Towson University
Department of Economics
Working Paper Series

Working Paper No. 2025-01

The Effects of ACA-Medicaid Expansion on Maternal and Infant Health Outcomes in the American South

By Juergen Jung and Vinish Shrestha

July 2025

The Effects of ACA-Medicaid Expansion on Maternal and Infant Health Outcomes in the American South*

Juergen Jung† Vinish Shrestha‡
Towson University Towson University

July 22, 2025

Abstract

Maternal and infant health outcomes in the American South lag behind those in other U.S. regions. This study estimates the causal impact of Medicaid expansion to 138 percent of the Federal Poverty Level on maternal and infant health outcomes, using county-level data from the National Vital Statistics System (2010–2017). We employ a difference-in-differences design, comparing counties in expansion and non-expansion states, and use inverse propensity score weighting to improve covariate balance. To flexibly adjust for baseline differences, we extend the analysis with a machine learning–based Doubly Robust DiD estimator. Our findings show that Medicaid expansion significantly reduced pregnancy-associated mortality among non-Hispanic Black mothers, primarily by lowering pregnancy-related deaths, with no effect on incidental causes. No significant impact is found for non-Hispanic White mothers. While infant mortality rates remained unchanged for both groups, the expansion modestly improved birthweight outcomes among Black infants. These results highlight the potential of Medicaid expansion to reduce racial disparities in maternal and infant health in the American South.

JEL: M51, J71, J62, J63, J16

Keywords: The Patient Protection and Affordable Care Act (ACA), Infant Health Outcomes, American South, Health Inequality.

*We would like to thank participants at the Development & Research Conference at the College of Business and Economics at Towson University for helpful comments. Juergen Jung received grant support from the School of Emerging Technologies (SET) at Towson University. Beyond that he has no further relevant or material financial interests that relate to the research described in this paper.

†Department of Economics, Towson University, *Email:* jjung@towson.edu

‡Corresponding author: Department of Economics, Towson University, *Email:* vshrestha@towson.edu

1 Introduction

Maternal and infant health are widely regarded as core indicators of a nation's well-being. Yet, the United States performs poorly on both measures compared to other industrialized countries. As shown in Figure 1, Panel A, the U.S. ranked second to last in maternal mortality in 2019. Even more troubling is the sharp increase in maternal mortality over time, rising from 17.4 deaths per 100,000 live births in 2018 to 32.9 in 2021 (Panel B). Infant mortality follows a similar trajectory (Panel C), marked by persistent racial disparities—particularly between Black and White infants. These trends are compounded by declining birthweight and a growing incidence of low birthweight (Hoyert, 2022; Pollock, Gennuso, Givens and Kindig, 2021).

While socioeconomic disparities significantly contribute to these poor outcomes, the burden is unevenly distributed across regions. Areas with higher proportions of African Americans, for example, face disproportionately worse outcomes. This is driven by the fact that minorities face worse health outcomes in general. For instance, in 2021, Black, non-Hispanic women experienced a maternal mortality rate of 69.9 per 100,000 live births, compared to 26.6 for White, non-Hispanic women (Hoyert, 2022). However, the racial disparity in outcomes is not adequately explained by socioeconomic factors including income as infant and maternal health outcomes among Black families at the top of the income distribution are discernibly worse than those of Whites at the bottom of the distribution (Kennedy-Moulton, Miller, Persson, Rossin-Slater, Wherry and Aldana, 2022). These patterns underscore the intersection of systemic inequities and geographic disparities in health outcomes (Valerio, Downey, Sgaier, Callaghan, Hammer and Smittenaar, 2023).

This study focuses on the American South to evaluate how expanding Medicaid income eligibility to 138 percent of the Federal Poverty Level (FPL) under the Affordable Care Act (ACA) affects pregnancy-related mortality (PRMR), infant mortality rates (IMR), and infant health outcomes such as birthweight and low birthweight. The South merits particular attention for three key reasons. First, as shown in Panel D of Figure 1, the region consistently records some of the worst health outcomes in the country (Arias, Bastian, Xu and Tejada-Vera, 2021; RWJF, 2020).[1] Second, despite having some of the highest rates of uninsured adults and medical debt, southern states have been among the most resistant to Medicaid expansions under the ACA. Notably, 7 of the 10 states that have not expanded Medicaid are located in the South. Third, the South remains home to approximately 58% of the Black population in the U.S. (Frey, 2021). Accordingly, we place special emphasis on the reform's differential impact on Black and White populations, given the region's persistent racial disparities in health outcomes.

Medicaid plays a vital role in maternal and infant health, covering more than 40 percent of

[1]For example, the South averages 1.8 more infant deaths per 1,000 live births than other U.S. regions (Hirai, Sappenfield, Kogan, Barfield, Goodman, Ghandour and Lu, 2014), and Mississippi's IMR is more than twice the OECD average.

births nationwide (Valenzuela and Osterman, 2023). Although income eligibility thresholds for pregnant women under the Children's Health Insurance Program (CHIP) are generally higher than those set by the ACA's Medicaid expansion, a substantial share of reproductive-age women in the American South were uninsured prior to the ACA—25 percent of White women and 33 percent of Black women, based on 2010 data from the American Community Survey. This coverage gap is further exacerbated by insurance churn—frequent changes or losses in insurance coverage—which is especially common during pregnancy. Daw, Hatfield, Swartz and Sommers (2017) find that 58 percent of women experience month-to-month changes in coverage during pregnancy, with the rate rising to 73 percent among those covered by Medicaid at the time of delivery.

Although the ACA's Medicaid expansion did not specifically target pregnant women, it can influence maternal and infant health outcomes through changes in insurance coverage before and during pregnancy. The expansion introduces several economic mechanisms that may affect these outcomes. From the outset, women in expansion states are more likely to be insured at the time of conception compared to those in non-expansion states. For non-pregnant women with incomes between 100 percent and 138 percent of the Federal Poverty Level—who were previously ineligible under many state-level Medicaid thresholds—the expansion significantly increases the likelihood of obtaining insurance.[2]

This can benefit health outcomes in the following ways. First, newly acquired insurance coverage can improve preconception health and increase access to early prenatal care. Second, the Medicaid expansion reduces the risk of losing coverage during pregnancy by extending eligibility both before and after childbirth. This improves continuity of care and reduces insurance instability for women of reproductive age. Third, by lowering the financial burden and health risks associated with pregnancy and childbirth, the expansion may encourage higher birth rates, as observed in the RAND Health Insurance Experiment (Leibowitz, 1990). At the same time, greater access to preconception care and family planning services may lead to more planned pregnancies. Women gaining coverage may choose to delay pregnancy to improve their health, financial security, or social circumstances—reflecting a quantity–quality trade-off in fertility decisions (Becker, 1960). By prioritizing fewer but healthier pregnancies, families may allocate more resources per child, potentially improving maternal and infant health outcomes while reducing overall fertility. Finally, the Medicaid expansion may also affect fertility indirectly through labor market channels. Pub-

[2] While non-disabled adults and non-pregnant women without children did not qualify for Medicaid in the South prior to the ACA, Medicaid eligibility criteria for parents remained well below 100 percent of the FPL for most of the Southern states. The two exceptions are Delaware and Tennessee with eligibility thresholds greater than 100 percent of the FPL prior to the ACA. By April 1990, all states were required to provide Medicaid to pregnant women with family incomes up to at least 138 percent of the FPL. Some states set higher thresholds up to 185 percent. Given that pregnancy was considered a preexisting condition in the individual health insurance market which often entailed waiting periods or required costly additional coverage, Medicaid was de facto the only pathway to insurance for uninsured, low-income, pregnant women (Gifford, Walls, Ranji, Salganicoff and Gomez, 2017).

lic health insurance access can influence employment choices, including reducing participation in stressful or physically demanding jobs during pregnancy.

The primary data for this study come from the National Vital Statistics System (NVSS), covering the years 2010–2017, with outcomes aggregated at the county level, including PRMR, IMR, birthweight, and low birthweight. We employ a difference-in-differences (DiD) framework to compare counties in Medicaid expansion states with those in non-expansion states. A key challenge in applying the standard DiD approach is its reliance on the unconditional parallel trends assumption—that in the absence of Medicaid expansion, outcomes in both groups would have followed similar trajectories. This assumption is problematic because pre-treatment characteristics differ substantially between expansion and non-expansion counties, likely influencing outcome dynamics even without the policy change. Additionally, other components of the ACA, such as subsidized private insurance through the exchanges, affected both groups. These elements may have altered health outcomes in non-expansion states in ways that differ from expansion states, further complicating the identification of causal effects.

In order to address these concerns, we adopt a two-step approach. First, we use a double LASSO procedure to select covariates for estimating propensity scores that predict treatment assignment. These scores are then used to weight control group observations, improving comparability between treated and control units in the regression framework. The goal is to enhance balance in pre-treatment covariates between expansion and non-expansion states (Belloni, Chernozhukov and Hansen, 2014; Urminsky, Hansen and Chernozhukov, 2016).[3] This approach leverages high-dimensional, pre-treatment county-level characteristics to identify key predictors of treatment. Propensity scores are estimated using LASSO with 10-fold cross-fitting to reduce overfitting.[4] In the final regression model linking ACA Medicaid expansion to health outcomes, we include interactions between the selected covariates and the post-treatment indicator. While this approach controls for influential covariates, it imposes linearity in their effects. As a result, the estimate of the average treatment effect on the treated (ATT) may still be biased if treatment effects vary with covariates included in the model (Meyer, 1995).

Second, in order to mitigate the issue caused by linear controls, we next employ a more flexible approach to account for pre-treatment covariates and estimate the ATT using a Doubly-Robust (DR) estimator tailored to DiD (Sant'Anna and Zhao, 2020). This approach combines the inverse probability weighting (IPW) approach from Abadie (2005) with outcome regression (OR) of Heckman, Ichimura and Todd (1997) and ensures valid estimates even if only one of the models—the propensity score or the outcome model—is correctly specified (Sant'Anna and Zhao, 2020).

[3]The double LASSO framework selects covariates in two stages: Step 1 uses LASSO regression to predict the treatment variable (expansion status), and Step 2 uses LASSO to predict the outcome variable (e.g., infant mortality rate). The union of covariates selected in both steps is used in a probit model to estimate propensity scores.

[4]Propensity score weighting substantially lowers the absolute standardized mean difference (ASMD) between treated and control units, improving covariate balance.

The DR approach involves modeling the propensity score and the conditional expectation function (so-called nuisance parameters). Importantly, the score function used for the DR-DiD estimator satisfies Neyman's orthogonality condition, which allows plugging-in estimates of nuisance parameters using ML methods following the general inference theory of the double/debiased machine-learning framework established in Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey and Robins (2018). We rely on ML methods to obtain estimates of the nuisance parameters using high-dimensional county level features measured prior to the treatment, that is prior to the expansion of Medicaid under the ACA. The propensity scores are estimated using LASSO, while the outcome regression is modeled using random forest (or boosted regression forest) algorithms. In order to further alleviate concerns of overfitting, we employ cross-fitting, where nuisance parameters are estimated using ML models trained on all folds (k^{-} $folds$) except the fold reserved for ATT estimation (k^{th} $fold$). We refer to this approach as ML driven DR-DiD (ML-DR-DiD).

Our results suggest that while the ACA Medicaid expansion did not significantly affect maternal mortality-related outcomes among White women, the pregnancy-associated mortality ratio (PAMR) among Black women decreased significantly by 32 percent a year after the expansion. This reduction is primarily driven by lower pregnancy-related mortality ratios (PRMR) due to the expansion. We find no detectable changes in incidental deaths. Regarding infant health, we find that the reform did not impact infant mortality rates for either Black or White populations. However, the reform led to modest improvements in birthweight and subsequently reduced the prevalence of low/very-low birthweight. For instance, birthweight increased by 13.049 grams, and the prevalence of low birthweight decreased by 0.4 percentage points, respectively. Event-study results show no significant pre-treatment trends, but reveal improvements in maternal mortality and infant health outcomes among Black mothers following the expansion. Overall, the findings underscore the differential impacts of the reform and its potential role in mitigating health disparities in the region.

The study is organized as follows. Section 2 presents a comprehensive literature review and highlights the contributions of this study. We discuss the data in Section 3. Section 4 outlines the empirical strategy, and Section 5 discusses the results. Section 6 concludes the study. A technical online appendix is made available with additional graphs and tables as well as further details about the estimation procedures.

2 Literature review and contributions

Past studies have examined the effects of the ACA on a wide range of outcomes, including access to healthcare, health and labor market outcomes, crime, education, and marriage (Courtemanche, Marton, Ukert, Yelowitz and Zapata, 2017; Frean, Gruber and Sommers, 2017; Hampton and Lenhart, 2019; Jung and Shrestha, 2018; Miller, Johnson and Wherry, 2021; Peng, Guo and

Meyerhoefer, 2020; Wen, Hockenberry and Cummings, 2017). Frean, Gruber and Sommers (2017) estimate that 60 percent of insurance gains linked to the ACA are driven by the expansion of Medicaid, underscoring Medicaid's importance.

Research examining the impact of the ACA on health outcomes using self-reported health measures yields mixed results, particularly in the early years following implementation (Cawley, Soni and Simon, 2018; Courtemanche, Marton, Ukert, Yelowitz and Zapata, 2018; Miller and Wherry, 2017). However, studies with longer follow-up periods find improvements in self-reported health, especially among subpopulations that experienced larger gains in insurance coverage (Gruber and Sommers, 2019; Sommers, Maylone, Blendon, Orav and Epstein, 2017).[5]

More recent research has examined the ACA's impact on mortality rates. Miller, Johnson and Wherry (2021) find that the Medicaid expansion reduced mortality in expansion states by approximately 9.4 percent. Borgschulte and Vogler (2020) estimate a decline of 11.36 deaths per 100,000 individuals aged 20–64, corresponding to a 3.6 percent reduction. Similarly, Wyse and Meyer (2025) report that the Medicaid expansion lowered mortality among low-income adults by 2.5 percent.

Prior studies on Medicaid's impact on fertility largely focus on pre-ACA expansions that primarily targeted pregnant women and show mixed results. Early research, using either natality data or data from the Consumer Population Survey, finds that Medicaid expansions between 1980–1990 increased birth rates by 3-5 percent (Baughman, 2001; Joyce, Kaestner and Kwan, 1998; Yelowitz, 1995). In contrast, more recent studies such as Zavodny and Bitler (2010) and DeLeire, Lopoo and Simon (2011) exploit state and time variations in Medicaid expansions and report small overall effects on birth rates. These studies, however, only examine expansions affecting pregnant women, overlooking fertility decisions made by women already covered by Medicaid prior to pregnancy. Focusing on the ACA, Abramowitz (2018) finds that the dependent coverage provision reduced the likelihood of giving birth. More recently Palmer (2020*a*) finds that subsidized private health insurance is associated with increased birth rates, while no significant relationship is observed between Medicaid and birth rates. To our knowledge, the impact of ACA-related Medicaid expansion on fertility has yet to be explored more comprehensively.

More recent studies examine the impacts of the ACA Medicaid expansion on perinatal insurance coverage, healthcare utilization, and health outcomes. Bellerose, Collin and Daw (2022) provide a systematic review of this literature, highlighting increases in preconception and postpartum insurance coverage, though the effects on birth outcomes remain mixed. Another group of studies finds no significant impact on overall infant health outcomes (Brown, Moore, Felix, Stewart, Mac Bird, Lowery and Tilford, 2019; Clapp, James, Kaimal, Sommers and Daw, 2019; Margerison, Kaestner, Chen and MacCallum-Bridges, 2021; Palmer, 2020*b*; Wiggins, Karaye and Horney, 2020). Using

[5] For a review of ACA's Medicaid expansion and health outcomes, see (Allen and Sommers, 2019).

data from Oregon, Harvey, Gibbs, Oakley, Luck and Yoon (2021) find that the reform reduced low birthweight rates among multiparous mothers, while Bhatt and Beck-Sagué (2018) associate the expansion with a decline in infant mortality rates. Using data from Oregon, Conversely, Cook and Stype (2021) apply a DiD approach and find no overall effect on infant mortality. Similarly, Chatterji, Glenn, Markowitz and Montez (2022) show little to no evidence that Medicaid expansions influence maternal morbidity. Finally, two studies focusing on the differential impact of the reform on infant health outcomes by racial groups suggest improvements among minority sub-groups (e.g., Blacks and Hispanics) (Clapp et al., 2019; Constantin and Wehby, 2023).

Most of the aforementioned studies focus on linking Medicaid expansions to maternal and infant health outcomes and remain largely agnostic about concerns that expansion and non-expansion states may differ systematically in their pre-treatment characteristics. For instance, many non-expansion states are concentrated in the Deep South. This raises several issues. First, the treatment and control units differ significantly in their pre-treatment characteristics, which can cause differential dynamics in outcomes even in the absence of the ACA implementation. Second, the DiD framework relies on the parallel trends assumption. Large disparities between the expansion and non-expansion groups in the pre-treatment period undermine the credibility of this assumption, thus weakening the basis for causal identification. Third, other ACA-related reforms, such as private health insurance subsidies and employment mandates, were rolled out concurrently starting in 2014. These reforms affected both expansion and non-expansion states. Differences in baseline characteristics likely influenced the effects of these auxiliary reforms across the treatment and control units. Hence, the effects of those reforms can potentially be incorrectly attributed to the ACA Medicaid expansion. Addressing these issues is essential for drawing credible causal inferences.

We improve upon the existing empirical approach in two important dimensions. First, we address the concern of systematic differences between expansion and non-expansion states by combining the DiD framework with propensity score weighting. This method reweights the control units using estimated propensity scores to better balance the treatment and control groups (Imbens, 2004). Second, and more critically, we flexibly account for pre-treatment covariates using the ML-DR-DiD approach. This combines the doubly robust difference-in-differences (DRDiD) framework presented in Sant'Anna and Zhao (2020) with the double/debiased machine learning (DML) approach of Chernozhukov et al. (2018). To our knowledge, this is the first study to approach the estimation of ATT by embedding the DRDiD estimator within the ML framework. Although the literature has made impressive headway on the theoretical aspects of DiD as well as DML (Chang, 2020; Chernozhukov et al., 2018; Sant'Anna and Zhao, 2020), the empirical application of this approach is still in its infancy.

As previously noted, the unconditional parallel trends assumption is likely violated for health-related outcomes. To address this, our empirical strategy relies on a weaker assumption—conditional parallel trends. In other words, we assume that, absent Medicaid expansion, health outcomes would have followed similar trajectories in areas with comparable baseline characteristics. Our analysis focuses on estimating year-specific average treatment effects on the treated (ATTs) relative to the treatment year (2013), following an event-study framework as outlined in Callaway and Sant'Anna (2021).

Additionally, we focus specifically on the American South and examine outcomes separately for non-Hispanic White and Black populations. This regional and demographic emphasis enhances the policy relevance of our findings, particularly given the current context in which most Southern states continue to resist Medicaid expansion for adults—despite persistently high uninsured rates and worse health outcomes compared to other regions.

3 Data

In this study we primarily use data from the National Vital Statistics System (NVSS, years 2010–2017) to extract the outcome variables. In addition, we use data from several different sources to estimate auxiliary parameters such as propensity scores and coefficients of outcome regression models. In addition, we also use data from Small Area Health Insurance Estimates (SAHIE, years 2010–2017) and the American Community Survey (ACS, years 2014–2017). [6]

3.1 Dependent variables

We are interested in four primary outcome variables, all aggregated at the county level: (*i*) the uninsured rate, (*ii*) pregnancy-related mortality ratio (PRMR), (*iii*) the infant mortality rate (IMR), (*iv*) infant natality measures based on birth weight, and (*v*) the fertility rate. The uninsured rate is the "first-stage" outcome variable from the Medicaid expansion. Outcome variables (*ii*)–(*v*) are "second-stage" outcomes that may be influenced by a prospective mother's insurance status.

