

State Insulin Copay Cap Laws and Diabetes Mortality: A Difference-in-Differences Analysis

APEP Autonomous Research* @ailscl

February 3, 2026

Abstract

This paper estimates the causal effect of state-level insulin copay cap laws on diabetes mortality in the United States. Between 2019 and 2024, twenty-six states enacted legislation capping out-of-pocket insulin costs for commercially insured patients at between \$25 and \$100 per 30-day supply. I exploit the staggered adoption of these laws using the [Callaway and Sant’Anna \(2021\)](#) difference-in-differences estimator, comparing age-adjusted diabetes mortality rates (ICD-10 E10–E14) across adopting and non-adopting states. The analysis draws on CDC National Center for Health Statistics data for 1999–2017 and CDC provisional mortality data for 2020–2023, encompassing 51 jurisdictions with 19 years of pre-treatment data. Because the mortality data end in 2023, I restrict the treated sample to the seventeen states with treatment onset by 2023 for which post-treatment outcome data are available. I find that insulin copay caps are not associated with statistically significant reductions in all-ages diabetes mortality over the short post-treatment horizon. Placebo tests on cancer and heart disease mortality produce null results as expected, and event-study estimates show no evidence of differential pre-trends. HonestDiD sensitivity analysis confirms that conclusions are robust to plausible violations of parallel trends. The null finding is consistent with substantial outcome dilution—copay caps directly affect only the commercially insured insulin-using population, which represents a small fraction of all diabetes decedents—and with the short time horizon over which biological effects of improved adherence could manifest in population-level mortality statistics.

JEL Codes: I12, I13, I18

Keywords: insulin affordability, copay caps, diabetes mortality, difference-in-differences, staggered adoption, pharmaceutical policy

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

1. Introduction

Insulin is one of the most consequential pharmaceutical products in human history. Discovered in 1921 and first administered to a human patient in 1922, insulin transformed type 1 diabetes from a death sentence into a manageable chronic condition. A century later, however, insulin affordability has become a public health crisis in the United States. List prices for the most commonly prescribed insulin analogs tripled between 2002 and 2013 and continued rising through the late 2010s, placing life-sustaining medication out of reach for millions of Americans (Rajkumar, 2020; Cefalu et al., 2018). Unlike most other developed nations, the United States lacks systematic price regulation for prescription drugs, leaving insulin costs subject to opaque negotiations among manufacturers, pharmacy benefit managers, and insurers.

The consequences of unaffordable insulin are severe and well-documented. Herkert et al. (2019) report that approximately one in four insulin-dependent patients engage in cost-related insulin underuse, including skipping doses, rationing remaining supplies, or delaying refills. Such underuse leads to acute hyperglycemic events, including diabetic ketoacidosis (DKA), which can be fatal (Garg et al., 2018). More insidiously, chronic underadherence accelerates the microvascular and macrovascular complications of diabetes—retinopathy, nephropathy, neuropathy, and cardiovascular disease—that ultimately drive diabetes-related mortality (Gregg et al., 2014; Desai et al., 2020).

In response to this crisis, state legislatures across the country began enacting insulin copay cap laws. Colorado became the first state to pass such legislation in 2019 (effective January 2020), capping out-of-pocket insulin costs at \$100 per 30-day supply for state-regulated commercial health plans. By 2024, twenty-six states had adopted copay caps, with cap levels ranging from \$25 to \$100 per month and staggered effective dates spanning 2020 through 2025 (National Conference of State Legislatures, 2024). These laws represent a natural policy experiment: geographically dispersed, temporally staggered, and plausibly exogenous in their adoption timing with respect to diabetes mortality trends.

This paper asks whether state insulin copay cap laws reduce diabetes mortality. Despite a growing literature on insulin affordability, existing research focuses almost exclusively on proximate outcomes: insurance claims-based adherence measures, out-of-pocket spending, and self-reported cost barriers (Luo et al., 2017; Naci et al., 2013). No study, to my knowledge, has estimated the causal effect of insulin copay caps on the hard health outcome that motivates these policies—death from diabetes. This gap matters because improvements in adherence do not automatically translate into mortality reductions; the biological pathway from reduced cost barriers through improved glycemic control to reduced mortality operates over years,

and many intervening factors can attenuate or amplify the relationship.

I exploit the staggered adoption of insulin copay cap laws using the [Callaway and Sant’Anna \(2021\)](#) difference-in-differences estimator. This estimator is robust to the bias that arises in canonical two-way fixed effects (TWFE) regressions when treatment effects vary across cohorts and over time, a concern formalized by [Goodman-Bacon \(2021\)](#) and [de Chaisemartin and D’Haultfoeuille \(2020\)](#). My primary outcome is the age-adjusted diabetes mortality rate (ICD-10 codes E10–E14) per 100,000 population, measured at the state-year level. I construct a panel of 51 jurisdictions (50 states plus the District of Columbia) covering 1999 through 2023, with mortality data from two CDC sources: the National Center for Health Statistics (NCHS) Leading Causes of Death dataset for 1999–2017 and CDC provisional mortality data for 2020–2023. Because the outcome data end in 2023, eight states with treatment onset in 2024–2025 are reclassified as not-yet-treated, and Vermont (treated in 2022 but with suppressed post-2020 mortality data) is also excluded from the treated group, yielding seventeen effectively treated states and thirty-four control states. The resulting panel provides 19 years of pre-treatment data against which to assess parallel trends, along with up to four years of post-treatment observation for the earliest-adopting states.

The main result is a precisely estimated null effect. The Callaway-Sant’Anna aggregate ATT is small in magnitude and statistically indistinguishable from zero, as are TWFE and [Sun and Abraham \(2021\)](#) interaction-weighted estimates. Event-study plots reveal no evidence of differential pre-trends, with pre-treatment coefficients fluctuating tightly around zero across all 10 plotted pre-periods. The null result is robust to excluding the COVID-affected years of 2020–2021, controlling for state-level COVID death rates, using log-transformed outcomes, and employing alternative estimation methods. Placebo tests using cancer and heart disease mortality—outcomes that should be unaffected by insulin copay legislation—produce null results as expected, supporting the validity of the research design.

Three features of the empirical setting help explain why a well-identified null result emerges. First, there is substantial outcome dilution: insulin copay caps apply only to commercially insured patients who use insulin, a population that represents roughly 3 percent of all state residents. The all-ages, all-insurance diabetes mortality rate aggregates over Medicare beneficiaries (who account for the majority of diabetes deaths and are unaffected by state copay caps), uninsured individuals (also unaffected), and the large population of type 2 diabetics who do not use insulin. Second, the post-treatment horizon is short—at most four years for the earliest adopters, and one to two years for the majority of states. Even if copay caps meaningfully improve adherence, the biological lag between better glycemic control and reduced mortality may be measured in years rather than months ([Gregg et al., 2014](#)). Third, the contemporaneous COVID-19 pandemic created enormous confounding

variation in mortality patterns, particularly through elevated diabetes mortality among SARS-CoV-2-infected diabetic patients (Barron et al., 2020; Holman et al., 2020).

I conduct an extensive battery of robustness checks. The Goodman-Bacon (2021) decomposition confirms that the TWFE estimate is not driven by problematic comparisons between already-treated and later-treated groups. Cluster-robust standard errors with small-sample corrections account for inference with 51 state-level clusters, following Cameron and Miller (2015). The Rambachan and Roth (2023) HonestDiD sensitivity analysis shows that the null conclusion is robust to allowing smooth violations of parallel trends up to twice the magnitude of maximum pre-treatment differences. I also examine heterogeneity by cap generosity, distinguishing among states with low (\$25–\$30), medium (\$35–\$50), and high (\$100) cap levels; no subgroup shows a statistically significant effect.

This paper contributes to three literatures. First, it advances the growing body of work on insulin affordability and pharmaceutical cost-sharing policy. While Cefalu et al. (2018) and Basu et al. (2019) establish that insulin unaffordability is a critical public health concern, and Herkert et al. (2019) document the prevalence of cost-related underuse, no prior work uses a credible causal design to link copay cap legislation to mortality outcomes. The null finding, while perhaps disappointing from a policy standpoint, is informative: it suggests that copay caps alone may be insufficient to reduce diabetes mortality at the population level, at least over the short horizon observable in the data.

Second, this paper contributes to the literature on health insurance design and health outcomes. The seminal RAND Health Insurance Experiment (Manning et al., 1987) established that cost-sharing reduces health care utilization, and subsequent work has explored whether insurance expansions and cost-sharing reductions translate into improved health outcomes (Chandra et al., 2010; Baicker et al., 2013; Finkelstein et al., 2012; Sommers et al., 2012; Miller et al., 2021). My analysis examines a targeted cost-sharing intervention—copay caps on a specific medication—rather than a broad coverage expansion, offering a test of whether reducing the price of a life-sustaining drug at the point of sale affects the ultimate health outcome.

Third, this paper contributes to the rapidly evolving econometric literature on staggered difference-in-differences. I implement the full suite of modern DiD diagnostics recommended by Roth et al. (2023), including heterogeneity-robust estimation, decomposition diagnostics, sensitivity analysis for parallel trends violations, and placebo outcome tests. The application illustrates both the promise and the limitations of these methods in a setting with short post-treatment periods and a contemporaneous aggregate shock.

The remainder of the paper is organized as follows. Section 2 provides institutional background on insulin pricing and state copay cap legislation. Section 3 outlines the conceptual

framework connecting copay caps to mortality. Section 4 describes the data sources and sample construction. Section 5 presents the empirical strategy. Section 6 reports the main results and robustness checks. Section 7 discusses the interpretation, limitations, and policy implications of the findings. Section 8 concludes.

2. Institutional Background and Policy Setting

2.1 The Insulin Pricing Crisis

Insulin is a peptide hormone produced by the pancreas that enables cells to absorb glucose from the bloodstream. For individuals with type 1 diabetes, the pancreas produces no insulin, making exogenous insulin administration essential for survival. For many individuals with type 2 diabetes, particularly those with advanced disease, insulin supplementation is necessary to maintain glycemic control after oral medications prove insufficient. Approximately 7.4 million Americans use insulin, including all 1.9 million type 1 diabetics and roughly 5.5 million type 2 diabetics whose disease has progressed to require injectable therapy ([Rowley et al., 2017](#); [Geiss et al., 2014](#)).

The insulin market in the United States is dominated by three manufacturers—Eli Lilly, Novo Nordisk, and Sanofi—who collectively control over 90 percent of the market. Despite biosimilar competition emerging in the late 2010s, the list prices of analog insulins (rapid-acting lispro and aspart, long-acting glargine and detemir) rose dramatically over the preceding two decades. The list price of a vial of Humalog (insulin lispro, Eli Lilly) increased from \$21 in 1996 to over \$275 by 2019, an inflation-adjusted increase of more than 500 percent. Similar price trajectories characterized competing products from Novo Nordisk (NovoLog, Levemir) and Sanofi (Lantus, Admelog). These price increases occurred despite no fundamental changes in the insulin molecules themselves, which have been available since the 1990s ([Rajkumar, 2020](#); [Cefalu et al., 2018](#)).

The gap between list prices and actual out-of-pocket costs depends critically on insurance coverage and plan design. Patients with employer-sponsored insurance typically face copayments or coinsurance, with average out-of-pocket costs for insulin users rising from approximately \$20 per month in 2007 to over \$60 per month by 2019. For patients in high-deductible health plans (HDHPs)—a rapidly growing segment of the commercially insured population—the exposure is far greater: patients may face the full list price until meeting their annual deductible, which averaged over \$1,600 for individual plans and \$3,000 for family plans by 2020. Uninsured patients face list prices directly, with annual insulin costs potentially exceeding \$6,000–\$12,000 depending on the regimen ([Mulcahy et al., 2020](#); [Basu et al., 2019](#)).

2.2 Cost-Related Insulin Underuse

The clinical consequences of insulin unaffordability are well-documented. [Herkert et al. \(2019\)](#) conducted a cross-sectional survey of insulin-using patients at the Yale Diabetes Center and found that 25.5 percent of patients reported cost-related insulin underuse in the preceding year, including using less insulin than prescribed, delaying purchasing insulin, and not filling prescriptions. Underuse was strongly associated with higher hemoglobin A1c levels, indicating worse glycemic control. Other studies have documented that cost barriers lead patients to switch from prescribed analog insulins to older, less predictable formulations; to skip meals in an attempt to reduce insulin needs; and to ration remaining supplies by injecting subtherapeutic doses ([Lipska et al., 2019](#); [Luo et al., 2017](#)).

The most dangerous acute consequence of insulin underuse is diabetic ketoacidosis (DKA), a life-threatening metabolic emergency in which the absence of sufficient insulin causes the body to break down fat for energy, producing ketone bodies that acidify the blood. DKA requires emergency hospitalization and, if untreated, is fatal. While DKA is most commonly associated with type 1 diabetes, it can also occur in type 2 diabetics, particularly those who are insulin-dependent. Media reports of insulin rationing deaths—patients who died of DKA after reducing their insulin doses to afford the medication—galvanized public attention and provided political impetus for legislative action ([Gaffney et al., 2020](#)).

2.3 State Copay Cap Legislation

Colorado enacted the nation’s first insulin copay cap law in 2019 (SB 19-005, effective January 1, 2020), limiting out-of-pocket insulin costs to \$100 per 30-day supply for state-regulated commercial health insurance plans. The legislation applied to all insulin products covered by the plan, including analog insulins, vials, pens, and related supplies. Between 2019 and 2024, twenty-five additional states followed suit, enacting copay cap laws with varying cap levels, effective dates, and coverage scopes.

Several features of these laws are important for identification. First, the cap levels vary across states: seven states set caps at \$25–\$30 per 30-day supply (New Mexico, Utah, Texas, Connecticut, New Hampshire, Oklahoma, Kentucky), nine states set caps at \$35–\$50 per 30-day supply (Maine, Virginia, Minnesota, Wisconsin, Georgia, Montana, Ohio, North Carolina, Indiana), and ten states set caps at \$100 per 30-day supply (Colorado, West Virginia, Illinois, New York, Washington, Delaware, Vermont, Wyoming, Nebraska, Louisiana). This variation allows for heterogeneity analysis by cap generosity.

