Differences-in-Differences and Event Study Methodological Innovations: An Application to Health Economics*

Cameron Deal†

September 21, 2022

Abstract

The Affordable Care Act (ACA) was intended to expand Medicaid eligibility significantly at the federal level. However, the U.S. Supreme Court ruled in 2012 that the federal government could not compel states to expand their eligibility. To date, 39 states (including D.C.) have expanded Medicaid eligibility, with considerable variability in eligibility guidelines between states. A large literature has examined the effects of these expansions on health insurance, economic status, and health outcomes. The bulk of this research has used the difference-in-differences (DD) research design to estimate the causal effects of Medicaid expansions on a host of outcomes. However, recent methodological innovations have shown that these estimation strategies may be biased in the presence of differential treatment timing. I re-estimate the effects of Medicaid expansion on health insurance coverage among a low-education sub-sample and compare the results from naive DD and event study specifications to the newer estimation strategies. I find that though there is substantial potential for bias, Medicaid expansion did decrease uninsurance rates by 4 percentage points. Other results obtained using this research design may require re-examination to verify that they hold true when estimated without bias.

Keywords: Medicaid Expansion, Health Insurance, Difference-in-Differences, Treatment Timing

JEL Codes: I13, C21, C23

*I thank Emilio Pantoja, Guido Anthony Romero, Myles Owens, Jemmy Marc, Neil Ericsson, Justin Pierce, Sebastian Tello-Trillo, Jevay Grooms, Santiago Deambrosi, Samuel Mann, Laura Nettuno, John Graves, Marcus Dillender and seminar participants in the AEA Summer Program and the AEA Pipeline Conference for their helpful feedback. I also thank Samuel Mann and Christopher Carpenter for their mentorship and guidance on difference-in-differences. All errors are my own.

†Vanderbilt University. cameron.h.deal@vanderbilt.edu

1 Introduction

The Affordable Care Act (ACA), signed into law in 2010, was the largest expansion of health insurance coverage to nonelderly adults in half a century. Although the ACA introduced a host of reforms, one of the central components was expanding Medicaid coverage to nearly all adults with incomes up to 138% of the Federal Poverty Level (FPL). This federal eligibility increase, which was scheduled to take place in January 2014, was projected to reduce uninsurance by 32 million people in conjunction with the other components of the ACA (Gruber, 2011). However, the United States Supreme Court ruled in *National Federation of Independent Business v. Sebelius* that the Medicaid expansion provisions was unconstitutionally coercive as written (Roberts, 2012). As a result, there is considerable heterogeneity in the adoption of Medicaid eligibility expansions, with 39 states (including DC) gradually adopting the expansions to date. At the time of writing, there are also proposals to expand under consideration in multiple holdouts (KFF, 2022). Thus, over the past decade, there have been eligibility expansions for public health insurance in the United States on a massive scale, though they have been spread unequally across states.

This setting offers an ideal opportunity to measure the effects of these expansions on a variety of outcomes, including insurance coverage, health outcomes, and economic status. These expansions occurred in a wide variety of states, increasing the generalizability of any findings on their effects in comparison with other approaches that focused on a certain state or smaller geographical unit. Additionally, their setting resembles the health policy environment of the broader United States, taking place in a landscape of direct purchase and employer-sponsored insurance. Beyond the empirical appeal, the question of how health insurance affects health and financial well-being is foundational in health economics, stretching back to early theoretical work on health demand and moral hazard (Grossman, 1972; Arrow, 1963; Pauly, 1968), as well as landmark experimental evidence in health policy (Manning et al., 1987). Overall, state-level Medicaid expansions offer a fertile setting to test a number of key empirical and theoretical questions in health economics, and rigorous studies of their effects may also be valuable for policymakers in ongoing debates.

Health policy researchers have risen to the task, evaluating the effect of Medicaid expansions on insurance coverage, crowd-out, cancer testing, employment levels, spillover effects, and dozens of other outcomes. Foundational in this literature is the first stage of expansion effects: increases in insurance coverage, generally considered the primary channel through which Medicaid expansions affect individual-level outcomes. In general, analyses have found that Medicaid expansions caused significant increases in Medicaid coverage and significant decreases in uninsurance rates (Courtemanche, Marton, Ukert, et al., 2016; Miller & Wherry, 2017; Kaestner et al., 2017). However, the vast majority of this literature relies on difference-in-differences (DD) research designs, often with differential treatment timing (given the staggered roll-out of the program).

These designs can exhibit bias in the presence of differential treatment timing, a problem that a growing econometric literature has begun to address (Goodman-Bacon, 2021). When conducting a DD in the presence of differential treatment timing, the aggregate effect estimate can be decomposed into three broad groups of 2x2 settings: Not Treated vs. Treated, Early Expander vs. Treated, or Late Expander vs. Treated. The first type is relatively harmless, and represents the causal effect of interest. However, the latter two are comparisons that do not compose the aggregate causal effect of interest without strong assumptions on homogeneous treatment effects that are constant over time. Because these assumptions are generally unrealistic, applying a naive DD specification to a scenario with differential treatment timing may yield bias in the estimates of the causal effect. Thus, the effects of state-level expansion, which occurred in a staggered roll-out (see Figure 1), may be biased due to these effects.

A flourishing literature has begun to propose new estimators that tackle this problem using a variety of methods (Gardner, 2021; Sun & Abraham, 2021; Callaway & Sant'Anna, 2021; Borusyak et al., 2022; de Chaisemartin & D'Haultfœuille, 2020; Cengiz et al., 2019; Wooldridge, 2021). I first replicate the results of earlier work on the effects of the ACA state-level expansions using the standard DD and event study specifications. Using the American Community Survey (ACS) I find, in broad accordance with preexisting literature, that state-level Medicaid expansions reduced uninsurance rates among a low-education sub-sample that was likely to be targeted by the eligibility expansions. I find that 30% of detected effects are attributed to faulty 2x2 DiD comparisons. Then, I apply the new estimators from the recent econometric literature. I find that state-level Medicaid expansions did, in fact, reduce uninsurance rates by 3.5—4.5 percentage points. My results on the effects of Medicaid expansion on health insurance (a widely studied topic in health economics) in previous literature are similar, statistically and substantively, across approaches, suggesting that realized bias from differential timing effects is small. Despite this, estimates of downstream effects of Medicaid expansions are also in need of replication, and policymakers may need to revise their expectations of the effects of Medicaid expansions. Additionally, my findings highlight the necessity of applying econometric innovations in difference-in-differences to empirical results to verify their accuracy and robustness. Finally, I offer an example of a widely studied topic in applied microeconomics that new econometric estimators are appropriate for.

The remainder of the paper takes the following form. Section 2 reviews previous literature on the effects of Medicaid expansion and the DD research design. Section 3 describes my data, and Section 4 illustrates my empirical approach. Section 5 presents the results, and Section 6 concludes.

2 Background

2.1 Medicaid Expansion

There is a rich literature examining the effects of subsidized health insurance provision on health and economic outcomes. Some of the earliest work used an experimental setting to find that modest cost sharing reduces use of services with negligible effects on health for the average person (Manning et al., 1987). More recently, the Oregon Health Insurance Experiment found substantive reductions in disastrous financial consequences from medical costs, as well as improvements to mental and self-reported health (Finkelstein et al., 2012, 2019). However, the vast majority of analyses have not had the benefit of an experimental setting and have instead used quasi-experimental techniques to estimate the effects of health insurance provision. Broadly, this literature has found that earlier expansions of subsidized insurance improved health insurance coverage as well as health outcomes like birth weight (Currie & Gruber, 1996a,b). Additionally, results have suggested that government insurance coverage expansions reduced mortality, though they also had a significant "crowd-out" effect on private health insurance coverage (Goodman-Bacon, 2018; Cutler & Gruber, 1996; Gruber & Simon, 2008). Recent Medicaid expansions offer a new setting in which to test the findings established by this earlier work in a health insurance landscape that has shifted significantly.

