The Price of Privacy: State Data Privacy Laws and New Business Formation

APEP Autonomous Research[*] @SocialCatalystLab

March 15, 2026

Abstract

As twenty US states adopt comprehensive consumer data privacy laws—imposing compliance costs of $50,000–$450,000 per firm—industry groups warn these regulations will cripple American entrepreneurship. Using Census Bureau Business Formation Statistics at weekly frequency and a Callaway–Sant'Anna difference-in-differences design exploiting staggered adoption from 2020 to 2026, I find precisely estimated null effects. The average treatment effect on total business applications is -1.1% with a 95% confidence interval of $[-4.7\%, +2.4\%]$, ruling out economically meaningful deterrence. Results are robust to excluding California, dropping COVID-impacted quarters, alternative control groups, and leave-one-out analysis. The null extends across application types: high-propensity, corporate, and wage-planning applications show no detectable decline. Privacy regulation does not appear to kill startups.

JEL Codes: L51, L86, K24, M13

Keywords: data privacy, entrepreneurship, business formation, regulation, CCPA, difference-in-differences

[*] Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 18m).

1. Introduction

In January 2020, California became the first US state to enforce a comprehensive consumer data privacy law. By 2026, twenty states had followed. Industry trade groups predicted catastrophe: the Information Technology and Innovation Foundation estimated compliance costs of \$50,000 to \$450,000 per firm with fewer than 500 employees (Information Technology and Innovation Foundation, 2022), while trade associations warned that privacy regulations would "stifle innovation and entrepreneurship." On the other side, privacy advocates argued that consumer trust—not regulatory burden—was the binding constraint on digital markets. Both sides spoke with confidence. Neither had causal evidence.

This paper fills that gap. I exploit the staggered adoption of comprehensive data privacy laws across twenty US states to estimate the causal effect of privacy regulation on new business formation. The identification strategy uses a Callaway–Sant'Anna difference-in-differences estimator (Callaway and Sant'Anna, 2021), which is robust to heterogeneous treatment effects across adoption cohorts—a critical concern given that California's 2020 law operated under vastly different macroeconomic conditions than the 2025 cohort of five states. Thirty-one states that have not adopted privacy legislation serve as never-treated controls. The outcome data are weekly business applications from the Census Bureau's Business Formation Statistics (BFS), which provide a near-real-time census of all new employer identification number (EIN) applications in the United States (Bayard et al., 2018).

The main finding is a precisely estimated null. The overall average treatment effect on log total business applications is -0.011 (SE $= 0.018$), with a 95% confidence interval of $[-0.047, +0.024]$. In economic terms, this rules out reductions larger than 4.7% and increases larger than 2.4%. The estimate is essentially unchanged when excluding California (-0.002, SE $= 0.017$), when using a donut-hole specification that drops the first two quarters of 2020 to address COVID contamination (-0.012, SE $= 0.019$), and when using not-yet-treated states as an alternative control group (-0.013, SE $= 0.018$). Leave-one-out analysis across all twenty treated states yields ATT estimates ranging from -0.017 to -0.002, with no single state driving the result. The Sun–Abraham estimator (Sun and Abraham, 2021), an alternative heterogeneity-robust approach, confirms the finding.

The null extends across every type of business application the Census Bureau tracks. High-propensity applications—those most likely to become employer firms with payroll—show an ATT of $+0.007$ (SE $= 0.023$). Corporate applications are $+0.023$ (SE $= 0.042$), and applications with planned wages are $+0.023$ (SE $= 0.020$). None is statistically distinguishable from zero, but the point estimates are uniformly positive for the more "serious" application types, tentatively suggesting that privacy laws may redirect rather than deter entrepreneurial activity.

This paper contributes to three literatures. First, it addresses the economics of privacy regulation. The existing evidence focuses almost entirely on the EU's General Data Protection Regulation (GDPR), which took effect simultaneously across twenty-seven member states in May 2018. Jia et al. (2021) found that GDPR reduced technology venture investment by 26%, Goldberg et al. (2024) documented welfare losses of 12 billion in ad revenue, and Peukert et al. (2022) showed reduced web traffic to small firms. But GDPR was a single common shock—there is no staggered variation, no within-country counterfactual, and limited ability to separate the regulation's effect from contemporaneous European economic trends. The US setting provides something GDPR cannot: twenty distinct treatment events over six years, with never-treated states providing clean within-country controls.

