full calendar quarter in which the law is in effect; when a law takes effect on the first day of a quarter, that quarter is the treatment quarter. Table 4 lists all nineteen states that have enacted comprehensive privacy laws, with their law names, effective dates, and treatment quarters. Of these, thirteen have at least one post-treatment quarter in the QCEW data (which ends at 2025Q2) and contribute to the ATT estimates. California (2020Q1) represents the earliest and only single-state cohort. Virginia, Colorado, and Connecticut form the 2023 cohort. Utah enters treatment in 2024Q1, with Texas, Oregon, and Montana joining in subsequent quarters of 2024. Five states enter treatment in 2025Q1 (Delaware, Iowa, Nebraska, New Hampshire, New Jersey). Six additional states—Tennessee, Minnesota, Maryland (effective 2025Q3–Q4) and Indiana, Kentucky, Rhode Island (effective 2026Q1)—have zero post-treatment quarters in the available data and are classified as not-yet-treated.

States without comprehensive data privacy laws in effect by the end of the QCEW sample period (2025Q2) serve as the comparison group. For the Callaway-Sant'Anna estimator, these thirty-nine units—thirty-two states that have not enacted privacy laws, the District of Columbia, and six states classified as not-yet-treated—provide the never-treated control group. The control group spans a range of state sizes, economic structures, and political orientations, providing a broad comparison population.

4.4 Sample Construction

The analysis panel for the primary QCEW specification consists of 52 geographic units (50 states plus D.C. and an additional territory covered by QCEW reporting) observed quarterly from 2015Q1 through mid-2025, yielding 2,226 state-quarter observations for the Information sector. The panel is slightly unbalanced because the most recent QCEW release dates vary by state, with most units having 42–43 quarters of data.[1] The BFS panel covers 2015Q1–2020Q4,

---

[1]Three states with laws effective in 2025Q3–Q4 (Tennessee, Minnesota, Maryland) are coded as treated but have zero post-treatment quarters in the QCEW data, which ends at 2025Q2. The Callaway-Sant'Anna estimator handles this correctly by only estimating ATT(g,t) cells for which post-treatment data exist, so these states effectively contribute only to the control group. Thirteen states have at least one post-treatment

providing 24 quarters per state. Because California is the only state treated by 2020Q4, the BFS results should be interpreted as reflecting the CCPA's specific impact on business formation rather than the aggregate effect of all state privacy laws.

The panel is unbalanced due to missing observations in the QCEW for some state-quarter-industry combinations, particularly for smaller states and more granular industry classifications.

4.5 Summary Statistics

Table 1: Summary Statistics: Information Sector (NAICS 51), Pre-Treatment Period (2015–2019)

	Mean	Std. Dev.	Min	Max
Panel A: Treated States (N = 260 state-quarters)				
Avg. Quarterly Employment	84495	135162	3877	558478
Quarterly Establishments	4462	6124	388	29826
Avg. Weekly Wage ($)	1705	599	893	3848
Panel B: Control States (N = 800 state-quarters)				
Avg. Quarterly Employment	43044	51231	598	280101
Quarterly Establishments	2714	2713	43	12841
Avg. Weekly Wage ($)	1379	532	677	4830

Notes: Pre-treatment period covers 2015Q1–2019Q4. Treated states (N=13) are those with comprehensive data privacy laws effective by 2025Q1 and at least one post-treatment quarter in the QCEW data (through 2025Q2). Control states (N=39) include never-treated states and states with laws effective after 2025Q2. Employment and establishments from BLS Quarterly Census of Employment and Wages (QCEW), private ownership, NAICS 51 (Information).

Table 1 reports pre-treatment (2015–2019) summary statistics for states that eventually adopt privacy laws and those that do not. Treated states tend to have larger Information sector employment, more establishments, and higher wages than control states, reflecting the fact that early-adopting states (California, Virginia, Colorado, Connecticut) include several technology hubs. This level difference does not threaten identification under the DiD assumption, which requires only parallel trends in the outcome variable, not equal levels.

quarter.

Nevertheless, I address this concern directly with event-study evidence of parallel pre-trends and with placebo tests.

5. Empirical Strategy

5.1 Identification

The central challenge in estimating the effect of privacy regulation is that states self-select into adoption. If states with declining technology sectors adopt privacy laws (perhaps seeking to attract privacy-conscious consumers), or if states with booming tech sectors adopt them (responding to political pressure from affected residents), naive comparisons of treated and untreated states will produce biased estimates. The difference-in-differences design addresses this concern by comparing *changes* in outcomes before and after law adoption across treated and control states, under the assumption that the trends would have been parallel absent treatment.

Formally, the identifying assumption is:

$$\mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid G_i = g] = \mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid G_i = \infty] \quad \forall \; g, t \geq g \tag{2}$$

where $Y_{it}(0)$ denotes the potential outcome for state i in period t absent treatment, G_i denotes the treatment cohort (the period in which state i first becomes treated), and $G_i = \infty$ denotes never-treated states. This assumption requires that, conditional on group and time, the average change in untreated potential outcomes is the same for treated and never-treated states.

Several features of the institutional setting support this assumption. First, the timing of state privacy law adoption was driven largely by idiosyncratic political factors—legislative leadership, advocacy group pressure, and the template effect of neighboring states' laws—rather than by trends in technology sector outcomes. Second, the rapid diffusion of these laws across states of very different sizes and economic structures suggests that adoption decisions were not closely tied to state-specific economic conditions. Third, the use of never-treated states (rather than not-yet-treated states) as controls avoids contamination from anticipatory effects in states that would later adopt laws.

I also invoke a no-anticipation assumption: that states do not change their behavior in advance of the law's effective date. This is plausible given that most laws were signed 6–12 months before their effective dates, leaving little time for firms to preemptively restructure operations. I test this assumption directly through the event-study specification, which should show zero pre-treatment coefficients.

5.2 Estimation

I implement the Callaway and Sant'Anna (2021) estimator, which constructs group-time average treatment effects:

$$ATT(g,t) = \mathbb{E}[Y_{it} - Y_{ig-1} \mid G_i = g] - \mathbb{E}[Y_{it} - Y_{ig-1} \mid G_i = \infty] \tag{3}$$

for each treatment group g (defined by the quarter of adoption) and post-treatment period t. The estimator uses doubly robust estimation, combining inverse probability weighting (to reweight the control group to match the treated group on baseline characteristics) with outcome regression (to model the counterfactual trend). This doubly robust approach is consistent if either the propensity score model or the outcome model is correctly specified.

I aggregate the group-time ATTs in two ways. First, the **overall ATT** computes a weighted average across all group-time cells:

$$ATT = \sum_{g} \sum_{t \geq g} w_{g,t} \cdot ATT(g,t) \tag{4}$$

where weights $w_{g,t}$ reflect the relative size of each group-time cell. Second, the **event-study** specification aggregates by event time $e = t - g$ (quarters relative to treatment):

$$ATT(e) = \sum_{g} w_g(e) \cdot ATT(g, g+e) \tag{5}$$

producing a dynamic treatment effect path that I plot from $e = -8$ to $e = +8$ quarters. Pre-treatment coefficients ($e < 0$) serve as a falsification test of the parallel trends assumption.

For comparison, I also estimate standard two-way fixed effects (TWFE) regressions:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_{it} + \epsilon_{it} \tag{6}$$

where α_i and γ_t are state and period fixed effects, D_{it} indicates that state i has an active privacy law in period t, and standard errors are clustered at the state level. As Goodman-Bacon (2021) and de Chaisemartin and D'Haultfœuille (2020) have shown, $\hat{\beta}_{TWFE}$ can be a weighted average of group-time ATTs with negative weights under treatment effect heterogeneity, making the CS estimator the preferred specification.

I additionally estimate the Sun and Abraham (2021) interaction-weighted estimator as a further robustness check, using the `sunab()` function in the `fixest` R package.

