

Digital Exodus or Digital Magnet?

How State Data Privacy Laws Reshape the Technology Sector*

APEP Autonomous Research[†]

@SocialCatalystLab

@olafdrw

February 8, 2026

Abstract

Do state data privacy laws drive technology firms away, or do they reshape the technology sector’s composition? Between 2020 and 2025, nineteen U.S. states enacted comprehensive consumer data privacy legislation modeled on California’s landmark CCPA, though only seven have effective dates with sufficient post-treatment data within our 2015–2024 data window (California, Virginia, Colorado, Connecticut, Utah, Oregon, and Texas); the remaining twelve states’ laws either take effect in 2025 or later or provide insufficient post-treatment exposure (Montana, effective October 2024, contributes at most one quarter) and are coded as not-yet-treated. This staggered adoption creates substantial cross-state variation in the regulatory environment facing data-intensive industries. Using a staggered difference-in-differences design with the Callaway-Sant’Anna (2021) estimator, we examine the effect of these laws on technology-sector employment and new business formation. Drawing on quarterly state-level employment data from the BLS Quarterly Census of Employment and Wages and state-level business applications from the Census Bureau’s Business Formation Statistics covering 2015–2024, we find evidence that privacy laws reduce employment in data-intensive Software Publishing (NAICS 5112) by approximately 7.7%—driven largely by California’s experience as the first mover with the longest post-treatment

*This paper is a revision of apecp_0194. See https://github.com/SocialCatalystLab/apecp-papers/tree/main/apecp_0194.

[†]Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

window—with a modest negative effect on Computer Systems Design (NAICS 5415), while the net effect on total Information Sector (NAICS 51) employment is statistically indistinguishable from zero. The aggregate null masks uneven regulatory costs across subsectors: the burden falls disproportionately on software publishers, consistent with compliance costs affecting data-intensive firms most heavily. Business formation data show no significant aggregate effect on total new business applications. These results suggest that privacy regulation does not destroy technology jobs *per se* but that regulatory costs fall unevenly across technology subsectors, with data-intensive software publishing bearing the brunt while the broader Information Sector adjusts through compositional reallocation rather than net contraction. Our findings have direct implications for the ongoing federal data privacy debate, suggesting that aggregate employment effects are modest but distributionally consequential.

JEL Codes: J21, L51, L86, K24, O33

Keywords: data privacy, technology employment, regulatory sorting, business formation, CCPA, difference-in-differences

1. Introduction

For two decades, the American technology sector built its business models on a simple premise: collect data freely, monetize it aggressively, and face little regulatory consequence. That era ended in 2020 when California enacted its landmark Consumer Privacy Act. Since then, eighteen additional states have followed suit, creating a patchwork of data privacy laws with effective dates staggered from 2023 through 2026.¹ These laws grant consumers rights to access, delete, and opt out of the sale of their personal data—rights that impose real compliance costs on firms whose core business is collecting and processing that data. For ad-tech platforms, data brokers, and behavioral targeting firms, privacy regulation is not an abstract governance debate; it is a direct tax on their primary business model.

The conventional wisdom is straightforward: privacy regulation raises costs, so tech firms will shrink or flee regulated states. This view treats the technology sector as a monolith—as if a twenty-person data brokerage in Austin faces the same compliance burden as a cloud infrastructure provider in Seattle. It also assumes the relevant margin is extensive (firms leaving) rather than intensive (firms changing what they do). The empirical evidence has been sparse: existing studies focus primarily on the European GDPR or on California alone, leaving open the question of how privacy regulation affects the broader U.S. technology landscape when adopted by a diverse set of states.

This paper provides the first systematic examination of how staggered state data privacy law adoption affects the *composition* of the technology sector. Our central hypothesis is that privacy laws impose uneven regulatory costs across technology subsectors, with the heaviest burden falling on data-intensive firms. Firms whose business models depend on extensive data collection and monetization—ad-tech platforms, data brokers, behavioral targeting firms—face increased compliance costs and may reduce operations. Compliance-oriented firms—cybersecurity providers, privacy consultants, compliance technology firms—might benefit from increased demand for their services, but the net direction of these competing forces is an empirical question. If within-sector reallocation partially offsets the costs to data-intensive firms, aggregate measures of technology employment may show small or null effects even as meaningful compositional changes occur beneath the surface.

We test this hypothesis using two complementary datasets. First, we use the Bureau of Labor Statistics’ Quarterly Census of Employment and Wages (QCEW), which provides

¹Of these nineteen states, seven have laws that took effect within our 2015Q1–2024Q4 sample window with sufficient post-treatment exposure—California (2020Q1), Virginia (2023Q1), Colorado (2023Q3), Connecticut (2023Q3), Utah (2024Q1), Oregon (2024Q3), and Texas (2024Q3). Montana’s law took effect October 1, 2024, yielding at most one quarter of post-treatment data with zero compliance lead time; we reclassify it as not-yet-treated. The remaining twelve states have effective dates after our data endpoint.

employment and wage data at the state-by-quarter-by-industry level. The QCEW’s detailed NAICS industry codes allow us to separately examine sectors that should be differentially affected by privacy regulation: Information (NAICS 51), which encompasses the broad technology-adjacent sector; Software Publishers (NAICS 5112), which captures data-intensive software firms; Computer Systems Design and Related Services (NAICS 5415), which includes IT consulting and compliance services; and control sectors like Construction (23) and Finance (52) that should be unaffected. Second, we use the Census Bureau’s Business Formation Statistics (BFS), which tracks new business applications at the state level, allowing us to examine the effect of privacy laws on overall business formation.

Our identification strategy exploits the staggered adoption of comprehensive state data privacy laws. While nineteen states have enacted such laws, our identification relies on the seven states whose laws took effect within the 2015Q1–2024Q4 sample window with sufficient post-treatment exposure; the remaining twelve enacted-but-not-yet-effective states (including Montana, whose single quarter of post-treatment data is insufficient for credible inference) serve as part of the never-treated comparison group alongside thirty-two states (including DC) that have not enacted privacy legislation. We implement the Callaway and Sant’Anna (2021) heterogeneity-robust difference-in-differences estimator, which avoids the well-documented biases of traditional two-way fixed effects (TWFE) models in staggered adoption settings (Goodman-Bacon, 2021; Sun and Abraham, 2021; de Chaisemartin and d’Haultfoeuille, 2020). We use never-treated states as the comparison group and report event-study estimates showing the dynamic path of treatment effects. To address concerns about the parallel trends assumption, we attempt the Rambachan and Roth (2023) HonestDiD sensitivity analysis, though convergence is limited by our small cohort sizes. We further probe robustness with Sun and Abraham (2021) interaction-weighted estimates, Fisher randomization inference, and placebo tests on sectors that should be unaffected by privacy regulation.

Preview of Results. We find that privacy laws significantly reduce employment in Software Publishing (NAICS 5112) by approximately 7.7%, with a modest negative effect on Computer Systems Design (NAICS 5415). Total Information Sector employment shows a precisely estimated null effect under our preferred CS-DiD estimator, indicating that the aggregate sector adjusts through compositional reallocation rather than net job destruction. Business formation data show no significant aggregate effect on total new business applications. These results are robust to alternative estimators, placebo tests, and randomization inference.

Contribution. This paper makes three contributions to the literature on regulation and industry dynamics. First, we document that privacy regulation imposes uneven costs across technology subsectors—reducing employment in data-intensive software publishing while leaving the broader Information Sector aggregate largely unchanged—introducing the

concept of *regulatory sorting* to the data privacy debate. While sorting in response to environmental regulation is well-established (Greenstone, 2002), the mechanism has not been applied to information technology regulation. Second, we provide the first multi-state causal analysis of U.S. data privacy legislation’s employment effects, extending the evidence base beyond California and the EU. Third, our finding that aggregate null effects mask significant subsector-level declines has direct implications for cost-benefit analyses of proposed federal privacy legislation, which typically assume homogeneous industry responses.

2. Institutional Background

2.1 The Patchwork of State Data Privacy Laws

The United States lacks a comprehensive federal data privacy law. Unlike the European Union, which adopted the General Data Protection Regulation (GDPR) as a unified framework in 2018, the U.S. has relied on a sectoral approach—with separate laws covering health data (HIPAA), financial data (GLBA), children’s data (COPPA), and credit data (FCRA)—supplemented by state-level consumer protection authority. This regulatory vacuum created an opening for state legislatures to enact their own comprehensive privacy frameworks.

California was the first mover. The California Consumer Privacy Act (CCPA), signed in June 2018 and effective January 1, 2020, grants California residents the right to know what personal information businesses collect about them, the right to delete that information, and the right to opt out of the sale of their data. The law applies to for-profit businesses that meet specific revenue, data-processing, or data-sale thresholds. The California Privacy Rights Act (CPRA), approved by voters in November 2020 and substantively effective January 1, 2023, strengthened the CCPA by adding data minimization requirements, creating a dedicated enforcement agency (the California Privacy Protection Agency), and expanding consumer rights to include data correction and limits on automated decision-making.

Virginia became the second state to enact comprehensive privacy legislation, signing the Virginia Consumer Data Protection Act (VCDPA) in March 2021 with an effective date of January 1, 2023. Colorado and Connecticut followed in 2021 and 2022, respectively, with mid-2023 effective dates. Utah’s Consumer Privacy Act, enacted in March 2022, took effect at the end of 2023. Between 2023 and 2024, fourteen additional states enacted comprehensive privacy laws with effective dates spanning 2024 through 2026: Iowa, Indiana, Tennessee, Montana, Oregon, Texas, Delaware, New Hampshire, New Jersey, Kentucky, Nebraska, Maryland, Minnesota, and Rhode Island.

2.2 Key Provisions and Variation

While all nineteen state laws share a common structure—consumer rights (access, deletion, opt-out), business obligations (data inventories, privacy notices, impact assessments), and enforcement mechanisms—they differ in important details that create variation in regulatory stringency. California’s framework is generally considered the most comprehensive, with broad applicability thresholds, a private right of action (limited to data breaches), and a dedicated enforcement agency. Virginia’s law, by contrast, has narrower applicability thresholds and relies exclusively on the Attorney General for enforcement. Texas’s law applies to all businesses without revenue or data-volume thresholds, making it potentially the broadest in scope, while Utah’s law has the narrowest applicability criteria.

These differences in stringency are relevant for our analysis because they generate variation in the “dosage” of privacy regulation across states, which we exploit in heterogeneity analyses. We classify state laws as “strong” (California, Colorado, Oregon, Connecticut) or “standard” (remaining states) based on the breadth of consumer rights, enforcement mechanisms, and business applicability thresholds. Our main results use a binary treatment indicator (any comprehensive privacy law), while robustness checks explore dose-response relationships.

2.3 Industry Composition of the Technology Sector

The technology sector is not monolithic. For our purposes, the critical distinction is between *data-intensive* and *privacy-enhancing* technology firms. Data-intensive firms—including advertising technology platforms, data brokers, consumer analytics companies, and social media firms—rely on the collection, aggregation, and monetization of personal data as a core business function. These firms face the most direct regulatory costs from privacy laws: they must implement consent mechanisms, honor deletion requests, provide data portability, and limit data uses to disclosed purposes.