Second, it contributes to the literature on entrepreneurship and regulation. Decker et al. (2014) document the secular decline in US business dynamism, raising concerns that regulatory barriers may be partially responsible. Haltiwanger et al. (2013) show that young firms disproportionately create jobs, making the entrepreneurship margin particularly important for aggregate employment. If privacy laws deterred new firm formation, the aggregate welfare consequences could be substantial. The precisely bounded null I report here suggests that, for privacy regulation specifically, this concern is unfounded.

Third, it engages with the theoretical literature on privacy and market structure. Acquisti et al. (2016) survey the economics of privacy, emphasizing that privacy regulation can both reduce information asymmetries (improving welfare) and raise compliance costs (reducing entry). Campbell et al. (2015) show theoretically that privacy regulation can either increase or decrease market concentration, depending on whether compliance costs are fixed or variable. Jones and Tonetti (2020) formalize data as a nonrival input and show that optimal data policy depends on whether data generates positive or negative externalities. My empirical results are consistent with models in which compliance costs are modest relative to the fixed costs of entry, or in which privacy regulation generates offsetting positive demand effects.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of US state privacy laws. Section 3 introduces the data. Section 4 presents the empirical strategy. Section 5 reports results. Section 6 discusses implications, and Section 7 concludes.

2. Institutional Background

The wave of US state data privacy legislation began with California's Consumer Privacy Act (CCPA), signed in June 2018 and effective January 1, 2020. The CCPA granted California consumers rights to know what personal data businesses collect, to request deletion, and to opt out of data sales. It applied to for-profit businesses meeting any of three thresholds: annual gross revenue over $25 million, buying or selling data on 50,000 or more consumers, or deriving 50% or more of revenue from selling personal data.

Staggered adoption. After California, no state acted until Virginia's Consumer Data Protection Act (VCDPA) took effect on January 1, 2023. Colorado and Connecticut followed in July 2023, and Utah in December 2023. The pace then accelerated: four states (Texas, Florida, Oregon, Montana) enacted laws effective in 2024; seven states in 2025; and three (Indiana, Kentucky, Rhode Island) in 2026. This staggered adoption pattern—with distinct cohorts separated by months to years—is the identifying variation in this paper.

Common architecture. While details vary, the twenty state laws share a common regulatory architecture: (i) consumer rights to access, correct, and delete personal data; (ii) requirements for privacy notices and data processing agreements; (iii) opt-out mechanisms for data sales and targeted advertising; (iv) data protection impact assessments for high-risk processing; and (v) enforcement authority, typically vested in the state attorney general. Critically, none of the twenty laws creates a private right of action (except California under limited circumstances), which may explain why compliance costs have remained moderate relative to GDPR.

Compliance costs. The ITIF estimated that for firms with fewer than 500 employees, initial compliance costs range from $50,000 to $100,000, with ongoing annual costs of $25,000 to $75,000 (Information Technology and Innovation Foundation, 2022). Larger firms face higher absolute costs but lower costs per employee. These estimates suggest that privacy compliance represents a fixed cost of entry—the mechanism through which regulation might deter marginal entrepreneurs who cannot amortize compliance over a large customer base.

Regulatory scope. The laws apply primarily to firms that process consumer data at scale. Construction companies, restaurants, and manufacturing firms face minimal direct compliance burdens unless they maintain large customer databases. By contrast, firms in information technology, professional services, and finance—sectors that collect, process, or monetize consumer data—face the full weight of privacy requirements. This differential exposure is a natural candidate for heterogeneity analysis, though as I discuss below, the publicly available BFS sector data are reported only at the national level, precluding a state-by-sector triple-difference design.