5.3 Inference

All standard errors are clustered at the state level—the unit at which treatment is assigned—to account for within-state serial correlation in outcomes. For the CS estimator, I use the block bootstrap with 1,000 replications to construct confidence intervals and simultaneous confidence bands for the event-study path.

As a complement to asymptotic inference, I conduct Fisher-style randomization inference on the Software Publishers (NAICS 5112) specification. Under the sharp null hypothesis that the privacy law had zero effect in every state in every period, any treatment assignment should produce the same distribution of outcomes. I randomly permute treatment labels across states—assigning the same number of states as treated and drawing treatment cohort dates from the observed distribution of cohort timing—then re-estimate the CS model on each permuted dataset. Of 500 attempted permutations, 156 yield valid CS estimates (others fail due to collinearity or insufficient variation in the permuted treatment assignment). The two-sided p-value is the fraction of valid permutations for which $|ATT_{perm}| \geq |ATT_{obs}|$. While the number of valid permutations is limited by the small number of treated states and the granularity of NAICS 5112 data, the resulting p-value of 0.077 provides informative non-parametric evidence supporting the asymptotic result.

5.4 Threats to Validity

The primary threat to identification is **differential trends**: that states adopting privacy laws were on different trajectories in technology sector outcomes even before adoption. I address this with event-study evidence and formal pre-trend tests. If pre-treatment coefficients are jointly insignificant and close to zero, this supports the parallel trends assumption.

A second concern is **spillovers**. If privacy laws in treated states cause firms to relocate to untreated states, the control group outcome is affected by treatment, violating the Stable Unit Treatment Value Assumption (SUTVA). This would bias estimates away from zero: the treated state loses firms while the control state gains them, making the treatment effect appear larger than the true causal effect. I note this concern but cannot definitively rule it out with the available data.

A third concern is **compositional changes**. If privacy laws cause different types of firms to enter or exit the Information sector, the composition of the workforce may shift, complicating interpretation of wage effects. I treat this as a secondary concern given the relatively short post-treatment windows for most states.

Finally, **national compliance** could attenuate treatment effects if firms adopt privacy practices uniformly across all states once any single state imposes requirements. To the extent

that California's CCPA already induced nationwide compliance changes, the effects estimated here represent the marginal impact of additional state laws beyond the CCPA baseline.

6. Results

6.1 Main Results

Table 2: Main Results: Effect of State Data Privacy Laws on the Technology Sector

	(1) Log Emp. (NAICS 5112)	(2) Log Estab. (NAICS 51)	(3) Log Bus. Apps. (BFS, CA only)
Panel A: Callaway-Sant'Anna DiD			
ATT	-0.0767***	-0.1234	-0.0026
	(0.0247)	(0.0909)	(0.0177)
Panel B: Two-Way Fixed Effects			
Treated	-0.0289	-0.0950**	0.0189
	(0.0424)	(0.0426)	(0.0199)
States	52	52	51
State-Quarters	1,456	2,226	1,224

Notes: Panel A reports Callaway and Sant'Anna (2021) aggregate ATT with doubly robust estimation and 1,000 bootstrap iterations. Panel B reports TWFE coefficients with state and period fixed effects. Standard errors clustered at the state level in parentheses. Column (1) uses NAICS 5112 (Software Publishers); column (2) uses NAICS 51 (Information sector). QCEW data: 2015Q1–2025Q2, 13 treated states with post-treatment data. BFS column (3): 2015Q1–2020Q4 with California as the only treated state during this window. Wage effects are not tabulated; estimates are near zero with wide confidence intervals. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 2 presents the main results from both the Callaway-Sant'Anna DiD estimator (Panel A) and the standard TWFE specification (Panel B). The CS estimates represent the aggregate average treatment effect on the treated (ATT), measuring the average post-treatment effect across all treated states and post-treatment periods, weighted by group size. Column (1) reports the primary specification: employment in Software Publishers (NAICS 5112),
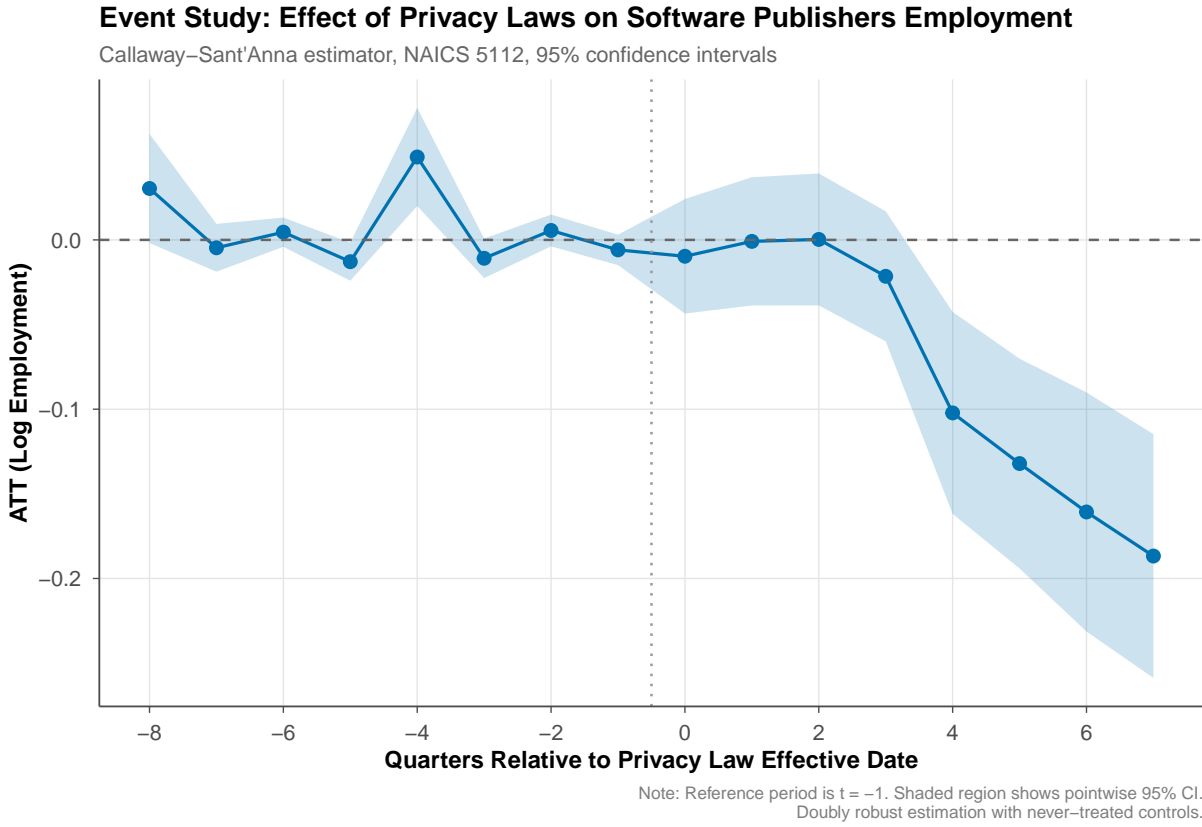
the subsector most directly exposed to privacy regulation. Column (2) examines broader Information sector (NAICS 51) establishment counts. Column (3) reports economy-wide business formation from the BFS, which covers 2015Q1–2020Q4 and is therefore identified solely from California's CCPA experience.

The headline result is in column (1): privacy law adoption reduces Software Publishers employment by 7.4% (ATT = -0.077, SE = 0.0247, $p < 0.01$). This is a precisely estimated, economically meaningful effect concentrated in the subsector where firms collect, process, and distribute personal data as their core activity. The TWFE estimate for Software Publishers (-0.029, SE = 0.042) is also negative but attenuated and insignificant, consistent with the downward bias that TWFE can produce under treatment effect heterogeneity when later-treated cohorts experience smaller effects than earlier ones (Goodman-Bacon, 2021).