The county-level uninsured rate data is obtained from the SAHIE. While SAHIE provides detailed estimates of uninsured rates by income groups relative to the Federal Poverty Level (FPL), it does not disaggregate these rates by race. Consequently, race-specific analyses of the uninsured rates are not possible. We therefore only use the uninsured rate to evaluate the *first-stage*.

The primary data source for both pregnancy-associated mortality and infant mortality are the mortality files from NVSS (mortality waves from 2010–2017). The pregnancy-associated mortality ratio (PARM) is defined as the number of deaths during pregnancy or within 1 year of the end of pregnancy per 1,000 live births. PARM differs from pregnancy-related mortality (PRM), which the Centeres for Disease Control and Prevention (CDC) has defined as deaths during pregnancy or

[6] ACS data include information on income, which we use to estimate the probability that individuals are near the Medicaid eligibility threshold. This estimated probability is then used to construct weights. A detailed discussion is provided in Section 4.

within 1 year of the end of pregnancy from "any cause related to or aggravated by the pregnancy" based on ICD-10 Codes A34 and O00-099. The PRM measure exempts deaths due to incidental causes such as suicide, violence, homicide, accidents, and overdose. The PARM measure is wider in the sense that it encompasses pregnancy-related deaths according to the definition of the CDC plus incidental deaths. This broader measure captures the full scope of maternal risk during and after pregnancy, thereby providing a comprehensive understanding of the potential impact of the Medicaid expansion.

Next, the infant mortality rate is defined as the number of infant deaths per 1,000 live births in a county. Following Borgschulte and Vogler (2020) we also distinguish between deaths due to "amenable" health conditions using ICD-10 codes from the CDC which is included in the mortality data and "non-amenable" health conditions. Amenable conditions are health conditions that are more affected by the availability and accessibility of healthcare than non-amenable conditions.[7]

In addition, we use infant natality measures connected to the birthweight of newborns. This information is available in the natality files of the National Vital Statistics System (NVSS Natality waves from 2010–2017) that records birth certificates. Birth certificates have been shown to record highly reliable information about birthweight outcomes (Northam and Knapp, 2006; Roohan, Josberger, Acar, Dabir, Feder and Gagliano, 2003). The first measure of interest is the average birth weight at the county level; the second and third measures are the fraction of infants with low birth weight of ($< 2,500$ *grams*) and very low birth weight ($< 1,500$ *grams*) at the county level. Birthweight itself is a strong predictor of adult well-being (Black, Devereux and Salvanes, 2007; Currie and Moretti, 2007). Although some states updated their birth certificate files between 2010–2017 to align with the 2003 Version of the U.S. Standard Birth Certificate, these updates did not affect the consistency of outcomes reported across both the older and newer versions.[8] Unfortunately, this prevents us from analyzing the impact of the reform on maternal morbidity as presented in Chatterji et al. (2022) since the morbidity-related variables such as maternal transfusion, perineal laceration, ruptured uterus, unplanned hysterectomy, intensive care admission, and unplanned operation were only added in the 2003 revision.

Although the NVSS provides rich demographic information on mothers, it does not include income, making it difficult to identify individuals likely affected by Medicaid expansion. To address this, we apply machine learning methods to estimate the probability of Medicaid eligibility. Using ACS data as the training sample, we focus on women aged 15–44 and train several machine learning models[9] to predict whether an individual would be eligible for Medicaid under the 138 percent FPL threshold introduced by the ACA expansion.

[7]Non-amenable conditions include accidents or sudden infant death syndrome (SIDS).

[8]Among the Southern states, Alabama, Arkansas, and West Virginia adopted the 2003 Revised Birth Certificate after 2013.

[9]These include Boosted Trees, Random Forests, Decision Trees, and Elastic Nets. See Appendix B for details.

We then use these trained models to predict Medicaid eligibility probabilities for women in the NVSS natality data. These probabilities are subsequently aggregated at the county level by race (White and Black), yielding an estimated share of NVSS observations likely affected by the expansion.

3.2 Variable of interest and control variables

Our primary variable of interest is an indicator variable for whether a county is in a state that expanded Medicaid under the ACA in 2014. Figure 2 shows maps according to treatment status. The top panel shows the map of the entire United States, while the bottom panel focuses on the American South.

Among the 16 Southern states, 5 (Arkansas, Maryland, Delaware, Kentucky, and West Virginia) expanded Medicaid in 2014. County observations from these states constitute the treatment group. County observations from states that did not expand Medicaid during this time form the control group. Louisiana, which implemented the expansion of Medicaid in July 2016, is excluded from the main analysis in order to maintain temporal consistency in treatment assignment and to avoid *forbidden comparisons* due to the staggered nature of the treatment.

Since our analysis is conducted at the county level, we augment our data with variables measuring important county-level features from different sources. These measures are particularly helpful for the ML components of the analysis, where detailed covariate measures are employed to improve prediction performance. These feature variables can be divided into five broad groups:

- **Demographic**: proportion of Whites in 2010; proportion of Blacks in 2010; population fertility among Blacks in 2013; population fertility among Whites in 2013; share of foreigners in 2009; and fraction of births by teenage mothers in 2013.
- **Socioeconomic**: per capita income in 2010; household median income in 2010; unemployment rate in 2013; number unemployed and labor force participants in 2013; percent high school and percent college educated in 2010; a list of race-specific variables divided by Black and White race groups for the year 2010: a) employment rate for Whites and Blacks, b) fraction of Whites and Blacks with college degrees, c) poverty rates for Whites and Blacks, d) fraction of females in a single-parent household by race, and e) household median income by race.
- **Geo-spatial/environmental**: longitude and latitude of county centroid; long-term average precipitation and temperature (1950-2000); elevation; PM 2.5 measure in 2010; cotton suitability measure; and county area.

- **Political**: indicator for whether a county is a Democrat county defined as whether a Democrat candidate won in the 2012 presidential election; proportion of votes for Obama in the 2008 and 2012 presidential elections; proportion of votes for Trump in the 2016 presidential election; and proportion of Whites voting for Obama in the 2008 presidential election.

- **Historical**: proportion enslaved in 1860; proportion of Blacks who were free in 1860; proportion of votes for Democrats in the 1912 presidential election; whether a county was close to water ways or rail road in 1860; cotton, tobacco, sugar and rice as a fraction of agricultural output; malaria stability index; and aggregated Hill Burton Fund expenses (years 1947-1971).

Note that all feature variables represent measures taken prior to the implementation of the Medicaid expansion in 2014 except for votes attributed to Trump in the 2016 presidential election. The data sources for these variables are listed in Table 2 along with the descriptive measures to maintain brevity.

4 Methods

Our estimation approach compares outcome measures of counties in expansion states with their non-expansion counterparts. In the first stage we use data from SAHIE to evaluate the impact of the Medicaid expansion on the county level fraction of uninsured individuals (i.e., uninsured rate) in the American South. In stage two we estimate the effect of the ACA Medicaid expansion on health-related outcome variables. We employ two estimation frameworks: i) parametric DiD with event study estimations; and ii) ML driven Doubly Robust DiD (ML-DR-DiD). We next describe these methods in more detail.

4.1 Parametric Difference-in-Differences

The event-study specification is written as:

$$Y_{ctr} = \alpha + \beta_j \sum_{j=-6}^{5} Medicaid_c \times I(t - \tilde{t} = j) + \delta \mathbf{X}_{ct} \times Post_t + \sigma_c + \gamma_t + \varepsilon_{ctr}, \quad (1)$$

where Y_{ctr} is the dependent variable (e.g., uninsured rate, pregnancy-associated mortality ratio, infant mortality rate, birthweight, incidence of low birthweight, fertility rate) pertaining to county c in year t and race group r (Black versus White Southerners). The indicator variable $Medicaid_c$ takes the value "1" if county c belongs to a state expanding Medicaid in 2014 via the ACA and "0" otherwise. The indicator $I(t - \tilde{t})$ turns on when the difference in birth year t and the expansion year $\tilde{t}$ ($\tilde{t} = 2014$) is equal to j. The omitted category is the year prior to the expansion year. The baseline specification includes county fixed effects σ_c and year fixed effects γ_t. The analysis is restricted to county/year observations from the 16 Southern states and the model is estimated separately for White and Black race groups. We present standard errors clustered at the state level to account for potential intra-state correlations.

As previously mentioned, a key challenge in comparing outcomes between units in expansion and non-expansion states are the inherent differences between the counties from these two groups, largely because the decision to expand Medicaid was left to the discretion of the states. For instance, none of the states in the Deep South opted to expand Medicaid.[10] Substantial differences in pre-treatment characteristics between expansion and non-expansion units undermine the credibility of the conditional parallel trends assumption that is key to identification in DiD. In other words, the trends in outcomes across these groups could have diverged even without the reform due to differences in baseline characteristics, calling into question the validity of DiD. Additionally, other ACA elements such as private insurance subsidies that have affected both groups, may have had differential effects on the trends of the outcome variables of the two groups. Therefore, improving baseline comparability between treatment and control units is essential.

Our first empirical approach addresses these concerns by adjusting the control units using propensity scores. More specifically, the specification estimates equation 1 using weighted least squares (WLS), where the "balancing weights", Ω_{crt}, are given as:

$$\Omega_{crt} = \omega_{crt} \times [W_c + (1 - W_c) \times \frac{\hat{e}(X_c)}{1 - \hat{e}(X_c)}], \tag{2}$$

where W_c is the treatment indicator and $\hat{e}(X_c)$ is the propensity score estimate of receiving the treatment. The weighting factor, ω_{crt}, adjusts the female population of child-bearing age scaled by the share of mothers/births eligible for ACA-Medicaid in county c, race group r, and birth-year t.

$$\omega_{crt} = female\ population\ (15 - 44)_{crt} \times \underbrace{share\ eligible_{crt}}_{\text{ML predictions on NVSS}} . \tag{3}$$

The share of mothers/births eligible for Medicaid is constructed from predictions on the NVSS data using ML models trained on the ACS sample.[11] This inverse probability weighting (IPW) scheme assigns a greater weight to control units with higher propensity scores that are likely similar to treated units, while reducing the weight of units that are highly unlikely to receive the treatment. By adjusting weights of control units to better match the treated units, this approach aims to improve balance in covariates across the two groups (Imbens, 2004).

Next, the specification incorporates important time invariant pre-treatment county-level charac-

[10] Alabama, Georgia, Mississippi, and South Carolina still have not expanded Medicaid. Louisiana expanded Medicaid in July 2016. These states are often characterized by their historical, cultural, and political ties to the Old South.

[11] The prediction approach is discussed in more detail in Appendix B.

teristics, X_{cs}, that are interacted with the post-treatment indicator. This approach uses a weighted version of the conditional parallel trend assumption. The covariates are selected using the Double LASSO method (Belloni, Chernozhukov and Hansen, 2014; Urminsky, Hansen and Chernozhukov, 2016), which involves two steps: *i*) fitting a LASSO regression that predicts the treatment indicator; and *ii*) fitting a separate LASSO regression that predicts the outcome variable. We then include the union of the selected covariates from each LASSO procedure into the event study and DiD regression frameworks, respectively. This data-driven approach ensures that the covariate selection is systematic rather than relying on an ad hoc variable selection approach chosen by the researcher.

The β_j coefficients for $j \geq 0$ measure the average change in outcome for the expansion counties compared to non-expansion counties in year j following the expansion year ($\tilde{t}$) relative to the year prior to the expansion year conditional on covariates. The validity of the results is based on the assumption that in the absence of the ACA Medicaid expansion, the trend in outcomes across counties in expansion vs. non-expansion states would not be systematically different. Although this assumption cannot be realistically tested, the coefficients β_j for $j \in \{-4, \ldots, -2\}$ provide suggestive evidence for the validity of the parallel trends assumption. If there are no changes in trend of outcomes between expansion vs. non-expansion counties prior to the expansion year, we expect estimates of β_j for $j \leq -1$ to be close to zero. Given this, the conventional viewpoint is to *assume* that trends would not systematically differ between these two groups in absence of the treatment.

Additionally, we present point estimates of the reform using the DiD model, which effectively averages the event-study estimates in the post-reform era. The DiD specification can be written as:

$$Y_{ctr} = \alpha + \beta Medicaid_c \times Post_t + \delta \mathbf{X}_{ct} \times Post_t + \sigma_c + \gamma_t + \varepsilon_{ctr}, \tag{4}$$

where $\hat{\beta}$ is the DiD estimate, representing the ATT estimate, given that the conditional parallel trend assumption is satisfied. A weighted version of specification 4 is estimated using similar weights to the ones used in estimating specification 1.

Propensity score estimation. The propensity score is estimated to predict a county's Medicaid expansion status based on a comprehensive set of pre-treatment features, both historical and contemporary, for all 3,000+ counties. Given the high dimensional setting, we implement a double LASSO procedure to select the relevant set of variables (Belloni, Chernozhukov and Hansen, 2014; Urminsky, Hansen and Chernozhukov, 2016) for the propensity score model. In the first step, the LASSO procedure selects covariates from a large list of county-level feature variables that explain the probability of expanding Medicaid. Second, we fit a LASSO model to predict the independent variable of interest (i.e., change in infant birthweight between 2013 and 2014). We then use the union of variables that the LASSO procedures selected in the two prior steps in a probit model to estimate propensity scores. In order to reduce the risk of overfitting, we implement 10-fold crossfitting during estimation. Specifically, we split the sample into roughly 10 equal folds, $k = 1, 2, , \ldots 10$. While data from the k^{-} fold is used to build the propensity score model, the predicted values are obtained using data from fold k^{th}.

The estimation is conducted at the county level rather than at the state level mainly to leverage richer variation which improves the chance to find comparable treatment and control units. For instance, while expansion and non-expansion states may differ significantly in aggregate measures, counties within these states can still have similar propensity scores. A detailed list of covariates used in the estimation, along with their data sources, is presented in Table 2.

The primary goal of constructing well-calibrated propensity scores is to enhance balance in pre-treatment characteristics between treatment and control units. In order to assess the balance, we first examine the Absolute Deviation of Standardized Means (ADSM) before and after adjustments with propensity scores. The unadjusted difference is simply calculated as the absolute value of the difference in means between the treated and control units divided by the average of the standard deviation: $ASDM_{unadjusted} = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{sd_1^2 + sd_0^2}}$. The adjusted difference, on the other hand, is calculated by weighting X_i in the control group with inverse propensity scores: $X_i.W_i + \frac{X_i.(1-W_i)}{(1-\hat{e}(X_i))}$. Figure 3 shows that the unadjusted differences are highly unbalanced, while the adjusted differences show substantial improvements, reducing the absolute standardized difference substantially. If we trim the sample and remove counties with propensity scores that are not within the common support across expansion and non-expansion states, the difference in adjusted means is not affected as shown in Figure 3. Additionally, in Appendix A, we present the histogram of the propensity scores of expansion versus non-expansion counties, as well as the distributional overlap of unweighted and propensity-score-weighted covariates in Figures A.1, and A.2–A.4, respectively. It is evident that propensity score weighting significantly improves overlap in baseline characteristics between expansion and non-expansion counties.

4.2 ML-Doubly Robust-Difference-in-Differences (ML-DR-DiD)

In the previous section we adopted the conditional parallel trends assumption by including interactions between time-invariant pre-treatment covariates and a post-ACA indicator as controls in the parametric DiD model specification. However, incorporating covariates in this manner may not yield valid estimates if there is heterogeneity in the ATT along the covariates as noted in Meyer (1995) and Abadie (2005). For example, one important covariate is the proportion of votes for Obama in the 2012 presidential election. If the ATT varies with support for Obama, the DiD estimate in a model that accounts for Obama's votes can be biased. In this case, the conditional parallel trends assumption requires more flexible modeling of covariates.

To overcome this challenge, we next incorporate ML within the DR-DiD framework. This estimation technique uses the Doubly Robust-DiD (DR-DiD) estimator, presented in Sant'Anna and Zhao (2020), as the "driver engine". The DR-DiD framework combines inverse probability weighting with propensity scores (Abadie, 2005) and outcome regression methods (Heckman, Ichimura and Todd, 1997) to form an estimator that yields valid estimates as long as either the propensity model or the outcome regression model is correctly specified. Both the propensity scores and the outcome regression models—used to adjust for differences in the baseline covariates between the treatment and control units—are carefully estimated using ML methods in the spirit of Chernozhukov et al. (2018).

The DR-DiD estimator for the panel data setting from Sant'Anna and Zhao (2020) using weights (ω) is given as:

$$\tau = E\left[\left(\frac{\omega.D_i}{E(\omega.D_i)} - \frac{\frac{\omega.(1-D_i)p(X_i)}{1-p(X_i)}}{E\left[\frac{\omega.(1-D_i)p(X_i)}{1-p(X_i)}\right]}\right)(Y_{i,t} - Y_{i,b} - E[Y_{i,t} - Y_{i,b}|X_i, D_i = 0])\right], \tag{5}$$

where D_i is the treatment indicator, $p(X_i)$ is the propensity score that depends on the covariates X_i, $Y_{i,t} - Y_{i,b}$ is the difference of the outcome variable at time t and the base period b (referred to as the first difference), and $E[Y_{i,t} - Y_{i,b}|X_i, D_i = 0]$ is the conditional expectation function (CEF) of the outcome differences given X_is and modeled using only the control units (outcome regression). The estimator works in two ways. First, it reweights the observations in the control group such that observations with higher propensity score receive more weight. Second, it attempts to correct the bias on first differencing using regression adjustments, $E[Y_{i,t} - Y_{i,b}|X_i, D_i = 0]$.

Under the assumption of no anticipation, we use year 2013—a year prior to the implementation of Medicaid—as the baseline year. For each survey year (2010–2012 and 2014–2017), the DR-DiD estimator is used to estimate the ATT. This process is similar to the estimation of the group-time ATT estimates in Callaway and Sant'Anna (2021), except that there are only two groups in our analysis: counties in states expanding Medicaid in 2014 and non-expansion counties.

The critical elements of the DR-DiD equation in 5 are the propensity scores and the CEF—the so-called nuisance parameters—which are referred to as the first-step in the estimation of ATT. Notably, the DR-DiD estimator satisfies the Neyman orthogonality condition, which ensures that small estimation errors in the nuisance parameters (e.g., the propensity score model or the outcome regression model) do not introduce additional bias into the ATT estimates. This alleviates regularization bias to some extent as discussed in Chernozhukov et al. (2018) and allows us to use ML methods to incorporate a rich set of covariates while modeling the propensity scores and the outcome regression.

Propensity score estimates are based on LASSO because of its ability to perform feature selection in high-dimensional settings. The outcome regression model is estimated with a random forest regression, a framework that is very flexible and able to capture non-linearities in an intuitive way. In order to avoid overfitting, we use 2-fold cross-fitting following Chernozhukov et al. (2018). More specifically, we split the sample into roughly 2 equal folds, $k = 1, 2$. While the k^{th} fold is used to train the propensity score and the outcome regression models, the predicted values are obtained for k^{-} or the auxiliary fold. The estimation of the ATT is then carried out using only the k^{-} folds. In this way, the estimation process avoids using the same observations for both estimating the nuisance parameters and estimating the ATT.

The algorithm used to estimate the ML-DR-DiD estimator as defined in equation 5 can be summarized as follows:

Algorithm 1 ML-DR-DiD Estimation algorithm

- Input: Data $(Y_t, Y_{b=2013}, X)$, $\underbrace{t}_{t \neq 2013} \in \{2010, .., 2017\}$; number of folds $K = 2$

1. Split the data into K folds consisting of roughly equal observations
2. For each fold k = 1 to K:
   a. Use the k^{-} fold to train propensity score model (ML using all counties in U.S.)
   b. Use the k^{-} fold to train outcome regression model (ML using Southern counties)
   c. Make predictions on the k fold
   d. Use plug-in approach on DR equation to get τ^{k}
3. After looping over 1:K folds, take mean: $\hat{\tau} = \frac{1}{K} \sum_{k=1}^{K} \tau_k$.
4. To account for variability due to sample-splitting, repeat 1-3 l $(l = 99)$ times. Take the median of $\hat{\tau}$s as the ML-DR-DiD estimate.