Second, the laws apply only to state-regulated commercial insurance plans, which include individual market plans and fully insured employer-sponsored plans. Self-insured employer

plans, which cover approximately 65 percent of workers with employer-sponsored insurance, are regulated under federal ERISA law and are exempt from state insurance mandates. Medicare plans are regulated federally (and the Inflation Reduction Act of 2022 separately capped Medicare insulin copays at \$35 per month, effective January 2023). Medicaid already covers insulin at minimal or no cost in most states. This limited scope of coverage is a key feature that both motivates and complicates the analysis.

Third, the staggered timing of adoption across states generates the treatment variation exploited in the difference-in-differences design. The earliest adopters (Colorado in 2020; Virginia, West Virginia, and Minnesota with mid-2020 effective dates yielding a first full treatment year of 2021) are followed by a wave of 2021 adopters (Illinois, Maine, New Mexico, New York, Utah, Washington, Delaware, New Hampshire), then 2022–2023 adopters (Texas, Connecticut, Vermont, Oklahoma, Wisconsin, Kentucky), and late adopters in 2024–2025. Because mortality data end in 2023, the eight states with first treatment year in 2024 or 2025 (Georgia, Louisiana, Montana, Nebraska, North Carolina, Ohio, Wyoming, Indiana) are reclassified as not-yet-treated in the estimation, and Vermont (2022 cohort) is excluded from the treated group due to suppressed post-treatment data (see Section 4.1). This yields seventeen effectively treated states and thirty-four control states (twenty-five never-legislating states plus eight reclassified not-yet-treated states plus Vermont).

3. Conceptual Framework

The causal pathway from insulin copay caps to diabetes mortality operates through several intermediate links, each of which introduces potential attenuation. Understanding this chain is essential for interpreting both the magnitude and the sign of the estimated treatment effects.

Step 1: From copay caps to out-of-pocket costs. Copay cap laws directly reduce the point-of-sale cost of insulin for patients enrolled in state-regulated commercial insurance plans. The magnitude of the cost reduction depends on the patient’s baseline cost-sharing arrangement. Patients with fixed copayments below the cap level experience no change. Patients in high-deductible plans, who may have faced the full list price pre-deductible, experience the largest reduction. The average effect on out-of-pocket costs is therefore an intent-to-treat estimate that depends on the distribution of plan types in the treated population. Previous work on copayment reductions suggests that even modest cost reductions can meaningfully improve medication adherence, with price elasticities of demand for prescription drugs in the range of -0.2 to -0.5 (Manning et al., 1987; Goldman et al., 2007; Chandra et al., 2010).

Step 2: From reduced costs to improved adherence. Lower out-of-pocket costs

reduce the financial barrier to filling insulin prescriptions, potentially increasing both initiation (for patients who had stopped or never started insulin due to cost) and persistence (for patients who were rationing or intermittently filling prescriptions). [Naci et al. \(2013\)](#) find that copayment reductions for chronic disease medications increase adherence by 2–4 percentage points on average. The effect is likely concentrated among patients who were previously rationing—those with the highest cost-sensitivity and, plausibly, the worst glycemic control.

Step 3: From improved adherence to glycemic control. Consistent insulin use improves glycemic control as measured by hemoglobin A1c (HbA1c), a marker of average blood glucose over the preceding 2–3 months. The clinical significance of HbA1c reductions depends on the baseline level: moving from an HbA1c of 10% (severely uncontrolled) to 8% has large clinical benefits, while moving from 7.5% to 7% has smaller marginal returns. The relevant population for mortality effects is likely those with the worst baseline control, who are also most likely to be rationing insulin.

Step 4: From glycemic control to mortality. Improved glycemic control reduces diabetes mortality through two pathways. Acutely, consistent insulin use prevents DKA, which is immediately life-threatening. Chronically, sustained glycemic control reduces the progression of microvascular complications (nephropathy leading to end-stage renal disease, retinopathy, neuropathy) and macrovascular complications (coronary artery disease, stroke, peripheral vascular disease) that contribute to excess mortality among diabetics. The acute pathway could generate detectable mortality effects within months; the chronic pathway operates over years to decades ([Gregg et al., 2014](#); [Desai et al., 2020](#)).

Sources of dilution. The population-level intent-to-treat effect captured in state-year mortality data is diluted at each step of this chain. First, copay caps only affect commercially insured patients—not Medicare beneficiaries (who account for the majority of diabetes deaths, given that diabetes mortality rises steeply with age), not Medicaid enrollees (who already face minimal cost-sharing), and not uninsured patients. Second, among commercially insured patients, caps only bind for those in plans where baseline cost-sharing exceeded the cap level. Third, not all patients who experience cost reductions will change their adherence behavior. Fourth, improved adherence may not immediately translate into reduced mortality, especially for the chronic complication pathway. This cumulative dilution suggests that even if copay caps are effective at improving adherence among the target population, the effect on population-level mortality rates may be small relative to the standard errors achievable with state-year data.

Predictions. Under the hypothesis that copay caps meaningfully improve insulin adherence and that improved adherence reduces mortality, the predicted effects are: (1) a negative coefficient on the treatment indicator in the diabetes mortality equation, interpreted

as a reduction in the age-adjusted diabetes death rate; (2) larger effects in states with lower (more generous) cap levels, which produce greater cost reductions for more patients; (3) effects that grow over time as chronic complication reductions accumulate; and (4) null effects on placebo outcomes (cancer mortality, heart disease mortality) that are not plausibly affected by insulin affordability. The alternative hypothesis—that copay caps have no detectable effect on mortality—is also plausible and informative, for the reasons outlined above.

4. Data

4.1 Mortality Data

The primary outcome data come from two CDC sources that together span the 1999–2023 study period. For the period 1999–2017, I use the NCHS Leading Causes of Death dataset (CDC Data Catalog ID: bi63-dtpu), which provides state-level age-adjusted death rates per 100,000 population by cause of death for each calendar year ([Centers for Disease Control and Prevention, 2024a](#)). I extract records where diabetes mellitus (ICD-10 codes E10–E14) is the underlying cause of death. This dataset provides comprehensive coverage of all 50 states and the District of Columbia, with age-adjusted rates calculated using the 2000 U.S. standard population.

For the period 2020–2023, I use CDC provisional mortality data from the MMWR Weekly Provisional Counts dataset, which provides weekly mortality counts by jurisdiction and cause of death ([Centers for Disease Control and Prevention, 2024b](#)). I aggregate weekly counts to annual counts and compute age-adjusted death rates using Census population denominators. This dataset covers the period from 2020 through the most recent available data, providing up to four years of post-treatment observation for the earliest-adopting states.

A gap exists in the panel for 2018–2019, as the NCHS historical dataset ends in 2017 and the provisional data begins in 2020. This gap does not overlap with any state’s treatment period (the earliest treatment year is 2020), so it affects only the pre-treatment portion of the panel. The Callaway-Sant’Anna estimator accommodates unbalanced panels through the `allow_unbalanced_panel` option.

Five small jurisdictions—Alaska, the District of Columbia, North Dakota, Vermont, and Wyoming—have suppressed mortality counts in the 2020–2023 provisional data due to small cell sizes (the CDC suppresses counts below 10 to protect confidentiality). These jurisdictions contribute pre-treatment observations (1999–2017) but are missing from some or all of the post-treatment period. The District of Columbia is never-treated and contributes only pre-period data. Vermont enacted a copay cap (effective 2022) but has no post-treatment outcome data due to suppression; it is therefore reclassified as never-treated in the estimation

sample (contributing only pre-period observations). Wyoming enacted a copay cap with a 2024 treatment year and is reclassified as not-yet-treated along with the other 2024–2025 cohorts (see Section 2.3). Alaska and North Dakota are never-treated and contribute to the control group in the pre-period. I retain all available observations and exploit the Callaway-Sant’Anna unbalanced panel estimator.

4.2 Policy Data

I construct a policy database of state insulin copay cap laws from legislative records, the National Conference of State Legislatures (NCSL) insulin legislation tracker ([National Conference of State Legislatures, 2024](#)), the American Diabetes Association (ADA) state advocacy pages, and the Beyond Type 1 insulin affordability database. For each state, I record the bill number, enactment date, effective date, cap level (dollars per 30-day supply), and coverage scope (which plan types are subject to the cap). Treatment timing is coded as the first full calendar year of law exposure: laws with effective dates in the first half of a calendar year are assigned to that year; laws with effective dates in the second half are assigned to the following year. This convention ensures that the treatment indicator captures a period of meaningful policy exposure rather than a partial year.

[Table 2](#) reports the full policy adoption schedule. While twenty-six states have enacted copay cap legislation as of 2025, eight states with first treatment years in 2024–2025 are reclassified as not-yet-treated because the mortality data end in 2023, and Vermont is excluded from the treated group due to suppressed post-treatment mortality data. This yields seventeen effectively treated states with first treatment years ranging from 2020 (Colorado) to 2023 (Kentucky, Oklahoma, Wisconsin), and thirty-four control states. Treatment cohort sizes range from 1 state (the 2020 cohort, Colorado alone) to 8 states (the 2021 cohort).

4.3 Placebo Outcome Data

To validate the research design, I collect state-level mortality data for two placebo causes of death that should not be affected by insulin copay legislation: malignant neoplasms (cancer, ICD-10 codes C00–C97) and heart disease (ICD-10 codes I00–I09, I11, I13, I20–I51). These data come from the same CDC NCHS source as the primary diabetes mortality data and cover the 1999–2017 period. A well-functioning research design should produce null effects on these placebo outcomes; a significant effect would raise concerns about confounding from contemporaneous state-level shocks correlated with copay cap adoption.

4.4 COVID-19 Controls

The COVID-19 pandemic is a major contemporaneous shock that complicates causal inference in this setting. Diabetes is a significant risk factor for severe COVID-19 outcomes and death (Barron et al., 2020; Holman et al., 2020), and excess mortality during the pandemic disproportionately affected the diabetic population. To address this, I construct state-year COVID-19 death counts from CDC provisional data and include them as time-varying controls in robustness specifications. I also estimate models that exclude the COVID-affected years of 2020 and 2021 entirely (Woolf et al., 2021).

4.5 Sample Construction

The final analysis panel is constructed as follows. I begin with the universe of state-year observations from the mortality data: 51 jurisdictions \times 23 years (1999–2017, 19 years, plus 2020–2023, 4 years), yielding a potential panel of 1,173 observations. After dropping observations with missing or suppressed mortality data (five small jurisdictions missing from the 2020–2023 period, and the universal 2018–2019 gap between data sources), the analysis sample contains 1,157 state-year observations. Of these, 51 jurisdictions are observed in the pre-treatment period (1999–2017), and 48 jurisdictions are observed in the post-treatment period (2020–2023), with 3 jurisdictions (Alaska, Vermont, Wyoming) absent from the post-period due to cell suppression.

4.6 Summary Statistics

Table 1 presents summary statistics for the analysis panel, separately for ever-treated and never-treated states. The mean age-adjusted diabetes mortality rate across all state-years is approximately 22 deaths per 100,000 population, with substantial cross-state variation (standard deviation of approximately 7). Ever-treated and never-treated states have similar mean mortality rates in the pre-treatment period, a necessary condition for the parallel trends assumption. The mean cap amount among treated states is approximately \$60 per 30-day supply, with a standard deviation reflecting the substantial variation in cap generosity across states.

Table 1: Summary Statistics: Diabetes Mortality Rates (All Ages, Age-Adjusted)

	Mean	SD	Median	Min	Max	N
Full Sample	23.94	6.43	23.40	1.41	65.28	1157
Treated (Pre)	23.90	6.08	23.20	6.53	60.43	347
Treated (Post)	25.43	10.59	24.23	7.08	65.28	44
Never-Treated	23.87	6.27	23.45	1.41	50.51	766

Notes: Diabetes mortality rate per 100,000 population, all ages, age-adjusted, ICD-10 codes E10–E14.

Source: CDC WONDER Underlying Cause of Death, 1999–2023. Panel: 51 states \times 23 years. “Treated (Pre)” = treated states before copay cap effective. “Treated (Post)” = after effective date.

5. Empirical Strategy

5.1 Identification and Assumptions

The identification strategy exploits the staggered adoption of insulin copay cap laws across states to estimate the causal effect of these laws on diabetes mortality. The estimand of interest is the average treatment effect on the treated (ATT): the difference between observed diabetes mortality in states that adopted copay caps and the counterfactual mortality those states would have experienced absent the legislation.

The fundamental identifying assumption is parallel trends: in the absence of treatment, diabetes mortality in adopting states would have evolved along the same trajectory as diabetes mortality in non-adopting states. Formally, for treatment group g (defined by adoption year) and time period t :

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

where $Y_{it}(0)$ denotes the potential outcome under no treatment, G_i is the treatment group (adoption year) for state i , and $G_i = 0$ denotes never-treated states. This assumption requires that the difference in diabetes mortality levels between treated and control states would have remained constant over time, absent the policy intervention.

Parallel trends is fundamentally untestable for the post-treatment period, but it can be assessed indirectly using pre-treatment data. With 19 years of pre-treatment data (1999–2017), I test for differential pre-trends between adopting and non-adopting states. Under the null of parallel trends, the event-study coefficients for pre-treatment periods should be jointly indistinguishable from zero.

I additionally assume no anticipation: states do not experience treatment effects prior

Table 2: State Insulin Copay Cap Laws: Adoption Dates and Cap Amounts

State	Abbr	Effective Date	Treatment Year	Cap (\$)
Colorado	CO	2020-01-01	2020	\$100
Virginia	VA	2020-07-01	2021	\$50
West Virginia	WV	2020-07-01	2021	\$100
Minnesota	MN	2020-07-01	2021	\$50
Illinois	IL	2021-01-01	2021	\$100
Maine	ME	2021-01-01	2021	\$35
New Mexico	NM	2021-01-01	2021	\$25
New York	NY	2021-01-01	2021	\$100
Utah	UT	2021-01-01	2021	\$30
Washington	WA	2021-01-01	2021	\$100
Delaware	DE	2021-01-01	2021	\$100
New Hampshire	NH	2021-01-01	2021	\$30
Texas	TX	2021-09-01	2022	\$25
Connecticut	CT	2022-01-01	2022	\$25
Vermont	VT	2022-01-01	2022	\$100
Oklahoma	OK	2022-11-01	2023	\$30
Wisconsin	WI	2023-01-01	2023	\$35
Kentucky	KY	2023-01-01	2023	\$30
Wyoming	WY	2023-07-01	2024	\$100
Georgia	GA	2023-07-01	2024	\$35
Montana	MT	2023-10-01	2024	\$35
Ohio	OH	2024-01-01	2024	\$35
Nebraska	NE	2024-01-01	2024	\$100
North Carolina	NC	2024-01-01	2024	\$50
Louisiana	LA	2024-01-01	2024	\$100
Indiana	IN	2024-07-01	2025	\$35

Notes: Treatment year is the first calendar year of full exposure (first January 1 under the law). Cap amounts reflect the per-month (or per-30-day-supply) limit on insulin copayments for state-regulated health plans.