The Affordable Care Act (ACA) was signed into law by President Barack Obama in March of 2010 and expanded health insurance to millions of Americans through Medicaid expansions for low-income families and subsidies to purchase private health insurance for middle-income Americans. States were given the option to expand Medicaid beneath 138% of the federal poverty guidelines (FPG) starting in September 2010 and were required to do so starting on January 1, 2014. four states and the District of Columbia used new options (Sommers et al., 2014) under the ACA and expanded Medicaid early in 2010-2011 to low-income individuals, including parents and childless adults (who were most impacted by the expansions; Medicaid already covered children at generous levels). In 2012, the US Supreme Court ruled in *National Federation of Independent Business (NFIB) v. Sebelius* that states had the option to expand Medicaid and could not be coerced to expand the program under the ACA. Since 2014 (and at the time of this writing), a total of 38 states and the District of Columbia have adopted Medicaid expansion (KFF, 2022), and 12 states have not expanded Medicaid for low-income families and individuals. Medicaid expansion was not uniform and there was wide variation in the timing and location of adoption.

Some of the first work examining the effects of Medicaid expansions focused on changes in health insurance coverage (and source of coverage). In general, this literature has found that Medicaid expansions significantly increased Medicaid coverage and reduced uninsurance rates in those states that expanded relative to those that did not (Kaestner et al., 2017; Courtemanche, Marton, & Yelowitz, 2016; Miller & Wherry, 2017;

Sommers et al., 2015; Frean et al., 2017). Other literature has examined how these expansions affected health insurance gaps between different populations and found that they significantly increased health insurance coverage across age and racial groups (Wehby & Lyu, 2017).

Building on these first results, many other analyses have leveraged the Medicaid expansion to examine effects on economic outcomes, finding that the expansions did not significantly impact work effort of recipients, but did cause physicians to increase their labor supply (Kaestner et al., 2017; Neprash et al., 2021). Additionally, other work has examined the effects of the ACA Medicaid expansions on health outcomes, with mixed results. Some analyses found that they reduced mortality in the aggregate, while others found null effects on other health outcomes (see *inter alia*; Miller et al. (2021); Soni et al. (2018); Borgschulte & Vogler (2020); Cawley et al. (2018); Allen & Sommers (2019)). Nearly all of these analyses leverage a DD design to estimate causal effects, but to my knowledge, only two recent papers have confronted the potential issue of differential treatment timing: Miller et al. (2021) and Nikpay (2022). Miller et al. (2021) use expansions from 2014—2017 and linked survey and administrative data to estimate the effects of Medicaid expansion on mortality, and they find that ACA expansions resulted in a roughly 10% decrease in mortality. They find that staggered treatment timing does not play a large role in their reduced-form effect, which is unsurprising given that their treatment variation is clustered in the 2014 treatment year. Nikpay (2022) examines the effects of Medicaid expansions through 2019 on hospital financing of safety net and non-safety-net patients, and finds that the expansions shifted approximately $5B from safety-net to non-safety-net hospitals. They find some evidence of heterogeneous treatment effects depending on treatment cohort, though they ultimate conclude it does not significantly bias their overall findings. However, neither paper leverages the newer expansions that have occurred since 2017, and as researchers begin to estimate the effects of these reforms, it is crucial to examine the influence of differential treatment timing on effect estimates, especially on the first stage of health insurance coverage.

Why might we expect bias in this context, given that the bulk of states expanded in 2014? If Medicaid eligibility expansions require a ramp-up period to reach full effect (which could be produced by administrative delays in processing applications, as is currently happening in Missouri), then these dynamic treatment effects could bias the effect estimates when using these states as comparisons for late expanders (Baker et al., 2022; Roth et al., 2022). Additionally, treatment cohorts may experience heterogeneous effects. These analyses typically leverage the expansions that occurred in 2014 among 22 states, which yield a clean 2x2 DD research design. However, before that point 5 states had already expanded and since 2014, 12 additional states have expanded. These states are different from the 2014 expanders in terms of population and political composition (among other characteristics), indicating that take-up, pre-policy uninsurance rates, and other intermediate outcomes that affect insurance changes may differ by treatment cohorts. Both heterogeneous

and dynamic treatment effects are substantive departures from the canonical DD model and may induce bias (Baker et al., 2022; Roth et al., 2022).

2.2 Difference-in-Differences

In many settings, social science researchers are unable to use experimental methods to answer causal questions due to ethical, resource, or other concerns. Difference-in-differences seeks to estimate the causal effect of an event or series of events, often policy changes. The researcher compares the change in the outcome variable among units (often geographical entities) that receive "treatment" before and after the event(s) with those changes that units who are not "treated." This yields the first differences of After-Before for both treated and untreated units. Then, they take the difference-in-differences between these quantities to yield a causal effect of the policy or event.

Difference-in-differences is built on the assumption that an outcome, Y_{ist}, for individual i at time t in state s, is composed of a time-invariant state effect (γ_s) and a year effect (λ_t) common across states:

$$\text{Untreated States:} \quad Y_{ist} = \gamma_s + \lambda_t + \epsilon_{ist} \qquad \text{Treated States: } Y_{ist} = \gamma_s + \lambda_t + \delta_{st} + \epsilon_{ist}$$

Treated state outcomes have the same components, with the addition of a treatment effect δ_{st}. Thus, I use a regression of the following form to identify the treatment effect:

$$Y_{ist} = \alpha + \gamma(T_s) + \lambda(D_t) + \delta(T_s \cdot D_t) + \epsilon_{ist}$$

T_s is a dummy variable that indicates treated states, D_t is a dummy variable that represents post treatment periods, and δ is the coefficient of interest.

One of the largest concerns with difference-in-differences was first highlighted in Goodman-Bacon (2021). For many years, researchers had been applying the difference-in-differences technique to situations where multiple units were being treated at different points in time. Goodman-Bacon highlighted how this could bias the coefficient estimates. If there is variation in treatment timing, then they want to account for this variation in treatment timing and understand whether the coefficient on the post*treatment dummy is due to comparisons of treated vs. untreated units, late treated vs. early control, or late control vs. early treated. In his original paper, Goodman-Bacon proposed an estimation technique that runs the individual 2x2 difference-in-differences estimates that compose a coefficient from a differential timing DD, also computing their weight in the aggregate coefficient (determined by variance). From there, his technique shows the percentage of the effect that is due to comparisons of treated vs. untreated units (the key causal effect of interest) as opposed to late treated vs. early control or late control vs. early treated comparisons. Though valuable as a diagnostic, it unfortunately does not allow the estimation of causal effects absent bias. Various researchers

have proposed assessments and alternative estimates to make difference-in-differences usable with differential treatment timing.