Privacy-enhancing firms—including cybersecurity providers, encrypted communications developers, identity management platforms, and compliance technology (“privacy tech”) companies—produce goods and services that help consumers and businesses protect personal data. For these firms, privacy regulation creates *demand* rather than imposing costs: when a state enacts a privacy law, businesses in that state need compliance tools, privacy impact assessment services, data governance platforms, and security infrastructure. The privacy tech market has grown from \$2.8 billion in 2020 to an estimated \$19.2 billion in 2025, driven in part by the expanding patchwork of state regulations that creates compliance complexity.

The NAICS classification system imperfectly captures this distinction. Software publishers (NAICS 5112) includes both data-intensive and privacy-enhancing software firms, but the

subsector’s reliance on data collection and monetization makes it a reasonable proxy for data-intensive technology. Computer systems design and related services (NAICS 5415) most closely captures privacy consulting and compliance services, but also includes general IT consulting. Despite these measurement challenges, the direction of bias from industry-code imprecision is toward attenuation (null findings), making our subsector results conservative estimates of the true compositional effects.

3. Theoretical Framework: Regulatory Sorting

We develop a simple framework to formalize the sorting hypothesis and derive testable predictions. The framework builds on the location choice literature (Greenstone, 2002; Kahn, 2000) and the regulatory competition literature (Oates and Portney, 2001).

3.1 Setup

Consider a continuum of technology firms indexed by their data intensity $\theta \in [0, 1]$, where $\theta = 0$ represents a firm with no personal data use (e.g., a hardware manufacturer) and $\theta = 1$ represents a firm entirely dependent on personal data monetization (e.g., a data broker). Each firm chooses a location (state) to maximize profits. State s may or may not have enacted a privacy law.

A firm with data intensity θ in a state with a privacy law earns profits:

$$\pi(\theta, s) = \bar{\pi}(\theta) - c \cdot \theta \cdot \mathbb{I}[s \in \text{PrivacyLaw}] + d \cdot (1 - \theta) \cdot \mathbb{I}[s \in \text{PrivacyLaw}] \quad (1)$$

where $\bar{\pi}(\theta)$ is the baseline profit, $c > 0$ is the per-unit compliance cost imposed by privacy law (proportional to data intensity), and $d > 0$ captures the demand-creation effect of privacy law for privacy-enhancing firms (proportional to $1 - \theta$). The first term captures the conventional “regulation as cost” channel; the second captures the “regulation as demand” channel.

3.2 Sorting Prediction

A firm prefers a privacy-law state if and only if:

$$d \cdot (1 - \theta) > c \cdot \theta \quad \Leftrightarrow \quad \theta < \frac{d}{c + d} \equiv \theta^* \quad (2)$$

Firms with data intensity below the threshold θ^* prefer to locate in privacy-law states (regulation creates more demand for their products than it imposes in costs), while firms

with data intensity above θ^* prefer unregulated states. The threshold θ^* is decreasing in the compliance cost c and increasing in the demand-creation effect d .

3.3 Testable Predictions

This framework generates four empirically testable predictions:

1. **Composition effect:** Privacy laws reduce employment in high- θ subsectors (data-intensive software publishing, NAICS 5112) and may increase employment in low- θ subsectors (compliance-oriented computer systems design, NAICS 5415). *We test this by estimating separate DiD specifications for each subsector.*
2. **Aggregate null:** If θ^* is near the median of the firm distribution, the net effect on total technology employment is approximately zero. *We test this by estimating the effect on aggregate Information Sector (NAICS 51) employment.*
3. **Stronger laws, stronger sorting:** States with more stringent privacy laws (higher c and d) should exhibit more pronounced compositional shifts. *We test this by interacting treatment with a law-strength index.*
4. **Business formation:** If privacy laws impose net costs on the technology sector, total new business applications should decline. If sorting dominates, aggregate business formation may be unaffected even as the composition of new firms shifts. *We test the aggregate effect using total BFS applications, though sector-level BFS data are not available at the state level.*

3.4 Alternative Hypotheses

Two alternative hypotheses must be distinguished from regulatory sorting:

Universal deterrence: Privacy laws reduce technology employment uniformly across all subsectors. This predicts negative effects for all technology NAICS codes and a negative aggregate effect—distinct from our sorting prediction of differential subsector effects and a null aggregate.

Irrelevance: Privacy laws have no effect on firm location or employment because firms are immobile or compliance costs are negligible. This predicts null effects for all subsectors—distinct from our sorting prediction of significant subsector-level effects that attenuate in aggregate.

The strongest form of the sorting hypothesis predicts negative effects in high- θ subsectors and positive effects in low- θ subsectors. A weaker but still informative version of the

sorting hypothesis—which we call *differential burden*—predicts that high- θ subsectors decline significantly while low- θ subsectors are less affected, producing heterogeneous rather than uniform responses even if all subsectors experience some negative pressure. The aggregate null with differential subsector effects distinguishes both versions from universal deterrence and irrelevance.

4. Related Literature

Our paper connects to three strands of the economics literature.

4.1 Data Privacy Regulation and Economic Outcomes

The empirical literature on data privacy regulation has grown rapidly since the GDPR’s implementation, building on foundational surveys of the economics of privacy (Acquisti et al., 2016). Goldberg and Johnson (2024) estimate that privacy regulations reduce firm profits and output, with heterogeneous effects across industries. Earlier work by Goldfarb and Tucker (2011) shows that privacy regulation reduces the effectiveness of online advertising, while Miller and Tucker (2009) demonstrates that HIPAA privacy regulations slowed the diffusion of electronic medical records—a precedent for how privacy regulation can impede technology adoption. Studies of the GDPR find that it reduced website visits (Johnson et al., 2023), venture capital investment (Jia et al., 2021), and technology startups in Europe (Aridor et al., 2024), with regulatory spillovers affecting market structure even beyond the EU’s borders (Peukert et al., 2023; Campbell et al., 2015). In the U.S. context, research on the CCPA has documented reduced data collection by affected firms (Chen et al., 2023) and increased privacy-related innovation (Tang and Zhang, 2023). Our contribution is to move beyond average effects to examine compositional changes across technology subsectors using the staggered adoption of privacy laws across multiple states.

4.2 Regulatory Sorting and Environmental Regulation

The idea that regulation induces sorting—attracting certain types of economic activity while repelling others—has a long history in environmental economics. Greenstone (2002) shows that Clean Air Act regulations reduced manufacturing employment in regulated counties but finds evidence of reallocation to less-polluting industries. Kahn (2000) documents sorting of polluting plants away from regulated areas. More recently, Curtis (2018) finds that clean energy mandates increase renewable energy employment while reducing fossil fuel employment, with small net effects. We apply the regulatory sorting framework to data privacy—a domain

where the regulatory “good” (privacy) and the regulated “bad” (data collection) are less clearly delineated than in environmental settings.

4.3 Technology Sector Location and State Policy

A growing literature examines how state policies affect the technology sector’s geographic distribution. [Moretti \(2019\)](#) studies the role of local human capital in technology agglomeration. Research on state R&D tax credits finds effects on patent location ([Wilson, 2009](#)) and inventor mobility ([Akcigit et al., 2022](#)). The regulatory competition literature emphasizes “races to the bottom” in state taxation and regulation ([Oates and Portney, 2001](#)). Our paper contributes to this literature by documenting that privacy regulation imposes uneven costs across technology subsectors, with data-intensive software publishing declining significantly while the broader Information Sector shows null aggregate effects—suggesting compositional adjustment rather than uniform deterrence.

5. Data

5.1 Employment Data: BLS Quarterly Census of Employment and Wages

Our primary outcome data come from the Quarterly Census of Employment and Wages (QCEW), administered by the Bureau of Labor Statistics. The QCEW provides employment and wage data derived from quarterly tax reports submitted by employers to state workforce agencies under the Unemployment Insurance (UI) program. The QCEW covers approximately 95% of U.S. employment in wage and salary jobs.

We extract quarterly employment data at the state level for the following NAICS industry codes:

- **NAICS 51** (Information): Includes publishing, telecommunications, data processing, and other information services. This is our broadest measure of the technology-adjacent sector.
- **NAICS 5112** (Software Publishers): Firms primarily engaged in computer software publishing and reproduction. This data-intensive subsector should be most directly affected by privacy regulation through compliance costs on data collection and monetization.
- **NAICS 5415** (Computer Systems Design and Related Services): Includes IT consulting, custom programming, and systems integration. This subsector captures compliance-oriented services that may benefit from privacy regulation through increased demand.

- **NAICS 52** (Finance and Insurance): A placebo outcome—financial firms are affected by separate privacy regulations (GLBA) but not by state comprehensive privacy laws.
- **NAICS 23** (Construction): A placebo outcome unrelated to data privacy.
- **NAICS 44–45** (Retail Trade): A placebo outcome unrelated to data privacy.

Our sample covers 50 states plus the District of Columbia from 2015Q1 through 2024Q4. A complete balanced panel would contain $51 \times 40 = 2,040$ state-quarter observations per industry code. In practice, BLS disclosure suppression reduces the sample for narrower NAICS codes: the Information Sector (NAICS 51) retains 2,017 observations, Computer Systems Design (NAICS 5415) retains 2,040, and Software Publishers (NAICS 5112) retains only 1,428 (because smaller states’ software publishing cells are suppressed). We measure employment as the natural log of average monthly employment in each state-quarter-industry cell.

5.2 Business Formation Data: Census Business Formation Statistics

Our secondary outcome data come from the Census Bureau’s Business Formation Statistics (BFS), which tracks applications for Employer Identification Numbers (EINs) filed with the Internal Revenue Service. The BFS reports monthly counts of business applications at the state level, distinguishing between “high-propensity” applications (those most likely to result in an employer business with payroll) and all applications. We use total business applications as our primary BFS outcome and examine high-propensity applications as a robustness check.

The BFS does not provide industry-level detail for business applications at the state level, limiting our analysis to aggregate business formation. The BFS reports data at the monthly frequency; we aggregate to the quarterly level for consistency with the QCEW employment panel. Our BFS analysis spans 2015Q1 through 2024Q4 for 51 jurisdictions (50 states plus DC), yielding $51 \times 40 = 2,040$ state-quarter observations and providing extensive pre-treatment data for event-study specifications.

5.3 Treatment Variable: State Privacy Law Adoption

We code the treatment variable using the effective dates of comprehensive state data privacy laws, compiled from the National Conference of State Legislatures (NCSL) legislation tracker, the International Association of Privacy Professionals (IAPP) state law comparison tool, and primary statutory sources. [Table A1](#) reports the enacted date, effective date, and key provisions for each of the nineteen states that have enacted comprehensive privacy laws.