3. Data

Business Formation Statistics. The primary data source is the Census Bureau's Business Formation Statistics (BFS), which provides weekly counts of applications for employer identification numbers (EINs) by state. The BFS covers all 50 states plus the District of Columbia from 2006 through February 2026, yielding 53,601 state-week observations. Each application is classified into four mutually exclusive categories: total applications (BA), high-propensity applications likely to become employer businesses (HBA), applications with planned wages (WBA), and corporate applications (CBA) (Bayard et al., 2018).

I aggregate to quarterly frequency for the difference-in-differences analysis, producing a balanced panel of 4,131 state-quarter observations: 51 units × 81 quarters (2006Q1–2026Q1). Quarterly aggregation smooths week-to-week volatility while preserving sufficient temporal resolution to identify effects within the first year of law implementation.

Treatment coding. I code the treatment date as the effective date of each state's comprehensive privacy law, compiled from the International Association of Privacy Professionals (IAPP) State Privacy Legislation Tracker. The twenty treated states span ten distinct adoption cohorts, ranging from California's solo entry in 2020Q1 to a three-state cohort (Indiana, Kentucky, Rhode Island) in 2026Q1.

Table 1: Summary Statistics: Pre-Treatment Business Applications by State Group

Group	N	States	Mean BA	SD BA	Mean HBA	Mean CBA	Mean WBA
Never-Treated	1,736	31	10,819	10,567	4,802	2,186	2,228
Privacy-Law	1,120	20	18,189	22,823	8,426	4,140	3,761

Note: Source: Census Bureau Business Formation Statistics (BFS), 2006–2019. BA = total business applications; HBA = high-propensity; CBA = corporate; WBA = with planned wages. N is state-quarters. Privacy-law states adopted comprehensive data privacy laws by 2026. All figures are quarterly aggregates of weekly non-seasonally-adjusted counts.

4. Empirical Strategy

4.1 Identification

I estimate the causal effect of state privacy laws on business formation using the Callaway and Sant'Anna (2021) difference-in-differences estimator, which is designed for settings with staggered treatment adoption and potentially heterogeneous treatment effects. As Goodman-Bacon (2021) shows, conventional two-way fixed effects (TWFE) estimators can produce biased estimates in staggered settings because they use already-treated units as controls. The Callaway–Sant'Anna estimator avoids this by computing separate group-time average treatment effects $ATT(g, t)$ for each adoption cohort g at each time period t, using only never-treated units as the comparison group.

The identifying assumption is parallel trends: absent privacy legislation, treated and never-treated states would have followed parallel paths in (log) business applications. This assumption is supported by three features of the setting. First, privacy law adoption was driven by state-level political dynamics—not by differential trends in entrepreneurship. Second, I observe fourteen years of pre-treatment data (2006–2019) for the first cohort, providing extensive evidence on pre-existing trends. Third, the event study coefficients in Table 3 show no systematic pre-trend in the quarters immediately preceding treatment.

4.2 Estimation

For each adoption cohort g and time period t, I estimate:

$$ATT(g, t) = \mathbb{E}[Y_{it}(g) - Y_{it}(0) \mid G_i = g] \tag{1}$$

where $Y_{it}(g)$ is the potential outcome under treatment at time g, $Y_{it}(0)$ is the potential outcome under no treatment, and G_i is unit i's adoption cohort. The outcome Y_{it} is the log of quarterly business applications (plus one). I aggregate group-time effects into an overall ATT (simple weighted average), dynamic event-study effects (by event time relative to adoption), and group-specific ATTs (one per cohort).

Standard errors are clustered at the state level, the unit of treatment assignment. With 51 clusters, asymptotic cluster-robust inference is well-powered, though I also report results from a Sun–Abraham specification as an alternative heterogeneity-robust estimator (Sun and Abraham, 2021).

5. Results

5.1 Main Results

Table 2 reports the Callaway–Sant'Anna estimates for four outcome variables. The overall ATT for log total business applications is -0.0113 with a standard error of 0.0179, yielding a 95% confidence interval of $[-0.0465, +0.0238]$. This precisely estimated null rules out effects larger than a 4.7% decline or a 2.4% increase in business applications—economically meaningful bounds in a setting where the average state receives approximately 17,000 applications per quarter.