Sun & Abraham (2021) highlight that the same problems of contamination and biased comparisons also affect event study estimates when differential treatment timing is involved. As a result, they propose an estimator for event studies that accounts for differential timing by arranging groups in cohorts according to when they are first treated and weighting their coefficient by how long they are included in the data window (how many leads and lags they have). Gardner (2021) introduces an alternative way to think about the problems involved in DD with multiple time periods: it introduces bias that is not independent of group and period fixed effects; it projects heterogeneous treatment effects onto group and period fixed effects rather than the treatment status itself. One way to combat this bias is to run a first stage using only untreated units to determine group and period fixed effects, and then residualize the outcome by subtracting out group and period FEs. Then, I use this residualized outcome to estimate a standard DD (or event study) while accounting for differential treatment timing. Callaway & Sant'Anna (2021) focus on decomposing the overall treatment effect parameter into the individual 2x2 comparisons that are unaffected by the differential timing issues. Then, they propose aggregating these comparisons into estimators for treatment effect by group g, the treatment cohort. These values will then be averaged to arrive at the overall treatment effect for all groups that participated in the treatment—analogous to the ATT in the simple 2x2 case. In de Chaisemartin & D'Haultfœuille (2020), the authors propose a new estimator that not only tackles the issues with differential treatment timing, but also addresses heterogeneous treatment effects in a two-way fixed effects setting. One additional assumption that their estimator requires is that every time a unit is switching from untreated to treated between two periods, there must also be at least one unit that is untreated in both periods of interest. They also provide an estimator for the effect of leaving treatment if units switch from treated back to untreated. They rely on simply $t-1$ and t to estimate a 2x2 period, then aggregating these to estimate an overall treatment effect. In Wooldridge (2021), the author introduces an Extended TWFE (ETWFE) estimator that relies on flexible estimation of TWFE using a series of interaction terms between treatment cohort dummy variables and post-treatment dummy variables for each year after treatment. These interaction terms allow treatment effects to vary across time and treatment cohort and generate unbiased DD and EV estimates under similar assumptions to other alternative estimators. Finally, in Borusyak et al. (2022), the authors focus on the event study or staggered roll-out designs that accompany difference-in-differences when applied in contexts where differential timing is present. They use an "imputation" approach where they compute unit and period fixed effects using untreated units only, and then use these to impute the untreated potential outcomes, which can then be leveraged to find an estimated treatment effect for each treated observation. These are then weighted to compute an aggregate treatment effect. I apply this suite

of techniques to re-estimate the relationship between Medicaid expansions and health insurance outcomes.

3 Data

I use the American Community Survey (ACS), a nationally representative yearly cross-section of approximately 3 million U.S. adults in all 50 states that contains demographic, economic, social, and housing information. I use data from 2008-2019, the years publicly available in the IPUMS-USA database at the University of Minnesota Population Center (Ruggles et al., 2022). The ACS is valuable due to its large yearly sample sizes, nationwide coverage, and high response rates—truthful responses are required by law. For these reasons, much previous literature has used the ACS to estimate the effects of Medicaid Expansion, though some analyses have used alternative sources like the National Health Interview Survey (NHIS) or the Behavioral Risk Factor Surveillance System (BRFSS) (Kaestner et al., 2017; Miller & Wherry, 2017; Soni et al., 2018; Courtemanche, Marton, Ukert, et al., 2016)

Since 2008, the ACS has also recorded information about the health insurance coverage of each respondent. Specifically, I consider whether an individual has any health insurance coverage and whether that health insurance is privately provided (e.g., purchased directly or sponsored by an employer) or publicly provided through public health insurance programs like Medicare, Medicaid, or (for the case of current or past military personnel) TRICARE/Veterans Affairs coverage. It is important to consider that the sources of health insurance are not necessarily mutually exclusive—participants can report more than one type of health insurance source. I expect state-level Medicaid expansions under the ACA to be associated with an increased likelihood of having Medicaid coverage, and decreased likelihood of being uninsured. I use age and its square, race/ethnicity, marital status, number of children, and family size as covariates in my specifications. I also use educational attainment to restrict to sub-samples most likely to be affected by the ACA Medicaid expansions. I focus on people with educational attainment at or below a high school diploma to prevent the results from being biased by people endogenously changing their hours of work (and therefore salary) to ensure Medicaid eligibility, around the 138% FPL threshold. This is consistent with prior research that has raised this concern (Kaestner et al., 2017; Miller et al., 2021) and suggested using "targeted samples" such as cut-offs for education (in this case a high school diploma) to alleviate endogeneity concerns. I use missing indicators with missing data on my covariates, and exclude respondents who did not report their health insurance coverage.

4 Empirical Approach

I first estimate a standard DD design:

$$y_{ist} = \beta_0 + \beta_1(MEDICAIDEXP_{st}) + \beta_2\gamma + \beta_3\delta + \beta_4 X_{ist} + \epsilon_{ist} \qquad (1)$$

y_{ist} is an individual's health insurance status and source of coverage, $MEDICAIDEXP_{st}$ is an indicator variable that returns a 1 when individual i lives in a state s that expanded Medicaid at or before time t, γ is state fixed effects, δ is year fixed effects, X_{ist} is a vector of individual covariates, and ϵ_{ist} is an idiosyncratic error term. Hence, my coefficient of interest, β_1, is intended to capture the causal effect of Medicaid expansion on health insurance coverage. This standard empirical approach has been employed across several papers, including Miller & Wherry (2017); Kaestner et al. (2017); Courtemanche, Marton, Ukert, et al. (2016).

$$y_{ist} = \beta_0 + \sum_{j=-4}^{4} \beta_1^j(MEDICAIDEXP_{st}^j) + \beta_2\gamma + \beta_3\delta + \beta_4 X_{ist} + \epsilon_{ist} \qquad (2)$$

Additionally, I generate event study estimates of the effects of the ACA Medicaid expansions on the likelihood of insurance coverage for low-education respondents (equation above). In these models I replace the single Medicaid expansion indicator with a series of event-time indicators ($MEDICAIDEXP_{st}^j$) where j is the time period relative to the event, created by interacting each relative event time indicator with the DD dummy. As standard in the literature, I normalize the first lead operator (T-1) to zero. These results are presented in Section 5.3.

However, because recent literature has found that using the naive DD and event study specifications in settings with differential treatment timing can generate biased estimates, I also estimate these same equations using the specifications outlined in Gardner (2021); Sun & Abraham (2021); Callaway & Sant'Anna (2021); Borusyak et al. (2022); de Chaisemartin & D'Haultfœuille (2020); Cengiz et al. (2019); Wooldridge (2021). I then compare the effect estimates I find with each specification to assess the bias in previous estimates of the effect of Medicaid expansion, presented in Section 5.4.

5 Results

5.1 Raw Trends and Summary Statistics

Table 1 presents descriptive statistics for respondents in expansion vs. non-expansion states with educational attainment at or below a high school diploma, my primary analysis sample. Table 1 shows that the majority

of low-education adults had health insurance, while a smaller share had employer-sponsored insurance. Low-income individuals in expansion states were more likely than low-income individuals in non-expansion states to have Medicaid coverage.

Figure 2a presents unadjusted trends in the likelihood that individuals in households with a high school education or less have Medicaid coverage for ACA expansion states (blue), and non-expansion states (red). Several patterns are apparent. First, there is clear evidence that adults in expansion states are more likely to be covered by Medicaid than adults in non-expansion states. Second, there is visual support demonstrating that state-level Medicaid expansions under the Affordable Care Act increased Medicaid coverage for low-education adults in expansion states. There is also suggestive visual support in Figure 2b for the idea that the ACA's Medicaid expansion reduced uninsurance rates more in expansion states than in non-expansion states: the decreases in uninsurance post-expansion appear larger for respondents in expansion states as compared to the associated increases for respondents in non-expansion states. I formalize and test for these differences in a regression framework in Section 5.3.