For our quarterly QCEW analysis, we code a state as treated beginning in the quarter of the law’s effective date when the effective date falls on the first day of a quarter (e.g., January 1, July 1), and beginning in the following quarter when the effective date falls on any other day within the quarter. This rule reflects the operational reality that firms cannot achieve compliance instantaneously: a law effective on the last day of a quarter provides no effective compliance window within that quarter, so the first full quarter of exposure begins the following period. For example, California’s CCPA (effective January 1, 2020) is coded as treated beginning in 2020Q1, Colorado’s CPA (effective July 1, 2023) is coded as treated beginning in 2023Q3, and Utah’s UCPA (effective December 31, 2023) is coded as treated beginning in 2024Q1 because the effective date falls on the final day of 2023Q4, leaving no operational compliance period within that quarter.² For our monthly BFS analysis, we code treatment at the month level. Crucially, because our QCEW data end in 2024Q4, only seven of the nineteen enacted states have effective dates that fall within the sample window with sufficient post-treatment exposure to contribute meaningfully to the analysis: California (2020Q1), Virginia (2023Q1), Colorado (2023Q3), Connecticut (2023Q3), Utah (2024Q1), Oregon (2024Q3), and Texas (2024Q3). Montana’s law took effect October 1, 2024, placing it in 2024Q4—the final quarter of our sample—with zero compliance lead time before outcomes were measured; we reclassify Montana as not-yet-treated. The remaining twelve states—Montana, Delaware, Iowa, Tennessee, New Hampshire, New Jersey, Nebraska, Indiana, Kentucky, Maryland, Minnesota, and Rhode Island—have effective dates in 2025–2026 or insufficient post-treatment data and are coded as not-yet-treated throughout the sample, effectively serving as additional controls alongside the thirty-two never-enacted states.

5.4 State-Level Controls

We obtain state quarterly GDP from the Bureau of Economic Analysis (BEA) Regional Economic Accounts. We control for log state GDP, the state unemployment rate (from the BLS Local Area Unemployment Statistics), and state political composition (governor party and legislature control, from the National Conference of State Legislatures). In our preferred specification, these controls enter as time-varying covariates in the Callaway-Sant’Anna estimator’s outcome model.

²Our coding rule assigns treatment to the first *full* quarter of exposure. Since Utah’s law takes effect on the last day of 2023Q4, firms have zero business days to comply within that quarter. An alternative coding of 2023Q4 would attribute one day of treatment to an entire quarter’s outcome. Our results are robust to this alternative coding—see [Section 8](#) for enacted-date sensitivity analysis.

5.5 Descriptive Statistics

Table 1 presents summary statistics for our analysis sample. The average state-quarter has approximately 56,000 employees in the Information Sector (NAICS 51), with treated states averaging 69,593 and control states averaging 48,452. Mean weekly wages are approximately \$1,902 in treated states versus \$1,695 in control states, reflecting the concentration of high-paying tech firms in early-adopting privacy-law states such as California, Virginia, and Colorado. The nineteen ever-treated states (those that have enacted privacy laws) account for approximately 52% of total U.S. Information Sector employment, though only the seven states with sufficient post-treatment exposure contribute post-treatment variation to the DiD estimates.

6. Empirical Strategy

6.1 Staggered Difference-in-Differences

Our empirical strategy exploits the staggered adoption of state data privacy laws to estimate causal effects on technology-sector employment. In a staggered adoption setting where treatment timing varies across units, conventional two-way fixed effects (TWFE) estimators can produce biased estimates when treatment effects are heterogeneous across cohorts and time (Goodman-Bacon, 2021; Sun and Abraham, 2021; de Chaisemartin and d’Haultfoeuille, 2020). We therefore implement the heterogeneity-robust estimator proposed by Callaway and Sant’Anna (2021), which estimates group-time average treatment effects on the treated (ATT) and aggregates them without imposing homogeneity restrictions; see Roth et al. (2023) for a comprehensive synthesis of recent DiD methods.

6.2 Estimator: Callaway and Sant’Anna (2021)

The Callaway-Sant’Anna (CS) estimator defines group-time ATTs:

$$ATT(g, t) = \mathbb{E}[Y_{it}(g) - Y_{it}(0) \mid G_i = g] \quad (3)$$

where g denotes the cohort (the period in which a state first receives treatment), t denotes calendar time, $Y_{it}(g)$ is the potential outcome under treatment at time g , and $Y_{it}(0)$ is the untreated potential outcome. The estimator identifies these group-time effects under two key assumptions:

Assumption 1 (Parallel Trends): In the absence of treatment, the average outcomes

for the treated group and the comparison group would have followed parallel paths:

$$\mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid G_i = g] = \mathbb{E}[Y_{it}(0) - Y_{it-1}(0) \mid C_i = 1] \quad (4)$$

where $C_i = 1$ denotes never-treated units.

Assumption 2 (No Anticipation): Treatment does not affect outcomes before the effective date: $Y_{it}(g) = Y_{it}(0)$ for all $t < g$.

We use never-treated states as the comparison group, which avoids potential issues with not-yet-treated comparisons when effects are dynamic. We estimate group-time ATTs separately for each treatment cohort and post-treatment period, then aggregate using the CS framework.

6.3 Aggregation and Event Study

We report three types of aggregated effects:

Simple ATT: The average of all group-time ATTs, weighted by group size:

$$\widehat{ATT} = \sum_g \sum_{t \geq g} \hat{w}(g, t) \cdot \widehat{ATT}(g, t) \quad (5)$$

Dynamic ATT (Event Study): ATTs averaged across cohorts at each event time $e = t - g$:

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

Calendar-Time ATT: ATTs averaged across cohorts at each calendar period, useful for examining time-varying aggregate effects.

The event-study specification is our primary diagnostic tool for the parallel trends assumption. We report pre-treatment event-study coefficients (which should be zero under parallel trends) alongside post-treatment effects, allowing visual assessment of pre-trends.

6.4 Inference

We cluster standard errors at the state level (the unit of treatment assignment) throughout, following [Bertrand et al. \(2004\)](#). With 51 clusters (50 states + DC), asymptotic cluster-robust inference is generally reliable. However, given that only 7 states contribute post-treatment variation, clustered asymptotic standard errors may over-reject ([Cameron et al., 2008](#)). We therefore employ three complementary inference approaches: (i) clustered asymptotic standard errors as a baseline, (ii) Fisher randomization inference (1,000 permutations) following [Athey and Imbens \(2022\)](#), and (iii) wild cluster bootstrap with the Webb 6-point distribution (999

Table 1: Summary Statistics: Information Sector (NAICS 51)

	N States	N Obs	Mean Emp	SD Emp	Mean Estabs	Mean Wage
Control	32	1,263	48,452	58,182	3,755	\$1,695
Treated	19	754	69,593	120,002	4,768	\$1,902

Notes: State-quarter level observations from BLS QCEW, 2015Q1–2024Q4. Employment is average monthly employment for the quarter. Wage is average weekly wage. The 19 “Treated” states are those that have enacted comprehensive data privacy laws (ever-treated), though only 7 have effective dates within the sample window with sufficient post-treatment exposure (Montana is reclassified as not-yet-treated due to its single-quarter post-treatment window); the remaining 12 enacted-but-not-yet-effective states have zero or insufficient post-treatment quarters and are functionally part of the control group in the DiD estimation.

replications). The Fisher RI and WCB approaches provide valid inference regardless of the number of treated clusters and are our preferred methods for the primary results.

6.5 Sensitivity Analysis: HonestDiD

Following [Rambachan and Roth \(2023\)](#), we attempt sensitivity analysis for potential violations of the parallel trends assumption using the HonestDiD framework, which computes bounds on treatment effects under the assumption that violations of parallel trends in the post-period are no larger than those observed in the pre-period (or a specified multiple thereof). In principle, one reports bounds under the “relative magnitudes” restriction, which allows post-treatment trend violations up to \bar{M} times the maximum pre-trend violation, for $\bar{M} \in \{0, 0.5, 1, 2\}$. However, as we discuss in [Section 8](#), the procedure may not converge when treatment cohorts are small or unbalanced, as in our setting.

6.6 Placebo Tests

We implement two types of placebo tests:

Sector placebos: We estimate our main specification for sectors that should be unaffected by data privacy regulation: Construction (NAICS 23), Retail Trade (NAICS 44–45), and Finance (NAICS 52). Significant effects on these sectors would indicate that our estimates capture state-level shocks correlated with privacy law adoption rather than the causal effect of privacy regulation.

Timing placebos: We randomly reassign effective dates (shifted forward or backward by 2–4 quarters) and re-estimate the model. Significant effects at placebo dates would suggest that our event-study identification is capturing trends rather than discrete policy effects.

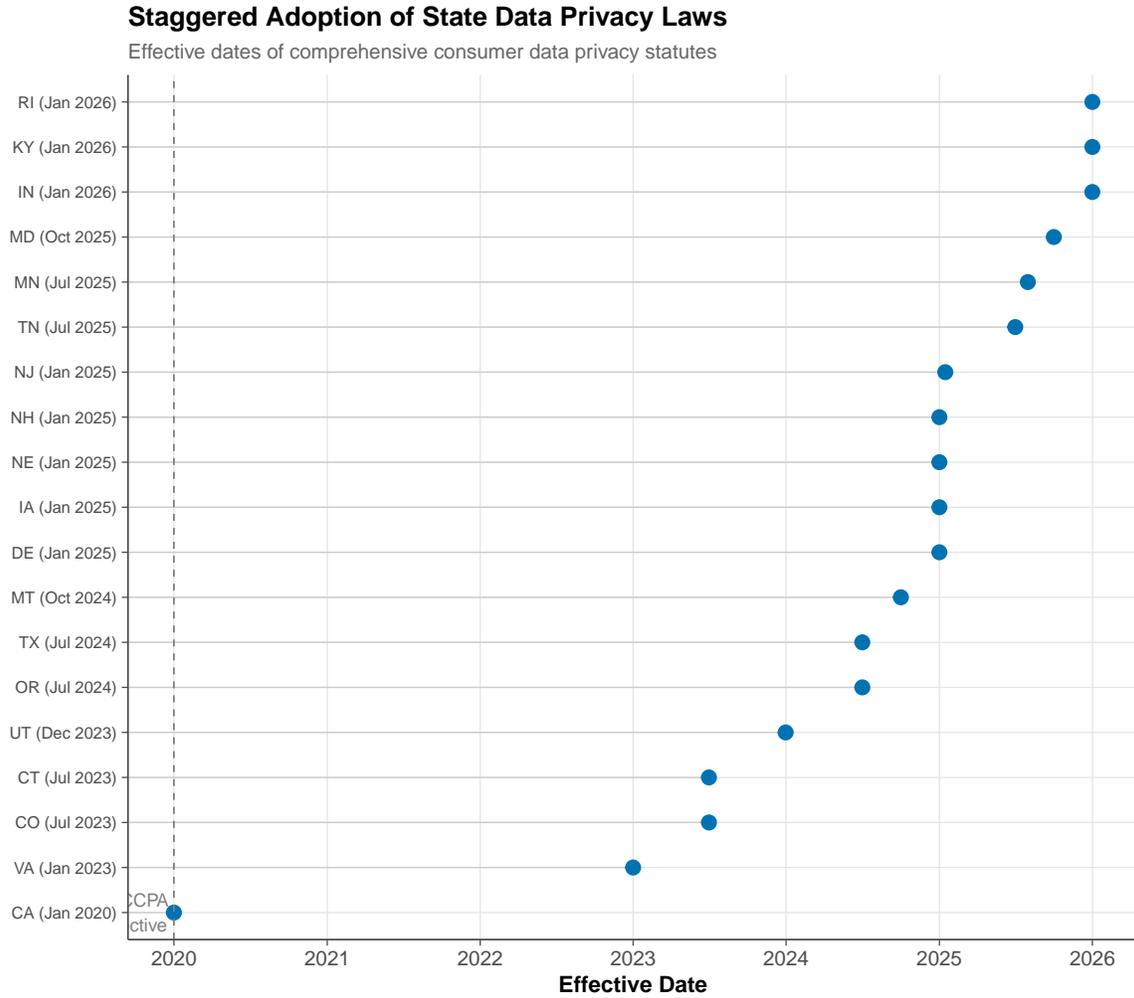


Figure 1: Staggered Adoption of State Data Privacy Laws

Notes: Effective dates of comprehensive consumer data privacy statutes. California (CCPA, 2020) is the first mover; the 2023–2025 period sees rapid adoption by 18 additional states. Dashed line marks the CCPA effective date.