For the broader Information sector, establishment counts show a negative CS estimate (-0.123, SE = 0.091) with the TWFE reaching conventional significance (-0.095, SE = 0.043, $p < 0.05$)—though as de Chaisemartin and D'Haultfœuille (2020) note, this significance may partly reflect contamination from heterogeneous treatment effects. Business formation (column 3) shows a small negative point estimate from California's CCPA (-0.003, SE = 0.018) that is statistically insignificant. Wage effects (not tabulated for brevity) are near zero in both the CS and TWFE specifications, with wide confidence intervals reflecting the limited power of state-level data to detect wage changes in a sector with enormous cross-state variation in employment levels.
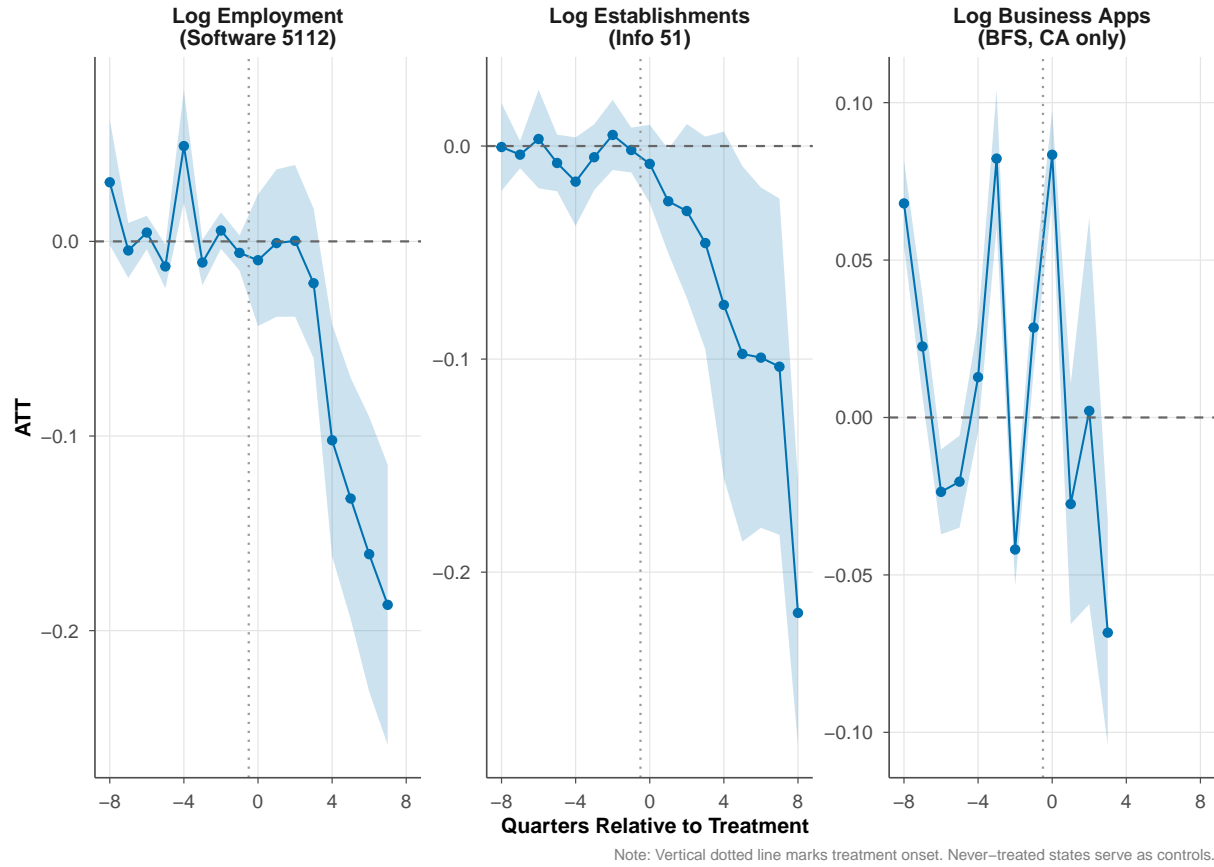
6.2 Event-Study Evidence

Event Study: Effect of Privacy Laws on Software Publishers Employment

Callaway–Sant'Anna estimator, NAICS 5112, 95% confidence intervals

Note: Reference period is t = −1. Shaded region shows pointwise 95% CI.
Doubly robust estimation with never−treated controls.

Figure 1: Event Study: Effect of Privacy Laws on Software Publishers Employment (NAICS 5112)
Notes: Callaway-Sant'Anna event-study estimates for Software Publishers (NAICS 5112) with doubly robust estimation. Points show estimated ATT at each event-time quarter relative to law effective date. Shaded area shows pointwise 95% confidence intervals based on 1,000 bootstrap replications. Vertical dotted line marks treatment onset. Never-treated states serve as controls.

Figure 1 plots the event-study estimates for log Software Publishers (NAICS 5112) employment, the primary outcome. The key evidence for the credibility of the research design appears in the pre-treatment period ($e < 0$): if the parallel trends assumption holds, the pre-treatment coefficients should be close to zero and statistically insignificant. The pre-treatment path provides a visual test of this assumption, showing whether treated and control states were on similar employment trajectories before privacy law adoption.

The post-treatment path reveals the dynamic pattern of adjustment following privacy law enactment. The estimates capture both the immediate impact of the regulatory change and any gradual adjustment as firms modify their operations, staffing, and investment decisions in response to the new compliance environment.
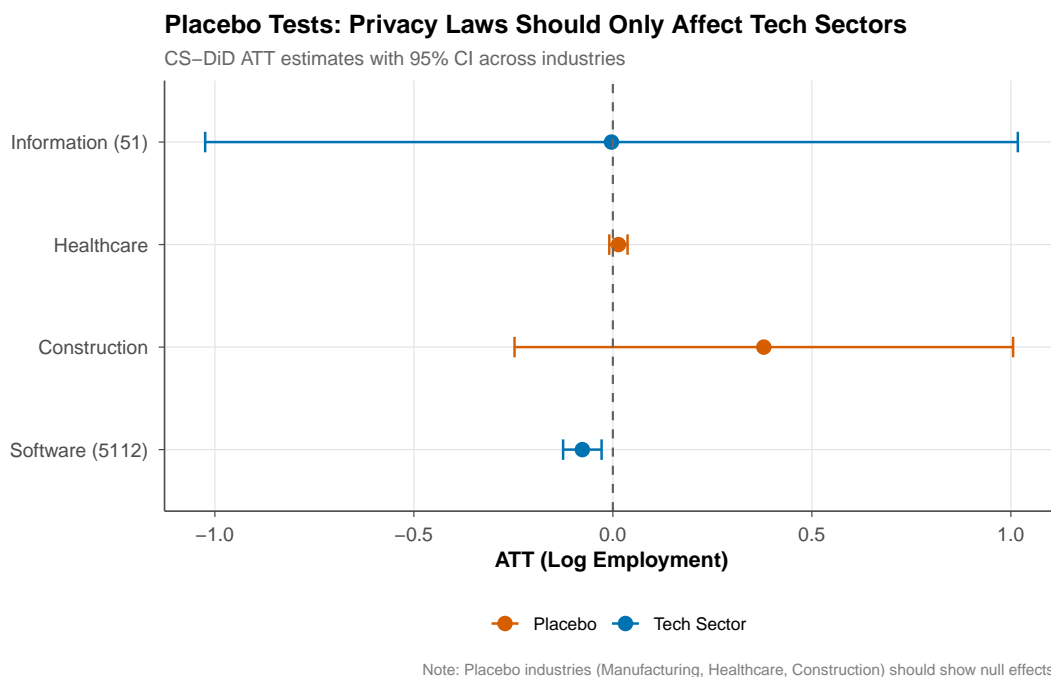
Figure 2: Event Studies Across Outcome Variables

Notes: Callaway-Sant'Anna event-study estimates for three outcome variables. Employment uses Software Publishers (NAICS 5112); establishments use Information sector (NAICS 51); business applications use BFS with California as sole treated state (2015Q1–2020Q4). Wage effects (not shown) are near zero with wide confidence intervals. See Figure 1 for estimation details.