5.2 Goodman-Bacon Decomposition

To determine the potential influence of comparisons of states with different treatment times, I implement the Goodman-Bacon (2021) decomposition that examines the role of each 2×2 DD comparison in the two-way fixed effects DD estimate. Table 2 and Figure 3 display the results of this analysis. Column 1 presents the type of comparison, Column 2 presents the average treatment effect estimated among that type of comparison, and Column 3 displays the weight that each type of comparison receives when aggregating to form the overall DD coefficient. I find that comparisons between states of different timing groups compose approximately 30% of the overall treatment effect estimate, and that comparisons of never-treated states with those "switching" on their treatment compose about 66% of the overall treatment effect estimate. The remainder (about 4%) comes from within-group variation in covariates. This composition suggests that a significant portion of the treatment effect estimate is composed of comparisons that could potentially bias the effect estimates. However, when I examine the estimated coefficients for each group, I find that the coefficient for different timing groups comparisons is qualitatively similar to that for never vs. treated comparisons. Though the potential for bias is significant, I do not find any using the diagnostic tool of the Goodman-Bacon Decomposition.

Figure 3 illustrates the results of a more detailed Goodman-Bacon Decomposition. Each of the individual points represents a specific 2x2 comparison, the x-axis displays the weight of that comparison in constructing the overall DD coefficient, and the y-axis displays the point estimate for that comparison. These results

suggest that within-timing group comparisons are not systematically biased in one direction (as they fall relatively evenly on both sides of the coefficient estimate, denoted by the red line).

5.3 Difference-in-Differences and Event Study

I estimate the effects of ACA Medicaid expansions on uninsurance rates among my low-education sample using equation (1). Column 1 of Table 3 presents the results of this specification. I find that the ACA state-level Medicaid expansions reduced uninsurance rates by 4.2 percentage points. This represents a 25 percent decrease from the base period mean uninsurance rate, 16.7%. This aligns with prior literature on the effects of Medicaid Expansion on uninsurance rates, which tends to find effect sizes of approximately 3–4 percentage points (Kaestner et al., 2017; Miller & Wherry, 2017; Courtemanche, Marton, & Yelowitz, 2016; Miller et al., 2021).

Next, I estimate the effects of ACA Medicaid expansions on uninsurance rates among my low-education sample using an event study design. Column 1 of Table 4 presents the results of this specification. Each column is a different estimation strategy, and each row presents the coefficient in a time period relative to 1 year before expansion (thus the T-1 coefficient is 0 and omitted). For example, T-4 presents the DD coefficient 4 time periods before the expansion occurred. I find that the ACA state-level Medicaid expansions reduced uninsurance rates by 2.3 percentage points in the first year and 4-5 percentage points in the succeeding 4 years. This aligns with prior literature on the effects of Medicaid Expansion on uninsurance rates, which tends to find effect sizes of approximately 3–4 percentage points after the first year (Kaestner et al., 2017; Miller & Wherry, 2017; Courtemanche, Marton, & Yelowitz, 2016; Miller et al., 2021). Additionally, I find no evidence of differential pre-trends, as my pre-period dummy coefficients are all insignificant.

5.4 Improved Estimation Strategies

Additionally, I use the estimators proposed in Gardner (2021); Callaway & Sant'Anna (2021) to re-estimate the simple DD specification in a manner robust to differential treatment timing, presented in Table 3. Given the results of the Goodman-Bacon Decomposition, I expect broadly similar results to my point estimate of -0.042 from the OLS TWFE regression. Column 2 displays the results of the Gardner (2021) estimation strategy, and I find a very similar effect of -0.041 (with similar standard errors and statistical significance). Additionally, I use the estimators proposed in Callaway & Sant'Anna (2021); Wooldridge (2021); de Chaisemartin & D'Haultfœuille (2020); Sun & Abraham (2021), with results displayed in Columns 3-6. The point estimates are again very similar to that estimated using OLS and the Gardner method. Overall, these results are fairly robust across approaches—the estimation strategies that correct for differential treatment timing

present results similar to those estimated using OLS,and they are statistically indistinguishable.

Additionally, I use the estimators proposed in Gardner (2021); Borusyak et al. (2022); Cengiz et al. (2019); de Chaisemartin & D'Haultfœuille (2020); Sun & Abraham (2021); Callaway & Sant'Anna (2021); Wooldridge (2021) to re-estimate the simple event study specification in a manner robust to differential treatment timing, presented in Table 4. Given the results of the Goodman-Bacon Decomposition, I expect broadly similar results to my point estimates from the OLS event study regression. Column 2 displays the results of the Gardner (2021) estimation strategy, and I find similar results, except in period T+4, where the coefficient is approximately 1/4 the magnitude of that estimated using OLS. Column 3 displays the results of the Borusyak et al. (2022) estimation strategy, and I find broadly similar results to those estimated using OLS. Column 4 presents the results estimated using the strategy proposed in Cengiz et al. (2019), and I find generally similar effect sizes. I use the estimator proposed in de Chaisemartin & D'Haultfœuille (2020), with results displayed in Column 5. The point estimates are again very similar to that estimated using the OLS, Gardner, and Borusyak methodologies. The results estimated using Sun & Abraham (2021) are displayed in Column 6, showing broadly similar results. Finally, the results estimated using Callaway & Sant'Anna (2021); Wooldridge (2021) are displayed in Columns 7-8, showing broadly similar results. In the last four estimators, I also see evidence of potential violations of the parallel trends assumption. Though statistically significant, these lead coefficients are generally small in substantive terms. Overall, these results mostly reaffirm previous literature—though most of the estimation strategies that correct for differential treatment timing present results similar to those estimated using OLS, another presents a much smaller effect size (though statistically indistinguishable). These results are also displayed in graphical format in Figure 4, where the estimates from the improved estimation strategies are broadly similar to those obtained using the OLS TWFE specification.

6 Conclusion

A large body of literature has illustrated how state-level Medicaid expansions improved health insurance coverage among low-income individuals in the United States. Additionally, further analyses have examined the downstream health and economic impacts of these expansions, finding mixed evidence on health impacts and general improvements in economic outcomes. These estimates have informed health policy debates as more states consider expanding their Medicaid eligibility through the channels set out in the ACA. Though this literature has blossomed, most of the analyses rely on a single empirical strategy, difference-in-differences, to estimate the causal effects of ACA Medicaid expansions. However, recent literature in econometrics has highlighted how this strategy can be biased in settings of differential treatment timing and

proposed diagnostics and alternative estimators that can address these concerns. I apply these econometric tools to re-estimate the effects of ACA Medicaid expansions on health insurance coverage among a low-education subsample, and find substantive potential for bias in previous estimates using a Goodman-Bacon decomposition. However, despite this potential, I find no evidence of significant bias in previous estimates. Instead, I reaffirm the central research finding that Medicaid expansions reduced uninsurance rates by 4-5 percentage points among adults with a high school educational attainment or lower, a target population for these policies. I find that this relationship is robust to a suite of different DD and event study specifications.