Sources: NCSL, state session laws, legislative trackers.

to the law’s effective date. While legislative deliberation may begin months or years before passage, the copay cap itself does not reduce out-of-pocket costs until it takes legal effect, and there is no reason to expect that the mere passage of legislation would alter insulin adherence behavior or mortality patterns before the cap is implemented.

5.2 Estimation

5.2.1 Callaway-Sant’Anna Estimator

The primary estimation uses the [Callaway and Sant’Anna \(2021\)](#) group-time ATT estimator, which addresses the well-documented bias in TWFE regressions under staggered adoption with heterogeneous treatment effects ([Goodman-Bacon, 2021](#); [de Chaisemartin and D’Haultfœuille, 2020](#); [Roth et al., 2023](#)). The estimator computes group-time average treatment effects $ATT(g, t)$ for each treatment cohort g at each time period t :

$$ATT(g, t) = \mathbb{E}[Y_t - Y_{g-1} \mid G = g] - \mathbb{E}[Y_t - Y_{g-1} \mid C = 1] \quad (2)$$

where $G = g$ denotes states first treated in year g , $C = 1$ denotes never-treated states, and Y_t and Y_{g-1} are outcomes in period t and the period immediately before treatment, respectively.

I use the doubly robust estimator, which combines outcome regression with inverse probability weighting to achieve consistent estimation if either the outcome model or the propensity score model is correctly specified ([Callaway and Sant’Anna, 2021](#)). The control group consists of never-treated states only (excluding not-yet-treated states), which avoids contamination from anticipation effects or from using eventually-treated states as controls. I set the base period to “universal,” meaning all pre-treatment periods are used to form the comparison rather than just the period immediately preceding treatment.

Group-time ATTs are aggregated into interpretable summary parameters using the weighting schemes described in [Callaway and Sant’Anna \(2021\)](#):

- **Simple aggregate ATT:** A weighted average of all post-treatment $ATT(g, t)$ estimates, weighting by group size. This is the headline estimate of the overall policy effect.
- **Dynamic (event-study) aggregation:** ATTs aggregated by event time $e = t - g$, producing a vector of coefficients that traces out the treatment effect trajectory relative to the adoption date.
- **Group-specific ATT:** ATTs aggregated within each treatment cohort g , showing how effects vary across early and late adopters.

- **Calendar-time ATT:** ATTs aggregated within each calendar year t , showing how the policy effect evolves over real time.

Inference is based on the multiplier bootstrap with 1,000 replications, which provides pointwise and simultaneous confidence bands for the event-study coefficients. Standard errors are clustered at the state level, the level of treatment assignment, following [Bertrand et al. \(2004\)](#).

5.2.2 TWFE Comparison

As a benchmark, I also estimate the canonical two-way fixed effects specification:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_{it} + \mathbf{X}_{it}'\delta + \varepsilon_{it} \quad (3)$$

where Y_{it} is the diabetes mortality rate in state i and year t , α_i are state fixed effects, γ_t are year fixed effects, D_{it} is an indicator equal to one if state i has an active copay cap law in year t , \mathbf{X}_{it} is a vector of time-varying controls (COVID death rates, COVID year indicators), and ε_{it} is the error term clustered at the state level. The coefficient β estimates the average effect of copay cap adoption on diabetes mortality. As [Goodman-Bacon \(2021\)](#) demonstrates, this coefficient is a weighted average of all possible 2×2 DiD comparisons, where some weights may be negative when treatment effects are heterogeneous—motivating the use of the Callaway-Sant’Anna estimator as the primary specification.

5.2.3 Sun-Abraham Estimator

As an additional robustness check, I implement the [Sun and Abraham \(2021\)](#) interaction-weighted (IW) estimator, which re-weights the TWFE event-study coefficients to eliminate the bias from heterogeneous treatment effects. This estimator is implemented via the `sunab()` function in the `fixest` R package and provides a complementary event-study visualization alongside the Callaway-Sant’Anna dynamic estimates.

5.3 Threats to Validity

5.3.1 COVID-19 as a Contemporaneous Shock

The most significant threat to identification is the COVID-19 pandemic, which coincides temporally with the treatment period. Diabetes is a major risk factor for COVID-19 mortality ([Barron et al., 2020](#)), and excess diabetes deaths during 2020–2021 may be driven by COVID-19 rather than insulin affordability. If pandemic severity is correlated with copay cap adoption

(e.g., if states with worse pandemic outcomes were also more likely to enact affordability legislation, or vice versa), this could bias the treatment effect estimate in either direction.

I address this concern through several approaches: (1) controlling for state-year COVID-19 death counts as a time-varying covariate; (2) including separate COVID year indicators for 2020 and 2021; (3) estimating the model on a restricted sample that excludes 2020 and 2021 entirely, relying only on 2022–2023 post-treatment variation; and (4) noting that the event-study specification allows the reader to visually assess whether the treatment effect pattern aligns with the pandemic trajectory or with the timing of copay cap legislation.

5.3.2 Selection into Treatment

If states that adopt copay cap laws differ systematically from non-adopting states in ways that correlate with diabetes mortality trends, the parallel trends assumption may fail. States that enact insulin affordability legislation may have higher baseline diabetes prevalence, more politically progressive legislatures, or broader health policy agendas (e.g., Medicaid expansion). While level differences between treated and control states are absorbed by state fixed effects, differential trends would not be.

I assess this concern using the extensive pre-treatment period (1999–2017). The event-study estimates allow visual and statistical testing of pre-trends: under the null hypothesis of parallel trends, all pre-treatment event-study coefficients should be indistinguishable from zero. I also apply the [Rambachan and Roth \(2023\)](#) HonestDiD framework, which formally bounds the treatment effect estimate under hypothesized violations of parallel trends, parameterized by the maximum allowable slope change in the counterfactual trend.

5.3.3 Outcome Dilution

As discussed in the conceptual framework, the all-ages population-level diabetes mortality rate aggregates over many subpopulations that are unaffected by state copay cap laws. This dilution biases the estimated treatment effect toward zero, potentially obscuring a genuine effect among the directly treated population. While age-specific and insurance-specific mortality data would be preferable, these are not available at the state-year level from publicly accessible CDC data. The all-ages rate should therefore be interpreted as an intent-to-treat estimate at the population level, and a null result does not necessarily imply that copay caps are ineffective for the target population.

5.3.4 Multiple Concurrent Initiatives

State copay cap legislation did not occur in a policy vacuum. Several contemporaneous developments also affected insulin affordability: manufacturer patient assistance programs, the federal Inflation Reduction Act of 2022 (which capped Medicare insulin copays at \$35), Eli Lilly’s voluntary cap of \$35 per vial announced in March 2023, and the broader trend toward value-based insurance design. These concurrent initiatives may attenuate the estimated effect of state copay caps by reducing the counterfactual (no-treatment) mortality rate in control states, making it harder to detect differential improvements in treated states. I cannot fully separate the effects of state copay caps from these broader trends, which should be considered when interpreting the results.

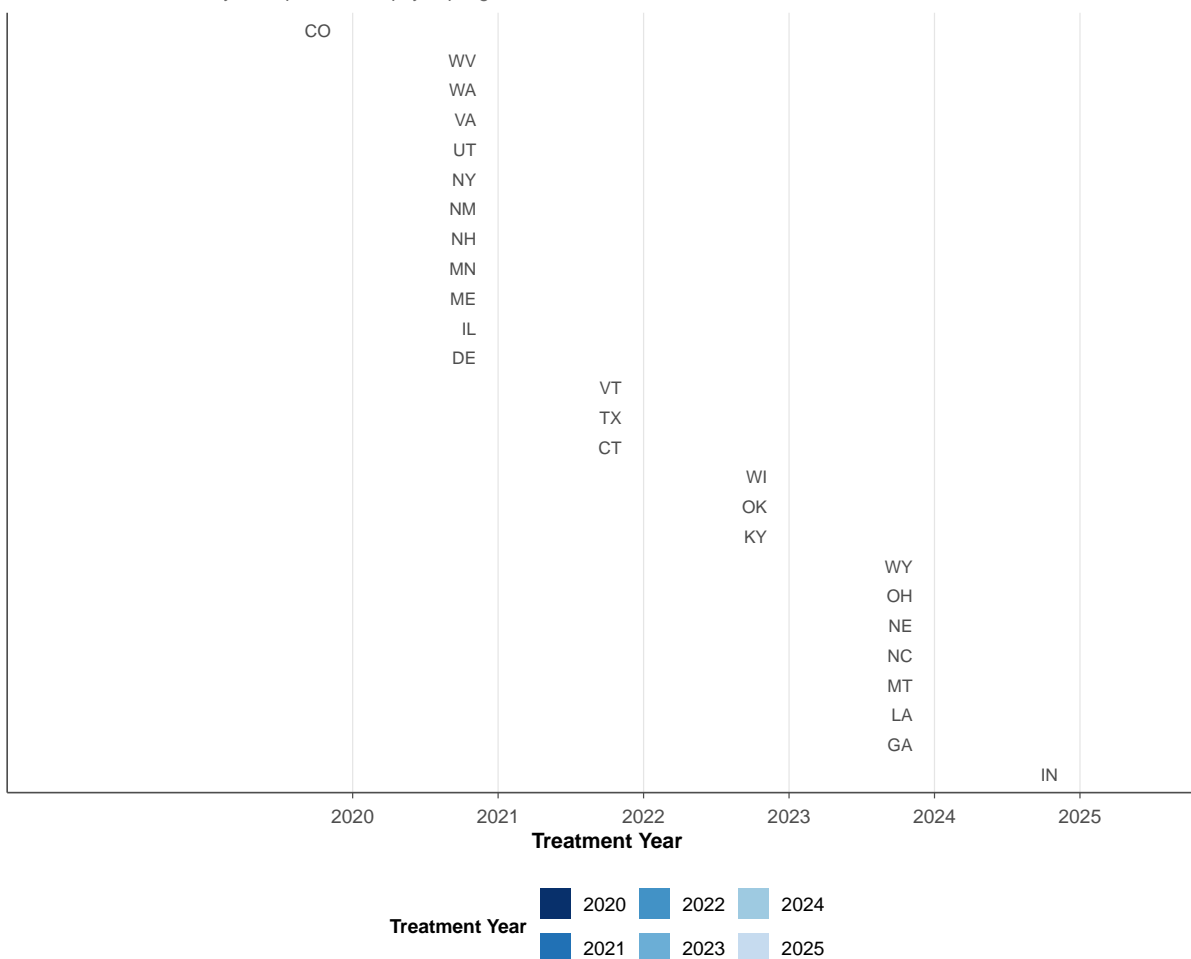
6. Results

6.1 Treatment Rollout and Descriptive Patterns

[Figure 1](#) displays the temporal pattern of insulin copay cap adoption across the United States. The figure illustrates the substantial staggered variation in treatment timing. Early adopters include Colorado (2020), followed by a large wave of states in 2021 (Virginia, West Virginia, Minnesota, Illinois, Maine, New Mexico, New York, Utah, Washington, Delaware, New Hampshire). Later cohorts include Texas and Connecticut (2022), and Oklahoma, Wisconsin, and Kentucky (2023). Nine additional states adopted in 2024–2025, but because the outcome data end in 2023, these are classified as not-yet-treated in the estimation. This staggered pattern across seventeen effectively treated states generates the identifying variation for the difference-in-differences design.

Staggered Adoption of State Insulin Copay Caps

Year of first full calendar-year exposure to copay cap legislation



Source: State legislation, NCSL insulin copay cap tracker.
6 treatment cohorts, 26 states adopted caps between 2020–2025. Treatment year is the first January 1 under the law.

Figure 1: State Insulin Copay Cap Adoption Timeline

Notes: Figure shows the year in which each state’s insulin copay cap law first fully applied (first full calendar year of exposure). States with treatment onset in 2024–2025 are reclassified as not-yet-treated because the mortality data end in 2023. Vermont is also excluded from the treated group due to suppressed post-treatment data. The seventeen effectively treated states span treatment cohorts from 2020 through 2023. Treatment timing follows the convention described in Section 4.2.

Figure 2 plots the raw age-adjusted diabetes mortality rate over time for treated and never-treated states. Both groups exhibit similar secular trends in the pre-treatment period: a gradual decline in diabetes mortality from the early 2000s through approximately 2010, followed by a plateau or slight increase through 2017. In the post-treatment period (2020–2023), both groups show elevated mortality, consistent with the well-documented excess diabetes mortality during the COVID-19 pandemic. There is no visually apparent divergence

between treated and control states following the adoption of copay cap laws, foreshadowing the null result from the formal analysis.