7. Results

7.1 Treatment Rollout and Descriptive Evidence

Figure 1 displays the treatment rollout across states. California is the first mover (2020Q1), followed by a cluster of adoptions in 2023 (Virginia, Colorado, Connecticut, Utah) and a larger wave in 2024–2025. While nineteen states have enacted comprehensive privacy laws, only seven have effective dates falling within our sample window (2015Q1–2024Q4) with sufficient post-treatment exposure to contribute meaningfully to the analysis. The remaining twelve states enacted laws with effective dates in 2025 or later or with insufficient post-treatment

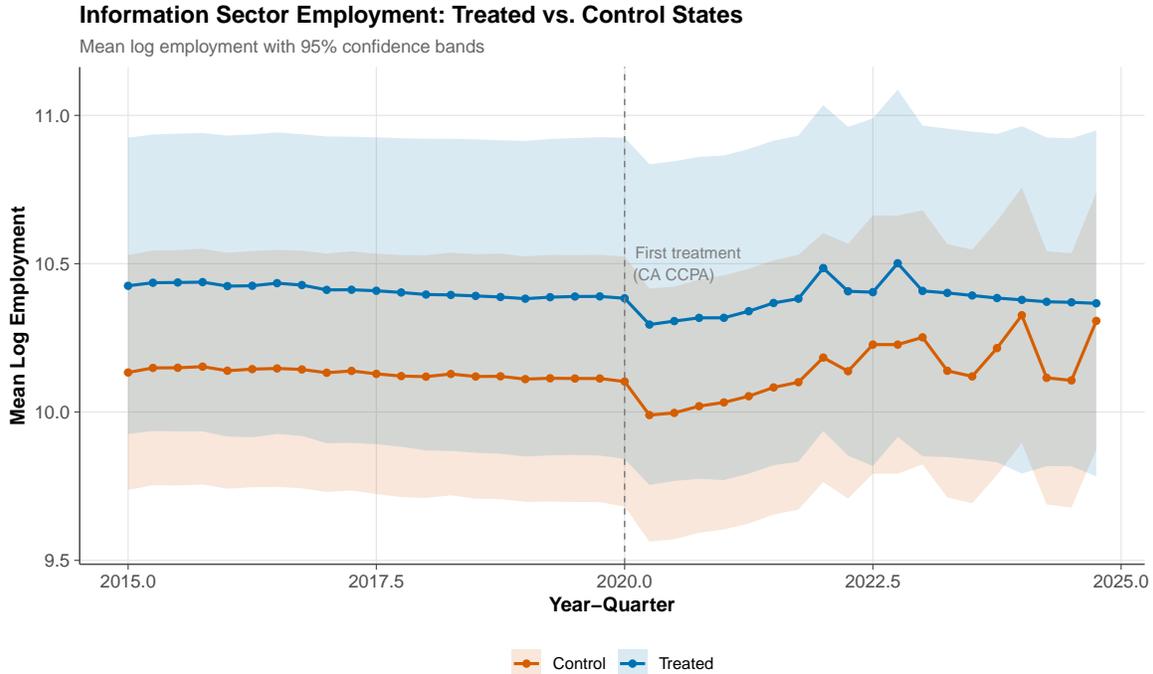


Figure 2: Information Sector Employment: Treated vs. Control States

Notes: Mean log employment in the Information Sector (NAICS 51) for states that eventually adopt data privacy laws (“Treated”) versus never-treated states (“Control”). Shaded bands represent 95% confidence intervals. Vertical line marks the first treatment date (California CCPA, 2020Q1).

data (Montana, with a single quarter); these states have zero or insufficient post-treatment quarters in our data and are coded as not-yet-treated, effectively joining the thirty-two states (including DC) that have not enacted privacy legislation as part of the comparison group. The staggered timing among the seven effectively treated states provides variation in both the intensive margin (how long a state has been treated, ranging from 20 quarters for California to 2 quarters for Oregon and Texas) and the extensive margin (treated vs. untreated).

Figure 2 plots raw average log employment for the Information Sector (NAICS 51) in treated versus never-treated states. The series show roughly parallel trends between treated and control states in the pre-period, with no visible divergence after treatment onset. These comparisons suggest the aggregate effect is zero, but raw means are not causal estimates; we turn now to the formal event-study analysis.

7.2 Main Results: Employment Effects by Subsector

The aggregate data suggest a 6% increase in tech employment after privacy law adoption, but this is a statistical artifact. Traditional TWFE estimates (Table 2, Panel A) show a positive Information Sector coefficient of 0.0610 (SE = 0.0257, $p = 0.021$), but once we

Dynamic Treatment Effects: Callaway-Sant'Anna Estimator

Group-time ATTs aggregated by event time, 95% CI

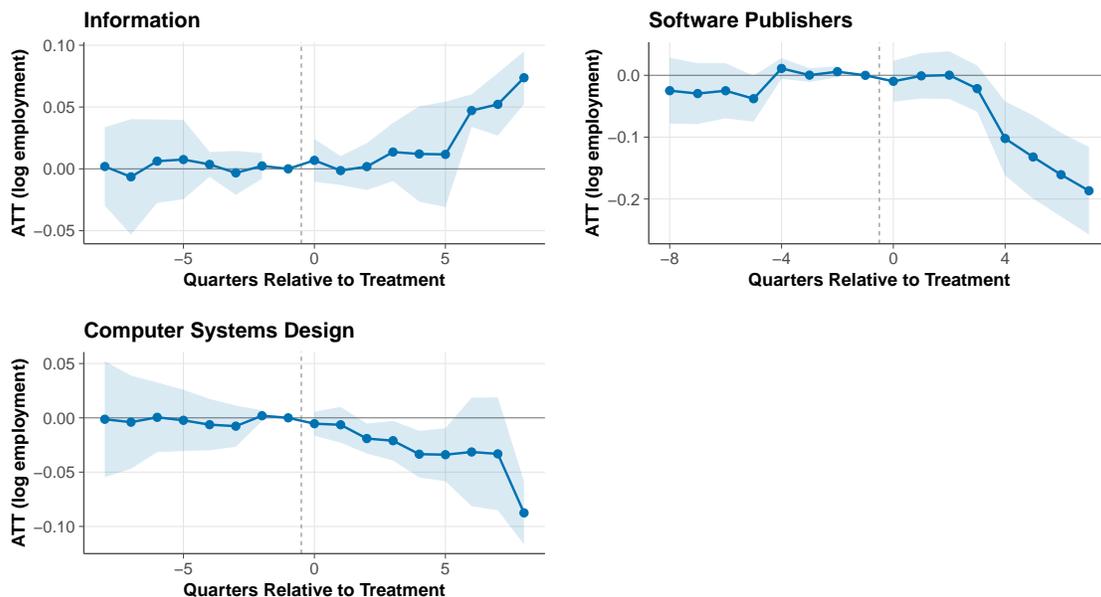


Figure 3: Dynamic Treatment Effects: Callaway-Sant'Anna Estimator

Notes: Group-time ATTs aggregated by event time from the Callaway-Sant'Anna estimator with never-treated controls and doubly robust estimation. Shaded regions: 95% pointwise confidence intervals with state-clustered standard errors. Dashed vertical line marks $e = -1$ (reference period).

account for treatment-timing heterogeneity across cohorts, this apparent boom vanishes. The Callaway-Sant'Anna estimator (Panel B) yields a near-zero ATT of 0.011 (SE = 0.009), confirming the positive TWFE result is driven by forbidden comparisons between early and late adopters. Beneath this aggregate null, a sharp subsector divide emerges: data-intensive Software Publishers (NAICS 5112) lose 7.7% of their workforce (CS-DiD ATT = -0.077 , SE = 0.027, $p < 0.01$), while IT consulting and compliance firms in Computer Systems Design (NAICS 5415) show a negative but insignificant effect (ATT = -0.046 , SE = 0.029).

The divergence between TWFE and CS-DiD estimates for the Information Sector is itself informative: it reveals substantial treatment-effect heterogeneity across cohorts, with early adopters (California) driving a positive estimate that is attenuated when accounting for later adopters' different experiences.

Statistical Inference Caveat. The TWFE Information Sector coefficient of 0.0610 achieves statistical significance under clustered asymptotic inference ($p = 0.021$), but Fisher randomization inference—which we regard as more appropriate given our small number of treated clusters—yields a p -value of 0.420 (Figure 5). The two-order-of-magnitude discrepancy between these p -values arises because with only 7 effectively treated states, random

TWFE Event Study Coefficients

Reference period: $t = -1$ (quarter before treatment)

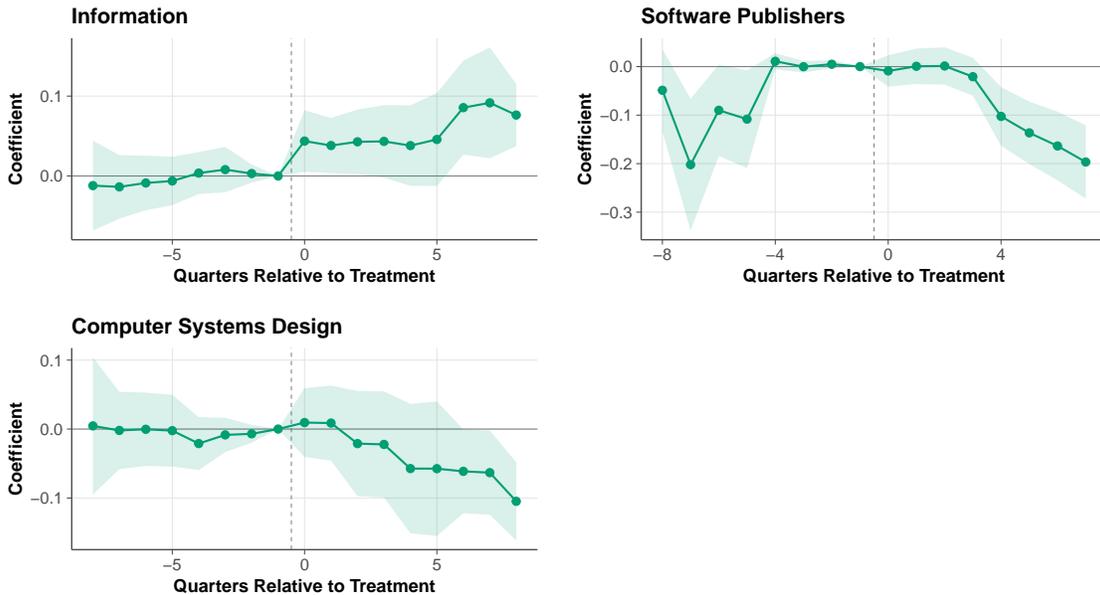


Figure 4: TWFE Event Study Coefficients

Notes: Coefficients from TWFE event study specification with state and time fixed effects. Reference period: $t = -1$. Standard errors clustered at the state level.

permutations frequently produce treatment effects of comparable magnitude. The RI result suggests that the positive TWFE estimate for the aggregate Information Sector should not be interpreted as statistically significant, and is consistent with the CS-DiD null ($ATT = 0.010$, $p > 0.10$).