- Output: DR-DiD estimate

It should be noted that the propensity score model uses all 50 states (including DC) while the outcome regression is only based on the southern states to alleviate geographic disparities in outcomes (e.g., differences in historical institutions, socioeconomic factors). The estimation produces $\widehat{\tau}_r$ specific to the relative time r where $r = \{-3, -2, ..., 4\}$ is based on survey data from 2010–2017, with 2013 as the base year. The post-reform ATT estimates (for $r = 1$ to $r = 4$) are averaged to obtain the aggregated ATT, which serves as the point estimate. Standard errors are computed using the score function.[12]

As an indirect test of the conditional parallel trends assumption, we examine the pre-treatment ATT estimates ($\widehat{\tau}$ for relative times $r < -1$). Estimates close to zero provide evidence of no "differential trends" in the pre-reform period which supports the conditional parallel trends assumption.

[12] Details regarding the estimation of standard errors are described in Appendix C.

5 Results

5.1 First stage: Effects of Medicaid expansion on uninsured

The first-stage regression results evaluating the effects of the Medicaid expansion on the proportion of the uninsured are based on county-level data from SAHIE for the years 2010–2017. We focus on individuals with incomes up to 138 percent of the federal poverty level (FPL) which is the eligibility threshold established by the ACA-related Medicaid expansion. Panel A in Figure 4 presents the dynamic effects based on the event-study specification in equation 1 whereas the estimates in Panel B are based on the ML-DR-DiD approach described in section 4.2.

The estimates in Panel A are parametrically controlled for the interaction between the post-treatment indicator and pre-treatment time invariant covariates selected with the double LASSO method. In addition, estimates are weighted with propensity score related weights, $\omega_{crt} \times [W_c + (1 - W_c) \times \frac{\hat{e}(X_c)}{1-\hat{e}(X_c)}]$. Panel B accounts for the covariates more flexibly using the ML-DR-DiD approach.

Our findings indicate that the ACA Medicaid expansion leads to significant improvements in access to health insurance for Medicaid eligible individuals with incomes between 0–138 percent of the FPL. On average, the uninsured rate in the U.S. South for the sub-group with income below 138 percent of the FPL decreases by about 8 percentage points in Panel A of Figure 4. Moreover, the pre-expansion coefficients are very close to zero. The estimates from the ML-DR-DiD framework in Panel B of Figure 4 are almost identical to those in Panel A.

The ACA-Medicaid expansion is widely recognized for improving access to health insurance in the United States (Miller and Wherry, 2019; Simon, Soni and Cawley, 2017). In the context of this literature our findings on the decrease in uninsured rates are somewhat modest. For example, Miller and Wherry (2019) report an increase of 12 percentage points in health insurance coverage by the fourth year post-expansion, while Simon, Soni and Cawley (2017) find a 10-percentage-point reduction in uninsured rates among childless adults. The smaller effects in our study may result from the exclusive focus on the American South, differences in data sources, and variations in model specifications.

We argue that accounting for relevant covariates and ensuring balance between treatment and control units is crucial for the evaluation of the impact of the Medicaid expansions on outcomes. For example, the proportion of White votes for Obama in 2012 emerges as an important covariate. Our results suggest that a one-percentage-point increase in White support for Obama is associated with a 0.15 percentage-point reduction in the uninsured rate.[13] Since expansion counties had higher levels of White support for Obama compared to non-expansion counties, it is plausible that the proportion of the uninsured would have declined more significantly in these counties even with-

[13] Results not shown, but available upon request.

out the Medicaid expansion, due to other ACA-related policies. This possibility underscores the importance of addressing potential violations of the parallel trends assumption.[14]

Our findings so far have shown that results based on parametric DiD—when properly specified with covariates and weights—are very similar to those based on the ML-DR-DiD framework. However, this observation should not (and cannot) be generalized to other outcome variables analyzed in this study. The parametric DiD approach, while simpler, relies on the stricter assumption that pre-treatment covariates influence outcome trends in both treatment and control units linearly. The proposed ML-DR-DiD framework, on the other hand, is much better equipped to flexibly account for baseline differences between the treatment and control units. For this reason, we treat the ML-DR-DiD estimates as the preferred estimates, while providing parametric DiD-based estimates alongside them.

5.2 Second stage: Maternal and infant health outcomes

The second stage results highlight how the reform affects maternal and infant health outcomes.

5.2.1 Maternal mortality outcomes

Table 3, Panel A, shows the results for pregnancy-associated mortality ratio (PAMR) at the county level for White women (Columns (1)–(3)) and Black women (Columns (4)–(6)), respectively. Columns (1) and (4) report the baseline parametric DiD estimates. Columns (2) and (5) present results from regression adjustments including the addition of relevant covariates and estimation of the model using balancing weights as depicted in equation 2. Finally, columns (3) and (6) display DiD estimates based on the ML-DR-DiD framework.

Columns (1)–(3) of Table 3 show null-effects of the Medicaid expansion on PAMR for White women. The DiD estimates are close to zero and statistically insignificant at any conventional levels. In contrast, the results for Black mothers reveal evidence of improvements in PAMR measure. While the baseline specification (Column 4) indicates a null effect, the regression-adjusted estimate in Column 5 suggests that the Medicaid expansion is associated with a reduction in PAMR among Black women of 0.54 pregnancy-associated deaths per 1,000 women of childbearing age (15–44).

The DiD estimate based on ML-DR-DiD (Column 6) are smaller than the result shown in Column 5, indicating a reduction of 0.36 deaths per 1,000 women of childbearing age. Taken at face value, this suggests a reduction in PAMR by about 32 percent following the reform, using the baseline measure of 1.189 deaths per 1,000 women of childbearing age in 2013. In summary, these

[14] Insurance outcomes in expansion and non-expansion states may have followed parallel trends absent the entire ACA. However, it is less plausible to assume parallel trends in the absence of the Medicaid expansion alone, while other ACA policies—such as individual subsidies and employer mandates—remained, due to potential heterogeneity in the effects of these auxiliary reforms.

findings suggest that while the Medicaid expansion has no significant impact on PAMR of White women, it has a significant effect in reducing pregnancy-associated deaths among Black women in the American South.

Panels B and C further categorizes PAMR into two components: pregnancy-related mortality ratio and maternal death rates due to incidental or accidental causes. The findings in Panel B pertaining to White mothers in Columns (1)–(3) show null-effects consistent with the overall PARM results in Panel A. In contrast, the estimates for Black mothers presented in Columns (4)–(6) in Panel B show that the Medicaid reform is associated with a reduction in PRMR. The baseline DiD estimate in Column 4 is negative but not statistically significant at the conventional levels. Including relevant controls and weighting increases the magnitude of the estimate and improves precision as shown in Column 5. The ATT estimate from ML-DR-DiD in Panel 6 is smaller than the estimate obtained from parametric DiD with regression adjustments in Column 5. Yet the estimate points to improvements in PRMR which decreases by 0.29 pregnancy-related deaths per 1,000 live births following the reform and is statistically significant at the 1 percent level.

In contrast, Panel C shows no evidence that the expansion of Medicaid affects maternal deaths due to incidental or accidental causes across either racial group. The null findings for incidental deaths are perhaps unsurprising, as these outcomes—such as those due to accidents, violence, or self-harm——are generally less affected by access to healthcare and are often beyond the scope of healthcare interventions. Overall, this indicates that the improvements in PAMR observed in Panel A are primarily driven by declines in pregnancy-related deaths among Black mothers.

Figure 5 presents the event-study results for maternal mortality. The estimates for White mothers are shown in the top row while the bottom row shows the estimates for Black mothers. Panels A, B, C and D display estimates based on the event-study model specification in equation 1, while Panels E–H display event-study-type estimates based on the ML-DR-DiD approach.

The top panels indicate that for White mothers the relative-time coefficient estimates remain consistently close to zero across both pre- and post-reform periods. These findings underscore that the ACA-Medicaid expansion has no measurable effect on their PARM and PRMR measures. In contrast, the results for Black mothers reveal a different pattern. Panel C shows that pre-expansion estimates are largely centered around zero, with the exception of a deviation in the fourth year prior to the reform (year 2010). Following the expansion, estimates indicate a significant reduction in PAMR. The pattern in Panel D, using PRMR as the outcome, mimics the pattern from Panel C. Similarly, the findings based on the ML-DR-DiD estimator in Panel D demonstrate stable pre-expansion estimates at zero, particularly near the reform period (years 2011 and 2012). The estimates then shift downwards post-reform. These findings offer compelling evidence that PAMR trends were not systematically different between the adjusted control and treatment groups before the policy implementation. Crucially, they highlight a meaningful improvement in pregnancy-associated deaths for Black women in states that expanded Medicaid. Panel H shows a similar pattern when using

PRMR as the outcome variable. This highlights that the improvements in pregnancy-associated deaths among Blacks are driven by the reduction in pregnancy-related mortality rates following the reform.

5.2.2 Infant outcomes

We next briefly consider how the finding that Medicaid expansion improves maternal health outcomes for Black mothers may affect estimates of infant health outcomes. Taken at face value, the results from the previous section suggest that in the absence of Medicaid expansion, some marginal mothers—those at high risk of mortality—might not have survived pregnancy or childbirth. If Medicaid helps these high-risk mothers survive, their infants may still face elevated health risks, potentially resulting in worse observed infant outcomes among the treated group. In this way, the expansion may introduce a form of selection bias, as the composition of births changes due to the reform.

Infant mortality. Panel A in Table 4 presents findings of White infant mortality rates (IMR) and the corresponding results for Blacks are shown in Panel B. Columns (1)–(3) present results for all causes mortality, whereas columns (4)–(6) show results for amenable mortality. The coefficient of interest is the interaction term $Expansion \times Post$, which measures the ATT effect of the ACA Medicaid expansion. In Column (1) of Panel A in Table 4 we find a null effect of the Medicaid expansion on the IMR using the baseline model which controls only for year and county fixed effects but does not include controls for county characteristics. Column (2) controls for relevant county characteristics—chosen with the double LASSO method—and estimates a WLS regression with weights described in equation 2.

The DiD coefficient estimate based on the described regression adjustments in Column (2) is not statistically different from zero. Column (3) presents the aggregated estimate based on the ML-DR-DiD approach which more flexibly accounts for pre-treatment controls and does not impose linearity. The ML-DR-DiD estimate is positive but not statistically significant at the conventional levels. The findings in Columns (4)–(6) pertain to amenable mortality and are structured similarly to Columns (1)–(3). Similar to all-causes mortality, these findings also provide null results. Furthermore, we find no effects for both all-causes and amenable-mortality among Black infants. The results presented in Panel B consistently show null effects across all specifications, including the ML-DR-DiD approach. In summary, the findings show that the Medicaid expansion did not affect the infant mortality rates of infants born to either White or Black mothers. The relative time coefficient estimates of the related event-study frameworks for child mortality outcomes provide similar evidence that there are no short or long-term discernible effects of the ACA-Medicaid expansion on infant mortality rates among White or Black infants.[15]

Overall, the finding that the ACA Medicaid expansion does not affect infant mortality in the American South aligns with findings in the literature. However, infant mortality is arguably an extreme outcome and while the ACA-Medicaid expansion may not have reduced the infant mortality rate, the reform may still contribute to improving infant health.

Infant health (Birthweight) outcomes. Next, we investigate the effects of the ACA-Medicaid expansion on arguably one of the most used infant health measures—birthweight. We focus on three birthweight-related measures: *i*) birthweight, *ii*) the prevalence of low birthweight ($< 2,500\ grams$), and *iii*) very low birthweight ($< 1,500\ grams$).

Table 5 shows the findings for infants born to White and Black mothers in Panel A and B, respectively. Columns (1)–(3), (4)–(6), and (7)–(9) show findings for birthweight, the fraction of infants with low birthweight, and the fraction with very low birthweight as the outcome variables, respectively. The coefficient estimates of columns (1), (4), and (7) are based on a regression framework that controls for county and year fixed effects but does not include any additional county-level controls. Columns (2), (5), and (8) add county-level characteristics that we selected with the double LASSO method. We use WLS using using *balancing weights* described in equation 2. Finally, columns (3), (6), and (9) show coefficient estimates based on the ML-DR-DiD approach.

The estimates in the top panel of Table 5 indicate that the ACA Medicaid expansion is not associated with improvements in birthweight among infants born to White mothers. The DiD estimate is positive but imprecisely estimated. Regression adjustments in Column (2) do not materially affect the magnitude or precision of the estimate. In Column (3), the ML-DR-DiD estimate is smaller in magnitude and remains imprecisely estimated. Similarly, the estimates for low birthweight and very low birthweight, shown in Columns (4)-–(6) and (7)-–(9), respectively, are generally close to zero and not statistically significant.[16] Overall, these results suggest that the ACA Medicaid expansion does not affect birthweight-related outcomes among infants born to White mothers.

In contrast, the findings in the bottom panel of Table 5 suggest that the reform improves health outcomes of infants born to Black mothers. The baseline estimate in Column (1) indicates that the reform is associated with an increase in birthweight among infants born to Black mothers by 9.659 grams on average and the estimate is statistically significant at the 10 percent level. The estimation coefficient based on regression adjustments in Column (2) shows a slight increase. The DiD estimate suggests that the ACA-Medicaid expansion is associated with an increase in birthweight among Black infants by 11.33 grams on average and the estimated effect is significant at the 1 percent level. The aggregated ML-DR-DiD estimate, presented in Column (3), is even larger (but

[15] A detailed discussion and the event-study figure for child mortality outcomes is provided in Appendix D.

[16] The exception is the estimate in Column (5), which suggests a reduction in the prevalence of low birthweight. However, the preferred ML-DR-DiD estimate in Column (6) is precisely estimated at zero.

less precisely estimated), suggesting that on average the reform is associated with an increase in birthweight among Black infants by 13.05 grams. This estimate is significant at the 5 percent level.

Columns (4)–(6) in the bottom panel in Table 5 show no detectable impacts of the Medicaid reform on the fraction of births categorized as low birthweight (below 2,500 grams). Although negative, the DiD estimates are imprecisely estimated, including the preferred Ml-DR-DiD estimate shown in Column (6). Finally, the estimates reported across Columns (7)–(9) show that the reform is associated with slight decreases in the prevalence of very low birthweights. The ML-DR-DiD estimate in Column (9) indicates that the Medicaid expansion is associated with a reduction in the prevalence of very low birthweight among Black infants by 0.4 percentage points (significant at the 1 percent level).

We next turn to the event-study models in order to investigate both pre-trends as well as the dynamic effects of the reform. The results for birthweight and low birthweight outcomes across both race groups are shown in Figure 6. The top and bottom rows show findings for infants born to White and Black mothers, respectively. Panels A–D show the coefficient estimates of the event-study specification in equation 1, whereas Panels E–H display event-study-type findings based on the ML-DR-DiD approach. The acronym "BW" refers to birthweight as the outcome and "LBW" refers to the fraction with low birthweight.

Panel A of Figure 6 presents event-study estimates for birthweight among White infants. The relative-time coefficients in the pre-reform period are close to zero and imprecisely estimated, indicating no evidence of differential pre-trends. Post-reform estimates also remain near zero, with the exception of the third year after expansion, where the estimate temporarily increases by approximately 20 grams. However, this effect dissipates in the following year, suggesting no sustained impact on birthweight for White infants. Similarly, the event-study estimates in Panel B, pertaining to low birthweight as the outcome measure, hover around zero in years before the reform; although the findings show a slight dip in trend in estimates following the reform, none are statistically significant at the 10 percent level.

Panel C, pertaining to Black infants, shows improvements in birthweight following the expansion period. The relative time coefficients prior to the expansion are not statistically different from zero, while the estimates increase in magnitude following the reform. With incidence of low birthweight as the outcome measure, Panel D shows that the pre-expansion relative time coefficients fluctuate around zero followed by a drop during the expansion year.

Next, the results from the ML-DR-DiD approach, shown in Panels E and F, generally mirror the findings from the regression-based event-study models in Panels A and B. For instance, Panel E demonstrates that the event-study estimates for White infants are close to zero in the pre-reform period, consistent with the absence of differential pre-trends. Similar to the regression-based event-study model, there is a temporary increase in birthweight in the third year post-reform, followed by a drop in the estimate to near-zero in the subsequent year. Similarly, the event-study-type estimates fluctuate around zero both pre and post-reform in Panel F, when using low birthweight measure as the outcome.

The relative time estimates from the ML-DR-DiD framework for Black infants provide evidence of no pre-trends and show improvements in infant outcomes following the reform. As shown in Panel G, which uses birthweight as the outcome measure, pre-reform estimates are close to zero. However, a noticeable break in the trend occurs following the reform. In the year of the expansion (2014) the average birthweight among Black infants increases by 16 grams. This estimate is statistically significant at the 10 percent level. These improvements persist in subsequent years, pointing to sustained impacts. Similarly, Panel H, which examines low birthweight as the outcome, shows a clear break in trend in estimates following the reform. In 2014, the prevalence of low birthweight declines nearly by a percentage point. Although the estimates show some fluctuation in the later years, the sustained effects are evident.

The improvements in health outcomes among infants born to Black mothers in the American South are modest. For instance, the ML-DR-DiD estimates during the third year following the reform in Figure 6, Panel G, suggest that birthweight increases by 0.66 percent, while the prevalence of low birthweight decreases by 7.33 percent relative to the baseline pre-ACA means $[\frac{ATT=1.1}{baseline=15} \times 100\%]$. While direct comparison across studies should be approached with caution due to methodological and contextual differences (e.g., overall U.S. vs. American South), our findings do partially align with those of Brown et al. (2019). The authors, employing a canonical DiD framework, find that while the Medicaid expansion did not significantly affect birthweight outcomes for White infants, the reform is associated with a 0.53 percentage point reduction in the prevalence of low birthweight among Black infants. However, as noted by the authors, this estimate should be interpreted with caution due to suggestive violation of the parallel trends assumption.

6 Conclusion

Over the past decade, maternal and infant health outcomes in the U.S. have continued to worsen. These adverse outcomes are disproportionately concentrated among the Black population and vary significantly across regions, with the American South bearing the greatest burden.

This study examines the impact of the ACA-Medicaid reform—which expanded coverage to adults with income below 138 percent of the FPL—on maternal and infant health outcomes, with a substantive focus on the American South. While the Medicaid expansion was not explicitly designed to target pregnant women, the reform's emphasis on improving the continuity of insurance coverage and enhancing outreach efforts may have generated positive spillover effects, benefiting not directly targeted groups.

A growing body of research evaluates the impact of the ACA-Medicaid expansion on infant and maternal health outcomes, yielding mixed results. While some studies report null effects (Brown et al., 2019; Chatterji et al., 2022; Clapp et al., 2019; Margerison et al., 2021; Palmer, 2020*b*; Wiggins, Karaye and Horney, 2020,2), others suggest improvements in infant health (Constantin and Wehby, 2023; Cook and Stype, 2021). A handful of studies provide analyses based on racial groups and mostly find that the positive effect of the reform is concentrated among Black and Hispanic minority groups (Clapp et al., 2019; Constantin and Wehby, 2023; Wiggins, Karaye and Horney, 2020). However, even when leveraging similar datasets and canonical DiD approaches, studies have reached divergent conclusions. For instance, Wiggins, Karaye and Horney (2020) find no significant impact on infant mortality among Black populations, while Constantin and Wehby (2023) report improvements for Blacks.

A fundamental challenge which is overlooked in much of the existing literature is the voluntary nature of the ACA-Medicaid adoption. Because of this expansion states differ from non-expansion states in baseline (observed) characteristics. For instance, none of the Deep Southern states chose to expand Medicaid. Hence, trends in outcomes between expansion and non-expansion units may have diverged due to dynamics shaped by both observed and unobserved (baseline) characteristics even in the absence of the reform. This would violate the parallel trends assumption which is crucial for identification in the DiD framework and leads to estimation bias.