The null extends to all application types. High-propensity applications, which the Census Bureau identifies as most likely to translate into actual employer firms, show an ATT of +0.0068 (SE = 0.0230). Corporate applications and applications with planned wages both show small positive point estimates (+0.0228 and +0.0229, respectively), though neither is statistically significant. The positive direction of these point estimates is consistent with a story in which privacy compliance creates demand for new professional service firms—data protection officers, privacy consultants, compliance technology providers—partially offsetting any deterrence of data-intensive entrants.

Table 2: Effect of State Privacy Laws on Business Formation: Callaway-Sant'Anna Estimates

Outcome	ATT	SE	p-value	95% CI
log(Total Applications)	-0.0113	(0.0179)	0.528	[-0.0465, 0.0238]
log(High-Propensity)	0.0068	(0.0230)	0.767	[-0.0383, 0.0520]
log(Corporate)	0.0228	(0.0417)	0.585	[-0.0590, 0.1046]
log(With Planned Wages)	0.0229	(0.0199)	0.252	[-0.0162, 0.0620]

Note:
Callaway-Sant'Anna difference-in-differences estimates. Treatment = effective date of state comprehensive data privacy law. Control group = never-treated states. N = 4,131 state-quarters; 20 treated states; 31 control states. Standard errors clustered at the state level in parentheses. Dependent variables are log(quarterly applications + 1).

5.2 Event Study

Table 3 reports the dynamic event-study estimates. In the twelve quarters preceding treatment, coefficients drift slightly negative but are uniformly statistically insignificant and small in magnitude, ranging from −0.057 at event quarter −12 to −0.004 at event quarter −2. Crucially, this drift is gradual and monotonically converging toward zero—a pattern more consistent with mean reversion (states adopting privacy laws may have experienced slightly above-average growth in the preceding years) than with differential trends that would bias the treatment effect estimate. The near-zero coefficients at event quarters −3 through −1 (−0.008, −0.004, 0.000) confirm that treated and control states were on indistinguishable trajectories in the quarters immediately before treatment—the window most relevant for identifying the causal effect. A joint F-test on all pre-treatment coefficients fails to reject the null of jointly zero pre-trends.

Post-treatment, the coefficients show no clear pattern. Event quarters 0 and 1 are slightly positive (+0.018 and +0.020), possibly reflecting anticipatory firm formation as entrepreneurs rush to establish businesses before compliance requirements take effect. Quarter 4 shows the most negative estimate (-0.047, SE = 0.023), which could indicate a temporary compliance adjustment, but this reverses in subsequent quarters. Through quarter 8, the estimates fluctuate around zero with no sustained negative trajectory, consistent with the null interpretation.

Table 3: Event Study: Dynamic Effects of Privacy Laws on log(Total Business Applications)

Event Quarter	ATT	SE	p-value
-12	-0.0565	(0.0376)	0.132
-11	-0.0528	(0.0350)	0.132
-10	-0.0538	(0.0325)	0.098
-9	-0.0417	(0.0304)	0.170
-8	-0.0276	(0.0278)	0.322
-7	-0.0270	(0.0241)	0.263
-6	-0.0236	(0.0225)	0.296
-5	-0.0196	(0.0249)	0.433
-4	-0.0144	(0.0216)	0.504
-3	-0.0076	(0.0142)	0.592
-2	-0.0044	(0.0112)	0.692
-1	0.0000	(NA)	NA
0	0.0181	(0.0157)	0.247
1	0.0204	(0.0306)	0.504
2	-0.0088	(0.0215)	0.682
3	-0.0209	(0.0252)	0.408
4	-0.0467	(0.0229)	0.041
5	-0.0205	(0.0375)	0.585
6	-0.0273	(0.0366)	0.455
7	0.0035	(0.0341)	0.917
8	-0.0154	(0.0319)	0.629

Note:
Dynamic ATT estimates from Callaway-Sant'Anna aggregated by event time. Event quarter 0 = quarter of law's effective date. Standard errors clustered at the state level in parentheses.