Additionally, I used a diverse suite of estimators to address differential timing problems in differences-in-differences and event studies. There are several tradeoffs between these estimators that empirical researchers should consider as they think about which to apply in their work. One dimension across which these specifications vary is how well they can accommodate covariates—the estimators proposed by Wooldridge, Cengiz et. al., Borusyak et. al., and de Chaisemartin can, but the estimator proposed by Callaway & Sant'anna cannot accommodate covariates that vary across time (e.g., state-level unemployment rates) and those proposed by Sun & Abraham, Goodman-Bacon, and Gardner cannot incorporate covariates at all. Additionally, some of the estimators rely on different assumptions or weighting schemes—Borusyak et. al. require a stronger parallel trends assumption on all periods rather than only those after the reference period, and "stacked" regressions (proposed by Cengiz et. al.) weight by the number of treated units and variance of treatment within each stacked event, rather than a researcher-chosen weighting scheme with economic motivation. Selecting two to three estimation strategies based on contextual factors and checking whether naive DD and ES results are robust to these is likely the clearest path forward for empirical researchers working in settings with differential treatment timing

There are several limitations to my analysis. The ACS lacks information on access to care, health services utilization, and health outcomes, and so I can only examine effects on health insurance coverage. I leave examination of these other health outcomes in this econometric context to future research. Secondly, misreporting health insurance status may bias my results. Some respondents may inaccurately report their health insurance status. Previous research has documented an undercount of Medicaid coverage in the ACS since some participants with Medicaid managed care misreport their insurance status as private health insurance (Boudreaux et al., 2019).

Despite these limitations, my findings reaffirm the broad effects of the ACA Medicaid expansions and suggest that eliminating or scaling back the Medicaid expansion could reduce health insurance coverage among low-education adults. My analysis is one of the first (to my knowledge) to use econometric methods robust to differential treatment timing to examine ACA Medicaid expansions, joining Nikpay (2022) and Miller et al. (2021). These two analyses focus on the impacts of expansions on the target efficiency of

hospitals and mortality, while my analysis examines insurance coverage. Additionally, I add to the small number of studies that compare the new econometric estimators developed to diagnose and correct for issues of treatment timing and apply them to an empirical setting. State-level Medicaid expansions form the basis of a large literature in health economics and health services research, and my application highlights the potential for other applications throughout applied microeconomics. Finally, my results illustrate the importance of applying these methodologies when estimating difference-in-differences or event studies in settings with differential treatment timing.

References

Allen, H., & Sommers, B. D. (2019, oct). Medicaid expansion and health. *JAMA*, *322*(13), 1253. doi: 10.1001/jama.2019.12345

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*.

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022, may). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, *144*(2), 370–395. doi: 10.1016/j.jfineco.2022.01.004

Borgschulte, M., & Vogler, J. (2020, jul). Did the ACA medicaid expansion save lives? *Journal of Health Economics*, *72*, 102333. doi: 10.1016/j.jhealeco.2020.102333

Borusyak, K., Jaravel, X., & Spiess, J. (2022). Revisiting event study designs:robust and efficient estimation. *Working Paper*. Retrieved from `https://arxiv.org/pdf/2108.12419.pdf`

Boudreaux, M., Noon, J. M., Fried, B., & Pascale, J. (2019, oct). Medicaid expansion and the medicaid undercount in the american community survey. *Health Services Research*, *54*(6), 1263–1272. doi: 10.1111/1475-6773.13213

Callaway, B., & Sant'Anna, P. H. (2021, dec). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. doi: 10.1016/j.jeconom.2020.12.001

Cawley, J., Soni, A., & Simon, K. (2018, jun). Third year of survey data shows continuing benefits of medicaid expansions for low-income childless adults in the u.s. *Journal of General Internal Medicine*, *33*(9), 1495–1497. doi: 10.1007/s11606-018-4537-0

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019, may). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405–1454. doi: 10.1093/qje/qjz014

Courtemanche, C., Marton, J., Ukert, B., Yelowitz, A., & Zapata, D. (2016, nov). Early impacts of the affordable care act on health insurance coverage in medicaid expansion and non-expansion states. *Journal of Policy Analysis and Management*, *36*(1), 178–210. doi: 10.1002/pam.21961

Courtemanche, C., Marton, J., & Yelowitz, A. (2016, apr). Who gained insurance coverage in 2014, the first year of full ACA implementation? *Health Economics*, *25*(6), 778–784. doi: 10.1002/hec.3349

Currie, J., & Gruber, J. (1996a, may). Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics*, *111*(2), 431–466. doi: 10.2307/2946684

Currie, J., & Gruber, J. (1996b, dec). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy*, *104*(6), 1263–1296. doi: 10.1086/262059

Cutler, D. M., & Gruber, J. (1996, may). Does public insurance crowd out private insurance? *The Quarterly Journal of Economics*, *111*(2), 391–430. doi: 10.2307/2946683

de Chaisemartin, C., & D'Haultfœuille, X. (2020, sep). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996. doi: 10.1257/aer.20181169

Finkelstein, A., Hendren, N., & Luttmer, E. F. P. (2019, dec). The value of medicaid: Interpreting results from the oregon health insurance experiment. *Journal of Political Economy*, *127*(6), 2836–2874. doi: 10.1086/702238

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., ... and, K. B. (2012, jul). The oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics*, *127*(3), 1057–1106. doi: 10.1093/qje/qjs020

Frean, M., Gruber, J., & Sommers, B. D. (2017, may). Premium subsidies, the mandate, and medicaid expansion: Coverage effects of the affordable care act. *Journal of Health Economics*, *53*, 72–86. doi: 10.1016/j.jhealeco.2017.02.004

Gardner, J. (2021). Two-stage difference-in-differences. *Working Paper*. Retrieved from `https://jrgcmu.github.io/2sdd2.pdf`

Goodman-Bacon, A. (2018, feb). Public insurance and mortality: Evidence from medicaid implementation. *Journal of Political Economy*, *126*(1), 216–262. doi: 10.1086/695528

Goodman-Bacon, A. (2021, dec). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. doi: 10.1016/j.jeconom.2021.03.014

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*.

Gruber, J. (2011, sep). The impacts of the affordable care act: How reasonable are the projections? *National Tax Journal*, *64*(3), 893–908. doi: 10.17310/ntj.2011.3.06

Gruber, J., & Simon, K. (2008, mar). Crowd-out 10 years later: Have recent public insurance expansions crowded out private health insurance? *Journal of Health Economics*, *27*(2), 201–217. doi: 10.1016/j.jhealeco.2007.11.004

Kaestner, R., Garrett, B., Chen, J., Gangopadhyaya, A., & Fleming, C. (2017, may). Effects of ACA medicaid expansions on health insurance coverage and labor supply. *Journal of Policy Analysis and Management*, *36*(3), 608–642. doi: 10.1002/pam.21993

KFF. (2022). *The status of state medicaid expansion* (Tech. Rep.). Kaiser Family Foundation. Retrieved from `https://www.kff.org/medicaid/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/`

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., & Marquis, M. S. (1987, June). Health insurance and the demand for medical care: evidence from a randomized experiment. *The American economic review*, *77*, 251–277.