To improve the estimation framework, we adopt more suitable empirical strategies with two main objectives: (i) to account for baseline covariates more carefully, and (ii) to improve balance between treated and control units. Our approach proceeds in two steps. First, we use a Double LASSO selection method to identify relevant baseline covariates, then estimate propensity scores and construct balancing weights to enhance comparability across groups within a regression framework. Second, and more importantly, we apply a machine learning–based Doubly Robust Difference-in-Differences (ML-DR-DiD) estimator, which allows for flexible relationships between covariates and outcomes.

This combined approach allows identification under weaker assumptions than the rather strong assumption of conditional parallel trends in the canonical DiD framework. We argue that in the context of Medicaid expansion, it is likely that the assumption of conditional parallel trends is violated.

We find no significant effects of the Medicaid expansion on health outcomes among White populations. Specifically, the expansion did not reduce pregnancy-associated mortality among White mothers, nor did it influence infant mortality rates or birthweight-related outcomes among White infants. In contrast, the results suggest notable improvements among Black populations. The Medicaid expansion significantly reduced pregnancy-associated deaths among Black mothers, largely driven by a decline in pregnancy-related (rather than incidental) causes. While no clear effect on infant mortality is detected among Black infants, the expansion is associated with modest improvements in birthweight, low birthweight, and very low birthweight outcomes.

Although comparisons to prior studies should be made cautiously given methodological and contextual differences, our findings underscore the role of the ACA Medicaid expansion in reducing racial disparities in maternal and infant health outcomes in the American South.

References

Abadie, Alberto. 2005. "Semiparametric difference-in-differences estimators." *The review of economic studies* 72(1):1–19.

Abramowitz, Joelle. 2018. "Planning parenthood: the Affordable Care Act young adult provision and pathways to fertility." *Journal of Population Economics* 31:1097–1123.

Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. "The political legacy of American slavery." *The Journal of Politics* 78(3):621–641.

Allen, Heidi and Benjamin D Sommers. 2019. "Medicaid expansion and health: Assessing the evidence after 5 years." *Jama* 322(13):1253–1254.

Althoff, Lukas and Hugo Reichardt. 2024. "Jim Crow and Black economic progress after slavery." *The Quarterly Journal of Economics* 139(4):2279–2330.

Arias, Elizabeth, Brigham Bastian, Jiaquan Xu and Betzaida Tejada-Vera. 2021. "National vital statistics reports." *National Vital Statistics Reports* 70(1).

Baughman, R. 2001. Medicaid, CHIP and fertility: Does expanding public health insurance encourage births? Technical report Unpublished PhD dissertation Syracuse University.

Becker, Gary S. 1960. An economic analysis of fertility. In *Demographic and Economic Change in Developed Countries*. Columbia University Press pp. 209–240.

Bellerose, Meghan, Lauren Collin and Jamie R Daw. 2022. "The ACA Medicaid Expansion And Perinatal Insurance, Health Care Use, And Health Outcomes: A Systematic Review." *Health Affairs* 41(1):60–68.

Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81(2):608–650.

Bhatt, Chintan B and Consuelo M Beck-Sagué. 2018. "Medicaid expansion and infant mortality in the United States." *American journal of public health* 108(4):565–567.

Black, Sandra E, Paul J Devereux and Kjell G Salvanes. 2007. "From the cradle to the labor market? The effect of birth weight on adult outcomes." *The Quarterly Journal of Economics* 122(1):409–439.

Borgschulte, Mark and Jacob Vogler. 2020. "Did the ACA Medicaid expansion save lives?" *Journal of Health Economics* 72:102333.

Brown, Clare C, Jennifer E Moore, Holly C Felix, M Kathryn Stewart, T Mac Bird, Curtis L

Lowery and J Mick Tilford. 2019. “Association of state Medicaid expansion status with low birth weight and preterm birth.” *JAMA* 321(16):1598–1609.

Callaway, Brantly and Pedro HC Sant'Anna. 2021. “Difference-in-differences with multiple time periods.” *Journal of econometrics* 225(2):200–230.

Cawley, John, Aparna Soni and Kosali Simon. 2018. “Third year of survey data shows continuing benefits of Medicaid expansions for low-income childless adults in the US.” *Journal of general internal medicine* 33(9):1495–1497.

Chang, Neng-Chieh. 2020. “Double/debiased machine learning for difference-in-differences models.” *The Econometrics Journal* 23(2):177–191.

Chatterji, Pinka, Hanna Glenn, Sara Markowitz and Jennifer Karas Montez. 2022. ACA Medicaid Expansions and Maternal Morbidity. Technical report National Bureau of Economic Research.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey and James Robins. 2018. “Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning.” *The Econometrics Journal* 21(1).

Clapp, Mark A, Kaitlyn E James, Anjali J Kaimal, Benjamin D Sommers and Jamie R Daw. 2019. “Association of Medicaid expansion with coverage and access to care for pregnant women.” *Obstetrics & Gynecology* 134(5):1066–1074.

Constantin, Joanne and George L Wehby. 2023. “Effects of recent Medicaid expansions on infant mortality by race and ethnicity.” *American journal of preventive medicine* 64(3):377–384.

Cook, Amanda and Amanda Stype. 2021. “Medicaid expansion and infant mortality: the (questionable) impact of the Affordable Care Act.” *J Epidemiol Community Health* 75(1):10–15.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz and Daniela Zapata. 2017. “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.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz and Daniela Zapata. 2018. “Early effects of the Affordable Care Act on health care access, risky health behaviors, and self-assessed health.” *Southern Economic Journal* 84(3):660–691.

Currie, Janet and Enrico Moretti. 2007. “Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight.” *Journal of Labor economics* 25(2):231–264.

Daw, Jamie R, Laura A Hatfield, Katherine Swartz and Benjamin D Sommers. 2017. “Women in the United States experience high rates of coverage ‘churn’ in months before and after childbirth.” *Health Affairs* 36(4):598–606.

DeLeire, Thomas, Leonard M Lopoo and Kosali I Simon. 2011. “Medicaid expansions and fertility in the United States.” *Demography* 48(2):725–747.

Di, Qian, Itai Kloog, Petros Koutrakis, Alexei Lyapustin, Yujie Wang and Joel Schwartz. 2016. “Assessing PM2. 5 exposures with high spatiotemporal resolution across the continental United States." *Environmental science & technology* 50(9):4712–4721.

Frean, Molly, Jonathan Gruber and Benjamin D Sommers. 2017. "Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act." *Journal of Health Economics* 53:72–86.

Frey, William H. 2021. "Mapping America's diversity with the 2020 census." https://www.brookings.edu/articles/mapping-americas-diversity-with-the-2020-census/.

Gifford, Kathy, Jenna Walls, Usha Ranji, Alina Salganicoff and Ivette Gomez. 2017. "Medicaid Coverage of Pregnancy and Perinatal Benefits: Results from a State Survey." Report, Kaiser Family Foundation(April 27).

Gruber, Jonathan and Benjamin D Sommers. 2019. "The Affordable Care Act's effects on patients, providers, and the economy: what we've learned so far." *Journal of Policy Analysis and Management* 38(4):1028–1052.

Hampton, Matt and Otto Lenhart. 2019. "The effect of the Affordable Care Act preexisting conditions provision on marriage." *Health economics* 28(11):1345–1355.

Harvey, S Marie, Susannah Gibbs, Lisa Oakley, Jeff Luck and Jangho Yoon. 2021. "Medicaid expansion and neonatal outcomes in Oregon." *Journal of Evaluation in Clinical Practice* 27(5):1096–1103.

Heckman, James J, Hidehiko Ichimura and Petra E Todd. 1997. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *The review of economic studies* 64(4):605–654.

Hirai, Ashley H, William M Sappenfield, Michael D Kogan, Wanda D Barfield, David A Goodman, Reem M Ghandour and Michael C Lu. 2014. "Contributors to excess infant mortality in the US South." *American Journal of Preventive Medicine* 46(3):219–227.

Hoyert, Donna L. 2022. "Maternal mortality rates in the United States, 2020.".

Imbens, Guido W. 2004. "Nonparametric estimation of average treatment effects under exogeneity: A review." *Review of Economics and statistics* 86(1):4–29.

Joyce, Theodore, Robert Kaestner and Florence Kwan. 1998. "Is Medicaid pronatalist? The effect of eligibility expansions on abortions and births." *Family Planning Perspectives* pp. 108–127.

Jung, Juergen and Vinish Shrestha. 2018. "The Affordable Care Act and college enrollment decisions." *Economic Inquiry* 56(4):1980–2009.

Kennedy-Moulton, Kate, Sarah Miller, Petra Persson, Maya Rossin-Slater, Laura Wherry and Gloria Aldana. 2022. Maternal and Infant Health Inequality: New Evidence from Linked Administrative Data. Technical report National Bureau of Economic Research.

Leibowitz, Arleen. 1990. "The Response of Births to Changes in Health Care Costs." *The Journal of Human Resources* 25(4):697.

Margerison, Claire E, Robert Kaestner, Jiajia Chen and Colleen MacCallum-Bridges. 2021. "Impacts of Medicaid expansion before conception on prepregnancy health, pregnancy health, and outcomes." *American Journal of Epidemiology* 190(8):1488–1498.

Meyer, Breed D. 1995. "Natural and quasi-experiments in economics." *Journal of business & economic statistics* 13(2):151–161.

Miller, Sarah and Laura R Wherry. 2017. "Health and access to care during the first 2 years of the ACA Medicaid expansions." *New England Journal of Medicine* 376(10):947–956.

Miller, Sarah and Laura R Wherry. 2019. Four years later: insurance coverage and access to care continue to diverge between ACA Medicaid expansion and non-expansion states. In *AEA Papers and Proceedings*. Vol. 109 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 pp. 327–333.

Miller, Sarah, Norman Johnson and Laura R Wherry. 2021. "Medicaid and mortality: new evidence from linked survey and administrative data." *The Quarterly Journal of Economics* .

Northam, Sally and Thomas R Knapp. 2006. "The reliability and validity of birth certificates." *Journal of Obstetric, Gynecologic & Neonatal Nursing* 35(1):3–12.

Olvera, Johabed G, Candis Watts Smith et al. 2023. "The effect of the affordable care act and racial dynamics on federal Medicaid transfers." *Journal of Public Policy* 43(3):533–555.

Palmer, Makayla. 2020*a*. "Does publicly subsidized health insurance affect the birth rate?" *Southern Economic Journal* 87(1):70–121.

Palmer, Makayla. 2020*b*. "Preconception subsidized insurance: Prenatal care and birth outcomes by race/ethnicity." *Health Economics* 29(9):1013–1030.

Peng, Lizhong, Xiaohui (Ronnie) Guo and Chad D. Meyerhoefer. 2020. "The Effects of Medicaid Expansion on Labor Market Outcomes: Evidence from Border Counties." *Health Economics* 29:245–260.

Pollock, Elizabeth A, Keith P Gennuso, Marjory L Givens and David Kindig. 2021. "Trends in infants born at low birthweight and disparities by maternal race and education from 2003 to 2018 in the United States." *BMC Public Health* 21(1):1117.

Roohan, Patrick J, Raina E Josberger, Janice Acar, Poornima Dabir, Harry M Feder and Patricia J Gagliano. 2003. "Validation of birth certificate data in New York State." *Journal of community health* 28:335–346.

RWJF. 2020. 2020 County Health Rankings Key Findings. Technical report University of Wisconsin Population Health Institute and Robert Wood Johnson Foundation.

Sant'Anna, Pedro HC and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of econometrics* 219(1):101–122.

Simon, Kosali, Aparna Soni and John Cawley. 2017. "The impact of health insurance on preventive care and health behaviors: evidence from the first two years of the ACA Medicaid expansions." *Journal of Policy Analysis and Management* 36(2):390–417.

Sommers, Benjamin D, Bethany Maylone, Robert J Blendon, E John Orav and Arnold M Epstein. 2017. “Three-year impacts of the Affordable Care Act: improved medical care and health among low-income adults.” *Health Affairs* 36(6):1119–1128.

Urminsky, Oleg, Christian Hansen and Victor Chernozhukov. 2016. “Using double-lasso regression for principled variable selection.” *Available at SSRN 2733374* .

Valenzuela, Claudia P and Michelle JK Osterman. 2023. “Characteristics of mothers by source of payment for the delivery: United States, 2021.” *National Center for Health Statistics: Data Brief No. 468* .

Valerio, Valerie C, Jordan Downey, Sema K Sgaier, William M Callaghan, Bettina Hammer and Peter Smittenaar. 2023. “Black-White disparities in maternal vulnerability and adverse pregnancy outcomes: An ecological population study in the United States, 2014–2018.” *The Lancet Regional Health–Americas* 20.

Wen, Hefei, Jason M Hockenberry and Janet R Cummings. 2017. “The effect of Medicaid expansion on crime reduction: Evidence from HIFA-waiver expansions.” *Journal of Public Economics* 154:67–94.

Wiggins, Alexandra, Ibraheem M Karaye and Jennifer A Horney. 2020. “Medicaid expansion and infant mortality, revisited: A difference-in-differences analysis.” *Health Services Research* 55(3):393–398.

Wyse, Angela and Bruce D Meyer. 2025. Saved by Medicaid: new evidence on health insurance and mortality from the universe of low-income adults. Technical report National Bureau of Economic Research.

Yelowitz, Aaron S. 1995. “The Medicaid notch, labor supply, and welfare participation: Evidence from eligibility expansions.” *The Quarterly Journal of Economics* 110(4):909–939.

Zavodny, Madeline and Marianne P Bitler. 2010. “The effect of Medicaid eligibility expansions on fertility.” *Social Science & Medicine* 71(5):918–924.

Figures

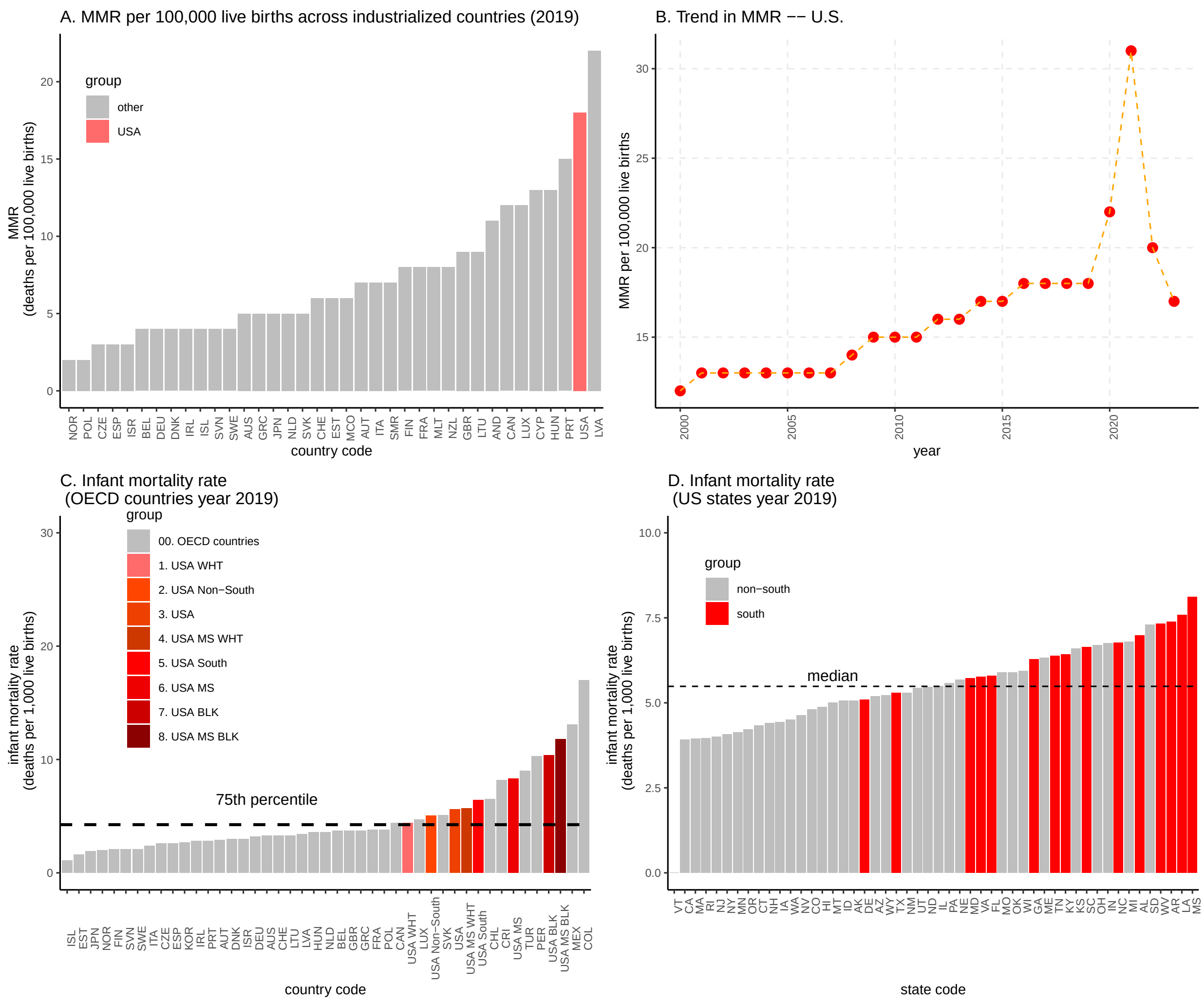


Figure 1: **Descriptive figures of maternal and infant mortality**

Notes: Based on authors' calculations. Data for IMR is from NVSS and maternal mortality is from UNICEF.