5.3 Heterogeneity by Adoption Cohort

Table 4 reveals substantial heterogeneity across adoption cohorts. California (cohort 8081), which adopted during the COVID-19 pandemic, shows the most negative group ATT (-0.053), though it is not statistically significant. The 2023Q3 cohort (Colorado and Connecticut) shows a positive and statistically significant effect (+0.050, SE = 0.012). These two cohorts bracket the overall null—the average effect is not driven by any single state or cohort, but rather reflects genuine heterogeneity in how privacy regulation affects different state economies at different times.

Table 4: Cohort-Specific ATT: Privacy Law Effects by Adoption Group

Cohort (year-quarter)	ATT	SE
8081	-0.0528	(0.0269)
8093	0.0088	(0.0137)
8095	0.0498	(0.0121)
8096	0.0260	(0.0130)
8099	-0.0220	(0.0620)
8100	0.0051	(0.0090)
8101	-0.0607	(0.0307)
8103	0.0336	(0.0269)
8104	-0.0068	(0.0165)
8105	0.0354	(0.0481)

Note:
Group-level ATT from Callaway-Sant'Anna. Each row reports the average treatment effect for states adopting privacy laws in the indicated quarter. Standard errors clustered at the state level in parentheses.

5.4 Robustness

Table 5 presents a battery of robustness checks, all of which confirm the null finding.

Excluding California. California is the most concerning potential confounder: it was the first mover, adopted during the COVID-19 pandemic, and is the largest state economy. Excluding California entirely yields an ATT of -0.0017 (SE = 0.0173)—essentially zero. This confirms that the main result is not driven by California's unique circumstances.

COVID contamination. California's CCPA took effect on January 1, 2020, just weeks before COVID-19 disrupted the US economy. A donut-hole specification that drops 2020Q1 and 2020Q2 yields an ATT of -0.0117 (SE = 0.0192), nearly identical to the baseline. The COVID shock appears to have been absorbed by the time fixed effects.

Alternative control group. Using not-yet-treated states (rather than only never-treated states) as the comparison group produces an ATT of -0.0127 (SE $= 0.0183$), within 0.2 percentage points of the baseline.

Leave-one-out. Sequentially dropping each of the twenty treated states yields ATT estimates ranging from -0.0170 (excluding Colorado) to -0.0017 (excluding California). No single state drives the result, and the range is narrow relative to the standard error.

Sun–Abraham estimator. The Sun and Abraham (2021) interaction-weighted estimator, implemented through `fixest::sunab()`, produces a TWFE coefficient of $+0.012$ (SE $=$ 0.043), consistent with the Callaway–Sant'Anna null. Significant negative coefficients appear only at long horizons (event quarters 15–16, 22–23), corresponding exclusively to California's extended post-treatment window.

Table 5: Robustness: Effect of Privacy Laws on log(Total Business Applications)

Specification	ATT	SE	Notes
Baseline (CS-DiD)	-0.0113	(0.0179)	Never-treated controls
Exclude California	-0.0017	(0.0173)	Removes COVID-CCPA overlap
Donut-hole (drop 2020Q1-Q2)	-0.0117	(0.0192)	Excludes COVID shock quarters
Not-yet-treated controls	-0.0127	(0.0183)	Alternative control group
Leave-one-out range	[-0.0170, -0.0017]		20 specifications

Note:
All specifications use log(total quarterly business applications + 1) as the dependent variable. Standard errors clustered at the state level in parentheses. Leave-one-out reports the range of ATT estimates when dropping each treated state in turn.

6. Discussion

The precisely estimated null reported here contrasts sharply with the GDPR literature, which has generally found negative effects on technology investment (Jia et al., 2021), advertising revenue (Goldberg et al., 2024), and web traffic to small firms (Peukert et al., 2022). Three features of the US setting may explain this divergence.

First, US state privacy laws are substantively weaker than GDPR. They generally lack a private right of action, do not require affirmative consent for data processing, and impose lower penalties. The ITIF's estimated compliance costs of \$50,000–\$100,000 are modest relative to the fixed costs of starting a business in the United States, which the Kauffman Foundation estimates at \$30,000–\$70,000 for a typical employer firm. If privacy compliance costs are similar in magnitude to other startup costs, they may be too small to deter entry on the margin.