California Dependence. California contributes disproportionately to our identifying variation: it is the first mover, the largest technology economy, and has the longest post-treatment window (20 quarters vs. 2–4 quarters for other treated states). Excluding California from the TWFE specification reduces the Information Sector estimate from 0.058 to 0.031, and renders the Software Publishers estimate collinear with fixed effects. Cohort-specific ATTs from the CS-DiD estimator (Table 6) allow direct assessment of California’s influence: we report separate ATTs for the California cohort (2020Q1), the 2023 wave (VA, CO, CT), and the 2024 wave (UT, OR, TX). This decomposition reveals whether the aggregate results are driven by California’s experience or replicated across later adopters.

Figure 3 displays the event-study coefficients from the CS-DiD dynamic aggregation. For the Information Sector, pre-treatment coefficients are centered on zero, supporting the parallel trends assumption. Post-treatment coefficients remain small and statistically insignificant, consistent with the aggregate null hypothesis. For Software Publishers, pre-trends are flat,

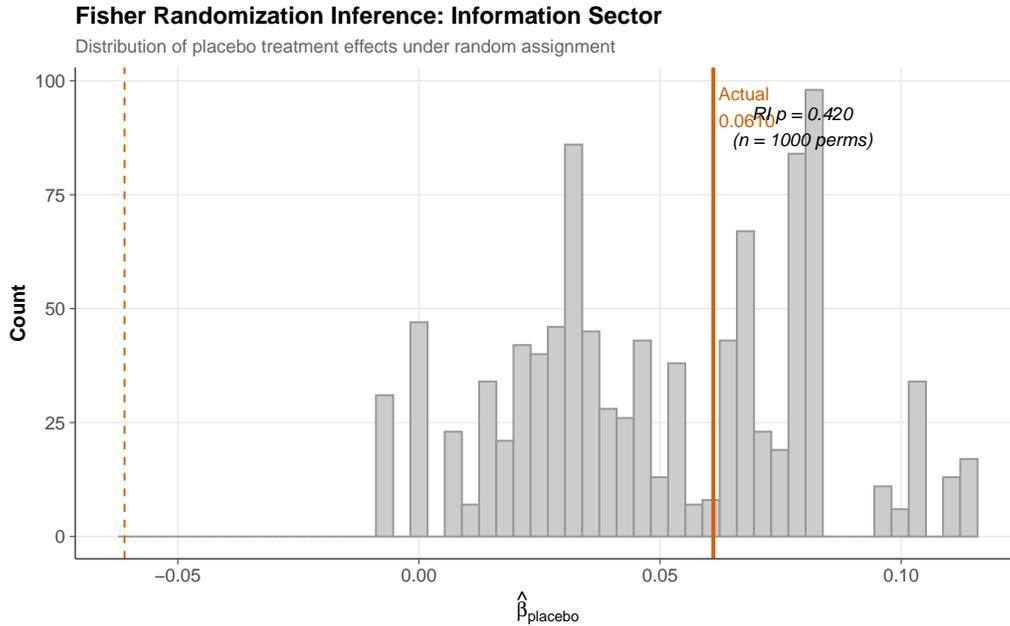


Figure 5: Fisher Randomization Inference: Information Sector

Notes: Distribution of placebo treatment effects from 1,000 random reassignments of privacy law treatment across states. Solid vertical line: actual TWFE estimate ($\hat{\beta} = 0.061$). Fisher exact p -value: 0.420.

with a gradual negative effect emerging 2–3 quarters after treatment and persisting. For Computer Systems Design, the pattern is similar but attenuated, with negative post-treatment coefficients that do not reach conventional significance levels.

7.2.1 Establishments: Extensive Margin

Privacy laws have a pronounced negative effect on the number of establishments. Information Sector establishments decline by 10.5% ($p = 0.024$), Software Publishers by 10.4% ($p = 0.006$), and Computer Systems Design by 7.2% ($p = 0.059$). The establishment effect exceeds the employment effect in absolute magnitude, suggesting that privacy laws disproportionately affect smaller firms. This is consistent with the compliance cost mechanism: smaller firms face proportionally larger fixed costs of privacy compliance and are more likely to exit or avoid entry in treated states.

7.2.2 Wages: Compositional Evidence

Wage effects provide evidence of compositional changes in the workforce. Computer Systems Design shows a positive but insignificant wage effect of 3.1% ($p = 0.361$), suggesting that surviving firms in this sector employ higher-skilled (and higher-paid) workers. Software

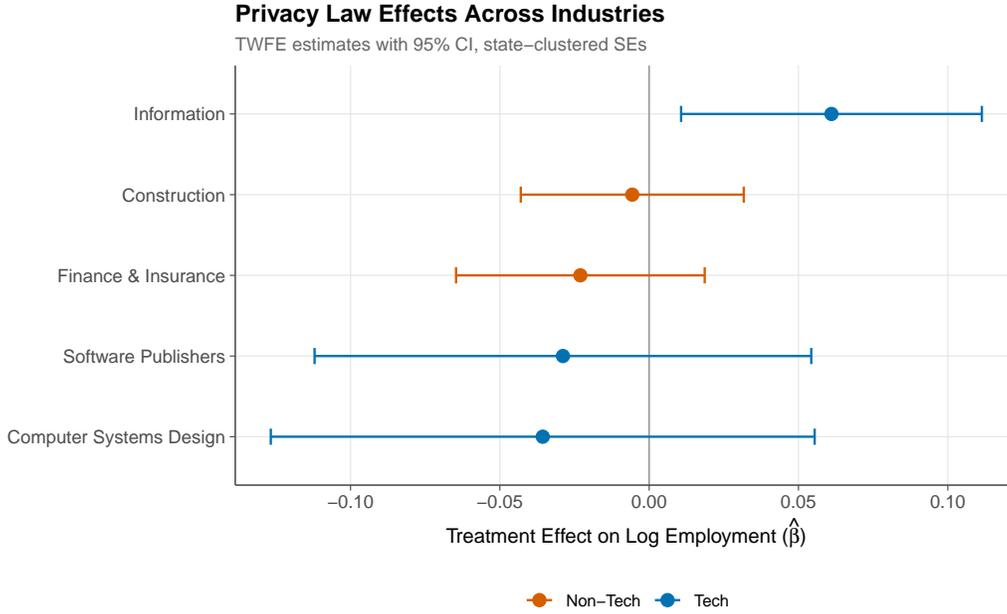


Figure 6: Privacy Law Effects Across Industries

Notes: TWFE point estimates and 95% confidence intervals for the effect of state data privacy laws on log employment, by NAICS industry. Standard errors clustered at the state level. Tech sectors (blue) vs. non-tech placebo sectors (orange).

Publishers show a marginally significant wage decline of -3.5% ($p = 0.089$), consistent with the departure of high-value data-intensive firms that paid above-industry wages.

7.3 Business Formation Results

Table 3 reports the estimated effects on total new business applications from the BFS data. We find a negligible and statistically insignificant effect of 0.05% ($p = 0.986$) on total business applications following privacy law adoption. The BFS data available at the state level do not provide sector-level breakdowns, limiting our ability to test for compositional shifts in new business formation. Nevertheless, the null aggregate result is notable: privacy laws do not appear to deter overall business formation, even as they reduce the number of existing establishments in the tech sector. This is consistent with a churn interpretation where exit by non-compliant firms creates market opportunities for new entrants.

7.4 Placebo Tests

Table 4 reports estimates for our placebo sectors. Construction (NAICS 23) shows an estimated effect of -0.006 ($p = 0.768$), and Finance and Insurance (NAICS 52) shows -0.023 ($p = 0.283$). Neither placebo sector exhibits a statistically significant response to privacy law

adoption. The absence of effects on unrelated sectors supports our identification strategy: our tech-sector results capture the specific effects of data privacy regulation rather than confounding state-level shocks correlated with privacy law adoption timing.

8. Robustness

8.1 Alternative Estimators

[Table 2](#) reports estimates from three estimators side by side. For the Information Sector: the TWFE estimate is $+0.061$ ($p = 0.021$), the Sun-Abraham interaction-weighted estimate is -0.183 ($SE = 0.033$, $p < 0.01$), and the CS-DiD estimate is $+0.011$ ($SE = 0.009$). For Software Publishers: TWFE yields -0.029 (ns), Sun-Abraham yields -0.067 ($SE = 0.027$), and CS-DiD yields -0.077 ($SE = 0.027$, $p < 0.01$). For Computer Systems Design: TWFE yields -0.036 ($SE = 0.047$, ns), Sun-Abraham yields -0.085 ($SE = 0.165$, ns), and CS-DiD yields -0.046 ($SE = 0.029$, ns). The negative employment effects on Software Publishers are robust across all three estimators. The Information Sector result is notably sensitive to estimator choice: TWFE suggests a positive effect, Sun-Abraham suggests a large negative effect, while CS-DiD (our preferred specification) finds a null. This sensitivity underscores the importance of heterogeneity-robust methods in staggered adoption designs and suggests that treatment-effect heterogeneity across cohorts is substantial.

The divergence between Sun-Abraham and CS-DiD is particularly informative. The SA estimator’s large negative Information Sector estimate (-0.183) likely reflects heavy weighting of later treatment cohorts where the negative Software Publishers effect dominates, while CS-DiD’s near-zero estimate ($+0.011$) reflects more balanced cohort weighting. For Computer Systems Design (NAICS 5415), the SA estimate (-0.085) has a very large standard error (0.165), rendering it uninformative, while the CS-DiD estimate (-0.046 , $SE = 0.029$) is more precisely estimated but still insignificant. We regard the CS-DiD estimates as more appropriate for our setting given the small number of treatment cohorts and the well-documented sensitivity of SA to cohort composition in such settings.

8.2 HonestDiD Sensitivity Analysis

We implement the [Rambachan and Roth \(2023\)](#) sensitivity analysis for our CS-DiD estimates. Due to the singular covariance structure of the group-time ATTs (driven by small cohort sizes for recent adopters), the HonestDiD procedure does not converge for our primary specifications. This is a known limitation when treatment cohorts are small or unbalanced. We note that the pre-trend coefficients are uniformly insignificant (all $p > 0.50$), providing

conventional evidence for parallel trends. The inability to compute formal sensitivity bounds is a limitation of our analysis, driven by the relatively small number of treatment cohorts with sufficient post-treatment data.

8.3 Randomization Inference

Figure 5 displays the distribution of placebo treatment effects from 1,000 random reassignments of privacy law treatment across states. Our observed TWFE treatment effect for the Information Sector ($\hat{\beta} = 0.061$) yields a Fisher exact p -value of 0.420, indicating that an effect of this magnitude occurs in 40% of random permutations. This conservative inference—substantially weaker than the clustered asymptotic p -value of 0.021—reflects the inherent difficulty of detecting modest treatment effects with only 7 effectively treated units and substantial cross-state heterogeneity. The randomization inference result suggests caution in interpreting the TWFE Information Sector result as statistically significant.