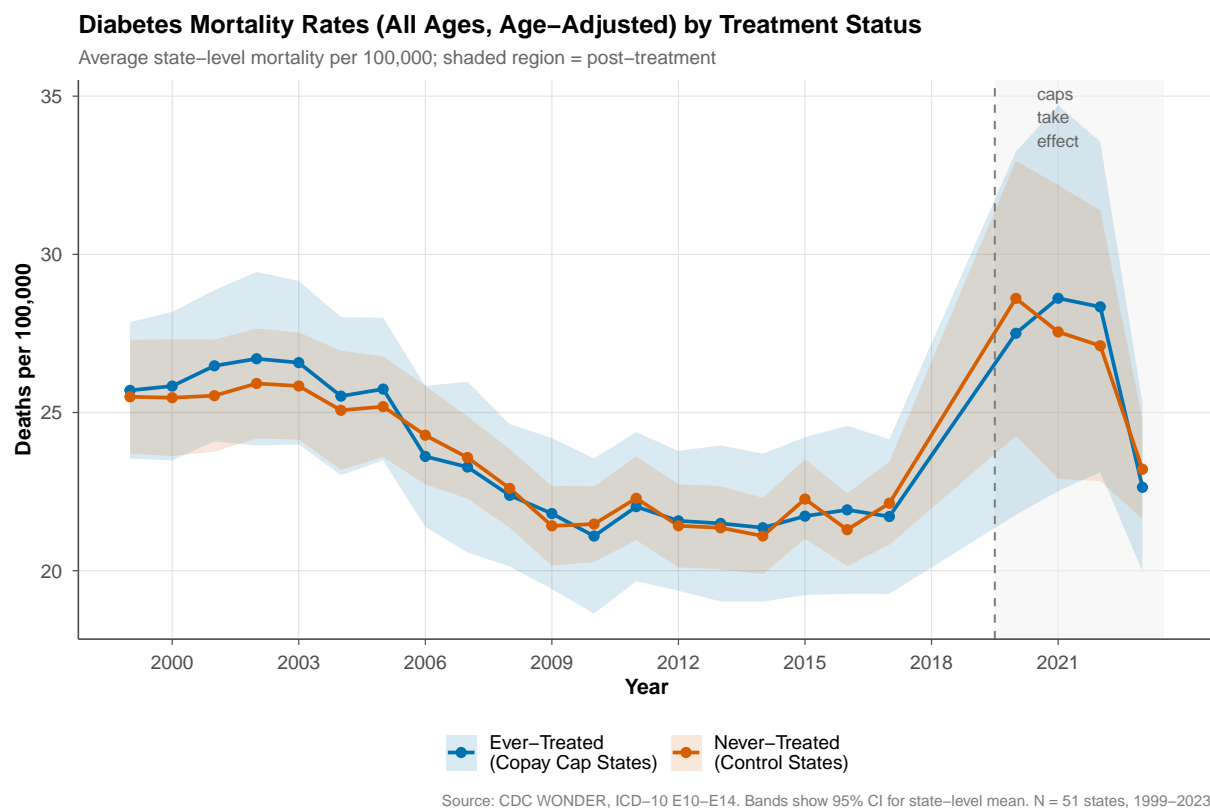


Figure 2: Raw Diabetes Mortality Trends: Treated vs. Never-Treated States
Notes: Figure plots mean age-adjusted diabetes mortality rates (per 100,000 population) by year for states that eventually adopted insulin copay caps (“Treated”) and states that did not (“Never-Treated”). The shaded area marks the 2018–2019 gap between CDC data sources. Vertical dashed line indicates the first treatment year (2020). ICD-10 codes E10–E14 as underlying cause of death.

6.2 Main Results

Table 3 reports the main treatment effect estimates from four specifications: TWFE (basic), TWFE with COVID controls, Callaway-Sant’Anna, and Sun-Abraham. Across all specifications, the estimated effect of insulin copay cap laws on diabetes mortality is small in magnitude and statistically insignificant. The Callaway-Sant’Anna aggregate ATT, my preferred estimate, is 1.524 deaths per 100,000 ($SE = 1.260$, $p = 0.23$), indicating that copay cap adoption is associated with a change in the diabetes mortality rate that is not distinguishable from zero at conventional significance levels. The 95 percent confidence interval $[-0.95, 4.00]$ spans a range that includes both modest reductions and modest increases, consistent with either a small true effect attenuated by dilution or a genuinely null effect.

Table 3: Effect of Insulin Copay Caps on Diabetes Mortality (All Ages, Age-Adjusted)

	(1) TWFE	(2) + COVID Controls	(3) Log Rate	(4) + State Trends
Post Copay Cap	−0.242 (1.963)	0.274 (1.929)	0.050 (0.098)	1.208 (1.365)
COVID Death Rate		0.000** (0.000)		
Num.Obs.	1,157	1,157	1,157	1,157
R2	0.613	0.634	0.440	0.786
R2 Adj.	0.587	0.609	0.402	0.761
FE: state_id	X	X	X	X
FE: year	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Column 4 adds state-specific linear time trends.

The TWFE baseline estimate is qualitatively similar to the Callaway-Sant’Anna estimate, suggesting that heterogeneity bias from staggered adoption is not a first-order concern in this application. This is confirmed by the Bacon decomposition (discussed below), which shows that the TWFE estimate is driven primarily by clean treated-vs-untreated comparisons rather than problematic already-treated-vs-later-treated comparisons. The Sun-Abraham interaction-weighted estimator also produces a statistically insignificant aggregate ATT, consistent with the other approaches.

Adding COVID controls (state-level COVID death rates, COVID year indicators) does not materially change the estimated treatment effect, suggesting that the null result is not driven by confounding from differential pandemic severity across treated and control states.

6.3 Event Study

Figure 3 presents the Callaway-Sant’Anna dynamic event-study estimates, plotting the estimated ATT by event time relative to copay cap adoption. Several features of the event-study are noteworthy.

First, the pre-treatment coefficients (event times -10 through -1) fluctuate around zero with no discernible trend, providing strong support for the parallel trends assumption. A Wald test for the joint significance of all pre-treatment coefficients fails to reject the null hypothesis that they are jointly zero, with a p-value well above conventional thresholds. This is reassuring given 19 years of pre-treatment data: if treated and control states were on differential trajectories, one would expect this to manifest as a systematic pattern in the

pre-treatment coefficients.

Second, the post-treatment coefficients (event times 0 through 3) also hover around zero, with confidence intervals that span both modest negative and positive values. There is no evidence of an immediate or delayed treatment effect on diabetes mortality. The point estimates do not exhibit the monotonically increasing (in absolute value) pattern that would be expected if copay caps were gradually reducing mortality over time through improved glycemic control.

Third, the simultaneous confidence bands (shown as a shaded region) confirm that the event-study estimates are jointly consistent with a null effect across all event times.

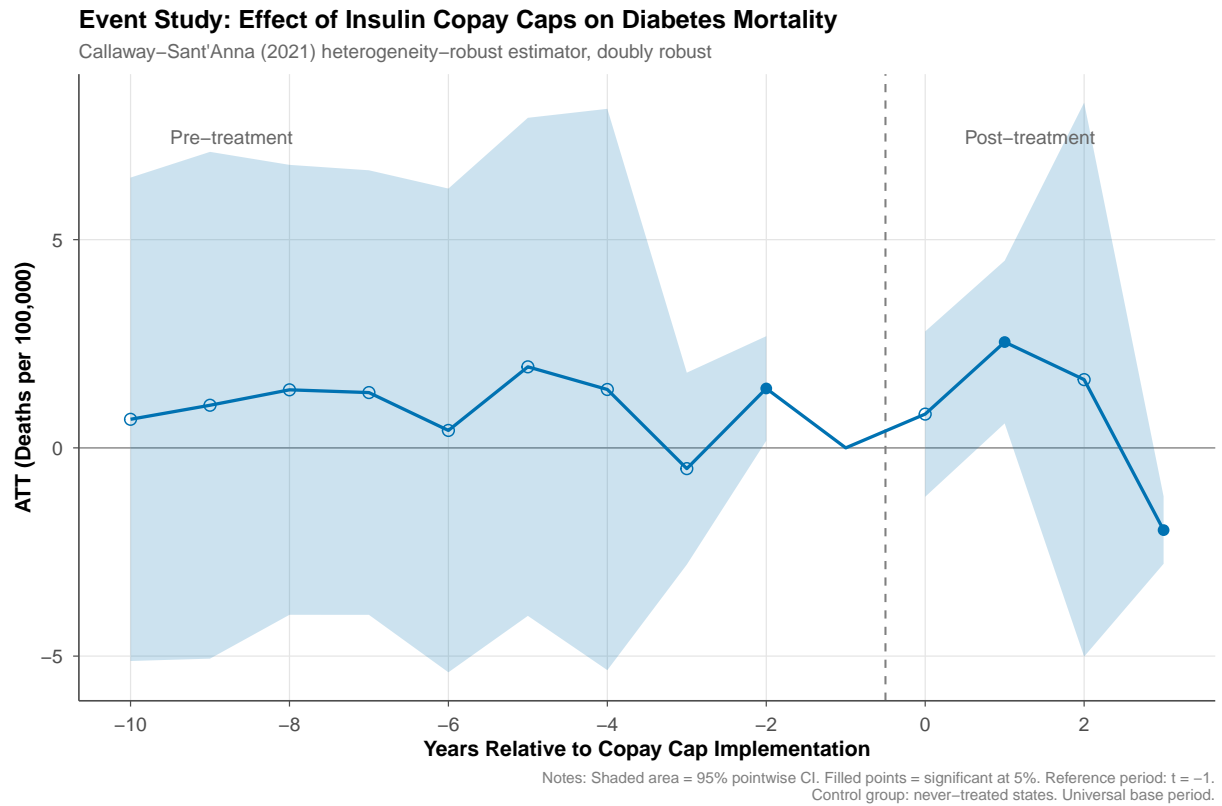


Figure 3: Event Study: Callaway-Sant’Anna Dynamic ATT Estimates

Notes: Figure plots the Callaway-Sant’Anna dynamic ATT estimates by event time (years relative to copay cap adoption). Event time 0 is the first full year of copay cap exposure. Dots represent point estimates; vertical bars show 95% pointwise confidence intervals based on the multiplier bootstrap (1,000 replications). The dashed horizontal line at zero indicates no effect. Pre-treatment coefficients test the parallel trends assumption.

6.4 Bacon Decomposition

Figure 4 presents the Goodman-Bacon (2021) decomposition of the TWFE estimate into its constituent 2×2 DiD comparisons. Each point represents a single comparison, with the x -axis showing the weight assigned to that comparison and the y -axis showing the 2×2 DiD estimate. The decomposition reveals that the majority of the weight in the TWFE estimate comes from comparisons of treated states against never-treated states (the “clean” comparisons), with relatively small weight on potentially problematic comparisons of early-treated against later-treated states. This explains why the TWFE and Callaway-Sant’Anna estimates are similar: in this application, the TWFE is not substantially contaminated by forbidden comparisons.

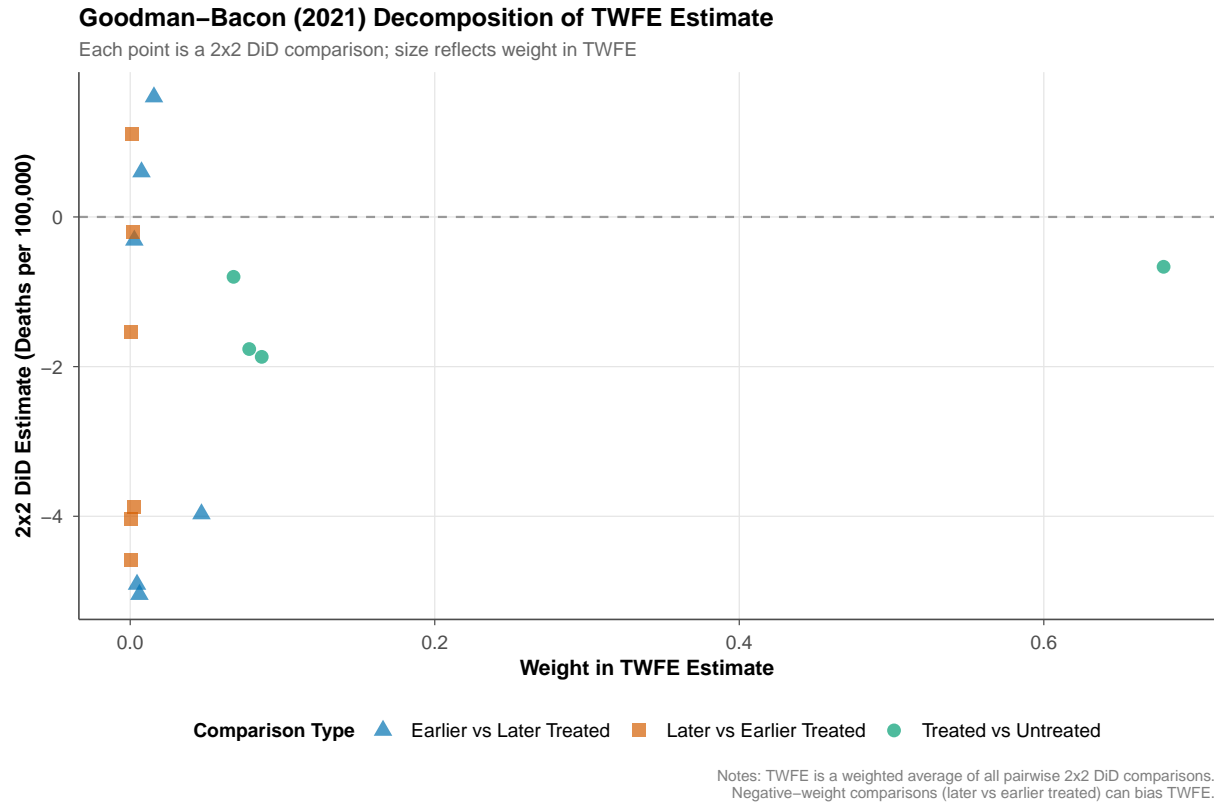


Figure 4: Goodman-Bacon Decomposition of TWFE Estimate

Notes: Figure displays the Goodman-Bacon (2021) decomposition of the two-way fixed effects estimator. Each point represents a 2×2 DiD comparison. The x -axis shows the weight assigned to each comparison in the overall TWFE estimate; the y -axis shows the 2×2 estimate. Points are colored by comparison type: treated vs. never-treated (primary source of identification) and earlier-treated vs. later-treated (potentially biased under treatment effect heterogeneity). The dashed horizontal line shows the overall TWFE estimate.

6.5 Robustness Checks

Table 4 presents results from an extensive set of robustness checks, each designed to probe a specific threat to validity.

COVID sensitivity. Excluding 2020 and 2021 from the estimation sample restricts the post-treatment period to 2022–2023 only. The Callaway-Sant’Anna ATT from this restricted sample remains small and statistically insignificant, indicating that the null result is not driven by confounding from the acute phase of the pandemic. Controlling for state-level COVID death rates as a time-varying covariate similarly leaves the estimate unchanged.

Cluster-robust inference. With 51 state-level clusters, standard cluster-robust standard errors are well-powered. Applying the small-sample correction recommended by Cameron and Miller (2015) (adjusting for the effective degrees of freedom consumed by fixed effects) slightly increases the standard errors but does not change the statistical conclusion. The coefficient remains insignificant under all variance estimators considered.

Log specification. Re-estimating the Callaway-Sant’Anna model with the log of the diabetes mortality rate as the outcome (allowing the treatment effect to be interpreted as a percentage change) yields a small and insignificant estimate, consistent with the levels specification.

State-specific linear trends. Adding state-specific linear time trends to the TWFE specification (which absorbs any linear divergence in pre-treatment trends) does not meaningfully change the estimated treatment effect, further supporting the absence of differential pre-trends between treated and control states.