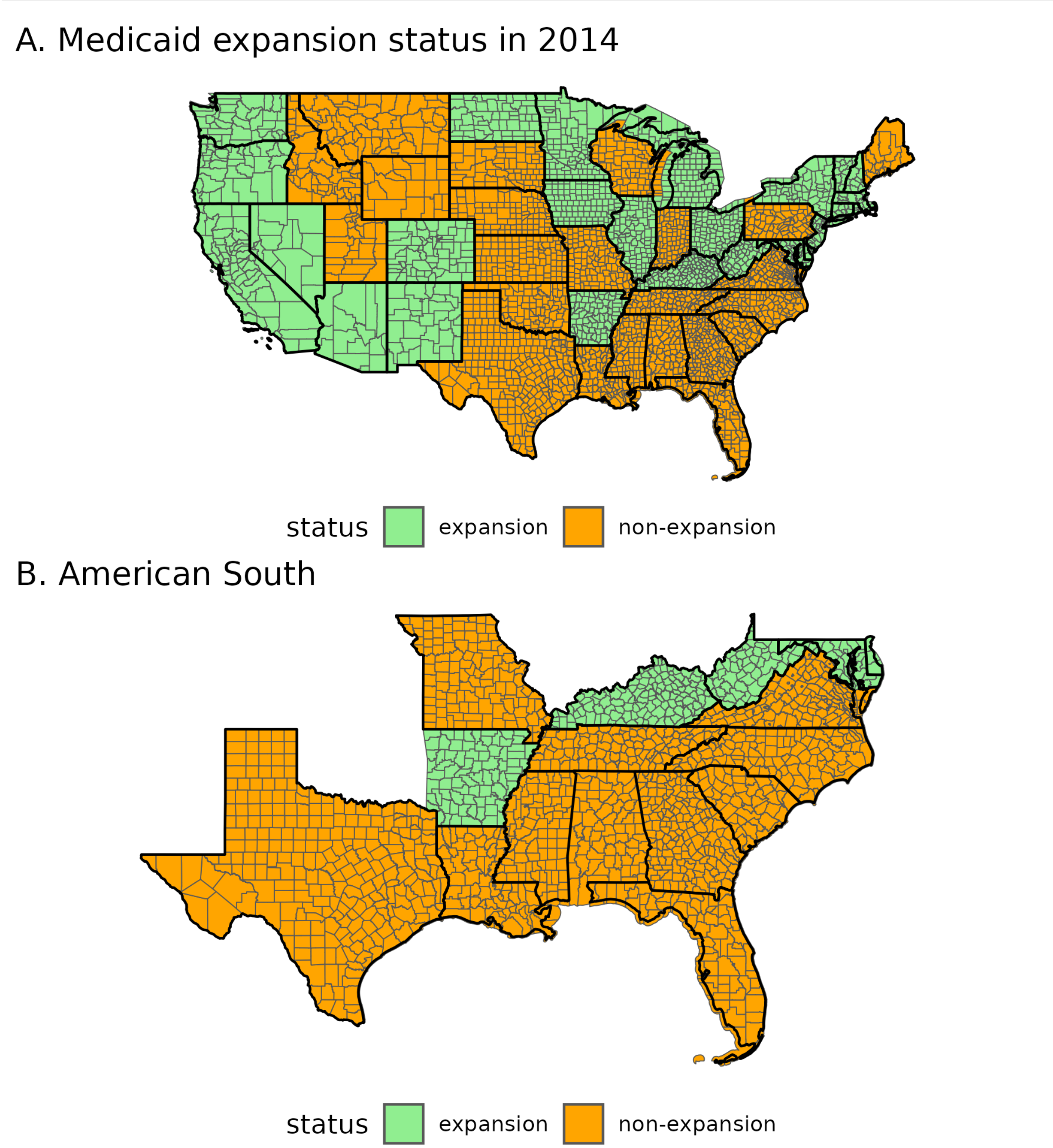


Figure 2: **Medicaid expansion status of states in 2014**

Notes: The map shows the expansion status by states in 2014. Panel B restricts the illustration to Southern states used in the analysis.
Data Source: KFF url: https://www.kff.org/status-of-state-medicaid-expansion-decisions/

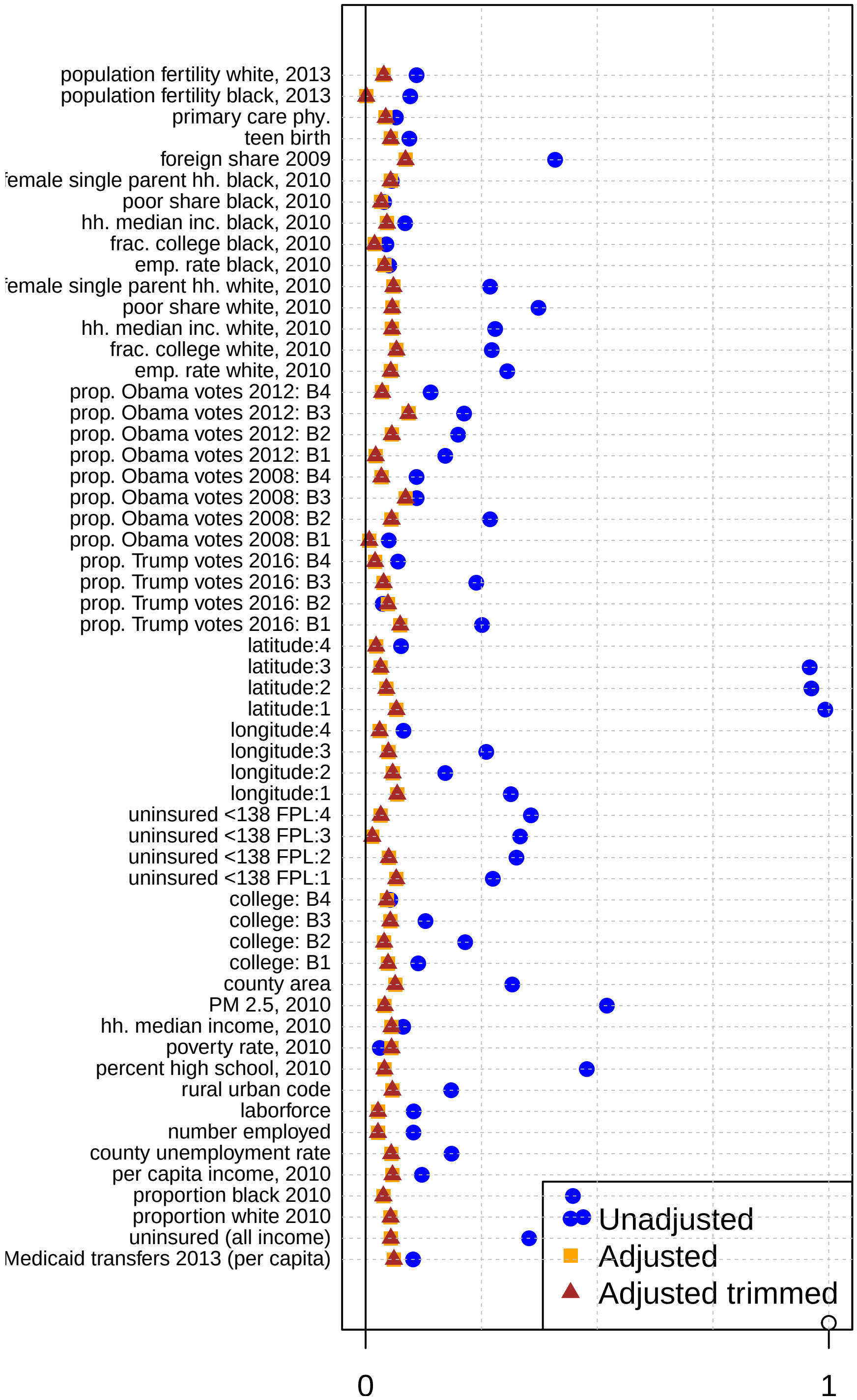


Figure 3: **Absolute Deviation of Standardized Means (ADSM)**

Notes: The graph displays the ADSM for the covariates used in the analysis, focusing on counties in the American South. Unadjusted values are indicated by blue circles. Yellow squares represent the ADSM after applying propensity score weights, while red triangles show the ADSM after additionally trimming counties with propensity scores outside the common support shared by treated and control units. Data source: Variables used in the propensity score model are reported in Table 2.

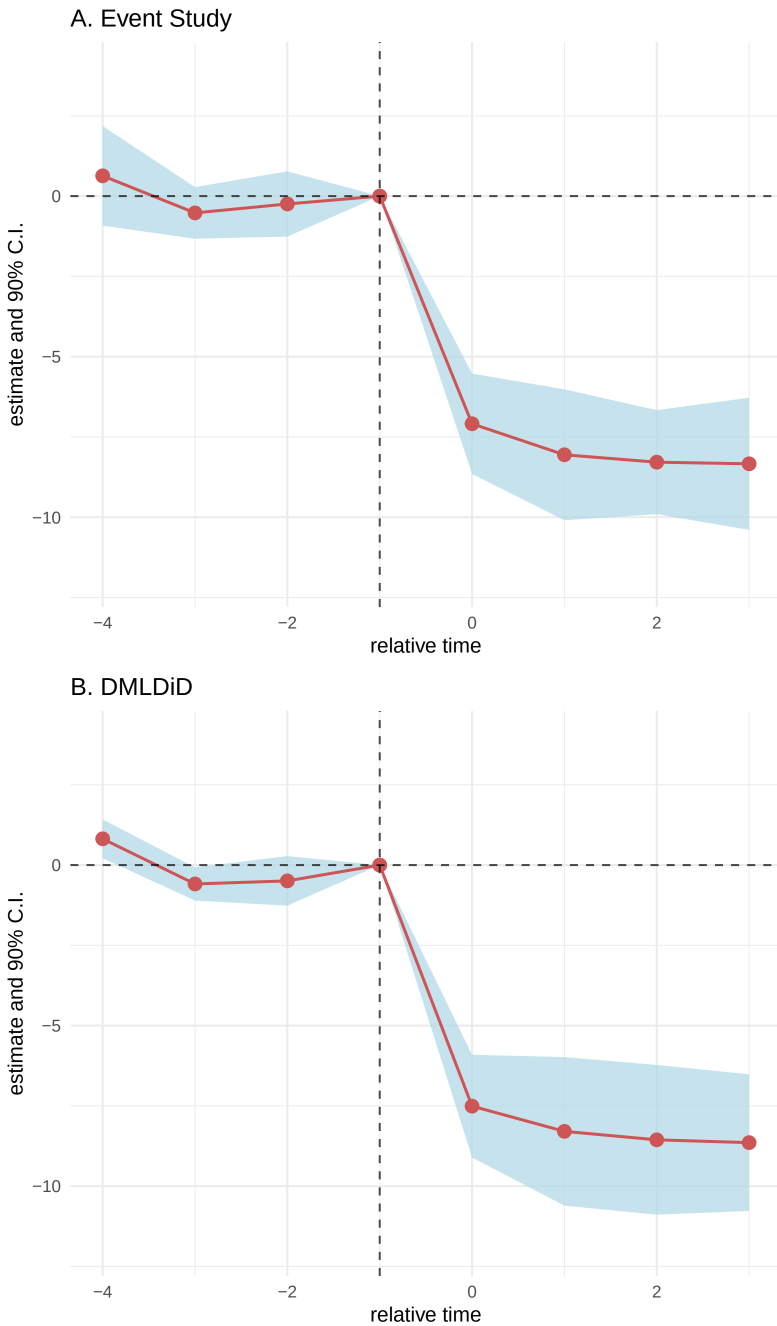


Figure 4: **First stage: Effects of Medicaid expansion on uninsured rates**

Notes: The dependent variable is the **county-level uninsured rate** for individuals with family income below 138 percent of the FPL. Panel A reports event-study estimates based on the event-study model in equation 1 which includes county and year fixed effects. All estimates are based on the weighting scheme defined in equation 2 and adjusted for the interaction between the post-ACA indicator and covariates selected with the double LASSO method. Confidence intervals are derived from state-clustered standard errors. Panel B provides event-study estimates obtained via the ML-DR-DiD framework. Data source: SAHIE 2010–2017 (primary source).

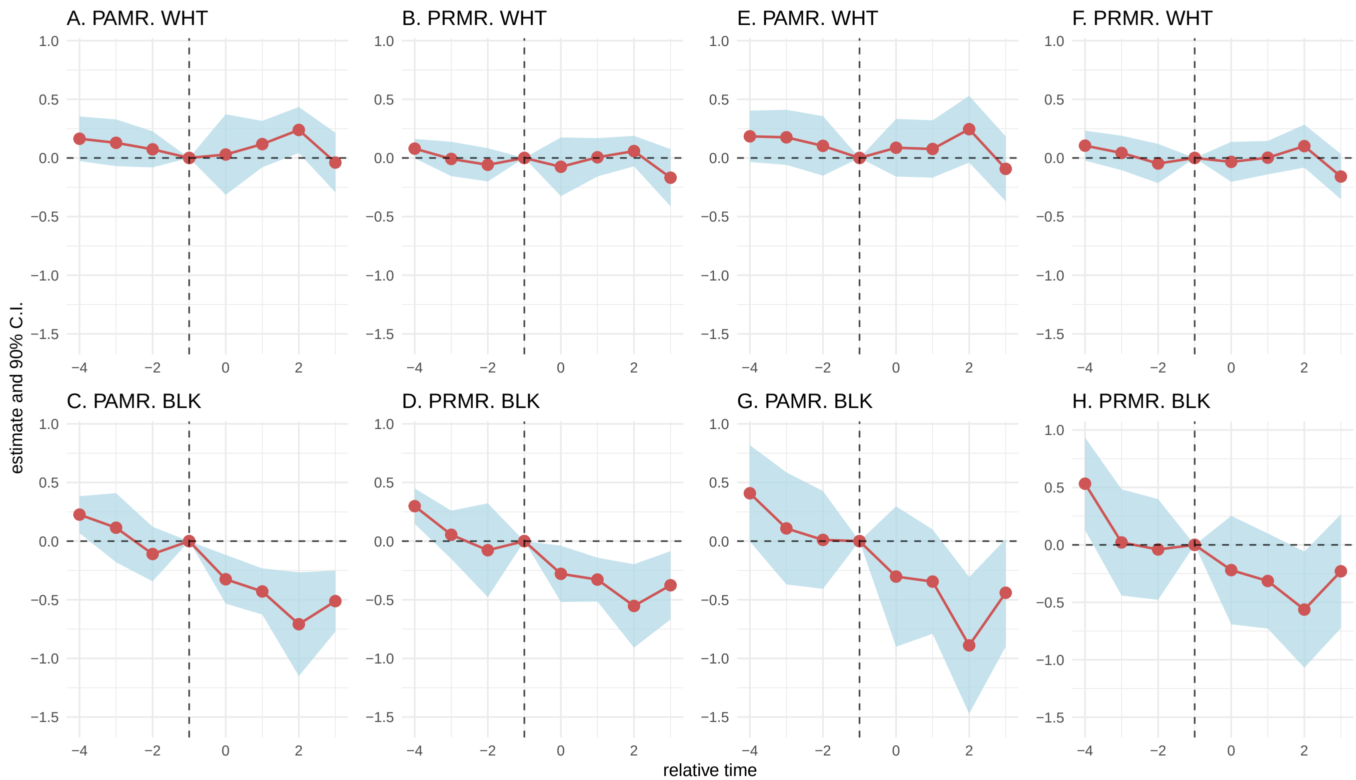


Figure 5: **Effects of Medicaid expansion on maternal mortality rate**

Notes: The dependent variables are the **pregnancy-associated mortality ratio (PAMR)** (Panels A, C, E, and G) and the **pregnancy-related mortality ratio (PRMR)** (Panels B, D, F, and H) per 1,000 women aged 15-44 for White (top row) vs. Black (bottom row) mothers. Panels A–D show results from the regression-based event-study specification in equation 1 which controls for county and year fixed effects. The estimates are weighted with the balancing weights described in equation 2 and include interactions between the Post-ACA indicator and county-level controls selected with the Double LASSO method. Standard errors are clustered at the state level. The results in Panels E–H are optained with the ML-DR-DiD approach. Data sources: NVSS Mortality files 2010–2017.

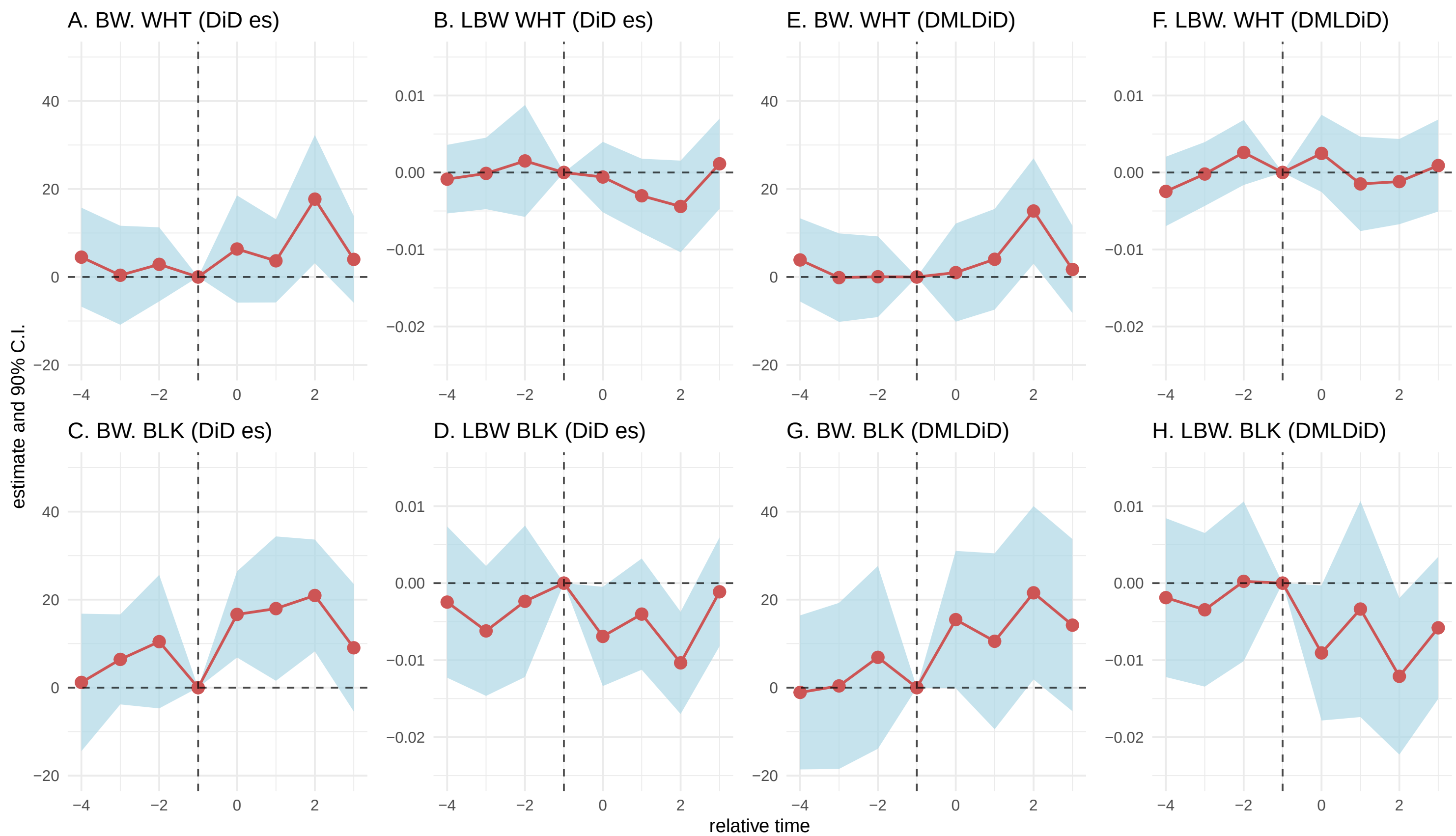


Figure 6: **Effects of Medicaid expansion on infants' health outcomes**

Notes: The county level dependent variables are **average birthweight (BW)** in Panels A, C, E, and G and **fraction of births with low birthweight (<2,500 grams) (LBW)** in Panels B, C, F, and H. The first row shows results for infants born to White mothers, while the second row shows results for infants born to Black mothers. Panels A–D present findings from the regression-based eventy-study design in equation 1 with controls for county and year fixed effects. The estimates are weighted using weights defined in equation 2. All specifications include interactions between the Post-ACA indicator and county-level controls selected with the Double LASSO method. The standard errors are clustered at the state level. Estimates in Panels E–H are based on the ML-DR-DiD framework. Data sources: NVSS Natality files 2010–2017.

Tables

Table 1: **Summary statistics (outcome variables)**

Characteristic	**Expansion**	**Non-expansion**	**p-value**[2]
Infant mortality rate (White)	7.01 (13.14)	5.93 (8.56)	<0.001
Infant mortality rate (Black)	14.70 (27.58)	12.32 (16.89)	0.11
Birthweight (White)	3,200.03 (96.85)	3,219.06 (109.25)	<0.001
Birthweight (Black)	2,986.99 (239.84)	2,967.18 (186.21)	<0.001
Low birthweight (White)	0.09 (0.04)	0.08 (0.05)	<0.001
Low birthweight (Black)	0.15 (0.13)	0.15 (0.10)	<0.001
Very low birthweight (White)	0.01 (0.02)	0.01 (0.02)	<0.001
Very low birthweight (Black)	0.03 (0.07)	0.03 (0.05)	<0.001
moms population (White)	172.71 (294.85)	226.02 (481.85)	0.018
moms population (Black)	107.74 (314.00)	121.95 (356.00)	<0.001
Characteristic	**Expansion**	**Non-expansion**	**p-value**[2]
PAMR per 1,000 live births (White)	0.247 (1.095)	0.361 (1.960)	0.11
PAMR per 1,000 live births (Black)	0.567 (4.205)	0.743 (6.111)	0.2
PRMR per 1,000 live births (White)	0.165 (0.901)	0.249 (1.664)	0.2
PRMR per 1,000 live births (Black)	0.359 (2.128)	0.600 (5.693)	0.5

[1]Mean (SD)
[2]Wilcoxon rank sum test
The Infant Mortality Rate (IMR), Pregnancy-Associated Mortality Rate (PAMR), and Pregnancy-Related Mortality Rate (PRMR) are each calculated per 1,000 live births.