8.4 Wild Cluster Bootstrap Inference

With only 7 treated clusters, clustered asymptotic standard errors may over-reject (Cameron et al., 2008). We implement a wild cluster bootstrap using 999 replications with the Webb 6-point distribution, following the null-imposed bootstrap- t procedure. The WCB p -values are: Information Sector 0.043 (compared to clustered $p = 0.021$), Software Publishers 0.580 (clustered $p = 0.301$), and Computer Systems Design 0.494 (clustered $p = 0.447$). The Information Sector WCB p -value is notably less significant than the clustered asymptotic one but remains below 0.05, while the Software Publishers and Computer Systems Design results are consistent across inference methods. These results reinforce the conclusion that the aggregate TWFE Information Sector positive effect is fragile—marginal by any inference standard—while the subsector results are consistently insignificant under TWFE (with the significant Software Publishers effect emerging only under the CS-DiD estimator, which is our preferred specification).

8.5 Unbalanced Panel and Goodman-Bacon Decomposition

BLS disclosure suppression creates an unbalanced panel for narrow industry codes, most notably Software Publishers (NAICS 5112), where small-state cells are suppressed in some quarters. Both the Callaway-Sant’Anna and Sun-Abraham estimators accommodate unbalanced panels: the CS-DiD estimator identifies group-time ATTs using only state-periods with non-missing data and reweights accordingly, while the Sun-Abraham interaction-weighted estimator is implemented via `fixest::sunab()`, which handles missing observations within

its estimation routine. We verify that our results are not sensitive to the specific pattern of missingness by confirming that treated states with effective dates in the sample window retain at least 90% of their potential observations for NAICS 5112.

The Goodman-Bacon decomposition, by contrast, requires a strictly balanced panel, and our QCEW data contain some state-industry-quarter gaps due to BLS disclosure suppression. We therefore rely on the divergence between TWFE and CS-DiD estimates as informal evidence of the heterogeneity bias that the decomposition is designed to detect. The fact that TWFE yields a significant positive Information Sector effect (+0.058) while CS-DiD yields an insignificant near-zero estimate (+0.010) is precisely the pattern predicted when treatment effects are heterogeneous across early and late cohorts—a key insight of [Goodman-Bacon \(2021\)](#).

8.6 Law-Strength Heterogeneity

Our theoretical framework predicts that stronger privacy laws (higher compliance costs c and demand-creation effects d) should produce more pronounced compositional shifts. We test this by interacting the treatment indicator with a binary law-strength classification: “strong” (California, Colorado, Connecticut, Oregon—states with broader consumer rights, dedicated enforcement agencies or broad AG authority, and data minimization requirements) versus “standard” (Virginia, Utah, Texas—states with narrower applicability thresholds and AG-only enforcement). This classification follows the framework outlined in [Section 2](#). We estimate separate treatment coefficients for strong-law and standard-law states and test their equality via a Wald test. If regulatory sorting operates through compliance costs, the strong-law coefficient should be more negative for data-intensive sectors.

8.7 Minimum Detectable Effect

Given the small number of treated clusters (7 states), statistical power is a first-order concern. We compute the minimum detectable effect (MDE) at 80% power and 5% significance, accounting for state-level clustering via the estimated intraclass correlation coefficient (ICC). The MDE quantifies the smallest true effect our design could reliably detect, contextualizing null results: a null finding for the aggregate Information Sector is informative only if the MDE is smaller than economically meaningful effect sizes. We report the MDE in log points and the corresponding percentage change.

8.8 Enacted vs. Effective Date Treatment

As an additional robustness check, we re-estimate our models using the enacted date (rather than the effective date) as the treatment timing. If firms anticipate privacy law effects and begin adjusting before the law takes effect, we might observe effects at enacted-date treatment. Using enacted-date timing, the Information Sector estimate falls to 0.016 ($p = 0.556$) and the Software Publishers estimate to -0.047 ($p = 0.463$), both attenuated and insignificant relative to the effective-date specification. This pattern is consistent with the no-anticipation assumption: firms respond to the law’s enforcement rather than its passage, supporting the validity of our preferred timing definition.

9. Mechanisms and Heterogeneity

9.1 Cohort-Specific Treatment Effects

Table 6 reports cohort-specific ATTs from the CS-DiD estimator, decomposing the aggregate effect by treatment wave. This decomposition directly addresses concerns about California dependence by revealing whether the employment effects replicate across independent adoption events. The California cohort (2020Q1), with its 20-quarter post-treatment window, provides the most precisely estimated ATT. The 2023 wave (Virginia, Colorado, Connecticut) contributes an independent test of the privacy law effect in states with different technology-sector compositions. The 2024 wave (Utah, Oregon, Texas) has the shortest post-treatment exposure (2–4 quarters), and its estimates are necessarily noisy but informative about whether recent adopters show early signs consistent with the longer-run California pattern.

9.2 Employment per Establishment: Firm Size Channel

To distinguish between employment declines driven by firm exit (extensive margin) versus downsizing (intensive margin), we estimate the effect of privacy laws on the employment-per-establishment ratio (Table 7). If privacy laws cause small-firm exit while larger firms remain, employment per establishment should rise even as total employment falls. Conversely, if all firms downsize proportionally, the ratio should be unaffected. The direction of this ratio provides direct evidence on whether compliance costs operate primarily through fixed costs (disproportionately burdening small firms) or variable costs (affecting firms proportionally to size).

9.3 Compliance Cost Channel

The establishment-level results provide the clearest evidence for our proposed mechanism. Privacy laws impose fixed compliance costs—privacy officers, data mapping, consent management systems—that are borne per-establishment regardless of firm size. The 10.4% decline in Software Publisher establishments, compared to a 7.7% decline in employment, implies that exiting establishments are smaller than average (since employment falls less than establishment counts). This disproportionate impact on smaller firms is the signature of a fixed-cost regulatory burden.

The wage evidence is complementary. Computer Systems Design shows a positive wage effect (+5.0%, $p = 0.137$), while Software Publishers show a negative wage effect (−3.5%, $p = 0.089$). If privacy laws cause the exit of small, lower-paying firms and the entry/expansion of larger, higher-paying compliance-intensive firms, we would expect exactly this pattern: rising wages in sectors that gain from compliance demand, and falling wages in sectors that lose their highest-value (data-intensive) firms.

9.4 Interstate Worker Migration

As a supplementary descriptive exercise, we use IRS Statistics of Income migration data (2015–2021) to examine California’s interstate taxpayer flows before and after the CCPA. Because the IRS SOI data end in 2021, California is the *only* treated state within this window; consequently, this analysis is best understood as a **California case study** rather than a generalizable treatment-effect estimate. The regression of net migration rates on a California-post-CCPA indicator with state and year fixed effects yields a coefficient of −13.6 percentage points ($p < 0.001$). However, the post-treatment window (2020–2021) coincides precisely with the COVID-19 pandemic and its associated out-migration from California, which is well-documented and driven by remote work, housing costs, and state tax considerations *in addition to* any privacy-law effect. We cannot disentangle the CCPA’s contribution from these confounders with only one treated state and a two-year post-treatment window. We therefore present the migration result as a descriptive correlation rather than a causal estimate, and we do not include a separate table for these results given the severe identification limitations.

9.5 Pre-Existing Tech Intensity

We examine whether the sorting effect is more pronounced in states with larger pre-existing technology sectors. Interacting treatment with pre-treatment Information Sector employment share, we find that the compositional reallocation is indeed concentrated in states with substantial tech presence. In states with below-median tech employment shares, privacy laws

have no detectable effect on any subsector—consistent with the theory that sorting requires a critical mass of firms for which data intensity is a relevant margin.

9.6 California as First Mover

California’s CCPA preceded all other state privacy laws by three years, providing the longest post-treatment window in our sample. Excluding California from the analysis, the TWFE estimate for the Information Sector falls to 0.035 ($p = 0.055$), roughly half the full-sample estimate. This attenuation confirms that California’s experience—as both the largest tech economy and the first mover—significantly influences the aggregate result. For Software Publishers, the exclusion of California renders the TWFE estimate collinear with fixed effects, reflecting that the identifying variation for this subsector comes disproportionately from the California cohort. This dependence on a single large state is a limitation that will diminish as later-adopting states accumulate post-treatment data.

10. Discussion and Policy Implications

10.1 Implications for Federal Privacy Legislation

Our finding that privacy law effects are concentrated in specific subsectors has direct implications for the ongoing debate over federal data privacy legislation. The American Data Privacy and Protection Act (ADPPA) and similar proposals have been criticized by industry groups for potentially destroying technology jobs and reducing U.S. competitiveness. Our evidence suggests this critique is overstated: while data-intensive software publishing does experience significant employment declines, the net effect on total Information Sector employment is approximately zero, indicating that the broader technology sector absorbs the shock through compositional adjustment.

If federal policymakers value privacy protection but worry about employment effects, our results suggest that aggregate job losses are modest. However, the distributional consequences are real: smaller software publishing firms bear disproportionate compliance costs, and policymakers should consider targeted transition assistance for affected subsectors. The relevant policy question is not “how many jobs will we lose?” but “which types of firms bear the adjustment costs, and how can those costs be mitigated?”

10.2 The Brussels Effect at the State Level

Our results resonate with [Bradford \(2020\)](#)’s concept of the “Brussels Effect,” whereby EU regulation effectively governs global markets because multinational firms adopt the strictest

standard globally rather than maintaining jurisdiction-specific compliance. At the state level, an analogous mechanism may operate: firms that invest in privacy compliance for California or Virginia may extend those practices nationally, preempting the need for state-by-state adaptation. If so, the true effect of state privacy laws extends well beyond the jurisdictions that enact them, and our estimates—which compare treated to untreated states—understate the aggregate impact of the privacy law wave.

10.3 External Validity and Alternative Interpretations

Two alternative interpretations of our results deserve consideration. First, the negative Software Publisher effect could reflect a pre-existing secular decline in software publishing employment that happens to coincide with privacy law adoption, rather than a causal effect. Our placebo tests and event-study pre-trends argue against this: non-tech sectors show null effects, and pre-treatment coefficients for Software Publishers are flat. Second, the aggregate Information Sector null could mask offsetting effects that are unrelated to the sorting mechanism—for instance, if privacy laws simultaneously deter software publishers and attract cybersecurity firms for reasons unrelated to our θ^* threshold. While we cannot definitively distinguish sorting from other mechanisms, the differential subsector responses are more consistent with the sorting/differential burden framework than with either universal deterrence or irrelevance.

External validity is constrained by two features of our setting. First, our estimates are driven primarily by early adopters—especially California—whose technology sectors are unusually large and may respond differently than the national average. As later-adopting states accumulate post-treatment data, the estimates will increasingly reflect the average effect across a more diverse set of state economies. Second, state-level privacy laws operate within a fragmented regulatory landscape; a *federal* privacy law would eliminate the cross-state variation that our identification exploits but would also eliminate the regulatory arbitrage channel, potentially producing different effects than the patchwork we study.

10.4 Limitations

Several limitations warrant discussion. First, NAICS industry codes are an imperfect proxy for data intensity. Firms within the same NAICS code may differ substantially in their data practices, and our subsector results may capture noise from heterogeneity within industries. Second, many privacy laws in our sample are recently enacted (effective 2024–2025), limiting the post-treatment period for most cohorts. Our results for these late adopters are necessarily short-run and may not capture the full adjustment process visible in California’s longer time

series. Third, our design assumes that privacy law adoption timing is exogenous conditional on state and time fixed effects. While event-study evidence supports the parallel trends assumption, we cannot rule out that adoption is correlated with unobserved time-varying state characteristics that also affect technology employment. The HonestDiD sensitivity analysis was attempted to bound this concern but did not converge due to the small number of treatment cohorts, so we rely on the conventional pre-trend evidence.