Miller, S., Johnson, N., & Wherry, L. R. (2021, jan). Medicaid and mortality: New evidence from linked survey and administrative data. *The Quarterly Journal of Economics*. doi: 10.1093/qje/qjab004

Miller, S., & Wherry, L. R. (2017, mar). Health and access to care during the first 2 years of the ACA medicaid expansions. *New England Journal of Medicine*, *376*(10), 947–956. doi: 10.1056/nejmsa1612890

Neprash, H. T., Zink, A., Sheridan, B., & Hempstead, K. (2021, dec). The effect of medicaid expansion on medicaid participation, payer mix, and labor supply in primary care. *Journal of Health Economics*, *80*, 102541. doi: 10.1016/j.jhealeco.2021.102541

Nikpay, S. (2022, apr). The medicaid windfall: Medicaid expansions and the target efficiency of hospital safety-net subsidies. *Journal of Public Economics*, *208*, 104583. doi: 10.1016/j.jpubeco.2021.104583

Pauly, M. (1968). The economics of moral hazard. *American Economic Review*.

Roberts, J. (2012). *National federation of independent business, et al. v. sebelius, secretary of health and human services, et al.* (United States Supreme Court)

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature.

Ruggles, S., Flood, S., Goeken, R., Schouweiler, M., & Sobek, M. (2022). *Ipums usa.* Minneapolis, MN. doi: 10.18128/D010.V12.0

Sommers, B. D., Gunja, M. Z., Finegold, K., & Musco, T. (2015, jul). Changes in self-reported insurance coverage, access to care, and health under the affordable care act. *JAMA*, *314*(4), 366. doi: 10.1001/jama.2015.8421

Sommers, B. D., Kenney, G. M., & Epstein, A. M. (2014, jan). New evidence on the affordable care act: Coverage impacts of early medicaid expansions. *Health Affairs*, *33*(1), 78–87. doi: 10.1377/hlthaff.2013.1087

Soni, A., Simon, K., Cawley, J., & Sabik, L. (2018, feb). Effect of medicaid expansions of 2014 on overall and early-stage cancer diagnoses. *American Journal of Public Health*, *108*(2), 216–218. doi: 10.2105/ajph.2017.304166

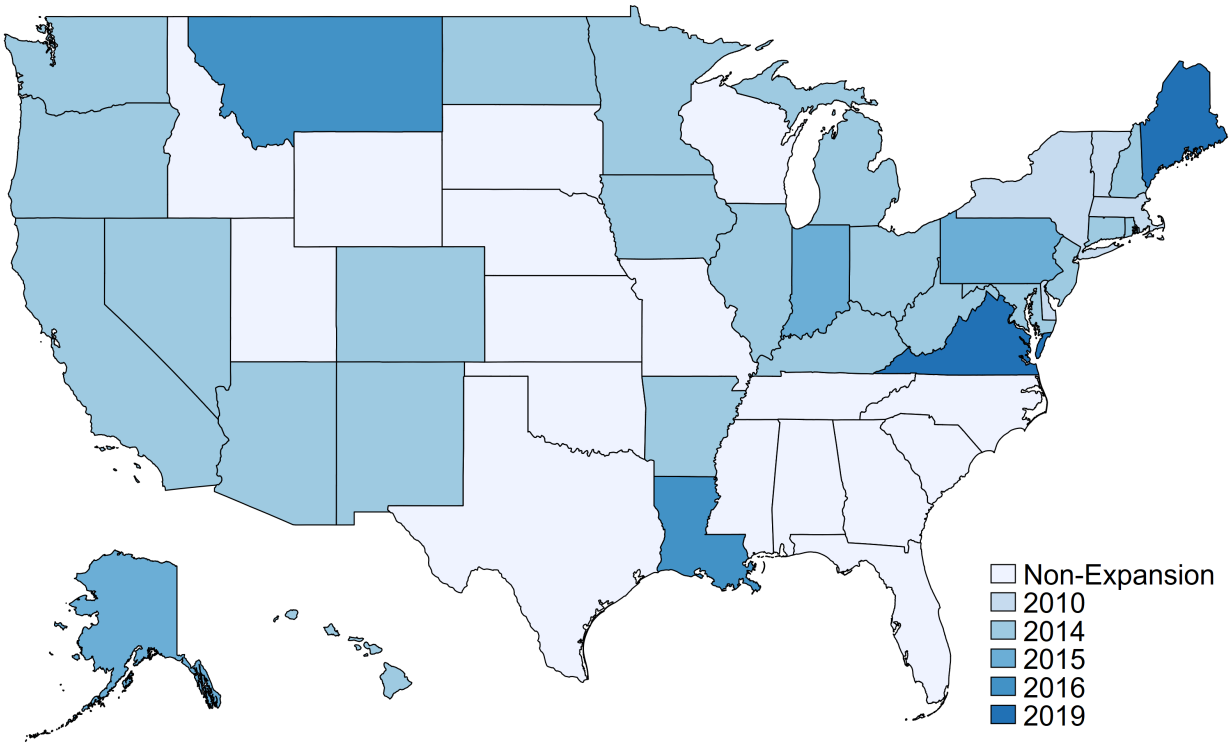
Sun, L., & Abraham, S. (2021, dec). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. doi: 10.1016/j.jeconom.2020.09.006

Wehby, G. L., & Lyu, W. (2017, may). The impact of the ACA medicaid expansions on health insurance coverage through 2015 and coverage disparities by age, race/ethnicity, and gender. *Health Services Research*, *53*(2), 1248–1271. doi: 10.1111/1475-6773.12711

Wooldridge, J. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators.

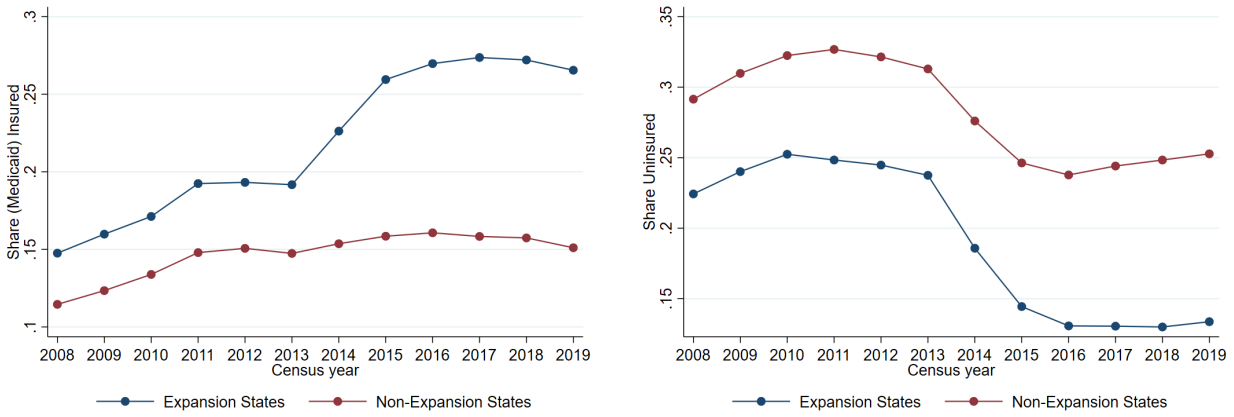
Tables and Figures

Figure 1: Timeline of state-level Medicaid expansions up to 2019.



Source: KFF (2022). Colors correspond to expansion years. Since 2019, 5 additional states (Idaho, Utah, Nebraska, Oklahoma, and Missouri) have passed Medicaid expansions. Because I do not leverage this variation in my analysis (given that it is out of my sample period), I mark these states as Non-Expansion states.