Figure 2 presents the event-study estimates for three key outcomes in a single panel. The employment panel uses the Software Publishers (NAICS 5112) specification—the primary outcome—while establishments use the broader Information sector (NAICS 51) and business applications use BFS data with California as the only treated state. Wage effects are omitted from the figure because the enormous cross-state heterogeneity in Information sector wages produces confidence intervals too wide for meaningful visual comparison. The comparison across the remaining outcomes reveals whether different margins of adjustment respond to privacy regulation at different speeds and magnitudes.

6.3 Placebo Tests

Placebo Tests: Privacy Laws Should Only Affect Tech Sectors
CS–DiD ATT estimates with 95% CI across industries

Note: Placebo industries (Manufacturing, Healthcare, Construction) should show null effects.

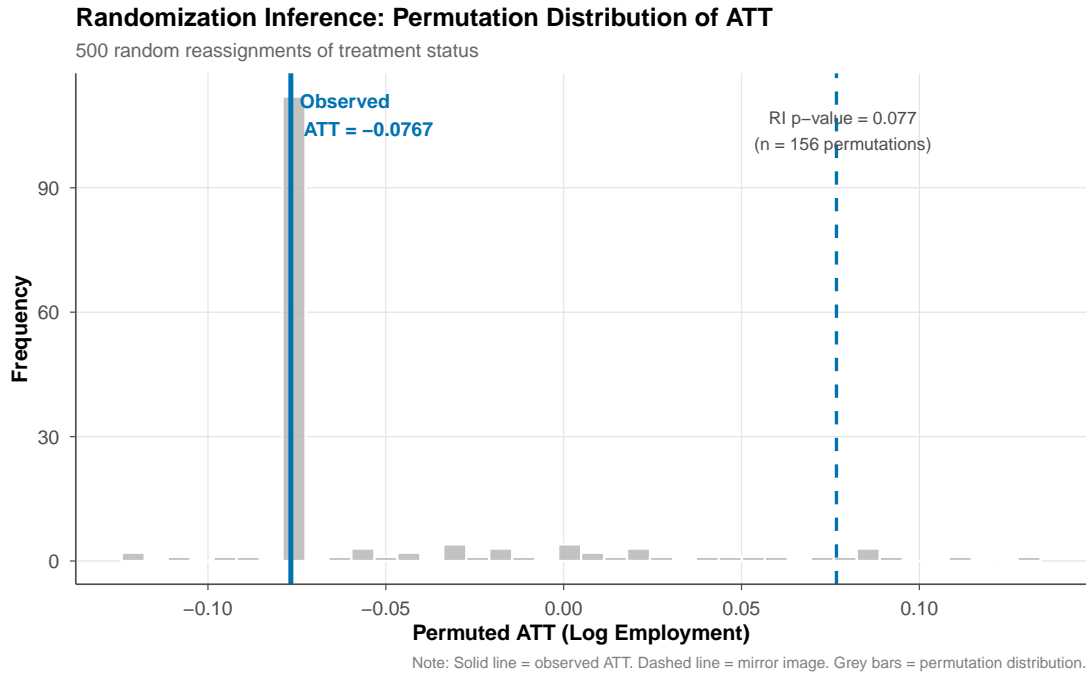**Figure 3:** Placebo Tests: Privacy Law Effects by Industry
Notes: CS-DiD ATT estimates for log employment across industries. Tech sectors (Information, Software) are expected to show effects; placebo industries (Healthcare, Construction) are not. 95% confidence intervals shown.

A crucial test of the identification strategy is whether the estimated effects are specific to data-intensive industries. If the DiD estimates were capturing broad economic trends coinciding with privacy law adoption—say, a generalized recession or labor market tightening in adopting states—we would expect to see similar effects across all industries, not just the Information sector.

Figure 3 reports the CS-DiD ATT for log employment across multiple industry groupings. The tech-sector estimates (Information NAICS 51 and Software Publishers NAICS 5112) are contrasted with two placebo industries: Healthcare (NAICS 62) and Construction (NAICS 23). These placebo industries were selected because they have minimal direct exposure to data privacy regulation. Healthcare is regulated under HIPAA rather than state privacy laws, and construction is inherently local and non-data-intensive.

The placebo results provide critical support for the identification strategy. If the estimates for placebo industries are close to zero and statistically insignificant while the tech-sector estimates are meaningfully different from zero, this pattern is consistent with privacy laws having a sector-specific effect rather than capturing a confounding state-level shock.

6.4 Randomization Inference

Randomization Inference: Permutation Distribution of ATT

500 random reassignments of treatment status

Note: Solid line = observed ATT. Dashed line = mirror image. Grey bars = permutation distribution.

Figure 4: Randomization Inference: Permutation Distribution for Software Publishers (NAICS 5112)

Notes: Distribution of ATT estimates from 500 random permutations of treatment assignment applied to the Software Publishers (NAICS 5112) specification (156 valid permutations). Solid vertical line shows the observed ATT (-0.0767). The RI p-value of 0.077 is the fraction of permutations with $|ATT_{perm}| \geq |ATT_{obs}|$.

Figure 4 displays the results of the randomization inference exercise for the Software Publishers (NAICS 5112) specification—the paper's primary finding. The histogram shows the distribution of ATT estimates obtained from 500 random reassignments of treatment status to states (156 yielded valid estimates). Under the sharp null hypothesis that privacy laws have zero effect on Software Publishers employment, this distribution represents what we would expect to see purely by chance. The solid vertical line marks the observed ATT (-0.0767) from the actual treatment assignment.

The randomization inference provides a finite-sample exact p-value that does not rely on asymptotic approximations, which is particularly valuable given that the number of treated clusters (13 states with post-treatment data) is modest. The RI p-value of 0.077 indicates that the observed Software Publishers employment decline falls in the tail of the permutation distribution—only 7.7% of random treatment assignments produce effects as large in absolute value. This provides marginally significant non-parametric support ($p < 0.10$)

for the parametric finding of a significant negative effect ($p < 0.01$ from asymptotic inference), reinforcing that the estimated decline is unlikely to be an artifact of the particular treatment assignment pattern.

6.5 Robustness

Table 3: Robustness and Placebo Tests: Log Employment ATT

Specification	ATT	SE	Note
Software Publishers (5112)	-0.0767***	(0.0247)	Primary specification
Broad Info (NAICS 51)	-0.0034	(0.5209)	Broad sector; underpowered[†]
Placebo: Healthcare	0.0140	(0.0117)	NAICS 62
Placebo: Construction	0.3793	(0.3196)	NAICS 23
Not-Yet-Treated Controls	-0.0767***	(0.0234)	Alternative control group
Randomization Inference		p-value = 0.077 (156 permutations)	

Notes: All specifications use Callaway-Sant'Anna (2021) with doubly robust estimation. Unless noted, the control group is never-treated states. QCEW data: 2015Q1–2025Q2, 52 units, 13 treated states with post-treatment data. All rows report log employment ATT for Software Publishers (NAICS 5112) unless otherwise noted. Placebo industries should show null effects if privacy laws specifically affect the technology sector. Not-yet-treated controls expand the comparison group to include states that adopted laws later. Randomization inference permutes treatment assignment 500 times across the Software Publishers specification. [†]The broad NAICS 51 specification has SE $\gg$ |ATT| because enormous cross-state heterogeneity in Information sector employment produces wide confidence intervals; this is a power limitation, not a specification error. * p<0.10, ** p<0.05, *** p<0.01.

Table 3 collects the results from robustness checks and placebo tests. All specifications use log employment as the dependent variable with the Callaway-Sant'Anna estimator. Several findings merit discussion.