Table 2: **Summary statistics (covariate/feature variables)**

Characteristic	**Expansion** N = 341[1]	**Non-expansion** N = 1,119[1]	**p-value**[2]	**Source**
Uninsured all income	22.84 (4.23)	24.79 (5.76)	<0.001	SAHIE
Medicaid transfers 2013 (per capita)	1.71 (0.74)	1.55 (0.68)	<0.001	BEA
Proportion white 2010	0.85 (0.17)	0.80 (0.19)	<0.001	Census
Proportion black 2010	0.13 (0.16)	0.18 (0.19)	<0.001	Census
Per capita income, 2010	30,834.88 (6,753.28)	31,158.06 (6,913.49)	0.2	Census
Democrat county, 2013	32 / 341 (9.4%)	223 / 1119 (20%)	<0.001	Olvera
County unemployment rate	8.36 (2.29)	7.92 (2.38)	0.003	BLS
Number employed	26,860.16 (58,879.56)	39,645.07 (114,269.94)	0.10	QCEW
Laborforce	28,885.82 (62,897.20)	42,657.58 (122,179.14)	0.11	QCEW
Rural urban code, 2013			0.003	NCHS
1	38 / 341 (11%)	188 / 1119 (17%)		
2	50 / 341 (15%)	155 / 1119 (14%)		
3	45 / 341 (13%)	132 / 1119 (12%)		
4	10 / 341 (2.9%)	72 / 1119 (6.4%)		
5	7 / 341 (2.1%)	24 / 1119 (2.1%)		
6	73 / 341 (21%)	247 / 1119 (22%)		
7	53 / 341 (16%)	112 / 1119 (10%)		
8	24 / 341 (7.0%)	96 / 1119 (8.6%)		
9	41 / 341 (12%)	93 / 1119 (8.3%)		
Percent high school, 2010	38.82 (5.65)	34.45 (6.62)	<0.001	NCES
Percent college, 2010	15.65 (7.57)	17.88 (8.70)	<0.001	NCES
Poverty rate, 2010	19.99 (6.34)	19.39 (6.41)	0.090	SAIPE
HH. median income, 2010	39,104.58 (11,202.60)	40,384.30 (10,703.85)	0.003	SAIPE
Longitude	-86.61 (5.38)	-87.64 (7.44)	0.3	ABS rep.
Latitude	36.10 (2.90)	34.04 (3.02)	<0.001	ABS rep.
Londitude square	7,530.21 (923.90)	7,735.69 (1,329.82)	0.3	ABS rep.
Latitude square	1,311.73 (202.40)	1,167.57 (204.00)	<0.001	ABS rep.
PM 2.5, 2010	10.91 (1.30)	9.98 (1.62)	<0.001	Di rep.
County area	0.13 (0.06)	0.16 (0.11)	<0.001	
Prop. Trump votes 2016: B1	0.12 (0.13)	0.16 (0.17)	0.049	ICPSR
Prop. Trump votes 2016: B2	0.34 (0.12)	0.33 (0.13)	0.13	ICPSR
Prop. Trump votes 2016: B3	0.49 (0.19)	0.44 (0.22)	0.013	ICPSR
Prop. Trump votes 2016: B4	0.04 (0.08)	0.05 (0.10)	0.5	ICPSR
HH. inc age 35, Whites	44,068.37 (6,286.65)	45,515.79 (6,398.51)	<0.001	Atlas
Prop. Obama votes, 2008	0.38 (0.11)	0.39 (0.14)	0.9	ICPSR
Prop. Obama votes 2012: B1	0.43 (0.17)	0.40 (0.21)	0.073	ICPSR
Prop. Obama votes 2012: B2	0.39 (0.11)	0.35 (0.13)	<0.001	ICPSR
Prop. Obama votes 2012: B3	0.14 (0.13)	0.18 (0.17)	0.2	ICPSR
Prop. Obama votes 2012: B4	0.01 (0.08)	0.03 (0.10)	<0.001	ICPSR
Cotton suitability	0.42 (0.16)	0.46 (0.17)	0.001	FAO

continued . . .

continued . . .

Characteristic	**Expansion** N = 341[1]	**Non-expansion** N = 1,119[1]	**p-value**[2]	**Source**
Proportion enslaved, 1860	0.23 (0.22)	0.30 (0.21)	<0.001	Census
White Obama votes	0.29 (0.12)	0.27 (0.13)	<0.001	ABS rep.
Free Blacks, 1860	0.02 (0.04)	0.01 (0.02)	0.005	Census
Cotton fraction	0.00 (0.01)	0.00 (0.00)	<0.001	AR
Tobacco fraction	0.50 (1.24)	0.23 (0.84)	<0.001	AR
Sugar fraction	0.03 (0.13)	0.00 (0.02)	0.14	AR
Rice fraction	0.00 (0.00)	0.00 (0.05)	<0.001	AR
Malaria stability	0.07 (0.04)	0.16 (0.41)	<0.001	AR
Hill Burton Funds (in 1,000 $)	17,964.57 (114,580.24)	14,922.05 (74,890.30)	0.3	HLW
Proportion Democrats, 1912	54.50 (15.91)	69.41 (18.13)	<0.001	ABS rep.
Water ways 1860	201 / 341 (59%)	396 / 1119 (35%)	<0.001	JA
Rail road 1860	58 / 341 (17%)	298 / 1119 (27%)	<0.001	JA
Average precipitation	102.80 (11.93)	95.04 (24.10)	0.001	CRU
Average temperature	14.43 (2.88)	16.25 (2.86)	<0.001	CRU
Elevation	0.21 (0.18)	0.25 (0.26)	0.6	CRU
Emp. rate white, 2010	0.63 (0.09)	0.66 (0.07)	<0.001	Atlas
Frac. college white, 2010	0.16 (0.08)	0.19 (0.10)	<0.001	Atlas
HH. median inc. white, 2010	59,007.01 (18,636.57)	63,903.41 (16,660.50)	<0.001	Atlas
Poor share white, 2010	0.16 (0.07)	0.14 (0.06)	<0.001	Atlas
Female single parent hh. white, 2010	0.29 (0.04)	0.27 (0.05)	<0.001	Atlas
Emp. rate black, 2010	0.51 (0.24)	0.53 (0.19)	0.5	Atlas
Frac. college black, 2010	0.11 (0.14)	0.11 (0.12)	0.6	Atlas
HH. median inc. black, 2010	41,130.39 (20,064.18)	40,808.11 (17,421.34)	0.4	Atlas
Poor share black, 2010	0.32 (0.22)	0.30 (0.17)	0.2	Atlas
Female single parent hh. black, 2010	0.59 (0.12)	0.58 (0.11)	0.14	Atlas
Foreign share, 2009	0.02 (0.03)	0.05 (0.05)	<0.001	Atlas
Frac. teen births	0.05 (0.02)	0.06 (0.02)	0.4	CHR
Primary care physicians	0.00 (0.00)	0.00 (0.00)	0.2	CHR
Population fertility black, 2013	2,947.97 (11,028.13)	3,996.60 (13,342.09)	<0.001	SEER
Population fertility white, 2013	8,482.34 (14,413.76)	12,858.82 (36,507.97)	0.6	SEER

[1]Mean (SD); n / N (%)
[2]Wilcoxon rank sum test; Pearson's Chi-squared test
Notes: Data sources: NVSS: National Vital Statistics System; NCES: National Center for Education Statistics; QCEQ: Quarterly Census of Employment and Wages; FAO: U.N. Food & Agriculture Organization; CRU: Climate Research Unit; CHR: County Health Ranking; BEA: Bureau of Economic Analysis; NCHS: National Center of Health Statistics; CHR: County Health Rankings and Roadmaps; Atlas: The Opportunity Atlas; SEER: Surveillance, Epidemiology, and End Results Program. Olvera replication files: Olvera, Smith et al. (2023); ABS replication files: Acharya, Blackwell and Sen (2016); AR: Althoff and Reichardt (2024); Di replication files: Di, Kloog, Koutrakis, Lyapustin, Wang and Schwartz (2016); HLW: Heidi L. Williams; JA: Jeremy Atack, downloaded from `https://my.vanderbilt.edu/jeremyatack/data-downloads/`.

Table 3: **Maternal mortality (PARM, PRMR, incidental)**

	PAMR per 1,000 live births					
	Whites (women 15-44)			Blacks (women 15-44)		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Expansion × Post	-0.0398	-0.0178	-0.079	-0.1881	-0.5375	-0.361
	(0.0761)	(0.0669)	(0.067)	(0.1226)	(0.0895)	(0.127)
Num. Counties	1,346	1,346	1,346	860	860	860
Year FE	✓	✓	NA	✓	✓	NA
County FE	✓	✓	NA	✓	✓	NA
controls		✓	✓		✓	✓
	PRMR per 1,000 live births					
	Whites (women 15-44)			Blacks (women 15-44)		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Expansion × Post	-0.0539	-0.0565	-0.083	-0.1278	-0.4137	-0.292
	(0.0719)	(0.0765)	(0.055)	(0.1034)	(0.0728)	(0.108)
Num. Counties	1,346	1,346	1,346	860	860	860
Year FE	✓	✓	NA	✓	✓	NA
County FE	✓	✓	NA	✓	✓	NA
controls		✓	✓		✓	✓
	Incidental maternal deaths per 1,000 live births					
	Whites (women 15-44)			Blacks (women 15-44)		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Expansion × Post	0.0141	0.0387	0.001	-0.0474	-0.0896	-0.076
	(0.0273)	(0.0218)	(0.034)	(0.0441)	(0.0223)	(0.07)
Num. Counties	1,346	1,346	1,346	860	860	860
Year FE	✓	✓	NA	✓	✓	NA
County FE	✓	✓	NA	✓	✓	NA
controls		✓	✓		✓	✓

Notes: The dependent variable in Panel A is the **pregnancy-associated mortality ratio (PAMR)**, defined as the number of deaths among women aged 15–44 during pregnancy or within one year after its termination, per 1,000 live births. Panel B uses the **pregnancy-related mortality ratio (PRMR)**, which includes only deaths directly related to pregnancy. A pregnancy-related death is one that occurs during pregnancy or within one year after its end, due to causes linked to pregnancy. Panel C examines **incidental deaths**, defined as maternal deaths unrelated to pregnancy. Columns (1)–(3) and (4)–(6) present estimates for White and Black women, respectively. Columns (1) and (4) show results from a baseline DiD model with county and year fixed effects but no additional controls. Columns (2) and (5) include covariates selected with the double LASSO method and use a weighted least squares (WLS) regression with balancing weights as shown in equation 2. Standard errors are clustered at the state level. Columns (3) and (6) report results from the ML-DR-DiD framework. Data sources: NVSS mortality files 2010–2017 and other sources as reported in Table 2.

Table 4: **Infant mortality at county level**

	infant mortality rate per 1,000: White					
	all causes			amenable		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Expansion × Post	0.2054	0.1700	0.357	-0.0483	-0.1226	-0.198
	(0.1399)	(0.2279)	(0.32)	(0.0996)	(0.1110)	(0.283)
Num. Counties	1,346	1,346	1,346	1,346	1,346	1,346
Year FE	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
controls × Post		✓			✓	

	infant mortality rate per 1,000: Black					
	all causes			amenable		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Expansion × Post	0.3265	1.116	0.556	-0.0843	0.4553	0.017
	(0.7255)	(0.8682)	(0.734)	(0.3374)	(0.4764)	(0.489)
Num. Counties	860	860	860	860	860	860
Year FE	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
controls × Post		✓			✓	

Notes: The dependent variable is the **infant mortality rate (IMR)** per 1,000 live births. Panels A and B pertain to infants born to White and Black mothers, respectively. Columns (1)–(3) show results for **all causes mortality**, whereas Columns (4)–(6) refer to **amenable mortality**. Columns (1) and (4) show results from a baseline specification that only accounts for county and year fixed effects. Columns (2) and (5) include the interactions between the Post-ACA indicator and relevant pre-treatment variables selected with the double LASSO method. A weighed least square (WLS) regression is estimated with weights defined in equation 2. The standard errors are clustered at the state level. Columns (3) and (6) show results based on the ML-DR-DiD approach. Data sources: NVSS mortality files 2010–2017 and other sources as reported in Table 2.

Table 5: **Infant birthweight outcomes at county level**

	infant from White mother								
	infant's birthweight			low birthweight			very low birthweight		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expansion × Post	6.405	6.714	2.912	-0.0018	-0.0021	0	-0.0002	-0.0001	0
	(3.914)	(4.094)	(3.44)	(0.0018)	(0.0011)	(0.002)	(0.0005)	(0.0006)	(0.001)
Num. Counties	1,362	1,362	1,362	1,362	1,362	1,362	1,362	1,362	1,362
Year FE	✓	✓	NA	✓	✓	NA	✓	✓	NA
County FE	✓	✓	NA	✓	✓	NA	✓	✓	NA
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

	infant from Black mother								
	infant's birthweight			low birthweight			very low birthweight		
Model	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD	base	controls	ML-DR-DiD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expansion × Post	9.659	11.33	13.049	-0.0038	-0.0026	-0.004	-0.0043	-0.0054	-0.004
	(5.298)	(3.751)	(6.335)	(0.0025)	(0.0021)	(0.003)	(0.0013)	(0.0017)	(0.001)
Num. Counties	914	914	914	914	914	914	914	914	914
Year FE	✓	✓	NA	✓	✓	NA	✓	✓	NA
County FE	✓	✓	NA	✓	✓	NA	✓	✓	NA
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variables are: (*i*) average infant birthweight at the county-level (denoted **infant's birthweight**) in Columns (1)–(3), (*ii*) the county-level fraction of births with a birthweight less than 2,500 grams (denoted **low birthweight**) in Columns (4)–(6), and (*iii*) the county-level fraction of births with a birthweight less than 1,500 grams (**very low birthweight**). Columns (1), (4), (7) show the baseline results with controls for county and year fixed effects. Columns (2), (5), (8) add in the relevant controls selected with the double LASSO method and uses weighted least square (WLS) with weights defined in equation 2. Standard errors are clustered at the state level. Columns (3), (6), (9) present results based on the ML-DR-DiD framework. Data sources: NVSS natality files 2010–2017 and other sources as reported in Table 2.

Appendix for:

"The Effects of ACA-Medicaid Expansion on Infant and Maternal Health Outcomes in the American South"

Juergen Jung[†] and Vinish Shrestha[‡]

[†]Department of Economics, Towson University, Email:jjung@towson.edu
[‡]Department of Economics, Towson University, Email:vshrestha@towson.edu

A Estimated propensity scores for counties in South

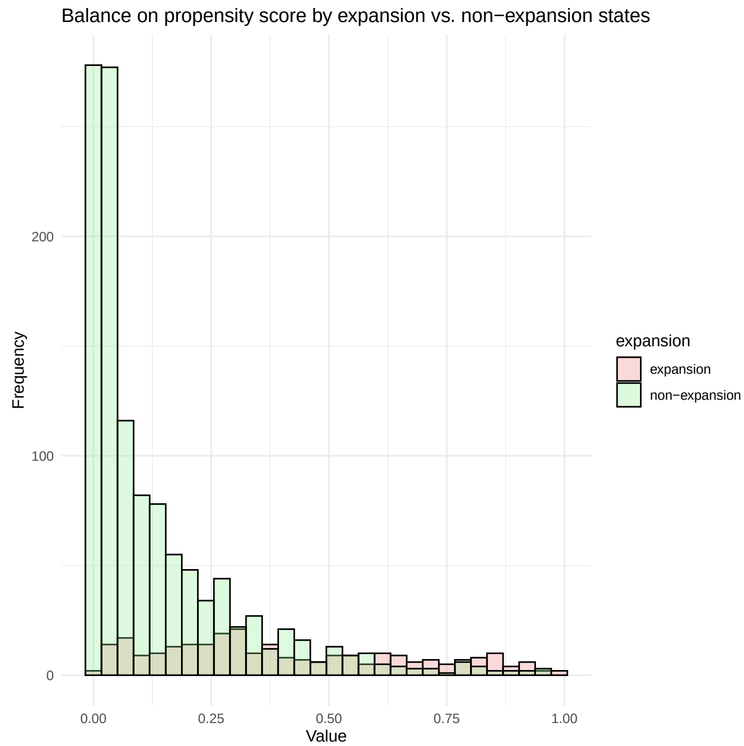


Figure A.1: **Propensity scores of expanding Medicaid in 2014 by expansion status**

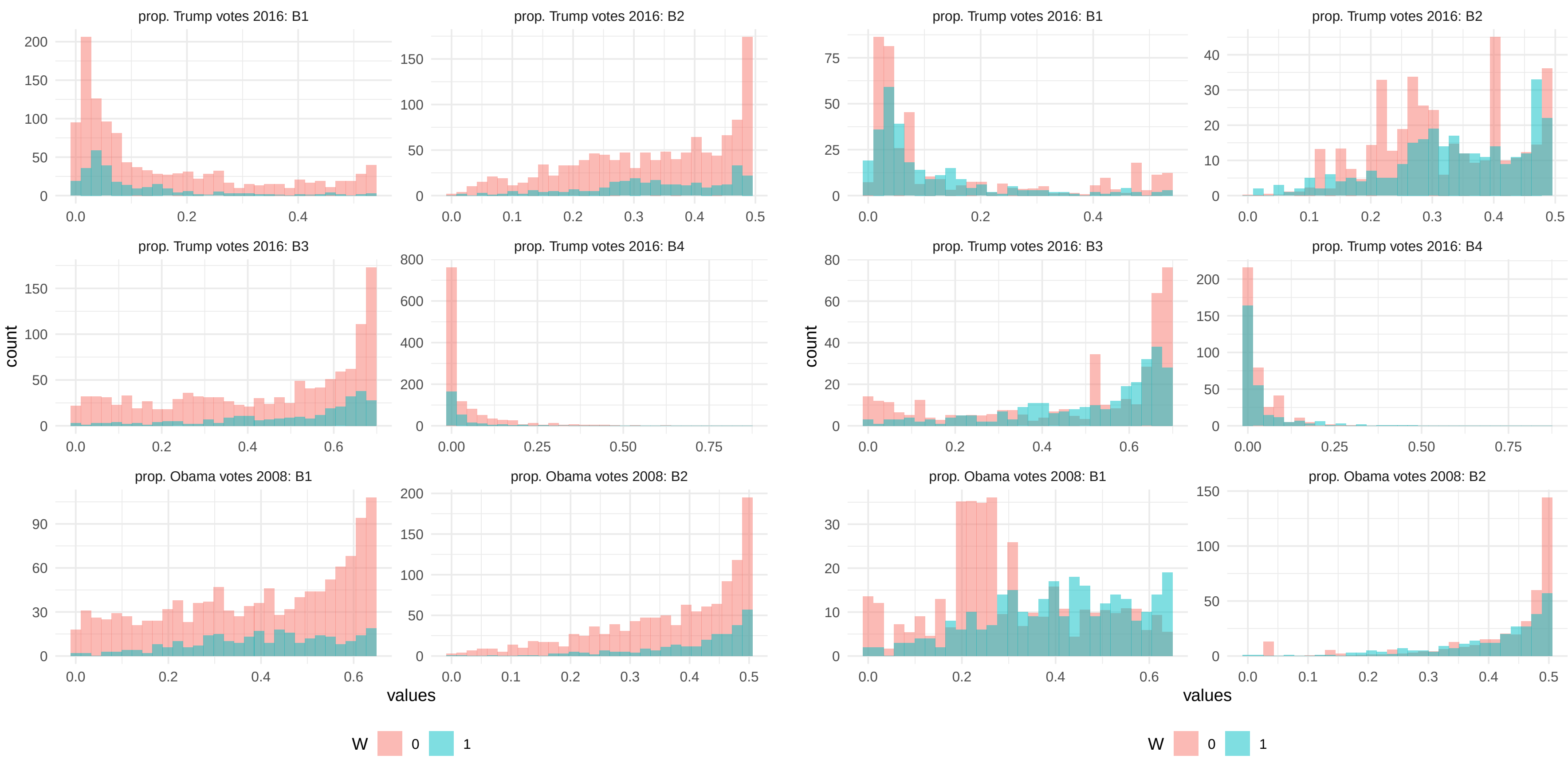
Notes: The graph shows the histograms of propensity scores (predicting whether a county expands Medicaid via the ACA) of counties that expanded Medicaid and counties that did not expand Medicaid in the American South. The variables used to build the propensity score model are determined using LASSO, which selects covariates from a list of high-dimensional county level characteristics. Following this, a probit model is used to estimate propensity scores using 10-fold cross-fitting.