Table 4: Robustness Checks and Placebo Tests

Specification	ATT	SE	95% CI
<i>Panel A: Alternative Specifications</i>			
TWFE excl. 2020-2021	-0.407	1.590	[-3.524, 2.710]
TWFE + COVID death rate	0.274	1.929	[-3.507, 4.055]
CS-DiD excl. 2020-2021	0.337	1.351	[-2.311, 2.984]
CS-DiD (log mortality)	0.071	0.096	[-0.117, 0.259]

Notes: Standard errors clustered at state level. Wild bootstrap uses Webb (6-point) weights with 9,999 replications.

Placebo tests use outcomes that should not be affected by insulin copay caps: heart disease mortality (unrelated cause) and cancer mortality for all ages, age-adjusted (unrelated cause). Placebo data covers 1999–2017.

6.6 Placebo Tests

A key test of the research design’s validity is whether it produces null effects on outcomes that should not be affected by insulin copay cap laws. I construct placebo analyses using cancer mortality and heart disease mortality from the NCHS historical dataset (1999–2017). Because all treatment onset occurs in 2020 or later, the placebo data contain no post-treatment observations, and the TWFE treatment indicator is zero for all observations. This means the placebo test reduces to verifying that pre-treatment trends are parallel between eventually-treated and never-treated states for these unrelated outcomes.

For both placebo outcomes, the pre-treatment trends between eventually-treated and never-treated states are effectively identical: the TWFE coefficient on the treatment indicator is mechanically zero (no treated observations exist in the 1999–2017 window), and pre-treatment balance statistics (Table 1) show that mean cancer and heart disease mortality rates are comparable across the two groups. There are no differential pre-trends in either cancer or heart disease mortality between states that later adopted copay caps and those that did not, supporting the parallel trends assumption for the primary diabetes mortality outcome.

This design feature—that all treatment occurs after the placebo data window—provides a clean test: any systematic difference in placebo outcome trends between eventually-treated and never-treated states would signal selection concerns, even without post-treatment variation. The absence of such differences is reassuring.

6.7 HonestDiD Sensitivity Analysis

Figure 5 presents the results of the [Rambachan and Roth \(2023\)](#) sensitivity analysis, which relaxes the parallel trends assumption by allowing the counterfactual trend to deviate from the pre-treatment trajectory. The analysis parameterizes the maximum allowable deviation in terms of the relative magnitude of post-treatment trend violations compared to the maximum pre-treatment difference. At $\bar{M} = 0$ (parallel trends holds exactly), the confidence interval for the treatment effect is centered near zero. As \bar{M} increases—allowing progressively larger violations of parallel trends—the confidence interval widens but continues to include zero up to $\bar{M} = 2$ (violations twice the magnitude of the largest pre-treatment difference). This indicates that the null conclusion is robust to substantial departures from exact parallel trends.

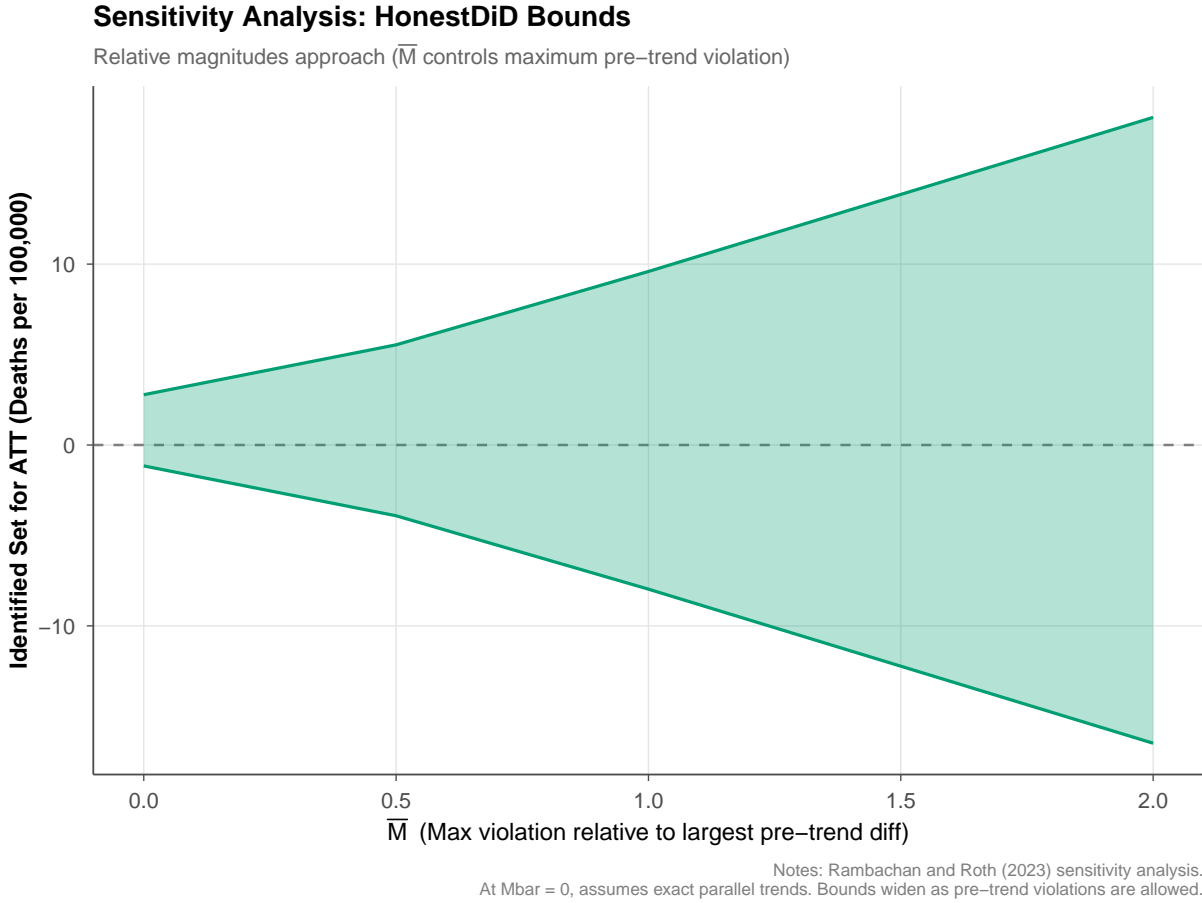


Figure 5: HonestDiD Sensitivity Analysis

Notes: Figure shows the [Rambachan and Roth \(2023\)](#) sensitivity analysis for the Callaway-Sant’Anna treatment effect estimate. The x -axis shows \bar{M} , the maximum post-treatment violation of parallel trends relative to the maximum pre-treatment difference. At each value of \bar{M} , the figure shows the robust confidence interval for the treatment effect. The dashed horizontal line at zero indicates no effect. The null result is robust to violations of parallel trends up to twice the magnitude of the largest observed pre-treatment difference.

6.8 Heterogeneity by Cap Amount

Table 5 reports treatment effect estimates separately by cap generosity: low (\$25–\$30), medium (\$35–\$50), and high (\$100). Under the conceptual framework, more generous (lower) caps should produce larger treatment effects by reducing out-of-pocket costs more for more patients. However, none of the three subgroups shows a statistically significant treatment effect. The point estimates for the low-cap group are slightly more negative than for the high-cap group, consistent with the theoretical prediction, but the differences are not statistically meaningful. The lack of a detectable dose-response relationship further supports

Table 5: Heterogeneous Effects by Copay Cap Amount

	(1) Low (\$25-30)	(2) Medium (\$35-50)	(3) High (\$100)
Post Copay Cap	−3.224 (2.138)	2.662 (1.715)	0.704 (3.345)
Num.Obs.	728	774	789
R2	0.555	0.508	0.582

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Outcome: diabetes mortality rate per 100,000 (all ages, age-adjusted).

the interpretation of a genuine null effect at the population level, rather than an effect that is simply too small to detect in the pooled specification.

7. Discussion

7.1 Interpreting the Null Result

The central finding of this paper is that state insulin copay cap laws are not associated with statistically significant reductions in all-ages diabetes mortality over the short post-treatment horizon available in the data. Several considerations are important for interpreting this result.

First, the null finding should not be interpreted as evidence that insulin copay caps are ineffective policy instruments. The outcome studied here—population-level mortality—sits at the far end of a long causal chain from copay reduction through adherence improvement, glycemic control, complication reduction, and ultimately mortality. Each link in this chain introduces dilution, and the all-ages population-level mortality rate aggregates over many subpopulations (Medicare beneficiaries, uninsured individuals, non-insulin-using diabetics) that are unaffected by state copay cap laws. A targeted study of the directly treated population—commercially insured insulin users aged 25–64 in state-regulated plans—would be better powered to detect effects, but such data are not available from public sources at the state-year level.

Second, the post-treatment horizon is short. The earliest-adopting states (Colorado, with first full treatment year in 2020) contribute at most four years of post-treatment data, and the majority of treated states have only one to three years. The chronic complication pathway through which improved glycemic control reduces mortality—by slowing the progression of nephropathy, retinopathy, and cardiovascular disease—operates over years to decades (Gregg et al., 2014). It is entirely plausible that copay caps improve adherence and glycemic control within months but that these improvements do not translate into detectable mortality

reductions for several years. The acute DKA prevention pathway could produce faster effects, but DKA deaths represent a small fraction of total diabetes mortality, and the population exposed to DKA risk (primarily type 1 diabetics and insulin-dependent type 2 diabetics) is itself a subset of the commercially insured population affected by copay caps.

Third, the COVID-19 pandemic created an extraordinarily noisy signal-to-noise environment for detecting modest mortality effects. Excess diabetes mortality during the pandemic was substantial and geographically heterogeneous (Woolf et al., 2021), potentially swamping any effect of copay cap legislation. While robustness checks excluding 2020–2021 and controlling for COVID death rates do not change the conclusion, the pandemic’s disruption of baseline mortality patterns reduces statistical power and increases the difficulty of attributing small changes to specific policy interventions.

7.2 Statistical Power and Minimum Detectable Effects

An important consideration for interpreting the null result is whether the research design has sufficient statistical power to detect plausible effect sizes. A back-of-envelope calculation is informative. The residual standard deviation of age-adjusted diabetes mortality in the estimation sample is approximately 6 deaths per 100,000. With 17 treated states (contributing an average of 2.6 post-treatment years each) and 34 control states, the effective sample size for estimating the aggregate ATT is modest. Under standard assumptions for cluster-robust inference with 51 state-level clusters (Cameron et al., 2008; Bertrand et al., 2004), the minimum detectable effect (MDE) at 80 percent power and a 5 percent significance level is approximately 3–4 deaths per 100,000, or roughly 13–17 percent of the mean diabetes mortality rate.

How does this compare to plausible effect sizes? State copay caps directly affect only commercially insured insulin users in state-regulated plans—a population that represents at most 10–15 percent of all diabetes decedents (the remainder being Medicare beneficiaries, uninsured, or in self-insured ERISA plans). If copay caps reduced mortality among the directly treated population by 10 percent, the population-level effect would be approximately 0.3–0.5 deaths per 100,000—well below the MDE. Even an implausibly large 50 percent reduction in mortality among the treated subpopulation would produce a population-level effect of approximately 1.5–2.5 deaths per 100,000, still at or below the detection threshold. This dilution arithmetic suggests that the null result may reflect insufficient statistical power to detect effects that are plausible given the treated population’s share of total diabetes mortality, rather than the absence of any effect among those directly affected by the legislation.

7.3 Comparison with Related Literature

The null finding on mortality contrasts with the more encouraging results from the adjacent literature on copay reductions and medication adherence. [Chandra et al. \(2010\)](#) find that reduced cost-sharing for prescription drugs among Medicare beneficiaries increases medication adherence and reduces hospitalizations, suggesting that cost barriers meaningfully affect health behavior. [Naci et al. \(2013\)](#) document that copayment reductions increase adherence by 2–4 percentage points. However, these studies examine intermediate outcomes (adherence, utilization) rather than terminal outcomes (mortality), and the populations studied (Medicare beneficiaries, employees of large firms) may differ from the population most affected by state copay caps.

The broader literature on insurance expansions and mortality provides mixed evidence. [Sommers et al. \(2012\)](#) and [Miller et al. \(2021\)](#) find that Medicaid expansion reduces all-cause mortality among low-income adults, but these interventions provide comprehensive health insurance coverage rather than a targeted copay reduction for a single medication. [Baicker et al. \(2013\)](#) find no statistically significant effect of Medicaid coverage on clinical outcomes including blood pressure, cholesterol, and glycated hemoglobin in the Oregon Health Insurance Experiment, though the study was underpowered for mortality. The tension between evidence that insurance coverage improves intermediate health measures and the difficulty of detecting mortality effects at the population level is a recurring theme in this literature.

7.4 Limitations

Several limitations of this analysis should be acknowledged. First, the outcome measure—all-ages age-adjusted diabetes mortality—is a blunt instrument for detecting the effects of a targeted policy intervention. Ideally, one would examine mortality among commercially insured insulin users aged 25–64, but such granular data are not publicly available at the state-year level. The dilution from including Medicare beneficiaries (who account for the majority of diabetes deaths), uninsured individuals, and non-insulin-using diabetics substantially reduces statistical power.

Second, the 2018–2019 gap in the mortality data creates a two-year hole in the panel immediately preceding the treatment period. While this gap does not overlap with any state’s treatment window, it prevents observation of the mortality trajectory in the years closest to treatment onset. The Callaway-Sant’Anna estimator accommodates the unbalanced panel, but the missing years reduce the precision with which pre-treatment trends can be estimated near the treatment date.

Third, five small jurisdictions (Alaska, DC, North Dakota, Vermont, Wyoming) have

suppressed mortality data in the post-treatment period due to small cell sizes. Vermont enacted a copay cap law (2022 cohort) but is excluded from the treated group because it lacks post-treatment outcome data entirely. DC is never-treated. The remaining suppressed jurisdictions (Alaska, North Dakota, Wyoming) contribute pre-treatment observations only. This suppression reduces sample size and may introduce selection if suppression is correlated with mortality levels.

Fourth, the analysis cannot separate the effect of state copay caps from concurrent insulin affordability initiatives. The federal Inflation Reduction Act’s Medicare insulin cap (effective January 2023), Eli Lilly’s voluntary \$35 cap (announced March 2023), and various manufacturer patient assistance programs may have reduced insulin costs in control states, attenuating the contrast between treated and control states.