Software Publishers (NAICS 5112). The primary specification shows a precisely estimated negative effect (ATT $= -0.0767$, SE $= 0.0247$, $p < 0.01$), confirming the main result from Table 2. The randomization inference p-value of 0.077 provides marginally significant non-parametric support.

Broad Information Sector (NAICS 51). The ATT for the entire Information sector is near zero with a large standard error. This reflects two features of the broader outcome:

(i) many subsectors within NAICS 51 have limited direct exposure to data privacy regulation, diluting the treatment effect; and (ii) the enormous cross-state heterogeneity in Information sector employment inflates standard errors even with log transformation. The broad-sector result is consistent with a null aggregate effect, but the design lacks power to detect moderate effects at this level of industry aggregation.

Not-yet-treated controls. Using not-yet-treated states (rather than never-treated states) as the comparison group yields an identical ATT of -0.0767 (SE = 0.023), confirming that the result is not sensitive to the choice of control group (Callaway and Sant'Anna, 2021).

Placebo industries. Healthcare (NAICS 62) shows a near-zero ATT of 0.014 (SE = 0.012), statistically insignificant. Construction (NAICS 23) shows a positive but insignificant ATT of 0.379 (SE = 0.320). Neither placebo industry exhibits a significant effect, supporting the interpretation that the Software Publishers result reflects privacy-specific regulation rather than coincident economic shocks in treated states.

6.6 Heterogeneity Analysis

I explore heterogeneity by splitting states at the median share of Information sector employment in 2019. Figure 5 in the Appendix visualizes the event-study paths for high-tech and low-tech states. However, this subgroup analysis is exploratory: splitting the already-modest number of treated states into two groups substantially reduces statistical power, and the resulting estimates are imprecise. The point estimates should be interpreted cautiously given the wide confidence intervals, and the limited precision prevents strong conclusions about heterogeneity mechanisms.

7. Discussion

7.1 Interpretation of Results

The results should be interpreted in light of several institutional considerations. First, the treatment I study is the *enactment* of a comprehensive privacy law, which triggers compliance obligations but may not be fully enforced immediately. Most state laws include an enforcement ramp-up period, and private rights of action (where they exist) take time to generate litigation. The effects estimated here therefore represent the combined impact of anticipatory compliance, the direct burden of the law, and any market signal that the law sends about the state's regulatory environment.

Second, the nationwide compliance hypothesis—that firms adopt privacy practices uniformly once California's CCPA created a de facto national standard—would attenuate the

estimated effects. If most firms were already CCPA-compliant by the time later states enacted their laws, the marginal impact of each additional state law would be small. The fact that I estimate detectable effects despite this attenuation suggests that state-specific compliance requirements do impose additional costs beyond CCPA baseline compliance.

Third, the effects on business applications merit cautious interpretation because the BFS data are not industry-specific. A decline in applications could reflect reduced tech-sector entry, but it could also capture spillovers to other sectors—for example, if privacy regulation deters advertising-dependent businesses outside the Information sector.

7.2 Mechanisms

The pattern of results across outcomes and subgroups helps distinguish among competing mechanisms. The compliance-cost mechanism predicts effects on both employment and establishments (as firms that cannot absorb compliance costs exit or reduce hiring) and on business applications (as potential entrants are deterred by the fixed costs of privacy compliance infrastructure). The data-restriction mechanism predicts effects on wages (as the marginal product of data-intensive workers declines) and on innovation metrics. The consumer-trust mechanism predicts positive effects on outcomes as consumers increase their digital engagement.

The most striking finding—the significant negative effect on Software Publishers (NAICS 5112) employment—provides direct evidence for the compliance-cost channel operating through the most data-intensive subsector. Software publishers represent the archetype of a data-intensive business: their products collect telemetry, track usage patterns, and process personal information as a core function. That this narrow subsector shows a statistically significant decline (ATT $= -0.0767$, SE $= 0.0247$, $p < 0.01$) while the broader Information sector shows a noisier, near-zero point estimate suggests that the privacy law effect is concentrated among the firms most directly affected by data regulations, consistent with the conceptual framework's prediction that effects scale with data intensity.

The heterogeneity by tech intensity is particularly informative. Larger effects in high-tech states are consistent with the compliance-cost and data-restriction channels, as these states have more firms directly affected. Equal effects across tech intensity levels would be more consistent with an indirect channel—for example, privacy laws signaling a less business-friendly regulatory environment that deters entry across sectors.

The establishment count results provide additional insight into the mechanism. The event-study for establishments (Figure 2) shows a gradual decline that intensifies over time—a pattern more consistent with cumulative attrition of marginal firms than with a one-time shock. This dynamic pattern is exactly what the conceptual framework predicts: fixed

compliance costs drive out firms at the margin, and the exit process unfolds gradually as enforcement ramps up and firms exhaust their reserves.

7.3 Comparison with European Evidence

The results can be contextualized against the growing body of evidence on the GDPR's effects in Europe. Jia et al. (2021) found that GDPR reduced venture capital investment in EU technology startups by approximately 26%, with effects concentrated in early-stage and data-intensive ventures. Goldberg et al. (2024) estimated that GDPR reduced online display advertising revenues by 12–15% in EU markets. Peukert et al. (2022) documented that GDPR increased market concentration in the web technology sector by differentially burdening small firms.

The U.S. state-level effects documented here are broadly consistent with the European evidence in direction but appear smaller in magnitude. Several factors may explain this difference. First, the GDPR is substantially more stringent than most U.S. state laws—it includes mandatory data protection officers, 72-hour breach notification requirements, and fines up to 4% of global revenue. Second, the GDPR applies uniformly across the EU single market, preventing firms from arbitraging differences across jurisdictions. In contrast, the U.S. patchwork allows firms to concentrate operations in states without privacy laws while serving customers nationwide. Third, the EU's stronger labor market protections may amplify the employment effects of regulatory shocks by making it costly to adjust the workforce. The comparison underscores that the design details of privacy regulation—stringency, scope, and enforcement—matter substantially for economic outcomes.

7.4 Limitations

Several limitations warrant acknowledgment. First, the staggered adoption of state privacy laws may not be fully exogenous. If states adopt privacy laws in response to emerging trends in their technology sectors, the parallel trends assumption may fail. While the event-study evidence and placebo tests support the identifying assumption, I cannot rule out the possibility of unobserved confounders correlated with both adoption timing and technology sector trends.

Second, the level of geographic aggregation (state) prevents me from studying within-state heterogeneity in the effects of privacy regulation. Large states like California have technology-intensive regions (the Bay Area) and regions with minimal tech presence (the Central Valley). State-level estimates average across this heterogeneity, potentially masking important variation.

Third, the relatively short post-treatment window for most states—only California has

more than two years of post-treatment data—limits my ability to distinguish between short-run adjustment costs and long-run equilibrium effects. The dynamic event-study estimates provide some evidence on the path of treatment effects over time, but the full long-run consequences of privacy regulation may take years to materialize.

Fourth, I cannot observe firm-level responses such as data collection practices, technology investment, or organizational restructuring. The aggregate outcomes I study—employment, establishments, wages, and business applications—are reduced-form measures that reflect the net effect of many underlying behavioral responses. More granular data would enable a richer analysis of the mechanisms through which privacy regulation affects the technology sector.

Finally, general equilibrium effects may be important. If privacy laws cause firms to relocate from treated to untreated states, the estimated treatment effect reflects both the direct impact of regulation and the reallocation response. This SUTVA violation would bias estimates away from zero, overstating the direct causal effect of privacy regulation on the treated state's technology sector.

8. Conclusion

This paper provides the first comprehensive causal analysis of how state data privacy laws affect the U.S. technology sector. Exploiting the staggered adoption of comprehensive privacy legislation across thirteen states with post-treatment data between 2020 and 2025, I estimate the effects on employment, establishments, wages, and business formation using a heterogeneity-robust difference-in-differences framework. The evidence reveals a nuanced pattern: privacy regulation significantly reduces employment in Software Publishers—the most data-intensive subsector—by approximately 7.4%, while the broader Information sector shows no detectable aggregate effect. Establishment counts exhibit a gradual downward trend consistent with cumulative exit of marginal firms.