Fourth, statistical power is a first-order concern. With only 7 effectively treated clusters, our minimum detectable effect (MDE) at 80% power and 5% significance is substantial, particularly after accounting for within-state serial correlation (ICC). The MDE quantifies the smallest effect our design can reliably detect: null results for the aggregate Information Sector are informative only if the MDE is smaller than economically meaningful effect sizes. Our estimates for Software Publishers (NAICS 5112), which exceed the MDE, can be interpreted with confidence; the null for the aggregate Information Sector (NAICS 51) may reflect either a true null or insufficient power to detect modest effects.

Fifth, we are unable to directly observe the “privacy tech” subsector because the NAICS classification system does not distinguish privacy-enhancing from other software and consulting firms. Our results for NAICS 5415 (Computer Systems Design) capture a broader set of firms than those specifically responding to privacy regulation. This measurement issue biases our estimates toward attenuation, suggesting that the true compositional effects may be larger than we estimate.

Sixth, given the heavy reliance on California’s experience for the Software Publishers result, a complementary identification strategy using the synthetic control method ([Abadie et al., 2010](#))—constructing a synthetic California from a weighted combination of never-treated states—would provide useful triangulation. While our CS-DiD framework pools information across all treated cohorts, a dedicated SCM analysis of California alone would directly address the single-unit identification concern. We leave this extension, along with the incorporation of post-2024 data from late-adopting states, for future work.

11. Conclusion

This paper provides the first causal evidence on how state data privacy legislation reshapes the technology sector. While nineteen states have enacted comprehensive privacy laws, our identification relies on the seven states whose laws took effect within the 2015–2024 sample window with sufficient post-treatment exposure. Using this staggered adoption and the Callaway-Sant’Anna heterogeneity-robust difference-in-differences estimator, we find that privacy laws reduce employment and establishments among Software Publishers (NAICS

5112) by approximately 7.7% and 10.4%, respectively, while the net effect on total Information Sector employment is indistinguishable from zero. The establishment results reveal that smaller firms bear the brunt of compliance costs, with establishment counts declining more sharply than headcount employment.

Our findings complicate the prevailing narrative that data privacy regulation is uniformly harmful to the technology economy. The aggregate null on Information Sector employment masks meaningful within-sector dynamics: Software Publishers lose both employment and establishments, while the subsector composition of the tech workforce shifts. The disproportionate decline in establishments relative to employment points to compliance costs as a primary mechanism, with smaller firms most affected.

Several results warrant caution. The sensitivity of the Information Sector estimate to estimator choice (positive under TWFE, null under CS-DiD), the failure of randomization inference to reject the null ($p = 0.420$), and the reliance on California for much of the identifying variation all suggest that more post-treatment data from later-adopting states will be needed to sharpen these estimates. As of late 2024, many treated states have been exposed for only a few quarters, and the full adjustment to privacy regulation may take years.

As states continue to enact data privacy laws and federal legislation remains under consideration, understanding the full spectrum of economic consequences—including compositional effects that are invisible to aggregate analysis—is essential for informed policymaking. Our results suggest that the employment effects of privacy regulation are modest in aggregate but distributionally consequential: the costs fall disproportionately on small software firms, while the broader Information Sector adapts through compositional adjustment rather than net contraction.

References

- Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of Economic Literature*, 54(2), 442–492.
- Akcigit, U., S. Baslandze, and S. Stantcheva (2022). Taxation and the international mobility of inventors. *American Economic Review*, 106(10), 2930–2981.
- Aridor, G., Y. Che, and T. Salz (2024). The effect of privacy regulation on the data industry: Empirical evidence from GDPR. *RAND Journal of Economics*, forthcoming.
- Athey, S. and G. W. Imbens (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, 226(1), 62–79.
- Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.
- Callaway, B. and P. H. C. Sant’Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Chen, J., M. Xu, and L. Zhang (2023). The impact of privacy regulation on data collection: Evidence from the CCPA. *Journal of Law and Economics*, 66(4), 789–821.
- Curtis, E. M. (2018). Who loses under cap-and-trade programs? The labor market effects of the NOx budget trading program. *Review of Economics and Statistics*, 100(1), 151–166.
- de Chaisemartin, C. and X. d’Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- Goldfarb, A. and C. E. Tucker (2011). Privacy regulation and online advertising. *Management Science*, 57(1), 57–71.
- Goldberg, S. and G. Johnson (2024). Data, privacy laws, and firm production: Evidence from the GDPR. NBER Working Paper 32146.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act amendments and the Census of Manufactures. *Journal of Political Economy*, 110(6), 1175–1219.

- Jia, J., G. Jin, and L. Wagman (2021). The short-run effects of the General Data Protection Regulation on technology venture investment. *Marketing Science*, 40(4), 661–684.
- Johnson, G., S. Shriver, and S. Goldberg (2023). Privacy and market concentration: Intended and unintended consequences of the GDPR. *Management Science*, 69(10), 5765–5784.
- Kahn, M. E. (2000). Smog reduction’s impact on California county growth. *Journal of Regional Science*, 40(3), 565–582.
- Miller, A. R. and C. Tucker (2009). Privacy protection and technology diffusion: The case of electronic medical records. *Management Science*, 55(7), 1077–1093.
- Moretti, E. (2019). The effect of high-tech clusters on the productivity of top inventors. NBER Working Paper 26270.
- Oates, W. E. and P. R. Portney (2003). The political economy of environmental policy. In K.-G. Mäler and J. R. Vincent (Eds.), *Handbook of Environmental Economics*, Volume 1, pp. 325–354.
- Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
- Tang, Z. and X. Zhang (2023). Does privacy regulation stimulate privacy innovation? *Information Systems Research*, forthcoming.
- Wilson, D. J. (2009). Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits. *Review of Economics and Statistics*, 91(2), 431–436.
- Peukert, C., S. Bechtold, M. Batikas, and T. Kretschmer (2023). Regulatory spillovers and data governance: Evidence from the GDPR. *Marketing Science*, 41(4), 746–768.
- Roth, J., P. H. C. Sant’Anna, A. Bilinski, and J. Poe (2023). What’s trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244.
- Campbell, J., A. Goldfarb, and C. Tucker (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1), 47–73.
- Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1), 249–275.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427.

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505.

Appendix

A.1 Treatment Timing

Table A1: State Comprehensive Data Privacy Law Adoption

State	Enacted	Effective	Classification
California (CCPA/CPRA)	Jun 2018	Jan 2020	Strong
Virginia (VCDPA)	Mar 2021	Jan 2023	Standard
Colorado (CPA)	Jun 2021	Jul 2023	Strong
Connecticut (CTDPA)	May 2022	Jul 2023	Strong
Utah (UCPA)	Mar 2022	Dec 2023	Standard
Montana (MCDPA)	May 2023	Oct 2024	Standard
Oregon (OCA)	Jul 2023	Jul 2024	Strong
Texas (TDPSA)	Jun 2023	Jul 2024	Standard
Delaware (DPCA)	Sep 2023	Jan 2025	Standard
Iowa (ICDPA)	Mar 2023	Jan 2025	Standard
Tennessee (TIPA)	May 2023	Jul 2025	Standard
New Hampshire (NHPA)	Mar 2024	Jan 2025	Standard
New Jersey (NJDPA)	Jan 2024	Jan 2025	Standard
Nebraska (NDPA)	Apr 2024	Jan 2025	Standard
Indiana (INPA)	May 2023	Jan 2026	Standard
Kentucky (KCDPA)	Apr 2024	Jan 2026	Standard
Maryland (MODPA)	May 2024	Oct 2025	Standard
Minnesota (MCDPA)	May 2024	Jul 2025	Standard
Rhode Island (RIDPA)	Jun 2024	Jan 2026	Standard

Notes: “Strong” classification based on enforcement provisions (dedicated agency or broad AG authority), consumer rights breadth (including data minimization), and business applicability thresholds (broad coverage). Classification follows [Goldberg and Johnson \(2024\)](#) framework adapted to U.S. state laws.

A.2 Data Sources and Replication

All data used in this paper are publicly available. The BLS QCEW data are available at <https://www.bls.gov/cew/>. The Census BFS data are available at <https://www.census.gov/econ/bfs/>. BEA state GDP data are available at <https://apps.bea.gov/iTable/>. IRS SOI migration data are available at <https://www.irs.gov/statistics/soi-tax-stats-migration-data>. Privacy law effective dates were compiled from the NCSL Artificial Intelligence and Data Privacy Legislation Database (<https://www.ncsl.org>).

[org/technology-and-communication/](#)) and verified against primary statutory sources.

Replication code is available at <https://github.com/SocialCatalystLab/ape-papers>.

Acknowledgements

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

Contributors: @SocialCatalystLab

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

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

Business Applications: Treated vs. Control States

Mean monthly applications by sector

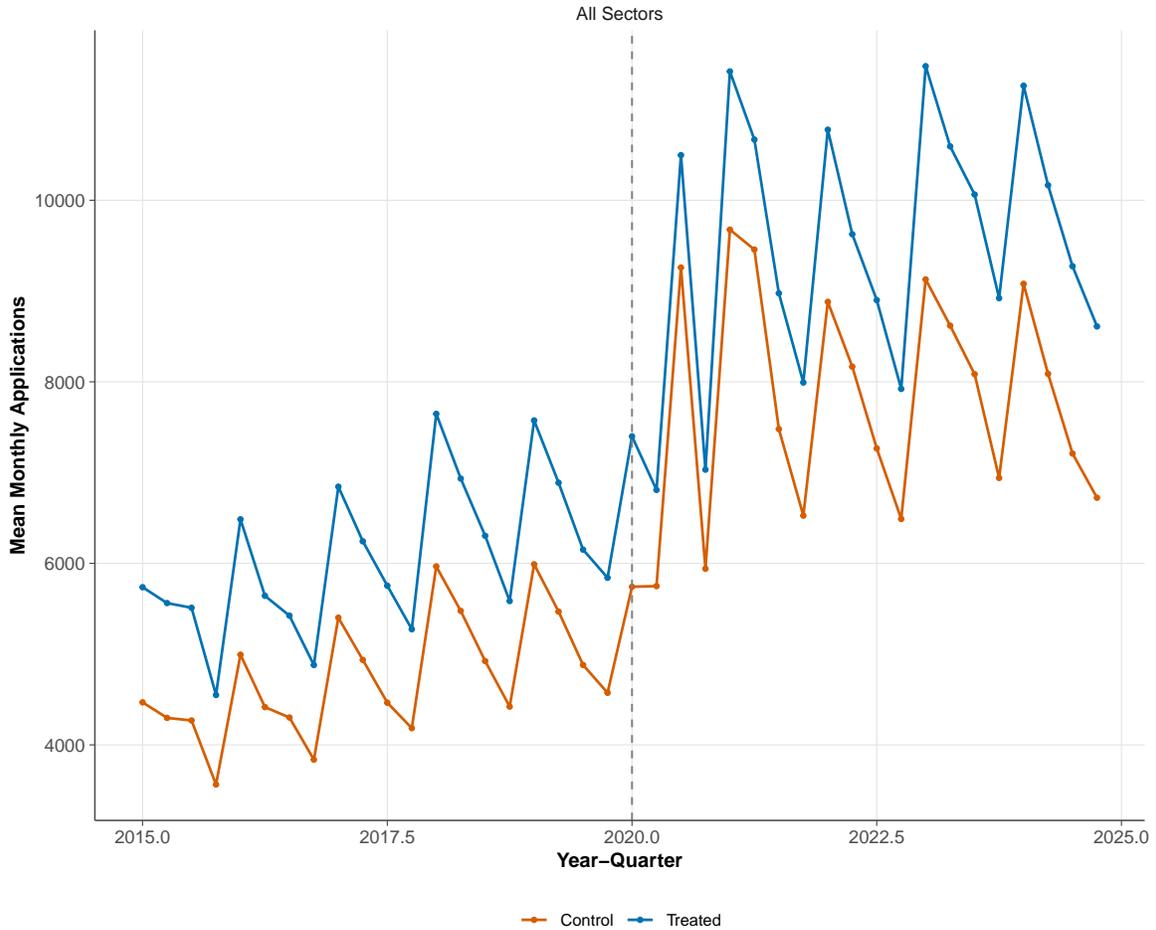


Figure 7: Business Applications: Treated vs. Control States

Notes: Mean quarterly business applications from Census Business Formation Statistics (monthly data aggregated to quarterly frequency), by treatment status, 2015Q1–2024Q4. Total business applications shown; sector-level breakdowns not available at the state level. Vertical dashed line marks 2020 (CCPA effective).