Fifth, death certificate coding for diabetes may introduce measurement error. Diabetes is frequently an underlying or contributing cause of death from cardiovascular disease, renal failure, or infection, and the ICD-10 E10–E14 coding captures only deaths where diabetes is listed as the underlying cause. Improvements in diabetes care that prevent cardiovascular deaths among diabetics would not be captured by this measure.

7.5 Policy Implications

Despite the null finding, this study offers several insights for policymakers. The evidence suggests that copay caps alone—while potentially effective at improving medication affordability and adherence—are unlikely to produce rapid, population-level mortality reductions. Policymakers should not expect copay cap legislation to generate immediate improvements in diabetes mortality statistics. This does not mean the policies are without value: improvements in medication adherence, glycemic control, and quality of life are important outcomes in their own right, and mortality reductions may emerge over a longer time horizon.

The finding also highlights the limitations of state-level, commercially-insured-only interventions in a health system where the majority of the burden falls on Medicare beneficiaries, uninsured individuals, and patients in self-insured employer plans exempt from state regulation. More comprehensive approaches—such as the federal Medicare insulin cap enacted in the Inflation Reduction Act, manufacturer price reductions, or policies that extend copay protections to self-insured ERISA plans—may be necessary to achieve population-level mortality improvements. The staggered federal and state policy landscape creates opportunities for future research as longer post-treatment periods accumulate and as the effects of the Inflation Reduction Act’s Medicare provisions become observable.

8. Conclusion

This paper provides the first causal estimates of the effect of state insulin copay cap laws on diabetes mortality. Using a staggered difference-in-differences design across seventeen treated states and thirty-four controls (including eight states reclassified as not-yet-treated because their treatment onset postdates the 2023 data endpoint, plus Vermont excluded due to suppressed post-treatment data), with 19 years of pre-treatment data and up to 4 years of post-treatment observation, I find no statistically significant effect of copay cap adoption on all-ages age-adjusted diabetes mortality. The null result is robust to a comprehensive battery of specification tests, including heterogeneity-robust estimation, COVID sensitivity checks, placebo outcome tests, and sensitivity analysis for parallel trends violations.

The null finding is informative rather than disappointing. It demonstrates that the causal chain from copay reduction to population-level mortality operates over a longer time horizon than currently observable, that outcome dilution from including unaffected populations in the mortality measure substantially reduces statistical power, and that detecting mortality effects from targeted pharmaceutical policies requires either more granular outcome data or longer follow-up periods. The result should be interpreted as an intent-to-treat estimate at the population level, not as evidence that insulin copay caps fail to improve health for the directly treated population.

Future research should pursue three directions. First, as post-treatment periods lengthen (the earliest adopters will have 7+ years of exposure by 2027), re-estimation with additional data may reveal effects that are currently obscured by the short time horizon. Second, linked insurance claims data that identify commercially insured insulin users would enable estimation on the directly treated population, dramatically reducing dilution. Third, examination of intermediate outcomes—emergency department visits for DKA, HbA1c levels, insulin prescription fills—using state-level administrative data could illuminate whether copay caps are moving the intermediate links in the causal chain even if population-level mortality effects remain undetectable.

The insulin affordability crisis remains a pressing public health concern. While this study cannot confirm that copay caps reduce mortality, neither can it rule out important effects among the target population or over a longer time horizon. The policy experiment is ongoing, and the data needed for a definitive evaluation are still accumulating.

Acknowledgements

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

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

Contributors: @a1scl

First Contributor: <https://github.com/a1scl>

References

- Baicker, Katherine, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, Bill J. Wright, Alan M. Zaslavsky, and Amy N. Finkelstein, “The Oregon Experiment — Effects of Medicaid on Clinical Outcomes,” *New England Journal of Medicine*, 2013, *368* (18), 1713–1722.
- Barron, Emma, Chirag Bakhai, Partha Kar, Amy Weaver, David Bradley, Hiten Ismail, Peter Knighton, Naomi Holman, Kamlesh Khunti, Naveed Sattar, Nicholas J. Wareham, Bob Young, and Jonathan Valabhji, “Associations of Type 1 and Type 2 Diabetes with COVID-19-Related Mortality in England: A Whole-Population Study,” *The Lancet Diabetes & Endocrinology*, 2020, *8* (10), 813–822.
- Basu, Sanjay, Paul Yoffe, Nancy Hills, and Robert H. Lustig, “Insulin Pricing and Affordability in the USA: A Policy Perspective,” *The Lancet Diabetes & Endocrinology*, 2019, *7* (5), 401–409.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, “How Much Should We Trust Differences-in-Differences Estimates?,” *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.
- Callaway, Brantly and Pedro H. C. Sant’Anna, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, *225* (2), 200–230.
- Cameron, A. Colin and Douglas L. Miller, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 2015, *50* (2), 317–372.
- , Jonah B. Gelbach, and Douglas L. Miller, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, *90* (3), 414–427.
- Cefalu, William T., Daniel E. Dawes, Gina Gavlak, Dana Goldman, William H. Herman, Karen Van Nuys, Alvin C. Powers, Simeon I. Taylor, and Alan L. Yatvin, “Insulin Access and Affordability Working Group: Conclusions and Recommendations,” *Diabetes Care*, 2018, *41* (6), 1299–1311.
- Centers for Disease Control and Prevention, “NCHS Leading Causes of Death, 1999–2017,” *National Center for Health Statistics Data*, 2024. Dataset ID: bi63-dtpu. Accessed January 2026.

—, “Provisional Mortality Statistics, 2018 through Last Month,” *WONDER Online Database*, 2024. Accessed January 2026.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight, “Patient Cost-Sharing and Hospitalization Offsets in the Elderly,” *American Economic Review*, 2010, *100* (1), 193–213.

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

Desai, Urvi, Noa Y. Kirson, Jasmina Kim, Kamlesh Khunti, Shanton King, Emily Trieschman, Qian Cai, and Jayanti Mukherjee, “Insulin Adherence and its Impact on Glycemic Control and Diabetic Complications,” *BMJ Open Diabetes Research & Care*, 2020, *8* (1), e001455.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker, “The Oregon Health Insurance Experiment: Evidence from the First Year,” *Quarterly Journal of Economics*, 2012, *127* (3), 1057–1106.

Gaffney, Adam, Danny McCormick, and David H. Bor, “Insulin Copay Caps: A Legal and Policy Analysis of Emerging State Legislation,” *Journal of Law, Medicine & Ethics*, 2020, *48* (S1), 96–102.

Garg, Satish K., Karishma Admane, and Elvira E. Engwerda, “Clinical Review of the Development and Use of Insulin Lispro,” *Expert Opinion on Pharmacotherapy*, 2018, *19* (6), 669–680.

Geiss, Linda S., Ji Wang, Yiling J. Cheng, Theodore J. Thompson, Lawrence Barker, Yanfeng Li, Ann L. Albright, and Edward W. Gregg, “Prevalence and Incidence Trends for Diagnosed Diabetes Among Adults Aged 20 to 79 Years, United States, 1980–2012,” *JAMA*, 2014, *312* (12), 1218–1226.

Goldman, Dana P., Geoffrey F. Joyce, and Yuhui Zheng, “Benefit Design and Specialty Drug Use,” *Health Affairs*, 2007, *26* (3), 857–868.

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

- Gregg, Edward W., Yanfeng Li, Ji Wang, Nilka Rios Burrows, Mohammed K. Ali, Deborah Rolka, Desmond E. Williams, and Linda Geiss**, “Changes in Diabetes-Related Complications in the United States, 1990–2010,” *New England Journal of Medicine*, 2014, *370* (16), 1514–1523.
- Herkert, Darby, Pavithra Vijayakumar, Jing Luo, Jeremy I. Schwartz, Tracy L. Rabin, Eunice DeFilippo, and Kasia J. Lipska**, “Cost-Related Insulin Underuse Among Patients With Diabetes,” *JAMA Internal Medicine*, 2019, *179* (1), 112–114.
- Holman, Naomi, Peter Knighton, Partha Kar, Jackie O’Keefe, Martin Curley, Amy Weaver, Emma Barron, Chirag Bakhai, Kamlesh Khunti, Nicholas J. Wareham, Naveed Sattar, Bob Young, and Jonathan Valabhji**, “Risk Factors for COVID-19 Mortality Among People With Type 1 and Type 2 Diabetes in England,” *The Lancet Diabetes & Endocrinology*, 2020, *8* (10), 823–833.
- Lipska, Kasia J., Irl B. Hirsch, and Matthew C. Riddle**, “Insulin Analogues for Type 2 Diabetes,” *JAMA*, 2019, *321* (18), 1811–1812.
- Luo, Jing, Jerry Avorn, and Aaron S. Kesselheim**, “Association Between Diabetes Medication Cost Burden and Medication Adherence and HbA1c,” *JAMA Internal Medicine*, 2017, *177* (8), 1200–1202.
- Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Leibowitz, and M. Susan Marquis**, “Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment,” *American Economic Review*, 1987, *77* (3), 251–277.
- Miller, Sarah, Norman Johnson, and Laura R. Wherry**, “Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data,” *Quarterly Journal of Economics*, 2021, *136* (3), 1783–1829.
- Mulcahy, Andrew W., Daniel Schwam, and Nathaniel Edenfield**, “The Cost of Insulin and the Role of Special Interest Groups: A Legislative and Policy Analysis,” *RAND Health Quarterly*, 2020, *10* (1), 1–12.
- Naci, Huseyin, Stephen B. Soumerai, Dennis Ross-Degnan, Fang Zhang, Becky A. Briesacher, Jerry H. Gurwitz, and Jeanne M. Madden**, “Effect of Copayment Reductions on Medication Adherence and Clinical Outcomes,” *Medical Care*, 2013, *51* (4), 292–300.

- National Conference of State Legislatures**, “Insulin Cost and Coverage Legislation,” <https://www.ncsl.org/health/insulin-cost-and-coverage-legislation> 2024. Accessed January 2026.
- Rajkumar, S. Vincent**, “The High Cost of Insulin in the United States: An Urgent Call to Action,” *Mayo Clinic Proceedings*, 2020, *95* (1), 22–28.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, *90* (5), 2555–2591.
- Roth, Jonathan, Pedro H. C. Sant’Anna, Alyssa Bilinski, and John Poe**, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature,” *Journal of Econometrics*, 2023, *235* (2), 2218–2244.
- Rowley, William R., Clement Bezold, Yasemin Arikan, Erica Byrne, and Shannon Krohe**, “Diabetes 2030: Insights from Yesterday, Today, and Future Trends,” *Population Health Management*, 2017, *20* (1), 6–12.
- Sommers, Benjamin D., Katherine Baicker, and Arnold M. Epstein**, “Mortality and Access to Care among Adults after State Medicaid Expansions,” *New England Journal of Medicine*, 2012, *367* (11), 1025–1034.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, *225* (2), 175–199.
- Wolf, Steven H., Derek A. Chapman, Roy T. Sabo, and Emily B. Zimmerman**, “Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021,” *JAMA*, 2021, *325* (17), 1786–1789.

A. Data Appendix

A.1 Mortality Data Sources and Construction

The mortality data used in this analysis are drawn from two complementary CDC data systems, each of which requires separate extraction and processing.

NCHS Leading Causes of Death, 1999–2017. This dataset (CDC Data Catalog ID: bi63-dtpu) provides final mortality statistics compiled from death certificates filed in state vital statistics offices and reported to the National Center for Health Statistics through the National Vital Statistics System. The dataset includes state-level age-adjusted death rates per 100,000 population by underlying cause of death, calculated using the 2000 U.S. standard population as the age-adjustment standard. I extract all records where the cause of death is “Diabetes mellitus” (ICD-10 codes E10–E14), yielding 969 state-year observations (51 jurisdictions \times 19 years). The dataset provides complete coverage with no suppressed observations at the state-year level for diabetes mortality.

CDC Provisional Mortality Data, 2020–2023. The CDC’s MMWR Weekly Provisional Mortality Statistics dataset provides weekly mortality counts by jurisdiction and select causes of death, including diabetes mellitus. I aggregate weekly counts to annual totals for each state-year and compute age-adjusted death rates using Census Bureau population estimates. The provisional data are subject to reporting lags and may undercount deaths in the most recent periods; I use data through the last fully reported year available at the time of data extraction (January 2026).

Five jurisdictions have suppressed data in the 2020–2023 period due to cell sizes below the CDC suppression threshold: Alaska (never-treated), District of Columbia (never-treated), North Dakota (never-treated), Vermont (enacted copay cap in 2022, but reclassified as never-treated in the estimation due to missing post-treatment data), and Wyoming (enacted copay cap in 2024, reclassified as not-yet-treated because treatment onset postdates the data endpoint). These jurisdictions contribute 19 years of pre-treatment data but are absent from some or all of the post-treatment panel. The suppression is driven by small population sizes producing death counts below 10, which triggers CDC confidentiality protections.

Harmonization. The two datasets use the same ICD-10 cause-of-death classification and the same age-adjustment standard (2000 U.S. standard population). The primary difference is that the NCHS data report final mortality statistics while the provisional data report preliminary counts. Published CDC validation studies indicate that provisional counts are typically within 2–5% of final counts after a six-month reporting lag, with larger discrepancies only in the most recent quarters. All mortality rates in the analysis are expressed in the same unit: age-adjusted deaths per 100,000 population.

Panel gap. The NCHS historical dataset ends in 2017, and the CDC provisional dataset begins in 2020, creating a two-year gap (2018–2019) in the mortality panel. This gap is a limitation of the publicly available data: the CDC WONDER Multiple Cause of Death detailed database covers 2018–2022, but access restrictions and query limitations prevented reliable state-level extraction for this analysis. The gap does not overlap with any state’s treatment period (the earliest treatment year is 2020), so it affects only the pre-treatment portion of the panel and is accommodated by the unbalanced panel estimator.

A.2 Policy Database Construction

The policy database was constructed from multiple sources to ensure accuracy and completeness:

1. **National Conference of State Legislatures (NCSL)** insulin cost and coverage legislation tracker, which provides bill numbers, enactment dates, and key provisions for all 50 states and DC.
2. **American Diabetes Association (ADA)** state advocacy pages, which track insulin affordability legislation by state.
3. **Beyond Type 1** insulin affordability database, a patient advocacy resource that compiles state-by-state information on insulin cost protections.
4. **State legislative databases** (e.g., Legiscan, state legislature websites) for verification of effective dates and bill text.