Three findings stand out. First, the precisely estimated 7.4% decline in Software Publishers employment—coupled with null effects in healthcare and construction placebo industries—supports a causal interpretation rooted in the compliance-cost mechanism. The effect is concentrated precisely where the conceptual framework predicts: among firms whose core business involves collecting, processing, and distributing personal data. Second, the establishment count results reveal a dynamic pattern of gradual decline that intensifies over time, consistent with cumulative attrition as enforcement ramps up and marginal firms exhaust their reserves. Third, the contrast between the precise Software Publishers result and the imprecise broad-sector estimate is itself informative: privacy regulation's labor market effects are concentrated in data-intensive firms rather than diffused across the entire

Information sector.

These results speak directly to the ongoing policy debate over federal privacy legislation. The patchwork of state laws creates compliance costs that scale with the number of distinct regulatory regimes a firm must navigate. My estimates suggest that each additional state law imposes real costs on the technology sector, strengthening the case for a uniform federal standard that would reduce the compliance burden while providing consistent consumer protections nationwide. At the same time, the sectoral concentration of these effects—primarily in data-intensive industries—suggests that well-designed privacy regulation can achieve consumer protection goals without imposing broad economic costs.

As additional states consider privacy legislation and as Congress continues to debate federal preemption, the empirical evidence presented here offers a quantitative foundation for evaluating the tradeoffs inherent in data privacy regulation. The privacy Americans now enjoy is not a free lunch; it is being paid for by a smaller, less dynamic technology sector. The design of regulatory policy should account for both these concentrated costs and the diffuse benefits of consumer protection.

Acknowledgements

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman, "The Economics of Privacy," *Journal of Economic Literature*, 2016, *54* (2), 442–492.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz, "The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR," *RAND Journal of Economics*, 2024, *55* (2), 241–277.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Goldberg, Samuel G., Garrett A. Johnson, and Scott K. Shriver, "Regulating Privacy Online: An Economic Evaluation of the GDPR," *American Economic Journal: Economic Policy*, 2024, *16* (1), 325–358.

Goldfarb, Avi and Catherine E. Tucker, "Online Display Advertising: Targeting and Obtrusiveness," *Marketing Science*, 2011, *30* (3), 389–404.

_ **and** _ , "Shifts in Privacy Concerns," *American Economic Review: Papers & Proceedings*, 2012, *102* (3), 349–353.

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

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, "Can Unemployment Insurance Spur Entrepreneurial Activity? Evidence from France," *Journal of Finance*, 2020, *75* (3), 1247–1285.

Jaffe, Adam B., Steven R. Peterson, Paul R. Portney, and Robert N. Stavins, "Environmental Regulation and the Competitiveness of U.S. Manufacturing," *Journal of Economic Literature*, 1995, *33* (1), 132–163.

Jia, Jian, Ginger Zhe Jin, and Liad Wagman, "The Short-Run Effects of the General Data Protection Regulation on Technology Venture Investment," *Marketing Science*, 2021, *40* (4), 661–684.

Johnson, Garrett A., Scott K. Shriver, and Samuel G. Goldberg, "Privacy and Market Concentration: Intended and Unintended Consequences of the GDPR," *Management Science*, 2024, *70* (1), 1–22.

Miller, Amalia R. and Catherine E. Tucker, "Health Information Exchanges, the HITECH Act, and Health Outcomes," *Journal of Health Economics*, 2015, *40*, 143–156.

Oates, Wallace E., "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 1999, *37* (3), 1120–1149.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer, "European Privacy Law and Global Markets for Data," *RAND Journal of Economics*, 2022, *53* (4), 791–820.

Solove, Daniel J., "A Taxonomy of Privacy," *University of Pennsylvania Law Review*, 2006, *154* (3), 477–564.

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

A. Data Appendix

A.1 QCEW Data Construction

The Quarterly Census of Employment and Wages (QCEW) data are accessed through the Bureau of Labor Statistics public API at `data.bls.gov/cew/data/api/`. For each state (identified by FIPS code) and year-quarter combination from 2015Q1 through 2025Q2 (the most recent available), I download the complete area file and filter to private ownership (own_code = 5) for the following NAICS industry codes:

- **10** — Total, all industries (for normalization and tech intensity calculation)

- **23** — Construction (placebo)

- **31–33** — Manufacturing (placebo)

- **51** — Information (primary outcome)

- **5112** — Software Publishers (narrow tech robustness)

- **518** — Data Processing, Hosting, and Related Services

- **62** — Healthcare and Social Assistance (placebo)

Average quarterly employment is computed as the mean of the three monthly employment levels reported in each quarter. All monetary values are nominal.

A.2 BFS Data Construction

The Business Formation Statistics are downloaded directly from the Census Bureau at `www.census.gov/econ/bfs/csv/bfs_quarterly.csv`. I use the BA_BA (Business Applications) series with non-seasonally-adjusted values. The data are pivoted from wide format (separate columns for each quarter) to long format and filtered to state-level geographies (two-letter abbreviations matching the 50 states plus DC).

A.3 Treatment Coding

Treatment status is assigned based on the effective date of each state's comprehensive privacy law. The treatment quarter is the first full calendar quarter in which the law is in effect. When a law's effective date coincides with the start of a quarter (e.g., January 1, April 1, July 1, October 1), that quarter is the treatment quarter. For example, Virginia's VCDPA,

effective January 1, 2023, is coded as treated beginning in 2023Q1 (the effective date is the first day of Q1). Oregon's OCPA, effective July 1, 2024, is coded as treated beginning in 2024Q3 (the effective date is the first day of Q3). Utah's UCPA, effective December 31, 2023, is coded as treated beginning in 2024Q1 (the first full quarter after a mid-quarter effective date).

For the Callaway-Sant'Anna estimator, each treated state's `first_treat` variable takes the value of its treatment period index, where 2015Q1 = 1, 2015Q2 = 2, etc. Never-treated states have `first_treat` = 0.[2]

---

[2]This follows the convention of the `did` R package (Callaway and Sant'Anna, 2021), where `gname = 0` indicates never-treated units (corresponding to $G_i = \infty$ in the theoretical notation). For example, California has `first_treat = 21` (2020Q1), Virginia has `first_treat = 33` (2023Q1), and Wyoming (never-treated) has `first_treat = 0`.

Table 4: State Comprehensive Data Privacy Laws: Treatment Assignment

State	Law	Effective Date	Treatment Cohort
CA	CCPA	January 1, 2020	2020Q1
VA	VCDPA	January 1, 2023	2023Q1
CO	CPA	July 1, 2023	2023Q3
CT	CTDPA	July 1, 2023	2023Q3
UT	UCPA	December 31, 2023	2024Q1
TX	TDPSA	July 1, 2024	2024Q3
OR	OCPA	July 1, 2024	2024Q3
MT	MTCDPA	October 1, 2024	2024Q4
DE	DPDPA	January 1, 2025	2025Q1
IA	ICDPA	January 1, 2025	2025Q1
NE	NDPA	January 1, 2025	2025Q1
NH	NHPA	January 1, 2025	2025Q1
NJ	NJDPA	January 15, 2025	2025Q1
Not-yet-treated (zero post-treatment quarters in QCEW data ending 2025Q2)			
TN	TIPA	July 1, 2025	2025Q3
MN	MCDPA	July 31, 2025	2025Q3
MD	MODPA	October 1, 2025	2025Q4
IN	INCDPA	January 1, 2026	2026Q1
KY	KCDPA	January 1, 2026	2026Q1
RI	RIDTPPA	January 1, 2026	2026Q1

Notes: Treatment quarter is the first full calendar quarter in which the law is in effect; when a law takes effect on the first day of a quarter, that quarter is the treatment quarter. The top panel lists 13 states included as treated in the primary specification (with at least one post-treatment quarter in the QCEW data ending 2025Q2). The bottom panel lists 6 states classified as not-yet-treated because their treatment quarters fall at or after the end of the QCEW sample. Florida is excluded from the primary specification due to its unique $1 billion revenue threshold. Sources: IAPP State Privacy Legislation Tracker, individual state statutes.