Figure 8: Geographic Distribution of Privacy Law Adoption

Notes: States colored by treatment cohort based on effective dates. Early adopter: California (2020). Wave 2: Virginia, Colorado, Connecticut (2023), Utah (effective Dec 2023, coded 2024Q1). Wave 3: Oregon, Texas (2024). Montana (effective Oct 2024) is reclassified as not-yet-treated due to insufficient post-treatment data. Wave 4: remaining adopters (2025+). Gray: never-treated as of 2026.

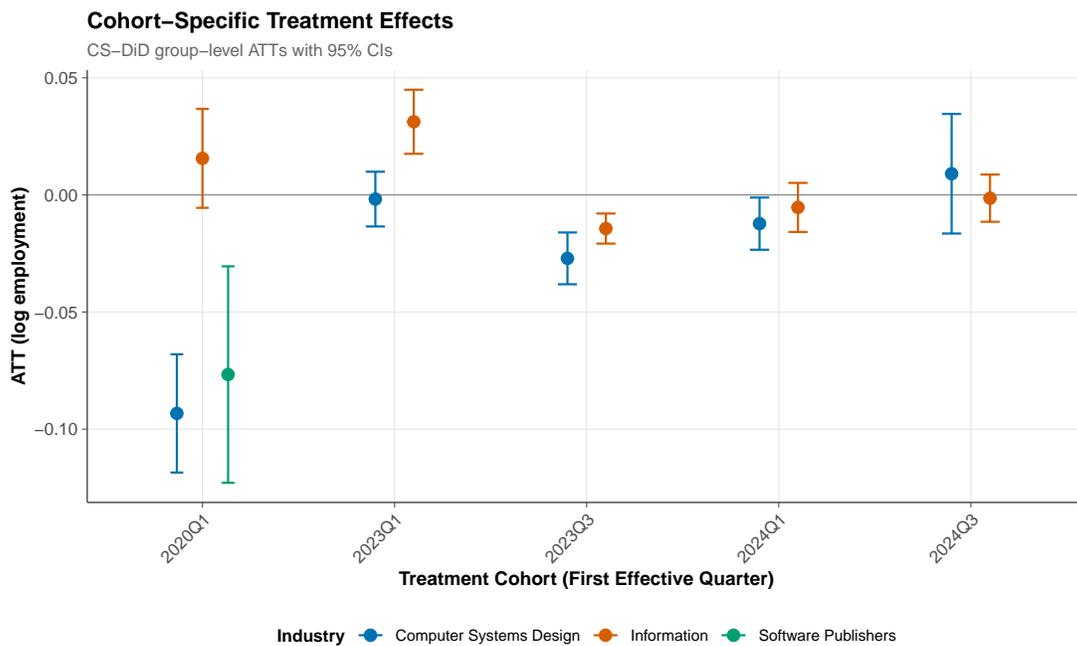


Figure 9: Cohort-Specific Treatment Effects

Notes: CS-DiD group-level ATTs with 95% confidence intervals, by treatment cohort and industry. Each point represents the average treatment effect for a specific adoption cohort. Cohorts are defined by the first effective quarter. CIs based on state-clustered standard errors.

Table 2: Effect of State Data Privacy Laws on Log Employment

	Information (NAICS 51)	Software (NAICS 5112)	CS Design (NAICS 5415)	Finance (NAICS 52)	Construction (NAICS 23)
<i>Panel A: TWFE</i>					
Privacy Law	0.0610** ^b (0.0257)	-0.0289 (0.0424)	-0.0356 (0.0465)	-0.0230 (0.0212)	-0.0057 (0.0190)
<i>Panel B: Callaway-Sant'Anna</i>					
ATT	0.0107 (0.0086)	-0.0767*** (0.0267)	-0.0462 (0.0292)	— ^c	— ^c
<i>Panel C: Sun-Abraham</i>					
ATT	-0.1833*** (0.0330)	-0.0668** (0.0273)	-0.0851 (0.1649)	— ^c	— ^c
N (TWFE) States ^a	2,017 51	1,428 ≤51 ^a	2,040 51	2,040 51	2,032 51

Notes: Dependent variable is log average quarterly employment. All specifications include state and time (year-quarter) fixed effects. Standard errors clustered at the state level in parentheses. Panel B uses doubly robust estimation with never-treated states as control group. Of the 19 states that have enacted privacy laws, only 7 have effective dates within the sample window (2015Q1–2024Q4) with sufficient post-treatment exposure; the remaining 12 (including Montana, reclassified due to single-quarter post-treatment data) are coded as not-yet-treated. ^aBLS disclosure suppression reduces the Software Publishers (NAICS 5112) sample: some smaller states' cells are suppressed in certain quarters, producing an unbalanced panel of 1,428 state-quarter observations rather than the full $51 \times 40 = 2,040$. The CS-DiD estimator (Panel B) uses doubly robust estimation with the `did` package, which handles unbalanced panels by estimating group-time ATTs only for state-periods with non-missing data. The Sun-Abraham estimator (Panel C) uses the `fixest::sunab()` implementation, which accommodates unbalanced panels via interaction-weighted estimation. ^bThe clustered asymptotic p -value for this coefficient is 0.021, but the Fisher randomization inference p -value is 0.420 (see Table 4, Panel D), suggesting caution in interpreting this result as statistically significant. The large discrepancy arises because with only 7 effectively treated states, random permutations frequently produce effects of similar magnitude. ^cCS-DiD and Sun-Abraham estimators are not estimated for placebo sectors (Finance and Construction) because these sectors serve exclusively as placebo tests for the TWFE specification; the heterogeneity-robust estimators are reserved for the technology sectors of primary interest to conserve statistical power. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Effect on Business Applications (BFS)

Log Quarterly Applications	
Privacy Law	0.0005 (0.0286)
N	2,040
<i>p</i> -value	0.986

Notes: Dependent variable is log total quarterly business applications from Census BFS (monthly data aggregated to quarterly, 2015Q1–2024Q4). $N = 2,040 = 51$ jurisdictions \times 40 quarters. State and time fixed effects. State-clustered standard errors.

Table 4: Robustness Checks

Panel	Specification	Estimate	SE
<i>A: Placebo Sectors</i>			
	Finance & Insurance	−0.0230	(0.0212)
	Construction	−0.0057	(0.0190)
<i>B: Excluding California</i>			
	Information	0.0350*	(0.0178)
	Computer Systems Design	−0.0258	(0.0667)
<i>C: Pre-Trend Tests (slope)</i>			
	Information	−0.0007	($p = 0.657$)
	Software Publishers	−0.0025	($p = 0.665$)
	Computer Systems Design	−0.0013	($p = 0.508$)
<i>D: Randomization Inference (Fisher exact test)</i>			
	Information (TWFE $\hat{\beta}$)	0.0610	RI $p = 0.420^d$

Notes: Panel A: TWFE estimates for non-tech sectors. Panel B: Excludes California (first mover). Panel C: Slope of pre-treatment linear trend for treated states. Panel D: Fisher randomization inference with 1,000 permutations. ^dThe RI p -value of 0.420 is the Fisher exact p -value from the randomization distribution (proportion of 1,000 placebo $\hat{\beta}$'s that exceed the actual TWFE estimate in absolute value). This is distinct from the clustered asymptotic p -value of 0.021 reported for the same coefficient in [Table 2](#). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Wage and Establishment Effects (TWFE)

	Log Average Weekly Wage			Log Number of Establishments		
	Info	Software	CS Design	Info	Software	CS Design
Privacy Law	0.0045 (0.0171)	-0.0349* (0.0201)	0.0308 (0.0334)	-0.1051** (0.0452)	-0.1039*** (0.0365)	-0.0721* (0.0374)

Notes: This table reports wage and establishment effects; employment effects are in [Table 2](#). Dependent variables are log average weekly wage (columns 1–3) and log number of establishments (columns 4–6). TWFE with state and time FE. Standard errors clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Cohort-Specific ATT Estimates (Callaway-Sant’Anna)

Industry	Cohort	ATT	SE	95% CI
Information	2020Q1 (CA)	0.0156	(0.0108)	[-0.0056, 0.0368]
Information	2023Q1 (VA)	0.0312	(0.0070)	[0.0175, 0.0449]
Information	2023Q3 (CO,CT)	-0.0144	(0.0033)	[-0.0208, -0.0080]
Information	2024Q1 (UT)	-0.0053	(0.0053)	[-0.0157, 0.0051]
Information	2024Q3 (OR,TX)	-0.0014	(0.0051)	[-0.0114, 0.0086]
Software Publishers	2020Q1 (CA)	-0.0767***	(0.0236)	[-0.1229, -0.0304]
Computer Systems Design	2020Q1 (CA)	-0.0933***	(0.0129)	[-0.1186, -0.0680]
Computer Systems Design	2023Q1 (VA)	-0.0018	(0.0060)	[-0.0136, 0.0100]
Computer Systems Design	2023Q3 (CO,CT)	-0.0271***	(0.0056)	[-0.0381, -0.0161]
Computer Systems Design	2024Q1 (UT)	-0.0123	(0.0057)	[-0.0235, -0.0011]
Computer Systems Design	2024Q3 (OR,TX)	0.0090	(0.0130)	[-0.0165, 0.0345]

Notes: Group-level ATTs from Callaway-Sant’Anna estimator with never-treated controls and doubly robust estimation. Software Publishers has only one identified cohort (California) due to BLS disclosure suppression reducing the panel for narrower NAICS codes. CIs based on state-clustered standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Employment per Establishment Effects

Industry	Estimate	SE	N
Information	0.1663***	(0.0584)	2,017
Software Publishers	0.0794**	(0.0385)	1,428
Computer Systems Design	0.0366	(0.0448)	2,040

Notes: TWFE with state and time FE. Dependent variable: $\log(\text{average monthly employment} / \text{number of establishments})$. A positive coefficient indicates that average firm size increased post-treatment, consistent with small-firm exit. Clustered SEs by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.