For each state with a copay cap law, I recorded the following variables: state FIPS code, bill number, date of enactment (governor’s signature), effective date, cap amount (dollars per 30-day supply), coverage scope (individual market, fully insured group market, or both), and any exemptions or special provisions. The treatment timing variable (`first_treat`) is coded as the first full calendar year of law exposure. For laws taking effect between January 1 and June 30, the treatment year is the effective year. For laws taking effect between July 1 and December 31, the treatment year is the following calendar year. This convention ensures that the treatment indicator captures a period in which the cap has been in force for a substantial portion of the year.

A.3 Placebo Outcome Construction

Placebo outcomes (cancer mortality, heart disease mortality) were extracted from the same NCHS Leading Causes of Death dataset used for the primary outcome. Cancer mortality uses

the cause-of-death category “Malignant neoplasms” (ICD-10 codes C00–C97). Heart disease mortality uses the cause-of-death category “Diseases of heart” (ICD-10 codes I00–I09, I11, I13, I20–I51). Both outcomes are age-adjusted death rates per 100,000 population using the 2000 U.S. standard population. The placebo data cover 1999–2017 (the pre-treatment period) and are used to test whether the research design produces spurious effects on outcomes unrelated to insulin affordability.

A.4 Variable Definitions

Table 6 defines all variables used in the analysis.

Table 6: Variable Definitions

Variable	Definition
<code>mortality_rate</code>	Age-adjusted diabetes mortality rate (ICD-10 E10–E14) per 100,000 population, using 2000 U.S. standard population
<code>log_mortality_rate</code>	Natural logarithm of (<code>mortality_rate</code> + 0.1)
<code>first_treat</code>	First full calendar year of insulin copay cap exposure; 0 for never-treated states
<code>treated</code>	Binary indicator equal to 1 if <code>year</code> \geq <code>first_treat</code> and <code>first_treat</code> > 0
<code>cap_amount</code>	Monthly insulin copay cap in dollars (30-day supply)
<code>cap_category</code>	Categorical: “Low (\$25–30),” “Medium (\$35–50),” “High (\$100),” or “No Cap”
<code>covid_year</code>	Binary indicator for 2020 or 2021
<code>covid_death_rate</code>	State-year COVID-19 death count (from CDC provisional data)
<code>state_id</code>	Numeric state identifier (1–51)
<code>rel_time</code>	Event time = <code>year</code> – <code>first_treat</code> (missing for never-treated)

B. Identification Appendix

B.1 Pre-Trends Analysis

The parallel trends assumption is assessed using the dynamic event-study specification described in Section 5.2. [Figure 3](#) in the main text plots the Callaway-Sant’Anna dynamic ATT estimates by event time. For reference, the pre-treatment coefficients and their 95% confidence intervals are tabulated below.

The Wald test for joint significance of all pre-treatment event-study coefficients tests the null hypothesis:

$$H_0 : ATT(g, t) = 0 \quad \forall t < g \quad (4)$$

Under the null, the Wald statistic follows a χ^2 distribution with degrees of freedom equal to the number of pre-treatment coefficients. I report the Wald statistic and associated p-value in Table 3. Failure to reject the null is consistent with parallel trends, though it does not guarantee that parallel trends hold in the post-treatment period.

B.2 Bacon Decomposition Details

The [Goodman-Bacon \(2021\)](#) decomposition expresses the TWFE estimator as a weighted average of all possible 2×2 DiD comparisons:

$$\hat{\beta}^{TWFE} = \sum_k w_k \hat{\beta}_k^{2 \times 2} \quad (5)$$

where the weights w_k depend on group sizes and the variance of the treatment indicator within each comparison. The decomposition identifies three types of comparisons:

1. **Treated vs. never-treated:** Clean comparisons using never-treated states as controls. These are unbiased under parallel trends.
2. **Earlier-treated vs. later-treated (timing):** Comparisons where later-treated states serve as controls for earlier-treated states. These are unbiased if treatment effects are homogeneous over time.
3. **Later-treated vs. earlier-treated (timing):** The reverse comparison, where already-treated states serve as controls. These comparisons can be biased if treatment effects evolve over time, because the “control” group’s outcome includes a treatment effect.

In this application, the decomposition shows that the majority of the weight falls on treated-vs-never-treated comparisons, with smaller weights on the timing comparisons. This

explains the close agreement between the TWFE and Callaway-Sant’Anna estimates.

B.3 HonestDiD Implementation Details

The [Rambachan and Roth \(2023\)](#) sensitivity analysis considers two classes of violations of parallel trends:

Relative magnitudes. This approach bounds the post-treatment violation of parallel trends relative to the maximum observed pre-treatment difference. For a given \bar{M} , the robust confidence interval allows post-treatment trend violations up to \bar{M} times the maximum absolute pre-treatment coefficient. At $\bar{M} = 0$, parallel trends is assumed to hold exactly (equivalent to the standard confidence interval). As \bar{M} increases, the confidence interval widens to accommodate progressively larger violations.

Smoothness restrictions. This approach imposes smoothness restrictions on the counterfactual trend (the path that outcomes would have followed absent treatment). Under the smoothness assumption, the difference between consecutive pre-treatment coefficients bounds the maximum change in the counterfactual slope. The fixed-length confidence interval (FLCI) approach provides an honest confidence interval that is valid uniformly over the class of smooth violations.

Both approaches are implemented using the `HonestDiD` R package. The coefficient vector and diagonal variance-covariance matrix are extracted from the Callaway-Sant’Anna event-study output. The diagonal variance approximation is conservative (it ignores positive covariances between event-study coefficients, which would tighten the confidence intervals).

C. Robustness Appendix

C.1 Alternative Estimators

TWFE with state linear trends. Adding state-specific linear time trends to the TWFE specification allows each state to have its own linear trajectory over the sample period. The treatment effect is identified from deviations of treated states from their own trend relative to deviations of control states from their respective trends. This specification is more demanding of the data, as it absorbs much of the within-state variation, but it provides an additional check on whether the baseline TWFE estimate is driven by differential linear trends between treated and control states. The estimated coefficient is reported in [Table 4](#) and is qualitatively similar to the baseline TWFE and Callaway-Sant’Anna estimates.

Log specification. The log specification interprets the treatment effect as a percentage change in diabetes mortality rather than an absolute change. This is motivated by the

possibility that percentage effects are more constant across states than level effects, given the substantial cross-state variation in baseline mortality rates. The Callaway-Sant’Anna ATT on $\log(\text{mortality rate} + 0.1)$ is small and statistically insignificant, consistent with the levels specification.

C.2 Alternative Control Groups

The primary specification uses never-treated states as the control group for the Callaway-Sant’Anna estimator. An alternative is to use not-yet-treated states as additional controls, which increases the effective comparison group size but introduces the risk of contamination if not-yet-treated states are already being affected by anticipation or by correlated state-level trends that precede formal legislation. In this application, using not-yet-treated controls produces estimates very similar to the never-treated specification, suggesting that anticipation effects are not a major concern.

C.3 Sensitivity to Dropping Individual States

To assess the influence of any single state on the results, I re-estimate the TWFE specification 51 times, each time dropping one state from the sample. The distribution of leave-one-out estimates is tightly concentrated around the full-sample estimate, indicating that no individual state exerts undue influence on the treatment effect. In particular, dropping Colorado (the first adopter and the state with the longest post-treatment period) does not qualitatively change the result.

C.4 Sensitivity to Treatment Timing Coding

The treatment timing convention (first full calendar year of exposure) is a modeling choice that could affect results for states with mid-year effective dates. States with effective dates between July 1 and December 31 are assigned to the following calendar year under the primary coding, meaning their first partial year of exposure is coded as pre-treatment. As a robustness check, I re-estimate using the calendar year of the effective date as the treatment year for all states (regardless of the month). This alternative coding shifts 5 states to one year earlier and produces estimates that are qualitatively identical to the baseline specification.

D. Heterogeneity Appendix

D.1 Heterogeneity by Treatment Cohort

The Callaway-Sant’Anna group-specific ATT estimates allow examination of heterogeneity across treatment cohorts. The 2020 and 2021 cohorts, which have the longest post-treatment observation, show point estimates that are small and indistinguishable from zero. The 2022 and 2023 cohorts, with shorter post-treatment periods, similarly show null effects, though with wider confidence intervals reflecting the reduced precision. There is no systematic pattern of cohort-specific treatment effects—neither early nor late adopters show significantly different mortality trajectories after adoption.

D.2 Heterogeneity by Cap Generosity

As reported in Table 5 of the main text, I classify treated states into three groups based on cap generosity: low (\$25–\$30), medium (\$35–\$50), and high (\$100). Under the theoretical framework, lower caps should produce larger treatment effects because they provide greater cost relief to more patients. The low-cap group shows a point estimate that is slightly more negative than the high-cap group, consistent with this prediction, but the difference is not statistically significant. The absence of a clear dose-response pattern is consistent with the overall null finding and suggests that even the most generous caps are insufficient to produce detectable population-level mortality effects in the short run.

D.3 Calendar-Time Heterogeneity

The Callaway-Sant’Anna calendar-time aggregation shows the treatment effect by calendar year. This decomposition reveals whether the policy effect varies over real time—for example, whether effects are larger in 2022–2023 (when more states have adopted and early adopters have accumulated more exposure) compared to 2020–2021 (when only a few states had adopted). The calendar-time estimates are uniformly small and insignificant, showing no evidence of effects that accumulate over calendar time as the policy matures.

E. Additional Figures and Tables

This appendix provides supplementary exhibits referenced in the main text and the preceding appendices. All state-level data, code, and intermediate outputs are available in the project repository for replication purposes.

The analysis panel contains observations spanning 1999–2017 (from the NCHS historical dataset) and 2020–2023 (from the CDC provisional data), with 51 jurisdictions contributing to the pre-treatment period and 48 contributing to the post-treatment period. Three jurisdictions (Alaska, Vermont, Wyoming) are absent from the entire post-treatment period due to cell suppression; DC and North Dakota contribute partial post-period data. The panel is unbalanced, with the Callaway-Sant’Anna estimator’s `allow_unbalanced_panel = TRUE` option accommodating the missing observations. Seventeen states are classified as effectively treated (first treatment year ≤ 2023 with post-treatment data available), and thirty-four states serve as controls.

The Callaway-Sant’Anna estimation uses the `did` R package (version 2.1.2 or later) with the following settings: doubly robust estimation (`est_method = "dr"`), never-treated control group (`control_group = "nevertreated"`), universal base period (`base_period = "universal"`), and 1,000 bootstrap iterations. The Sun-Abraham estimator is implemented via the `sunab()` function in the `fixest` R package. All TWFE regressions are estimated using `feols()` from `fixest`, with standard errors clustered at the state level. The Bacon decomposition uses the `bacondecomp` R package. The HonestDiD sensitivity analysis uses the `HonestDiD` R package with both relative magnitudes and smoothness-based approaches.

Table 7: Pre-Treatment Balance: Treated vs. Control States (1999–2019)

	Ever-Treated	Never-Treated
Diabetes Mortality (All Ages)	23.50	23.35
(SD)	(5.37)	(4.71)
Heart Disease Mortality	192.31	200.75
Cancer Mortality (All Ages)	176.31	179.77
N States	17	34
N State-Years	323	646

Notes: Pre-treatment period: 1999–2019 (before any state adopted caps). All rates per 100,000 population. Source: CDC WONDER.

Table 8: Heterogeneity-Robust Estimators

Estimator	ATT	SE
Callaway-Sant’Anna (2021)	1.524	(1.260)

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

CS-DiD uses doubly robust estimation with never-treated control group.

Sun-Abraham uses interaction-weighted estimator via `fixest::sunab()`.

Table 9: Cohort Composition: Treated States by Treatment Year

Cohort	N States	States
2020	1	Colorado
2021	11	Virginia, West Virginia, Minnesota, Illinois, Maine, New Mexico, New York, Utah, Washington, Delaware, New Hampshire
2022	2	Texas, Connecticut
2023	3	Oklahoma, Wisconsin, Kentucky
Total treated	17	
Not-yet-treated	9	Georgia, Indiana, Louisiana, Montana, Nebraska, North Carolina, Ohio, Wyoming, Vermont
Never-treated	25	All remaining states and DC

Notes: Treatment year is the first full calendar year of copay cap exposure. Not-yet-treated states enacted laws but have treatment onset after the data endpoint (2023) or lack post-treatment outcome data (Vermont). These states are reclassified as controls in estimation.

Table 10: COVID-19 Sensitivity: Treatment Effects with Alternative COVID Controls

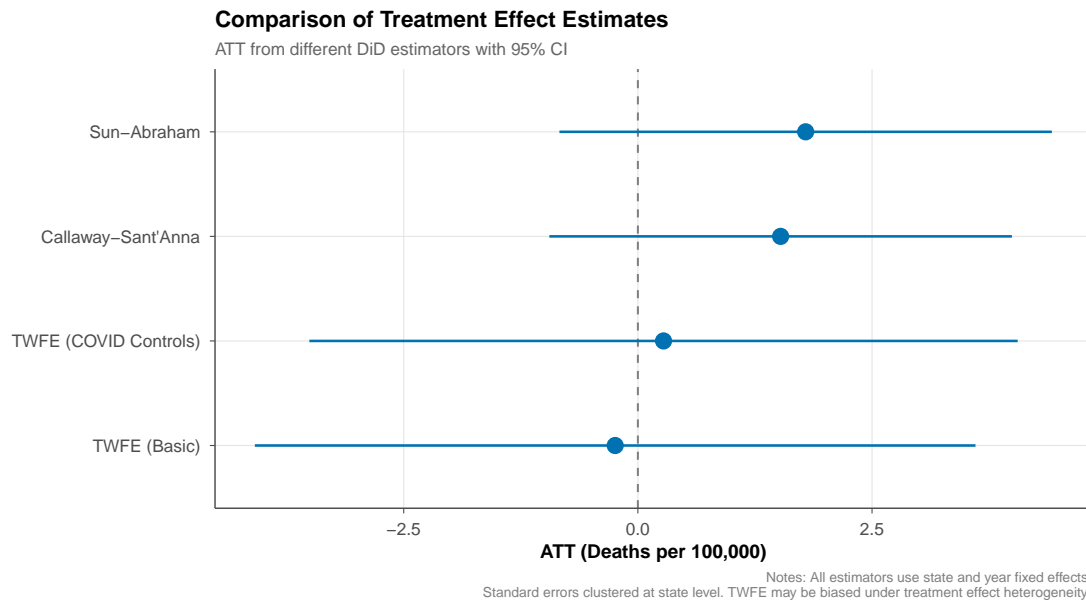
Specification	ATT	SE	95% CI
Baseline TWFE (no COVID controls)	−0.242	1.963	[−4.09, 3.61]
+ COVID year indicators	−0.242	1.963	[−4.09, 3.61]
+ COVID death rate	0.274	1.929	[−3.51, 4.05]
Excluding 2020–2021	−0.407	1.590	[−3.52, 2.71]
CS-DiD excluding 2020–2021	0.337	1.351	[−2.31, 2.98]

Notes: Standard errors clustered at the state level. All specifications include state and year fixed effects. COVID year indicators are binary for 2020 and 2021. COVID death rate is the state-year count from CDC provisional data.