(a) Unadjusted

(b) Adjusted

Figure A.2: **Balance of covariates by expansion status**

Notes: The graph shows the distribution of covariates before and after adjusting with propensity score weights.

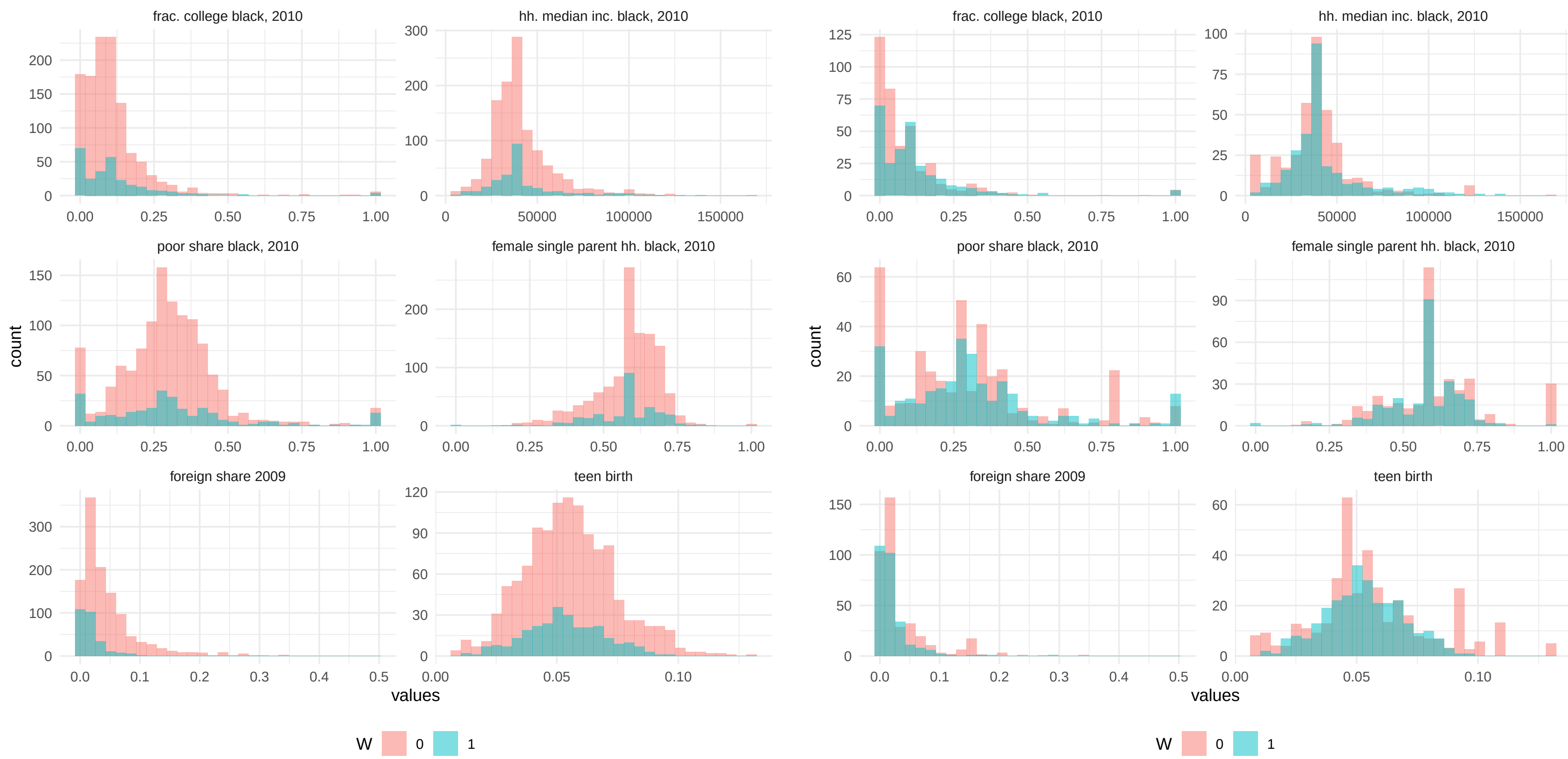


(a) Unadjusted

(b) Adjusted

Figure A.3: **Balance of covariates by expansion status**

Notes: The graph shows the distribution of covariates before and after adjusting with propensity score weights.



(a) Unadjusted

(b) Adjusted

Figure A.4: **Balance of covariates by expansion status**

Notes: The graph shows the distribution of covariates before and after adjusting with propensity score weights.

B Predicting who is a Medicaid eligible female

We follow the methods in Card and Krueger (1995) and Cengiz, Dube, Lindner and Zentler-Munro (2022) and build prediction models to explain the relationship between being eligible for Medicaid and various demographic variables. We then use the model to predict the likelihood of being eligible for Medicaid of young mothers in the NVSS (which lacks information about income) using the same set of demographic variables.[1]

B.1 Data

We use ACS data from 2014–2017 limited to US states that expanded Medicaid eligibility in either 2014, 2015, 2016 or 2017 via provisions introduced by the ACA. ACS data are available annually and include information about demographic characteristics as well as income. The latter is important as income is one of the key defining characteristics for Medicaid eligibility, especially for younger women of reproductive age 12–50 who are the focus of this study. In addition, the ACS data also includes information about the Public Use Microdata Areas (PUMAs) level.[2]

We next define an indicator variable equal to one if a woman qualifies for Medicaid based on the Medicaid income criteria of their state of residence. This eligibility is based on household income defined by the MAGI (Modified Adjusted Gross Income) eligibility rules that include criteria such as income, marital status, household size etc. but exclude asset tests.[3] This definition therefore includes women that are either eligible and on Medicaid, eligible but not on Medicaid, or within 25 percent of the residency state's Medicaid income eligibility threshold.

The ACS sample is not restricted to just mothers. It includes 1,522,777 person level observations of all females between ages 12–50 who live in 2014–2017 ACA expansion states in years 2014–2017. Table B.1 shows the summary statistic for the key variables used in training the machine learning algorithms that predict Medicaid eligibility. In our sample about 31 percent of females are eligible for Medicaid, the average age is 31 years, 39 percent are married, 31 percent have

[1]We subsequently focus on mothers with a high enough likelihood of Medicaid eligibility (say those in the 90th percentile) and collapse the NVSS birth data by county using only birth information of mothers with a high enough probability of qualifying for Medicaid. See main text for details about the estimation and identification strategy.

[2]PUMAs are used by the Census and define non-overlapping areas that cover the entire US. It is usually a combination of contiguous counties while large urban counties are often subdivided into multiple PUMAs. Each PUMAs includes at least 100,000 inhabitants. The Census Bureau redraws PUMA boundaries every 10 years based on population information from the most recent decennial census. The 2010-based PUMAs were first used in ACS/PRCS samples for 2012 respondents. There were 2,071 PUMAs in the 2010 Census and for the 2012 American Community Survey (ACS), there are 2,378 PUMAs. See https://usa.ipums.org/usa/volii/pumas10.shtml and https://www.census.gov/programs-surveys/geography/guidance/geo-areas/pumas.html for more information about PUMAs.

[3]For more details about Medicaid eligibility criteria compare: https://www.medicaid.gov/medicaid/eligibility-policy/index.html

at least college degree, 27 percent are non-white, 10 percent are black, 22 percent are teenagers and 14 percent are young adults.

The second source of data is the NVSS. We use two samples from the NVSS, the 2009–2017 NVSS-Natality sample (excluding births from 2010) and the 2010–2017 NVSS-Mortality sample. The natality sample comprises all US births between 2009–2017 (excl. 2010) and includes information such as birthweight, birth-height as well as information about the mother's age, education, marital status and race. We can therefore predict for every birth entry whether the mother was eligible for Medicaid based on the mother's demographic characteristics. Since the prediction models are trained on ACS data from ACA expansion states between 2014–2017, the generated predicted probabilities for Medicaid eligibility in the NVSS 2009-2017 sample reflect counterfactual eligibility probabilities for birth-observations of states that either have not expanded Medicaid in the 2014–2017 time frame or are from states before the expansion which would be the years 2009–2013. The probabilities are subsequently used to either reduce the sample to birth observations that have a high probability of Medicaid eligibility or to produce weights that up-weight birth observations that have a high likelihood of being Medicaid eligible if the ACA expanded in a given state. Table B.1 shows summary statistics of this sample. It contains 28 million birth observations between 2009–2017 (excluding 2010), about 4 million per year. The probability that a mother giving birth is eligible for Medicaid (post ACA expansion) ranges from 26 to 32 percent, depending on the ML model used to make the prediction. The average age of mothers is 28.3 years, 59 percent are married, 30 percent have a college or higher degree, 25 percent are birth by non-white mothers, 16 percent are births by black mothers, 7 percent of births are linked to teenage births and 27 percent of births are linked to young adult mothers.

Finally, the NVSS mortality sample records all deaths (male and female) in the US. Since we are interested in mother's mortality in connection with child birth, we limit this sample to female deaths at ages 12–50. The mortality data also contains information about age (at death), marital status, education, and race so that we can again use the trained ML models to predict the probability of Medicaid eligibility of every recorded female that died between the ages 12–50. Table B.1 shows that the mortality sample comprises 668,766 female deaths at ages 12–50 between the years 2010–2017. The average predicted probability of Medicaid eligibility ranges from 23 to 28 percent, depending on the particular ML model used to make the prediction. The average age is 39.2 years, 37 percent were married, 14 percent had a college or higher degree, 25 percent were non-white, 21 percent were black, 4 percent were teenagers and 7 percent were young adults.

B.2 Empirical strategy

We split the ACS sample into a training sample (80 percent of observations) and a test sample (20 percent of observations). We then fit various machine learning models (e.g., linear probabil-

Table B.1: **Summary statistics: ACS, NVSS natality, NVSS mortality**

	Obs.	Mean	Std	Min	Max
ACS Female 2014-2017					
Maid-eligible	1,522,777	0.31	0.46	0	1
Age	1,522,777	31.05	11.52	12	50
Married	1,522,777	0.39	0.49	0	1
No-HS	1,522,777	0.23	0.42	0	1
HS	1,522,777	0.17	0.38	0	1
Some college	1,522,777	0.29	0.45	0	1
College+	1,522,777	0.31	0.46	0	1
Nonwhite	1,522,777	0.27	0.44	0	1
Black	1,522,777	0.10	0.30	0	1
Teenager	1,522,777	0.22	0.41	0	1
Age 20-25	1,522,777	0.14	0.35	0	1
NVSS natality 2014-2017					
P1(Maid)	28,510,205	0.32	0.18	0	1
P2(Maid)	28,510,205	0.32	0.19	-0	1
P3(Maid)	28,510,205	0.30	0.17	-0	1
P4(Maid)	28,510,205	0.26	0.22	0	1
P5(Maid)	28,510,205	0.26	0.20	0	1
P6(Maid)	28,510,205	0.26	0.21	0	1
Age	28,510,205	28.25	5.94	12	50
Married	28,510,205	0.59	0.49	0	1
No-HS	28,510,205	0.16	0.37	0	1
HS	28,510,205	0.25	0.43	0	1
Some college	28,510,205	0.29	0.45	0	1
College+	28,510,205	0.30	0.46	0	1
Nonwhite	28,510,205	0.25	0.43	0	1
Black	28,510,205	0.16	0.36	0	1
Teenager	28,510,205	0.07	0.25	0	1
Age 20-25	28,510,205	0.27	0.45	0	1
NVSS mortality 2014-2017					
P1(Maid)	668,766	0.23	0.15	0	1
P2(Maid)	668,766	0.25	0.17	-0	1
P3(Maid)	668,766	0.27	0.17	-0	1
P4(Maid)	668,766	0.28	0.19	0	1
P5(Maid)	668,766	0.27	0.16	0	1
P6(Maid)	668,766	0.27	0.19	0	1
Age	668,766	39.22	9.53	12	50
Married	668,766	0.37	0.48	0	1
No-HS	668,766	0.20	0.40	0	1
HS	668,766	0.40	0.49	0	1
Some college	668,766	0.26	0.44	0	1
College+	668,766	0.14	0.35	0	1
Nonwhite	668,766	0.25	0.43	0	1
Black	668,766	0.21	0.40	0	1
Teenager	668,766	0.04	0.21	0	1
Age 20-25	668,766	0.07	0.26	0	1

*Notes:*The ACS sample comprises observations of females between the ages 12–50 from US states that expanded Medicaid in 2014, 2015, 2016, or 2017 and covers calendar years 2014–2017. The NVSS natality sample comprises observations of all US births between 2009–2017 (excluding 2010). The variables Age, Married, etc. concern the mothers connected to the births. The NVSS mortality sample comprises observations of all US deaths of females between the ages 12–50 between 2010–2017.

ity model similar to Card and Krueger (1995), a basic logistic model, elastic net, random trees, random forests, and tree boosting) using the training data and use the test sample to compare the performance (i.e., precision recall curves and ROC curves). We then use the trained model to make forecasts of new mother's Medicaid eligibility in the NSS natality sample as well as females Medicaid eligibility in the NVSS mortality sample.

We next describe the prediction models in more detail. We first use a basic logistic regression model with just age and education as explanatory variables so that the probability of being eligible for Medicaid can be written as

$$\Pr\left(y_{i,g}=1|X\right)=\frac{\exp\left(X\beta\right)}{1+\exp\left(X\beta\right)}$$

and the logit estimator solves

$$\hat{\beta}=\arg\min_{\beta}\sum_{i=1}^{N}\left(y_{i,g}\ln\left(\frac{1}{1+\exp\left(-X\beta\right)}\right)+\left(1-y_{i,g}\right)\ln\left(1-\frac{1}{1+\exp\left(-X\beta\right)}\right)\right), \tag{6}$$

where $y_{i,g}$ is an indicator outcome variable measuring whether a woman of reproductive age between 15–49 in PUMA g is eligible for Medicaid in her state of residence. The demographic variables of the model in matrix X only include age and education.

Second, we employ the linear probability model based on Card and Krueger (1995) which can be written as

$$y_{i,g}=\beta_0+X\beta+\varepsilon_{i,g},$$

where the demographic variables of the model in matrix X include age, quadratic and cubic terms of age, three-way interaction variables between teen (aged 15–19), nonwhite, and gender indicators , three-way interaction variables between young adult (aged 20-25), nonwhite, and gender indicators, three-way interactions of age, categorical education (i.e., less than high school, high school or GED, some college or associates degree, and college or more), and gender variables; and indicator variables for black and nonwhite individuals.[4]

Third, we use an elastic net estimator that introduces a penalizing function

$$\hat{\beta}_{\lambda,\alpha}=\arg\min_{\beta}\|y-X\beta\|_2^2+\lambda\left(\left(1-\alpha\right)\|\beta\|_2^2+\alpha\|\beta\|\right)$$

on top of a logistic regression which effectively combines Lasso and Ridge penalization terms with λ and α being tuning parameters determining the importance of the quadratic and absolute value

[4] Medicaid.gov shows the Medicaid eligibility based on income in relation to the federal poverty level at https://www.medicaid.gov/medicaid/national-medicaid-chip-program-information/medicaid-childrens-health-insurance-program-basic-health-program-eligibility-levels/index.html.

of the sum of estimation coefficients. Since the model introduces punishing terms for the number of coefficients, we use a wider set of explanatory variables that includes a cubic and quartic term of age as well as two-way and four-way interactions of nonwhite, black, age squared, married, and education.[5]

The final three models are tree-based methods.[6] More specifically, the fourth model is a decision tree model with the following variables as the feature space: age, education, race black, race nonwhite, marriage status. The decision tree will recursively split the sample so that the prediction errors from the resulting subsamples are minimized according to some loss criteria. Subsample predictions are typically averages for continuous outcome variables and modes or proportions for categorical outcome variables.[7]

Fifth, it has been shown that combining trees into random forests can significantly improve the predictive power of the model. Each tree is formed with a random selection of features (i.e., explanatory "x" variables) and the prediction is averaged over all trees. Cross validation of bootstrapped samples is used to evaluate the performance of the random forest model.[8]

Finally, the boosted trees method is implemented. Instead of combining and averaging over many random trees, tree boosting is a sequential method, where subsequent trees are used to improve the performance of the previous tree in a sequence.[9]

B.3 Model comparison

We next show the precision recall (PR) and receiver operating characteristic (ROC) curves of the various models based on predictions using the test data sample described above. The are diagnostic tools for binary classification models.

Figure B.1 shows the precision recall curves for the different models. Precision is a metric that shows the fraction of correctly identified women that are Medicaid eligible out of all the predictions

[5]The elastic net is implemented in Python using `ElasticNet` from the `sklearn.linear_model` library of the scikit-learn machine learning toolbox available at https://scikit-learn.org The implementation in R uses the `glmnet` library available at https://glmnet.stanford.edu/

[6]Compare Chapter 8 in James, Witten, Hastie, Tibshirani and Taylor (2023) for more details on decision trees.

[7]This method was implemented using `DecisionTreeClassifier` from the `sklearn.tree` library of the scikit-learn machine learning toolbox available at https://scikit-learn.org The R implementation is based on the `tree` package available at https://www.rdocumentation.org/packages/tree

[8]This method was implemented using `RandomForestClassifier` from the `sklearn.ensemble` library of the scikit-learn machine learning toolbox available at https://scikit-learn.org The R implementation is based on the `ranger` package available at https://www.rdocumentation.org/packages/ranger

[9]This method was implemented using `GradientBoostingClassifier` from the `sklearn.ensemble` library of the scikit-learn machine learning toolbox available at https://scikit-learn.org An R implementation of this algorithm is based on the `gbm` package available at https://www.rdocumentation.org/packages/gbm

that the model flagged as Medicaid eligible. In other words,

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}. \tag{7}$$

Recall on the other hand measures the fraction of correctly identified women that are Medicaid eligible out of all the Medicaid eligible women in the data. In other words recall is the True Positive Rate or

$$\text{True Positive Rate = Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}}. \tag{8}$$

A precision-recall curve plots precision on the y-axis and recall on the x-axis for various probability thresholds. A model with perfect predictions would result in a point at coordinate (1,1). A model with good predictive power is a curve that bows towards the coordinate of (1,1). A model without any predictive power is a horizontal line with a precision that is proportional to the number of positive examples in the dataset. For a dataset with an equal number of women qualifying for Medicaid and not qualifying for Medicaid this would be a horizontal line at a precision of 0.5.

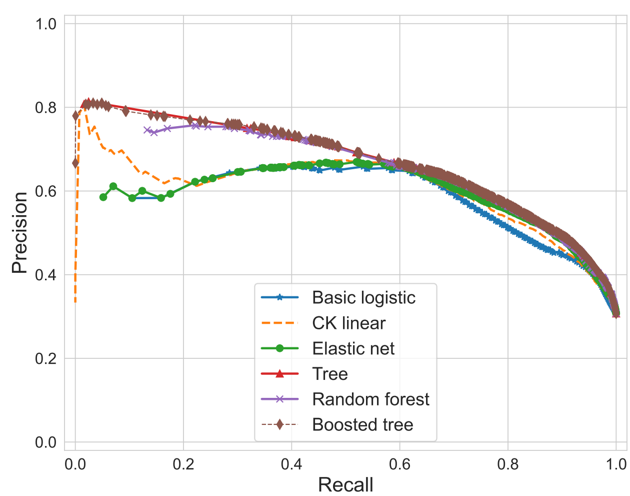


Figure B.1: **Precision-Recall curves for Females ages 12–50**

Notes: The figure plots the precision-recall curves for various prediction models described in section B. We use the prediction model to calculate the probability that someone is a person affected by Medicaid (either eligible and on Medicaid, eligible but not on Medicaid, or within 25 percent of the residency state's Medicaid income eligibility threshold. The areas below the precision-recall curves are 0.430 for the **basic logistic** model and 0.489 for the **Card and Krueger linear** probability model.