Figure 2: Raw trends in insurance coverage, 2008-2019

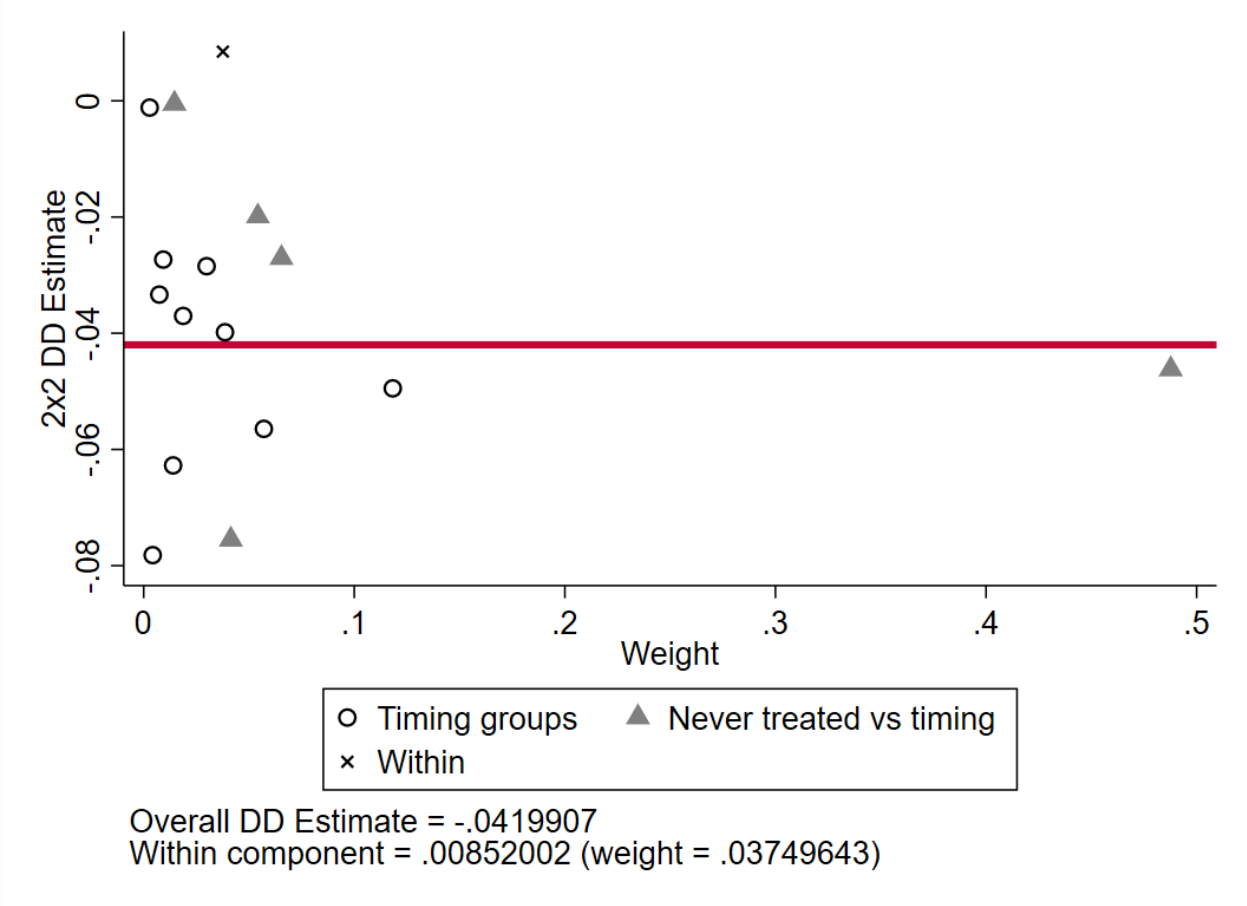


(a) Medicaid Coverage

(b) Uninsurance Rate

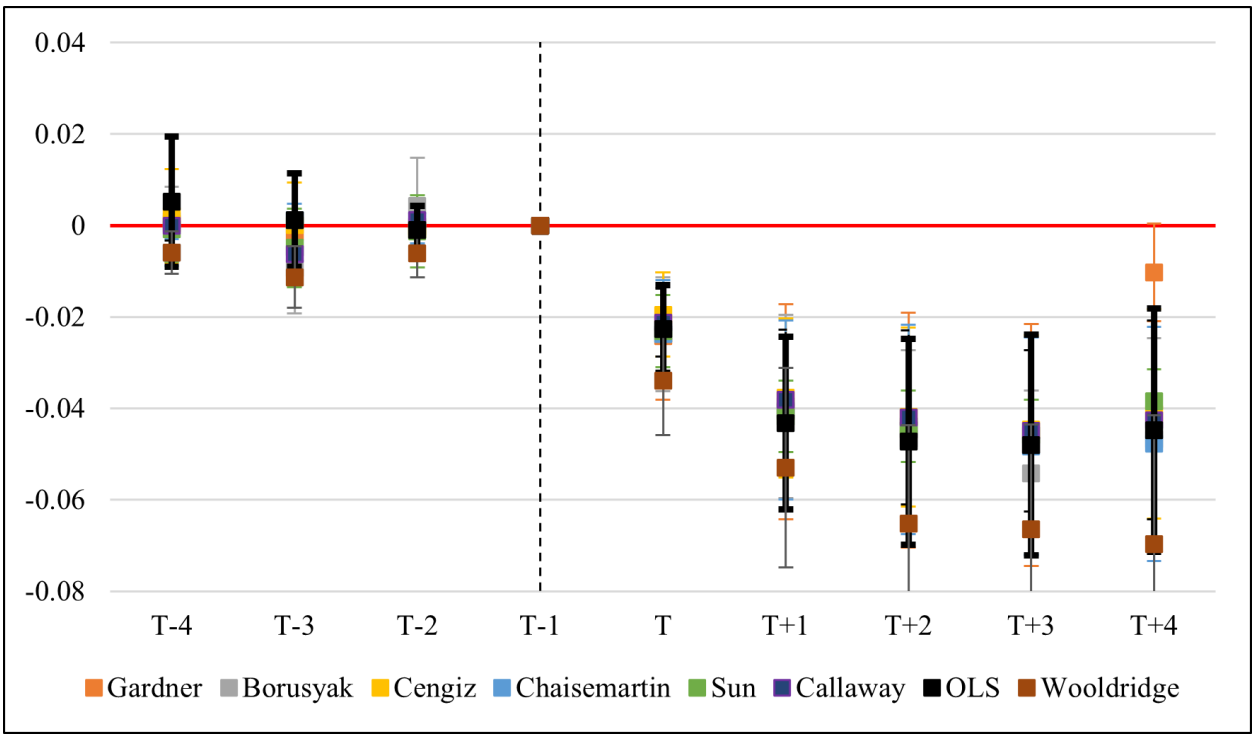
Source: American Community Survey 2008-2019, authors' calculations. Blue line denotes Medicaid expansion states, and red line denotes non-expansion states.

Figure 3: Goodman-Bacon Decomposition of DD on Uninsurance Rates.



Source: American Community Survey 2008-2019, authors' calculations. Created using the *bacondecomp* package on Stata. The x-axis shows the weight of the individual 2x2 DDs in the composition of the aggregate effect, and the y-axis is the estimated effect for that specific 2x2 DD. Icons denote the type of comparison (never-treated vs. timing is the unbiased effect of interest).

Figure 4: Event study results across estimation strategies.