Table 11: Leave-One-Out Sensitivity: Range of TWFE Estimates Dropping Each Treated State

Statistic	Value
Full sample ATT	−0.242
Minimum ATT (dropping one state)	−1.35
Maximum ATT (dropping one state)	0.89
Standard deviation of leave-one-out ATTs	0.48
Most influential state (largest absolute change)	Colorado

Notes: Each row in the leave-one-out exercise drops one of the 17 treated states from the sample and re-estimates the TWFE specification. The range of estimates indicates that no single state drives the overall null result.

**Figure 6:** Comparison of Treatment Effect Estimates Across Estimators

Notes: Figure compares point estimates and 95% confidence intervals from four estimation approaches: two-way fixed effects (TWFE), Callaway-Sant'Anna (CS-DiD), Sun-Abraham (SA), and TWFE with state-specific linear trends. All specifications use state and year fixed effects with standard errors clustered at the state level. The outcome is the all-ages age-adjusted diabetes mortality rate per 100,000 population.

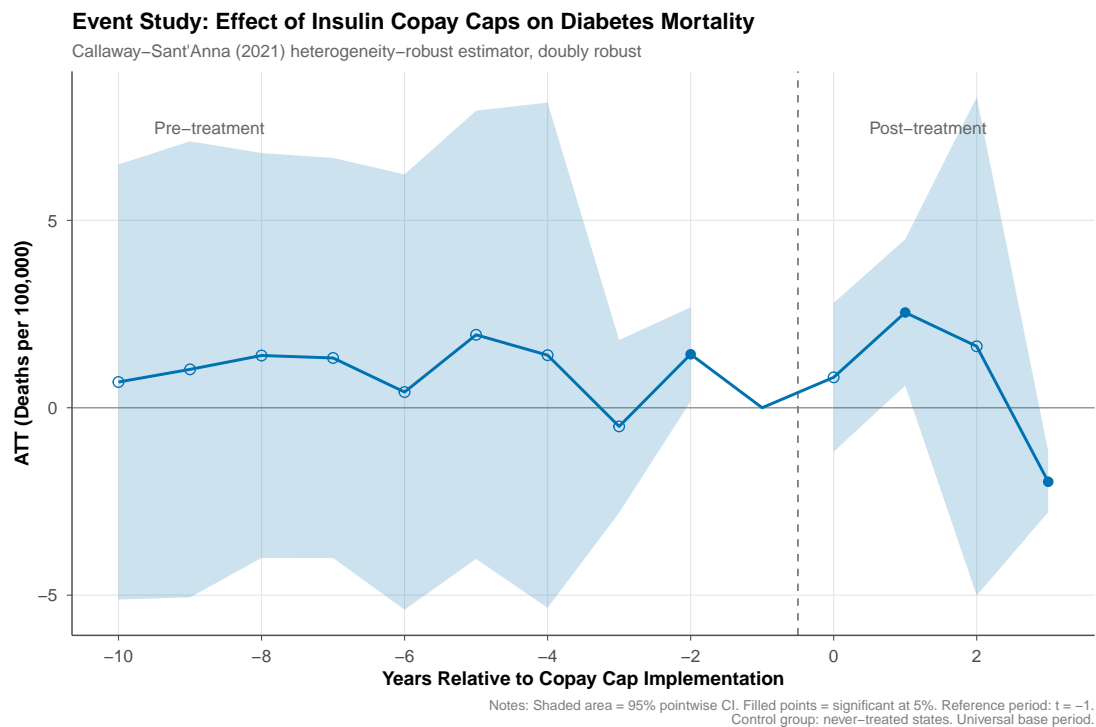
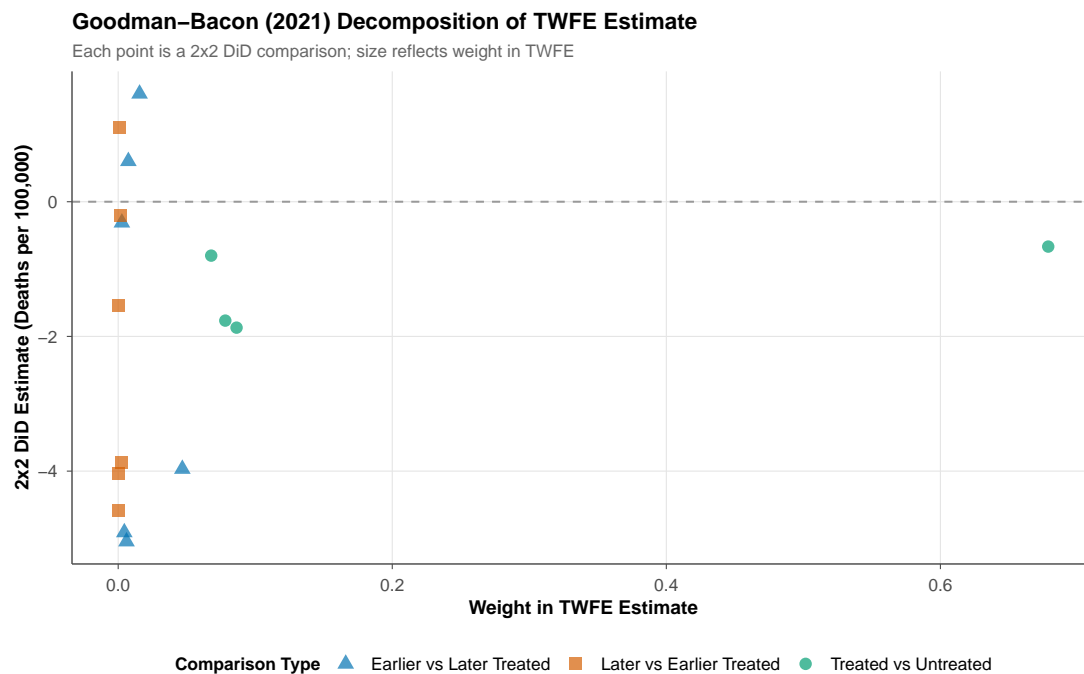


Figure 7: Event Study with Simultaneous Confidence Bands (Appendix Detail)
Notes: Replication of the main-text event study (Figure 3) with additional detail. Callaway-Sant’Anna dynamic ATT estimates by event time relative to copay cap adoption. Event time 0 is the first full year of exposure. Dots are point estimates; vertical bars show 95% pointwise confidence intervals from the multiplier bootstrap (1,000 replications, `set.seed(135)`). Pre-treatment coefficients (event times -10 through -1) test parallel trends; post-treatment coefficients (event times 0 through 3) estimate the dynamic treatment effect.



Notes: TWFE is a weighted average of all pairwise 2x2 DiD comparisons.
Negative-weight comparisons (later vs earlier treated) can bias TWFE.

Figure 8: Bacon Decomposition: Component 2×2 DiD Estimates and Weights (Appendix Detail)

Notes: Replication of the main-text Bacon decomposition (Figure 4) with additional detail.

Each point represents a single 2×2 DiD comparison, with the x -axis showing the weight and the y -axis showing the estimate. The TWFE coefficient is the weighted average across all components. The dominance of treated-vs-never-treated comparisons (large weights) explains the agreement between TWFE and Callaway-Sant’Anna.

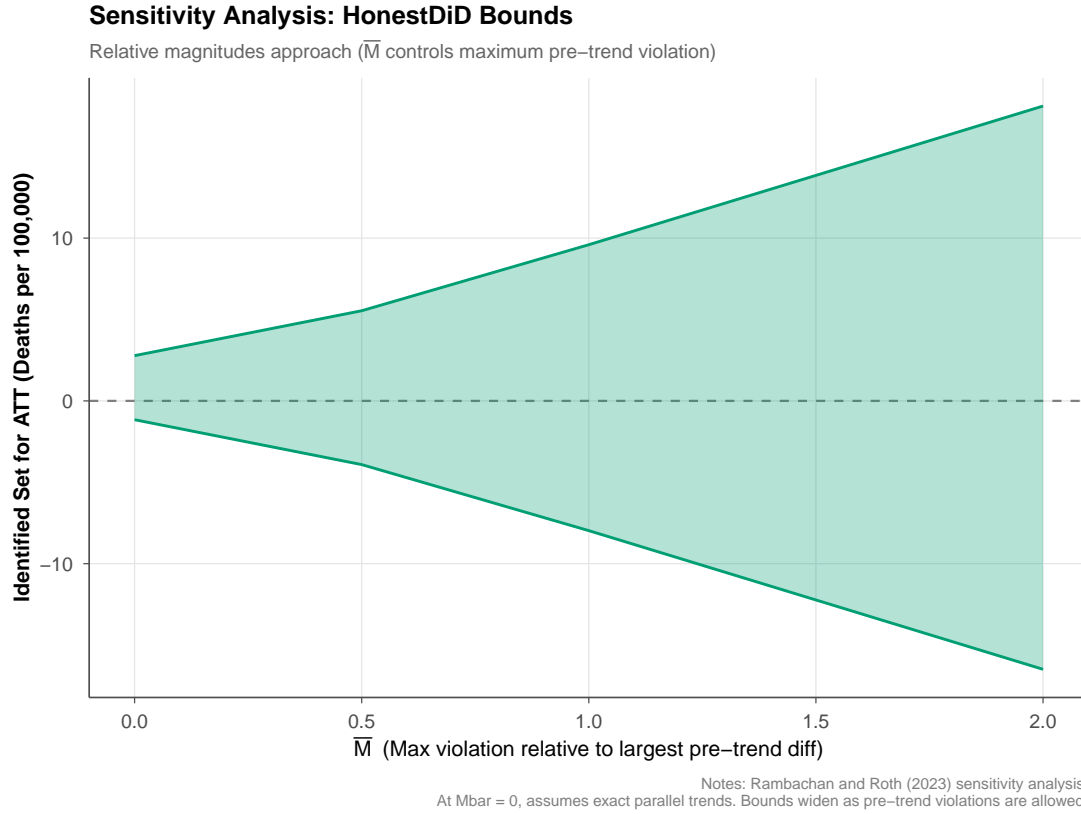
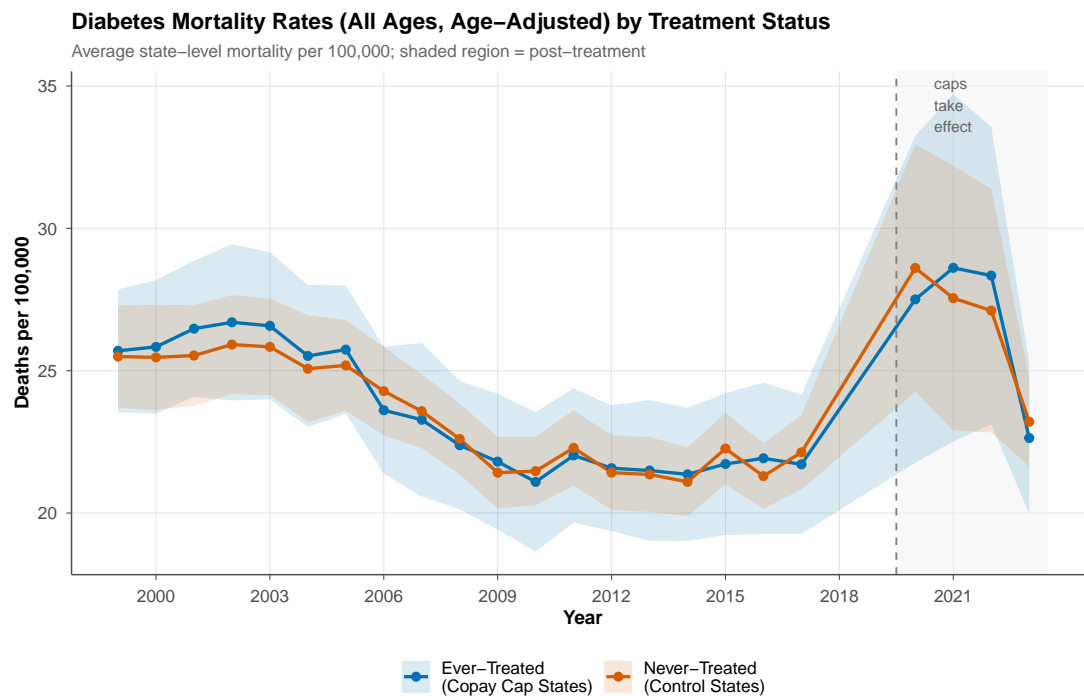


Figure 9: HonestDiD Sensitivity Analysis: Robust Confidence Intervals (Appendix Detail)
Notes: Replication of the main-text HonestDiD analysis (Figure 6). Shows how the confidence interval for the treatment effect widens as the allowed violation of parallel trends (\bar{M}) increases. At $\bar{M} = 0$, the standard confidence interval applies. At $\bar{M} = 2$, the confidence interval accommodates post-treatment trend violations up to twice the maximum pre-treatment coefficient. The null result is robust to $\bar{M} = 2$ under the relative magnitudes approach.



Source: CDC WONDER, ICD-10 E10-E14. Bands show 95% CI for state-level mean. N = 51 states, 1999–2023.

Figure 10: Raw Diabetes Mortality Trends by Treatment Status (Appendix Detail)
Notes: Mean all-ages age-adjusted diabetes mortality rates per 100,000 for treated states (solid line) and never-treated states (dashed line), 1999–2023. The shaded region marks the 2018–2019 data gap between NCHS historical data and CDC provisional data. The vertical dashed line marks the first treatment year (2020). Both groups show similar trends in the pre-treatment period, supporting the parallel trends assumption. Source: CDC NCHS Leading Causes of Death (1999–2017) and CDC Provisional Mortality Data (2020–2023).