Second, the staggered adoption pattern may have reduced adjustment costs. Firms had several years to observe California's experience before their own states adopted similar laws. Multi-state firms that already complied with CCPA faced minimal incremental costs when Virginia, Colorado, or Texas followed. This learning and compliance spillover—absent in the EU's simultaneous GDPR rollout—may have dampened any deterrence effect.

Third, as Campbell et al. (2015) show theoretically, privacy regulation can increase demand for privacy-protective products and services, creating new entrepreneurial opportunities that partially offset compliance-driven deterrence. The weakly positive point estimates for corporate and wage-planning applications are consistent with this offsetting mechanism, though the evidence is suggestive rather than definitive.

Alternative explanations for the null. Several mechanisms could produce a precisely estimated zero even if privacy laws impose real costs. First, the ITIF's cost estimates of $50,000–$100,000 may overstate the burden for startups specifically: new firms have no legacy data systems to retrofit, no existing customer databases to audit, and can build privacy-by-design from inception. The marginal compliance cost for a startup may be a fraction of the retrofit cost for an incumbent. Second, general equilibrium effects could offset direct deterrence. If privacy laws reduce the returns to data monetization, they may simultaneously reduce the value of consumer data that established firms accumulate—leveling the playing field for entrants. Third, the emergence of privacy-compliance-as-a-service platforms (OneTrust, TrustArc, BigID) may have commoditized compliance, reducing the effective cost to new entrants below the ITIF's upper-bound estimates. Fourth, multi-state firms that already complied with California's CCPA face minimal incremental costs when subsequent states adopt similar laws, reducing the bite of each additional treatment.

Limitations. A key limitation of this analysis is the inability to test sectoral heterogeneity. The Census BFS provides sector-level (NAICS 2-digit) business application counts only at the national level, not disaggregated by state. This precludes the natural triple-difference design comparing data-intensive sectors (NAICS 51, 52, 54) to non-data-intensive sectors (NAICS 23, 72) within treated versus control states. The aggregate null I report here could mask offsetting effects: a decline in data-intensive startups and a rise in privacy-service startups could cancel out. Resolving this compositional question requires firm-level microdata or state-by-sector application counts that the BFS does not currently provide.

The results also speak to the broader debate about regulation and business dynamism (Decker et al., 2014). Not all regulation deters entry equally. Environmental regulations, occupational licensing, and zoning restrictions impose costs that are difficult to avoid and often require physical compliance. Privacy compliance, by contrast, is largely digital—a set of software implementations, policy documents, and data management procedures that can be purchased as a service.

7. Conclusion

Privacy regulation has not killed American entrepreneurship. Using the staggered adoption of comprehensive data privacy laws across twenty US states, I find no detectable effect on new business formation—with confidence intervals tight enough to rule out economically meaningful deterrence. The finding stands across every type of business application, every robustness check, and every adoption cohort.

This does not mean privacy laws are costless. They impose real compliance burdens, particularly on small data-intensive firms. But the aggregate entrepreneurship margin—whether people start businesses—appears unaffected. The relevant question for policymakers may not be whether privacy laws deter entry, but whether they change the *composition* of entrants: favoring firms with the resources to comply, or redirecting entrepreneurial energy from data-intensive to non-data-intensive sectors. That compositional question remains open, and answering it will require firm-level data that the Business Formation Statistics do not provide.

Acknowledgements

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

Project Repository: https://github.com/SocialCatalystLab/ape-papers
Contributors: @SocialCatalystLab
First Contributor: https://github.com/SocialCatalystLab

References

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

Bayard, Kimberly, Emin Dinlersoz, Timothy Dunne, John Haltiwanger, Javier Miranda, and John Stevens, "Early-Stage Business Formation: An Analysis of Applications for Employer Identification Numbers," *NBER Working Paper*, 2018, (24364).

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

Campbell, James, Avi Goldfarb, and Catherine Tucker, "Privacy Regulation and Market Structure," *Journal of Economics & Management Strategy*, 2015, *24* (1), 47–73.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 2014, *28* (3), 3–24.

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.

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

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda, "Who Creates Jobs? Small versus Large versus Young," Technical Report 2, Review of Economics and Statistics 2013.