The ROC curve (or receiver operator characteristic) plots the true positive rate (or Recall) as already defined in equation 8 against the false positive rate defined as

$$\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives+True Negatives}}. \tag{9}$$

In other words it plots the fraction of correct predictions for the positive class (here it's Medicaid eligibility) again the fraction of errors in the negative class (here it's individuals not eligible for Medicaid). Figure B.2 shows the ROC curves based on predictions using the test data sample described above.

The prediction error rates depend on the probability threshold that is used by the classifier algorithm to predict whether a person qualifies for Medicaid or not. Changing this cut-off threshold produces a trade-off between the True-Positive-Rate (TPR) and the False-Positive-Rate (FPR) and changes the balance of predictions towards either improving the TPR at the expense of the FPR, or the other way around. The ROC curve is a representation of the two rates based on different threshold values. The ROC curve stretches from the bottom left to the top right and bows toward the top left. The ideal point is at coordinate 0,1 as here the fraction of correctly predicted individuals who qualify for Medicaid is 1 and the fraction of errors for individuals not qualifying for Medicaid is 0. This would be the perfect classifier. A classifier without discriminative power between the two classes (Medicaid eligible yes/no) will form a diagonal line.

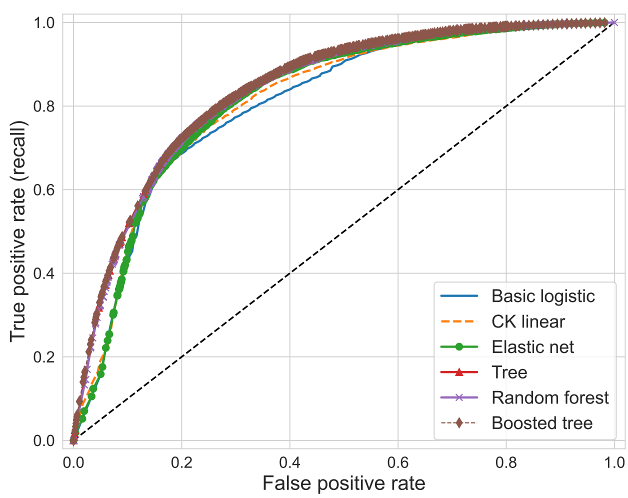


Figure B.2: **Receiver Operating Characteristic (ROC) curves: Females ages 12–50**

Notes: The figure plots the precision-recall curves for various prediction models described in section B. We use the prediction model to calculate the probability that someone is a person affected by Medicaid (either eligible and on Medicaid, eligible but not on Medicaid, or within 25 percent of the residency state's Medicaid income eligibility threshold. The areas below the precision-recall curves are 0.430 for the **basic logistic** model and 0.489 for the **Card and Krueger linear** probability model.

According to the PR curves in Figure B.1 ML decision tree, random forest and boosted tree models show better performance over basic logistic, linear regression or elastic net models. The ROCs curves in Figure B.2 further support this conclusion.

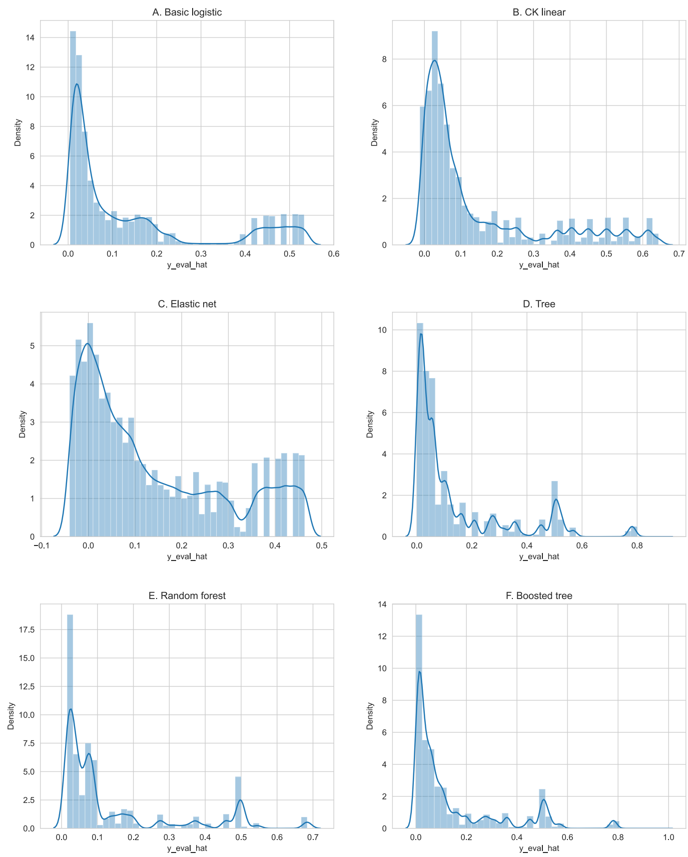


Figure B.3: **Predictions of Medicaid eligibility probabilities for females (ages 12–50) in the evaluation sample (ACS data)**

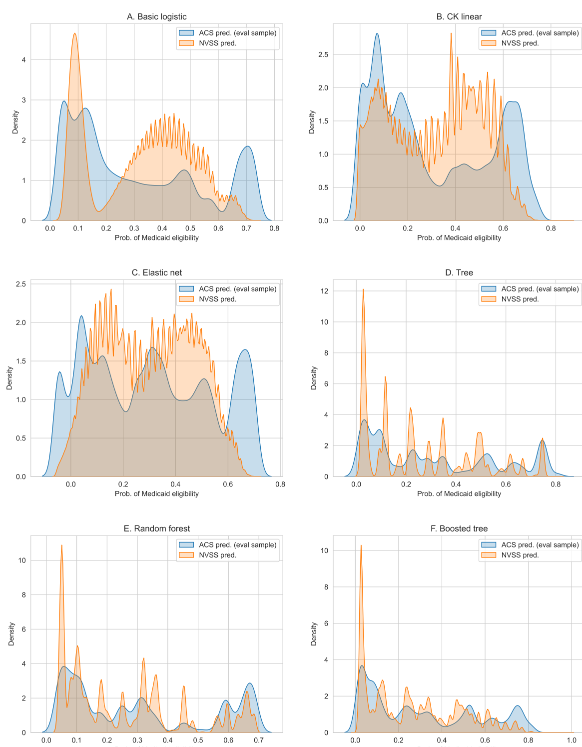
Notes: The figure plots the distribution of predicted Medicaid eligibility probabilities based on the 20 percent evaluation sample of the ACS 2014–2017 sample of females between 12–50 years old who lived in ACA expansion states. All models where trained on the corresponding 80 percent training sample. We use the different prediction models to calculate the probability that someone—here we limit it to females between the ages 12–50—is a person affected by Medicaid (i.e., either eligible and on Medicaid, eligible but not on Medicaid, or within 25 percent of the residency state's Medicaid income eligibility threshold).

Finally, Figure B.3 shows the histogram of predicted probabilities of being Medicaid eligible for the ACS test (evaluation) data. The blue line in the histograms is based on a kernel smoothing algorithm and approximates the density function. The pictures show the distribution of females' predicted probabilities of being eligible for Medicaid in US states that expanded Medicaid via the ACA in 2014, 2015, 2016, or 2017. The sample only includes females aged 12–50 and covers the years 2014–2017.

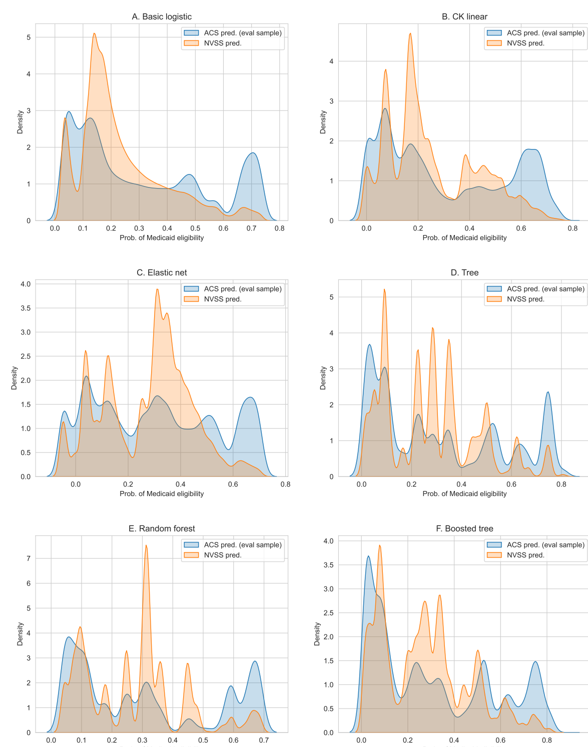
We see that models that lead to higher classification errors according to PR and ROC curves, show a larger mass of high probabilities of Medicaid eligibility (compare Panels A, B, and C) whereas models that exhibit fewer misclassification errors are more right skewed and show relatively fewer occurrence of very large probabilities of Medicaid eligibility.

B.4 Imputation of Medicaid eligibility in NVSS samples

We next use the trained ML models to predict Medicaid eligibility probabilities for females between the ages of 12–50 in the NVSS samples. The predictions are possible because we only used information about age, education, marital status and race in order to train the model in the ACS data. As these variables are also available in the NVSS samples, we are now able to predict Medicaid eligibility probabilities for the NVSS samples. We next show the predicted probabilities of Medicaid eligibility of 12–50 year old females in the ACS test (evaluation) sample and the NVSS natality sample on the left in Figure B.4a. One the right, in Figure B.4b we show the predicted probabilities of Medicaid eligibility of 12–50 year old females in the ACS test (evaluation) sample and the NVSS mortality sample.



(a) NVSS Natality: Mothers ages 12–50



(b) NVSS Mortality: Females ages 12–50

Figure B.4: **Predictions of Medicaid eligibility probabilities in evaluation sample (ACS data)**

Notes: The blue plots are the distributions of predicted Medicaid eligibility probabilities based on the 20 percent evaluation sample of 12–50 year old females that live in ACA expansion states in the ACS 2014–2017 sample. All models where trained on the corresponding 80 percent training sample. These females are not restricted to be mothers. The orange plots are the predicted Medicaid eligibility probabilities using the ACS-trained model but using predictor data from either the NVSS 2009–2017 (excluding 2010) natality or the 2010–2017 mortality samples. The predictors in the natality sample include a newborn's mother's age, marriage status, education, race and combinations thereof. The predictors in the mortality sample include a female's age, marriage status, education, race and combinations thereoff. We train different prediction models and present the predicted probability that a newborn's mother (in the natality sample) or a female (in the mortality sample) is a person affected by Medicaid (i.e., either eligible and on Medicaid, eligible but not on Medicaid, or within 25 percent of the residency state's Medicaid income eligibility threshold).

C Estimation of standard errors in ML-DR-DiD framework

The method-of-moments (MM) condition of the parameter of interest θ_0—ATT in our case—can be written as:

$$E[\psi(D;\, \theta_0,\, \eta_0)] = 0, \tag{10}$$

where $\psi(.)$ is the score function, D is the treatment indicator, and η_0 are the so-called nuisance parameters (e.g., propensity score and outcome regression). A critical aspect of the score function is that it needs to satisfy the Neyman orthogonality condition (Chernozhukov et al., 2018)

$$\partial E(\psi(D;\, \theta_0,\, \eta))|_{\eta=\eta_0} = 0, \tag{11}$$

which states that the moment conditions used to identify θ_0 are locally insensitive to the nuisance parameter η around the true parameter value η_0. This allows us to plug in noisy estimates of these nuisance parameters without strongly violating the moment condition.

Furthermore, the score functions of many double/debiased ML models are linear in parameter θ, so that

$$\psi(D;\, \theta_0,\, \eta_0) = \psi_a(D;\, \eta)\theta + \psi_b(D;\, \eta); \tag{12}$$

and the estimator can be written as:

$$\tilde{\theta}_0 = -\frac{\psi_b(D;\, \eta)}{\psi_a(D;\, \eta)}. \tag{13}$$

The score function for the case of DR-DiD is given as:

$$\psi(D;\, \theta,\, \eta) = \underbrace{-\frac{D}{E_n[D]}}_{\psi_a(.)}\theta + \underbrace{\left(\frac{D}{E_n(D)} - \frac{\frac{e(X)(1-D)}{1-e(X)}}{E_n\frac{e(X)(1-D)}{1-e(X)}}\right) \times \left(\Delta Y - m(X)\right)}_{\psi_b(.)}, \tag{14}$$

where $e(X) = P(D = 1|X)$ are the propensity scores and $m(X) = E(Y_1 - Y_0|D = 0, X)$ is the conditional expectation function estimated for non-expansion units. The variance estimator is given as:

$$\sigma^2 = J_0^{-2} E[\psi^2(D;\, \theta_0,\, \eta_0)] \tag{15}$$

and the sample analog is:

$$\hat{\sigma}^2 = \frac{1}{N}\hat{J}_0^{-2} \underbrace{\sum_{k=1}^{K}\sum_{i\in I_k}\psi^2(D;\, \tilde{\theta}_{0,k},\, \hat{\eta}_{0,k})}_{\text{summing up score vals across K folds}}, \tag{16}$$

with $\hat{J}_0 = \frac{1}{N}\sum_{k=1}^{K}\sum_{i\in I_k}\psi_a(D,\, \hat{\eta}_0)$. The standard error can then be calculated as: $\frac{\hat{\sigma}}{\sqrt{N}}$.

D Event-study results for infant mortality outcomes

In this section we briefly discuss the estimates of the event-study specification in Figure D.1. The top row panels A, B, E, and F pertain to the estimates of White IMR whereas the bottom panels refer to the estimates of Black IMR. Furthermore, the panels A–D show estimates based on the regression-based event-study design while panels E–H pertain to event-study-type estimates based on the more robust ML-DR-DiD framework. The first and third columns—panels A, C, E, and G—are the estimates for all-causes mortality whereas the second and last columns—panels B, C, F, and H—are the estimates for amenable mortality as outcome variable.

The regression-based event-study estimates for White infants, shown in panels A and B, are close to zero in magnitude and are statistically insignificant at the conventional levels for both, all-cause mortality as well as amenable mortality. The ML-DR-DiD estimates for both types of mortality (Panels E and F) are precisely estimated at zero. The event-study estimates for Black infant mortality are more noisily estimated as shown in Panels C and D compared to Whites. The ML-DR-DiD approach (Panels G and H) shows more stable estimates that are closer to zero than the regression-based event-study estimates. In summary, the relative time estimates of the event-study frameworks provide evidence that there are no short or long-term discernible effects of the ACA-Medicaid expansion on infant mortality rates among White or Black infants.

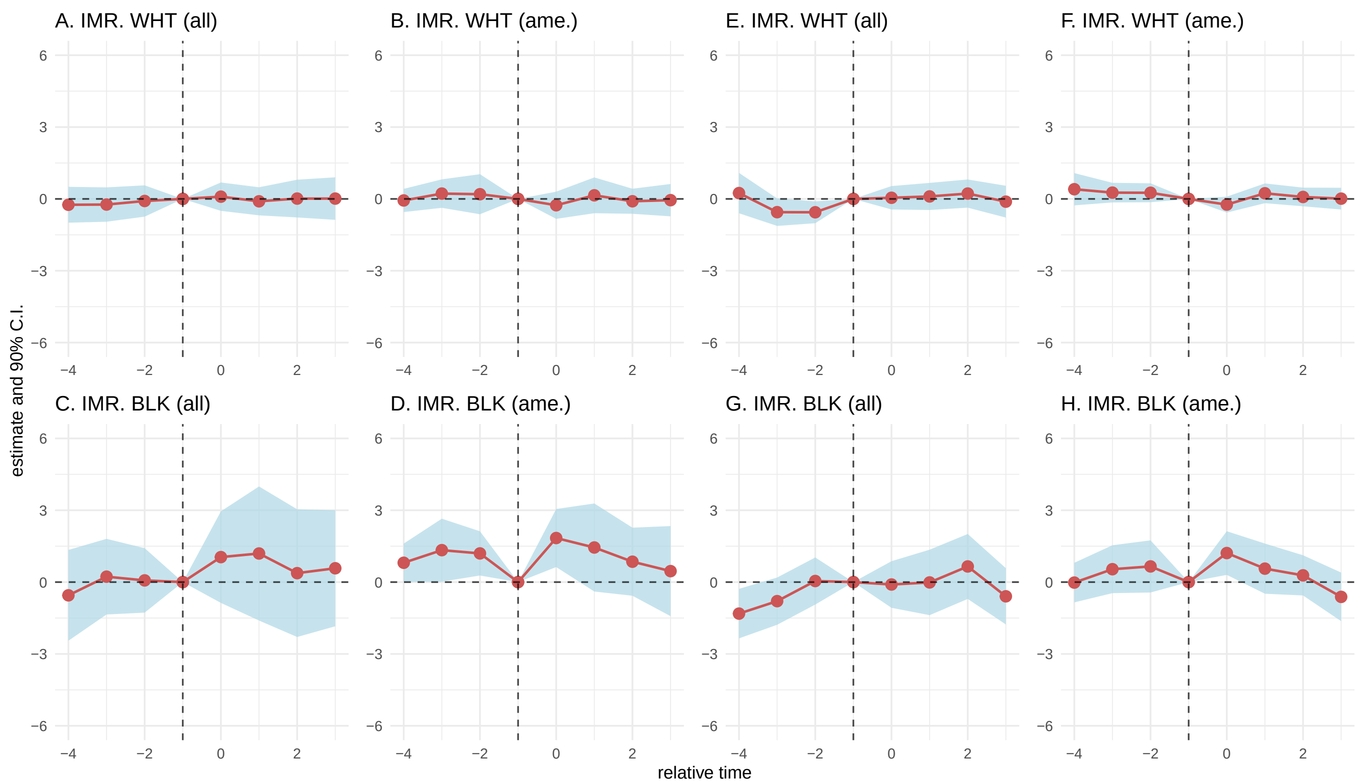


Figure D.1: **Effects of Medicaid expansion on infant mortality (per 1,000 infants)**

Notes: The dependent variable is the **infant mortality rate (per 1,000 live births)** of infants from White (top row) vs. Black (bottom row) mothers aggregated at the county level. Panels A, C, E, and G pertain to **mortality from all causes** while Panels B, D, F, and H refer to **amenable mortality**. Panels A–D show results from the regression-based event-study design in equation 1 with controls for county and year fixed effects. These specifications are weighted based on weights in equation 2. Additionally, the model accounts for the interactions between the Post-ACA indicator and controls selected with the double LASSO method. The standard errors are clustered at the state level. Panels E–H refer to the eventy-study-type findings using the ML-DR-DiD framework. Data sources: NVSS Mortality files 2010–2017.

Appendix References

Card, David and Alan B. Krueger. 1995. *The Effect of the Federal Minimum Wage on Low-Wage Workers: Evidence from Cross-State Comparisons*. Rev - revised ed. Princeton University Press pp. 113–151.

Cengiz, Doruk, Arindrajit Dube, Attila S. Lindner and David Zentler-Munro. 2022. "Seeing Beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes." *Journal of Labor Economics* 40(S1).

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning." *The Econometrics Journal* 21(1).

James, Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani and Jonathan Taylor. 2023. *An Introduction to Statistical Learning with Applications in Python*. 1 ed. Springer.