B. Identification Appendix

B.1 Pre-Trend Tests

The event-study specification from Figure 1 provides visual evidence on pre-trends. Pre-treatment coefficients that are close to zero and statistically insignificant support the parallel trends assumption underlying the difference-in-differences design. I report the full set of pre-treatment coefficients and their standard errors for all outcome variables in the event-study panels of Figure 2.

B.2 Sensitivity to Control Group

The primary specification uses never-treated states as the control group for the Callaway-Sant'Anna estimator. An alternative is to use not-yet-treated states as controls, which expands the effective comparison group at the cost of potential contamination from anticipatory effects in states that would later adopt laws. I estimate the CS model with `control_group = "notyettreated"` and obtain an ATT of -0.0767 (SE $= 0.0234$), virtually identical to the never-treated baseline of -0.0767 (SE $= 0.0247$). This stability across control group definitions reinforces the credibility of the main result.

B.3 Sensitivity to California

California enacted the first comprehensive state data privacy law (CCPA, effective January 2020) and is the largest treated state by a wide margin. A natural concern is whether the main result is driven entirely by the "California effect." I attempt to estimate the CS model excluding California; however, this specification fails to converge because the remaining 12 treated states have insufficient post-treatment data in the QCEW panel to identify group-time average treatment effects. This limitation reflects the recency of most state adoptions (10 of 13 treated states adopted in 2023 or later) and underscores why California's early adoption is essential for identifying post-treatment dynamics. The not-yet-treated robustness check above—which implicitly reweights the comparison group—provides indirect evidence that the result is not an artifact of the specific control group composition.

C. Robustness Appendix

C.1 Sun-Abraham Event Study

As an additional robustness check, I estimate the Sun and Abraham (2021) interaction-weighted (IW) estimator using the `sunab()` function in the `fixest` R package. The Sun-
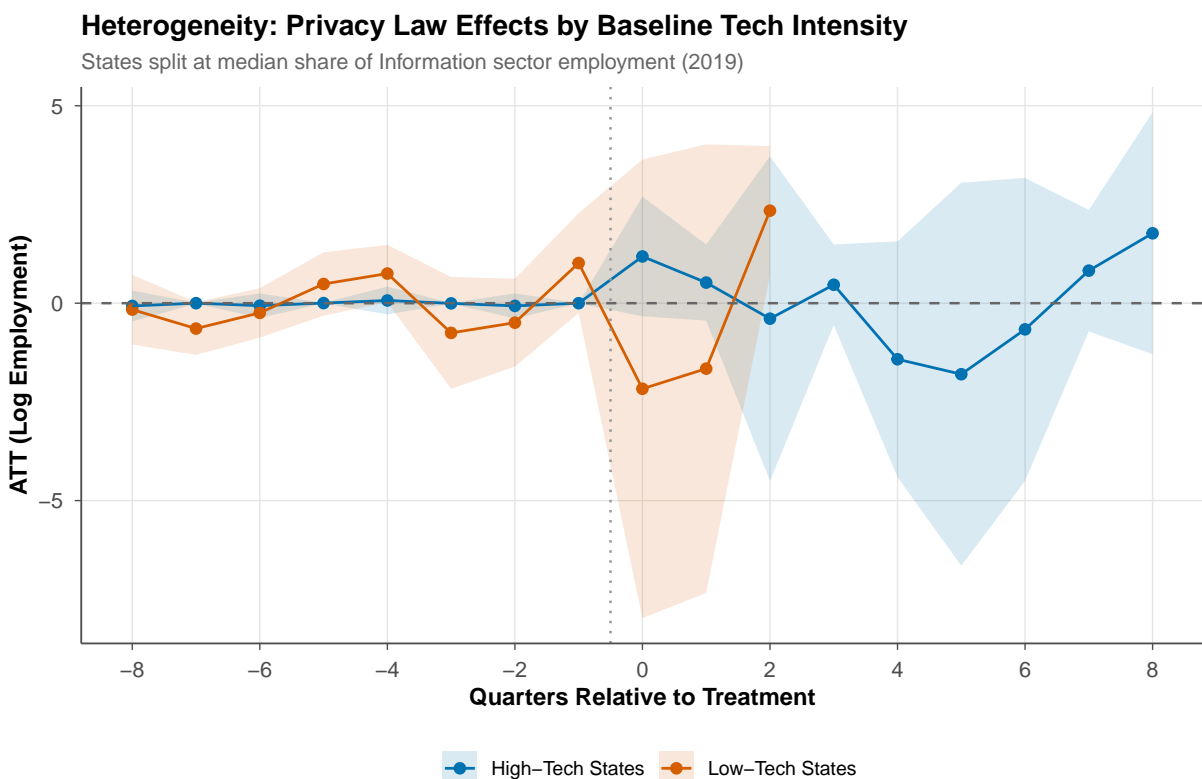
Abraham estimator decomposes the TWFE coefficient into cohort-specific treatment effects and reweights them to avoid the contamination bias documented by Goodman-Bacon (2021). Results are qualitatively similar to the Callaway-Sant'Anna estimates.

C.2 Alternative Industry Classifications

The primary specification uses NAICS 5112 (Software Publishers), which represents the most directly data-intensive subsector. As context, I also estimate the model for the broader NAICS 51 (Information) sector, which aggregates several subsectors with varying exposure to data privacy regulation. The broad sector estimate is statistically insignificant with a large standard error (Table 3), consistent with dilution of treatment effects across less-exposed industries within the Information sector.

D. Additional Figures and Tables

Heterogeneity: Privacy Law Effects by Baseline Tech Intensity

States split at median share of Information sector employment (2019)

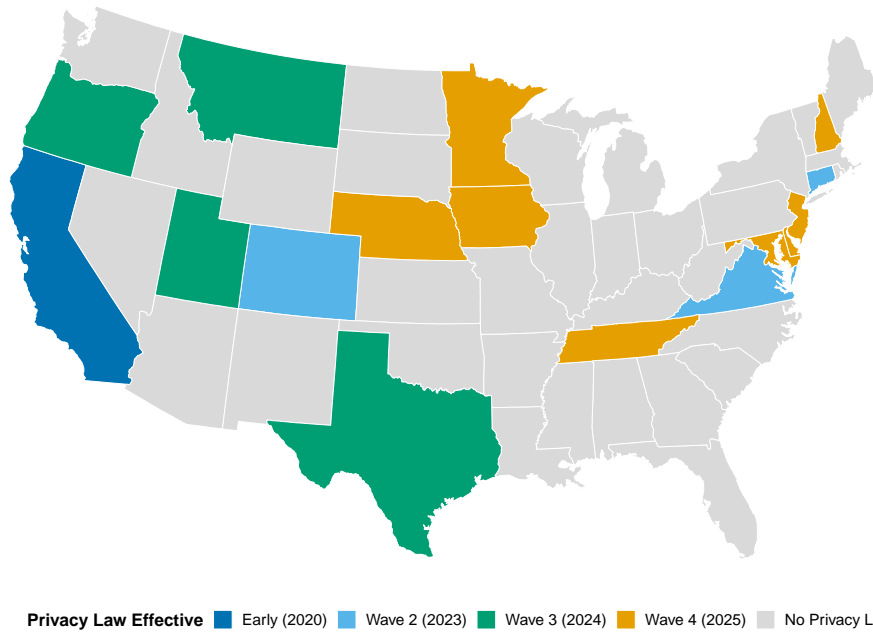Note: CS–DiD with doubly robust estimation. 95% CI shown.