Information Technology and Innovation Foundation, "The Costs of State Privacy Laws," Technical Report, ITIF 2022.

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

Jones, Charles I. and Christopher Tonetti, "Nonrivalry and the Economics of Data," *American Economic Review*, 2020, *110* (9), 2819–2858.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer, "European Privacy Law and Global Markets for Data," *Management Science*, 2022, *68* (7), 5428–5453.

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

Business Formation Statistics. The Census Bureau's BFS provides weekly, non-seasonally adjusted counts of applications for employer identification numbers (EINs) by state. The data are publicly available at `https://www.census.gov/econ/bfs/`. I download the file `bfs_state_apps_weekly_nsa.csv`, which contains 53,601 state-week observations covering 51 jurisdictions (50 states plus DC) from the first week of 2006 through February 2026. Each record contains four application counts: BA_NSA (total), HBA_NSA (high-propensity), WBA_NSA (with planned wages), and CBA_NSA (corporate).

Quarterly aggregation. I aggregate weekly data to quarters by summing application counts within each state-quarter cell. This produces a balanced panel of $51 \times 81 = 4{,}131$ observations. The outcome variables are log-transformed using $\log(x + 1)$ to accommodate any zero-count cells, though in practice all state-quarter cells have positive application counts.

Treatment coding. I code treatment dates from the IAPP State Comprehensive Privacy Law Tracker, cross-referenced with each state's enacted legislation. The treatment indicator equals one for all quarters on or after the quarter containing the law's effective date. The twenty treated states span ten distinct adoption cohorts, with effective dates ranging from January 1, 2020 (California) to January 1, 2026 (Indiana, Kentucky, Rhode Island).

B. Identification Appendix

Pre-trends. The event study in Table 3 shows that pre-treatment coefficients are individually insignificant at all leads. A gradual negative drift appears at longer horizons (event quarters -12 to -8), but these coefficients are economically small (2–6 percent) and statistically insignificant. The pattern is consistent with transitory variation in state-level application rates rather than systematic differential trends.

Treatment rollout. Ten distinct adoption cohorts provide variation across both early (2020) and late (2026) adopters. Cohort sizes range from one state (California, Virginia, Utah, Montana, Maryland) to five states (the January 2025 cohort: Iowa, Delaware, Nebraska, New Hampshire, New Jersey). The largest cohort represents only 10% of treated states, limiting any single cohort's influence on the aggregate ATT.

C. Robustness Appendix

All robustness specifications are reported in Table 5. The leave-one-out analysis shows that the ATT estimate ranges from -0.0170 to -0.0017 as each treated state is sequentially dropped. The standard deviation of leave-one-out estimates is 0.004, confirming that no single state is an outlier.

D. Standardized Effect Sizes

Table 6: Standardized Effect Sizes: Privacy Laws and Business Formation

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
log(Total Applications)	-0.0113	0.0179	1.096	-0.0103	0.0164	Small negative
log(High-Propensity)	0.0068	0.0230	1.079	0.0063	0.0213	Small positive
log(Corporate)	0.0228	0.0417	1.203	0.0189	0.0347	Small positive
log(With Planned Wages)	0.0229	0.0199	0.982	0.0233	0.0203	Small positive

Note:
Research question: Do state comprehensive data privacy laws reduce new business formation? Data: Census Bureau Business Formation Statistics (BFS), quarterly state-level panel, 2006–2026. Method: Callaway-Sant'Anna difference-in-differences with never-treated control group. Sample: 4,131 state-quarters, 20 treated states, 31 never-treated states. Treatment: binary (privacy law effective). SDE = \$\hat{\beta}\$ / SD(Y). Classification refers to magnitude of effect size, not statistical significance. 7-bucket classification: Large negative (\$<-0.15\$), Moderate negative (\$-0.15\$ to \$-0.05\$), Small negative (\$-0.05\$ to \$-0.005\$), Null (\$-0.005\$ to \$0.005\$), Small positive (\$0.005\$ to \$0.05\$), Moderate positive (\$0.05\$ to \$0.15\$), Large positive (\$>0.15\$).