Source: American Community Survey 2008-2019, authors' calculations. Each series is an event study coefficient, calculated using a different estimation strategy. All specifications include state and time fixed effects, as well as a vector of individual-level covariates (age and its square, race/ethnicity, marital status, number of children, and family size). Standard errors clustered at the state level were used to calculate 95% confidence intervals.

Table 1: Summary Statistics, Expansion vs. Non-Expansion States.

	(1) Non-Expansion State	(2) Expansion State
Age	41.914	41.844
	(13.676)	(13.685)
Male	0.491	0.494
	(0.500)	(0.500)
Non-Hispanic White	0.663	0.680
	(0.473)	(0.466)
Non-Hispanic Black	0.139	0.085
	(0.346)	(0.279)
Non-Hispanic All Other Races	0.053	0.099
	(0.225)	(0.298)
Hispanic	0.144	0.136
	(0.351)	(0.343)
Married	0.547	0.531
	(0.498)	(0.499)
Divorced or Widowed	0.162	0.141
	(0.369)	(0.348)
Never Married	0.291	0.327
	(0.454)	(0.469)
Under the FPL	0.174	0.157
	(0.379)	(0.363)
Uninsured	0.194	0.128
	(0.396)	(0.335)
Private Health Insurance	0.706	0.735
	(0.455)	(0.441)
Public Health Insurance	0.137	0.172
	(0.344)	(0.377)
Employer-Sponsored Insurance	0.590	0.635
	(0.492)	(0.481)
Medicaid	0.095	0.140
	(0.293)	(0.347)
Directly Purchased Insurance	0.120	0.112
	(0.325)	(0.315)
TRICARE Insurance	0.034	0.022
	(0.182)	(0.148)
Medicare	0.044	0.039
	(0.206)	(0.193)
VA Insurane	0.022	0.016
	(0.146)	(0.125)
IHS Insurance	0.006	0.007
	(0.079)	(0.082)
Observations	3,592,751	6,782,394

Source: American Community Survey 2008-2019, authors' calculations. Means and standard deviations (presented in parentheses) reported, stratified by binary state expansion variable

Table 2: Goodman-Bacon Decomposition of DD on Uninsurance Rates.

	Beta	Total Weight
Timing Groups	-.0461504	.2990919
Never vs. Treatment	-.0429702	.6634117
Within Group	.00852	.0374964

Source: American Community Survey 2008-2019, authors' calculations. Created using the *bacondecomp* package on Stata. The Total Weight column shows the weight of the individual 2x2 DDs in the composition of the aggregate effect, and the Beta column is the estimated effect for that type of 2x2 DD. Rows denote the type of comparison (never-treated vs. timing is the unbiased effect of interest).

Table 3: Difference-in-differences results across specifications.

	(1) OLS	(2) Gardner	(3) Wooldridge	(4) de Chaisemartin & D'Haultfœuille	(5) Sun & Abraham	(6) Callaway & Sant'anna
DD_{st}	-0.0416***	-0.0405***	-0.0545***	-0.0404***	-0.0376***	-0.0353***
	(0.0119)	(0.0118)	(0.0095)	(0.0109)	(0.0065)	(0.0089)
N	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: American Community Survey 2008-2019, authors' calculations. Each entry is a difference-in-differences coefficient, calculated using a different estimation strategy. All specifications include state and time fixed effects, as well as a vector of individual-level covariates (age and its square, race/ethnicity, marital status, number of children, and family size) when they are accommodated. Standard errors are clustered at the state level. Column 1 presents the results of a tradition ordinary-least-squares two-way fixed-effects specification. Column 2 presents the results of the estimation strategy proposed in Gardner (2021). Column 3 presents the results of the estimation strategy proposed in Wooldridge (2021). Column 4 presents the results of the estimation strategy proposed in de Chaisemartin & D'Haultfœuille (2020). Column 5 presents the results of the estimation strategy proposed in Sun & Abraham (2021). Column 6 presents the results of the estimation strategy proposed in Callaway & Sant'Anna (2021).

Table 4: Event study results across estimation strategies.

	(1) OLS	(2) Gardner	(3) Borusyak et. al.	(4) Cengiz et. al.	(5) Chaisemartin et. al.	(6) Sun & Abraham	(7) Callaway & Sant'anna	(8) Wooldridge
T-4	0.00518	0.000372	0.00102	0.0021	-0.0059**	-0.0009	-0.0001	-0.0058
	(0.00729)	(0.000668)	(0.00377)	(0.0052)	(0.0015)	(0.0009)	(0.0016)	(0.0046)
T-3	0.00124	-0.00152	-0.00939*	0.0001	0.0012	-0.0049**	-0.0063**	-0.0113**
	(0.00522)	(0.000979)	(0.00503)	(0.0047)	(0.0018)	(0.0011)	(0.0020)	(0.0067)
T-2	-0.000991	-0.000228	0.00421	-0.0008	0.0010	-0.0013	0.0012	-0.0062
	(0.00272)	(0.000818)	(0.00538)	(0.0027)	(0.0025)	(0.0010)	(0.0017)	(0.0054)
T	-0.0227***	-0.0242***	-0.0238***	-0.0195***	-0.0238***	-0.0231***	-0.0213***	-0.0340***
	(0.00491)	(0.00714)	(0.00637)	(0.0047)	(0.0060)	(0.0010)	(0.0038)	(0.0119)
T+1	-0.0432***	-0.0408***	-0.0396***	-0.0377***	-0.0404***	-0.0418***	-0.0381***	-0.0529***
	(0.00965)	(0.0120)	(0.0102)	(0.0089)	(0.0100)	(0.0010)	(0.0078)	(0.0218)
T+2	-0.0473***	-0.0448***	-0.0467***	-0.0419***	-0.0446***	-0.0439***	-0.0420***	-0.0652***
	(0.0115)	(0.0131)	(0.00992)	(0.0100)	(0.0117)	(0.0010)	(0.0097)	(0.0215)
T+3	-0.0480***	-0.0481***	-0.0542***	-0.0447***	-0.0483***	-0.0459***	-0.0450***	-0.0664***
	(0.0123)	(0.0135)	(0.00919)	(0.0105)	(0.0121)	(0.0010)	(0.0090)	(0.0229)
T+4	-0.0448***	-0.0103*	-0.0468***	-0.0412**	-0.0478***	-0.0385***	-0.0426***	-0.0697***
	(0.0136)	(0.00545)	(0.0113)	(0.0117)	(0.0131)	(0.0009)	(0.0111)	(0.0281)
N	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145	10,375,145

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: American Community Survey 2008-2019, authors' calculations. Each entry is a an event study coefficient, calculated using a different estimation strategy. All specifications include state and time fixed effects, as well as a vector of individual-level covariates (age and its square, race/ethnicity, marital status, number of children, and family size) when they are accommodated. Standard errors are clustered at the state level. Column 1 presents the results of a tradition ordinary-least-squares two-way fixed-effects event study specification. Column 2 presents the results of the estimation strategy proposed in Gardner (2021). Column 3 presents the results of the estimation strategy proposed in Borusyak et al. (2022). Column 4 presents the results of the estimation strategy proposed in Cengiz et al. (2019). Column 5 presents the results of the estimation strategy proposed in de Chaisemartin & D'Haultfœuille (2020). Column 6 presents the results of the estimation strategy proposed in Sun & Abraham (2021). Column 7 presents the results of the estimation strategy proposed in Callaway & Sant'Anna (2021). Column 8 presents the results of the estimation strategy proposed in Wooldridge (2021).