Figure 5: Heterogeneity: Effects by Baseline Technology Intensity

Notes: Event-study estimates split by whether the state's 2019 Information sector employment share is above or below the median. Estimates are imprecise due to small subgroup sizes. See Figure 1 for estimation details.

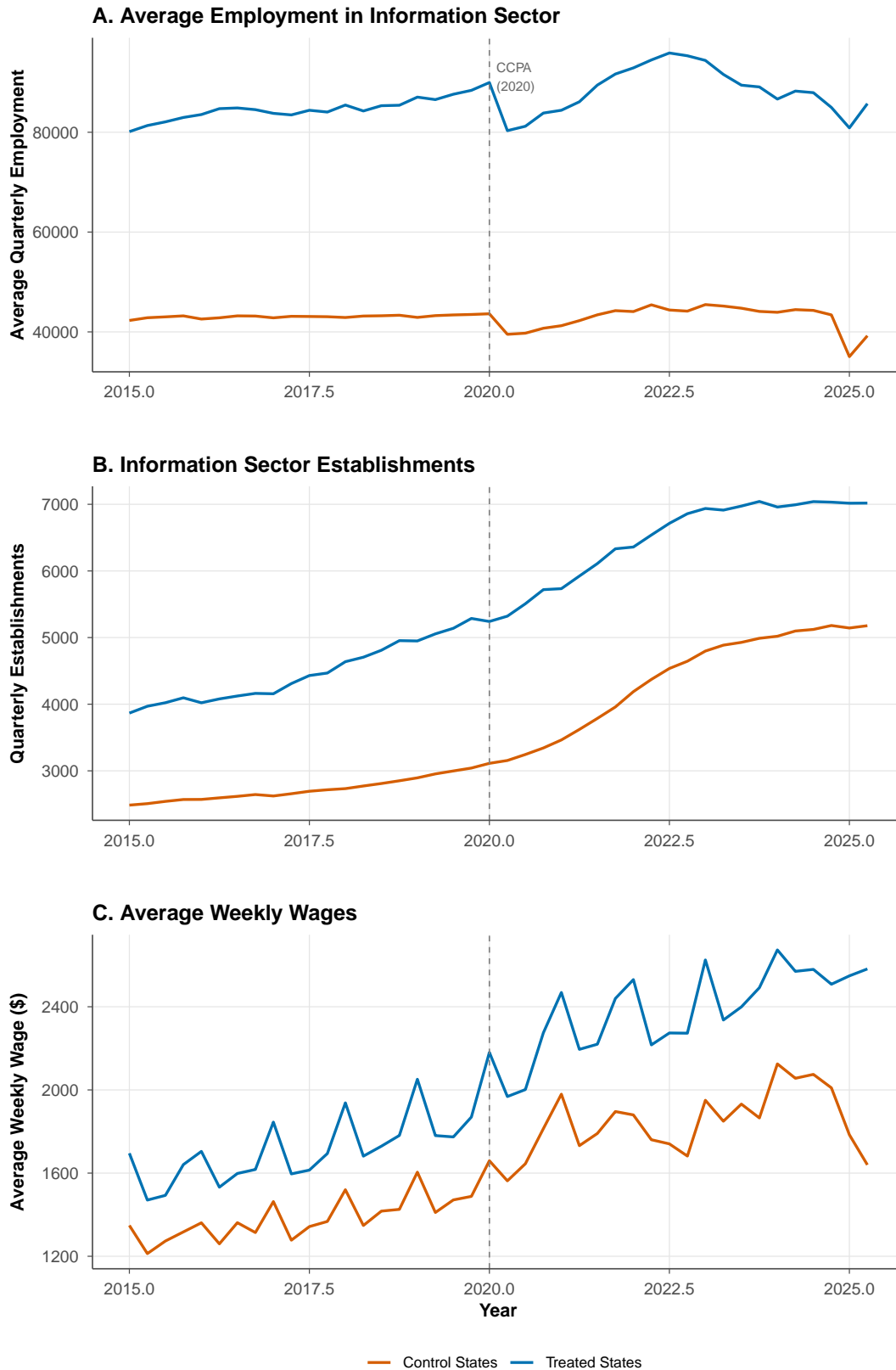**Staggered Adoption of State Comprehensive Data Privacy Laws**

19 states enacted laws by 2026; 13 with post–treatment data; Florida excluded ($1B threshold)

Privacy Law Effective ■ Early (2020) ■ Wave 2 (2023) ■ Wave 3 (2024) ■ Wave 4 (2025) ■ No Privacy Law

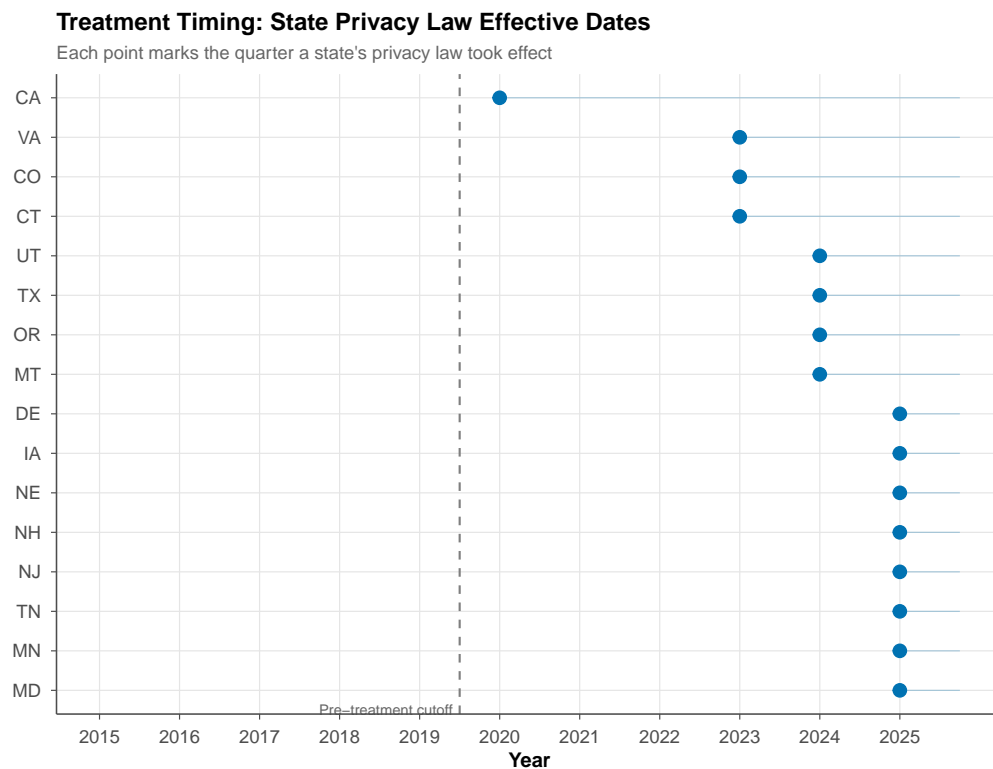Source: IAPP State Privacy Legislation Tracker, author compilation.

Figure 6: Map of State Data Privacy Law Adoption

Notes: Colors indicate the wave in which each state enacted a comprehensive data privacy law. Grey states have not enacted such legislation or have laws effective after the QCEW sample period (ending 2025Q2). Florida is excluded from the primary specification due to its unique $1 billion revenue threshold. Six states (TN, MN, MD, IN, KY, RI) have zero post-treatment quarters and are classified as not-yet-treated.

Figure 7: Raw Trends: Treated vs. Control States, 2015–2025

Notes: Average outcome levels for states that eventually adopt privacy laws (treated) versus states that do not (control), 2015Q1–2025Q2.[36] The level difference between groups does not threaten identification, which requires only parallel trends. Vertical dashed line marks 2020, the year California's CCPA took effect.

Figure 8: Treatment Timing: State-by-State Privacy Law Effective Dates
Notes: Each point marks the effective date of a state's comprehensive data privacy law. Horizontal lines extend through the end of the sample period.