Towson University
Department of Economics
Working Paper Series

Working Paper No. 2024-08

Heterogeneous Impacts of ACA-Medicaid Expansion on Insurance and Labor Market Outcomes in the American South

By Vinish Shrestha

June 2024

Heterogeneous Impacts of ACA-Medicaid Expansion on Insurance and Labor Market Outcomes in the American South

Vinish Shrestha*

June 3, 2024

Abstract

The expansion of Medicaid through the Affordable Care Act (ACA-Medicaid) has sparked debates on its impact on local labor markets. This study delves into the heterogeneous impacts of ACA-Medicaid expansion on insurance and labor market outcomes in the American South. Utilizing the modified version of Causal Forest approach, the research uncovers significant heterogeneity in treatment effects. Notably, counties ranking within the top 10% based on the conditional average treatment effect (CATE) estimates exhibit a reduction in the uninsured rate by approximately 13 percentage points more than the average during the year of the reform. Moreover, the estimated heterogeneity suggest evidence of increased unemployment rates, decreased total employment, and a contraction of labor force in the years following the expansion. Such effects are particularly concentrated in food/accommodation and retail/trade sectors. These findings offer insights for refining and optimizing healthcare reform strategies.

Keywords: Medicaid expansion, Labor market outcomes, Heterogeneous effects, Conditional average treatment effect (CATE), Causal Forests (CF), Causal Machine Learning

JEL codes: I10, I14, C10, C19

*Towson University, Stephens Hall Room 101B, Towson, MD 21252. Email: *vshrestha@towson.edu*. *Phone*: +1 410-704-2956.

1 Introduction

Ten states, to date, have abstained from expanding Medicaid under the Affordable Care Act (ACA), limiting the reach of its healthcare provisions.[1] Alongside the concerns over costs, state budgets and dependency on government assistance, opponents of ACA and Medicaid expansion frequently argue that the implementation of Medicaid expansion without any work requirement policy creates disincentives to work (Harris and Mok, 2015; Greibrok, 2023). This issue has emerged as a focal point of debate surrounding the expansion of Medicaid across the United States (Kaiser Family Foundation, 2023). Indeed, several states have engaged in discussions about the implemention of work requirements to determine Medicaid eligibility.[2]

Given the increasing debate between Medicaid expansion and labor market outcomes, it is important to attain nuanced understanding regarding the prospect of heterogeneous impacts of the policy. While the intention of the Medicaid expansion policy is to increase access to health insurance among the poor, labor market distortions are typically unintented by the reform. These adverse effects primarily are defined by individuals who would have otherwise been employed or actively participated in the labor market in absence of the reform. One proposed approach to mitigate adverse labor market outcomes is the implementation of work requirement programs, which link health insurance eligibility to labor force participation. However, such policies may conflict with the overarching goals of the reform by negatively impacting health insurance coverage for individuals unable to work due to underlying health issues or structural barriers, such as lack of employment opportunities. This underscores the importance of examining whether areas with certain characteristics are more susceptible to adverse labor market impacts than others, as such insights can inform targeted interventions to mitigate unintended labor market distortions. For example, policymakers may look to subsidize job creation in particular local labor markets through measures such as local job training initiatives, place-based jobs policies (often referred to as "economic-development"), local business tax incentives, or community development policies, to offset negative effects on employment prospects.

This study considers the heterogeneous effects of ACA-Medicaid expansion on insurance and labor market outcomes in the American South,[3] encompassing the uninsured rate, Medicaid transfers per capita, unemployment rate, the total employment changes, and changes in the labor force following the reform. While the average treatment effect (ATE) estimate can mask differential effects of the policy, the conditional average treatment effect (CATE) estimates can help detect

[1] See https://www.kff.org/affordable-care-act/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/

[2] Arkansas was the first state to implement the work requirement policy, followed by Kentucky. Both policies were halted by the federal judge in 2019. Compare Guth and Musumeci (2022) for an overview of Medicaid work requirements. Starting from July 1, 2023, Georgia's "Pathways to Coverage" legislation has instituted a Medicaid work requirement of 80 hours per month and expanded the eligibility threshold from a very low 31 percent to 100 percent of the Federal Poverty Level (FPL).

[3] Following Acharya et al. (2016), the Southern states under examination include Texas, Louisiana, Arkansas, Missouri, Tennessee, Mississippi, Kentucky, Alabama, Florida, Georgia, South Carolina, North Carolina, Virginia, West Virginia, Maryland, and Delaware.

heterogeneous effects. Standard analyses that rely on predetermined hand-picked variables (i.e., race, gender) may overlook critical characteristics explaining treatment effect heterogeneity. In contrast, CATE estimates derived using Machine Learning (ML) methods can identify mediating variables of heterogeneity in treatment effects that can otherwise remain undetected.

We use a modified version of the causal forest (CF) approach, developed within the Generalized Random Forest (grf) framework in Athey et al. (2019), to obtain CATE estimates.[4] Alongside the estimation of ATE (or intent to treat effects), we explore heterogeneity in treatment based on the CATE estimates. In other words, analyses detect whether some areas with certain characteristics are significantly more affected by the policy compared to the average effect. Next, to formally evaluate heterogeneous treatment effects, we employ the Rank Average Treatment Effects (RATE) by following Yadlowsky et al. (2021), which compares effects at the certain segments as determined by the distribution of priority ranking to the average effect.

There are multiple reasons why the study is centered on the American South. First, out of the ten states without Medicaid expansion, eight are situated in the South, highlighting a significant regional disparity. This is compounded by the irony of Southerners experiencing high uninsured rates and poor health indicators, including life expectancy rates (Arias et al., 2021; RWJF, 2020). The below-par health outcomes and access to insurance are tied to historically high poverty rates, reliance on certain industries such as agriculture and manufacturing, as well as the legacy of historical institutions (Baker, 2022; Shrestha, 2023). Second, discussions regarding work requirement programs to govern Medicaid eligibility are particularly prevalent in the Southern region.[5] One major criticism posed by opponents of the Medicaid program is that it is seen as an entitlement. Critics argue that the program creates a dependency on government assistance, leading to concerns about individuals relying on Medicaid rather than seeking employment or other means of healthcare coverage. Next, focusing specifically on the American South may help in addressing spatial differences within the labor market. In a standard estimation setup of treatment effects of the ACA Medicaid expansion, the treatment group is heavily weighted by non-Southern states, while most of the control group consist of the Southern states. This can introduce bias due to differences in labor market outcomes across the spatial landscape.

The findings summarized by the ATE estimates from the Causal Forest approach demonstrate significant improvements in insurance outcome following the reform, suggesting that individuals below 138% of the Federal Poverty Level (FPL) experienced a reduction in uninsured rate of approximately 12-15 percentage points between the years 2014-2018. Although, the intent to treat (ITT) effects on unemployment rates are non-detectable, the estimates show patterns of decreased

[4] The identification based on the causal forest relies on the conditional independence assumption, i.e., the treatment assignment is independent of the potential outcomes conditional on the covariates. In context of this study, the identification for the modified version of Causal Forest is based on the parallel trend assumption, similar to the difference-in-differences framework.

[5] Among Southern states that expanded Medicaid, Arkansas and Kentucky attempted to implement work requirement program. The former state implemented the program, but it was halted by a federal judge in 2019. Similarly, Kentucky's program was blocked by a federal judge in 2018. Tennessee, Mississippi, and Georgia are curently in discussion to implement work requirement program. Compare Guth and Musumeci (2022).

total employment and labor force following the reform. The findings suggest a reduction of close to 800 employed people during the year of the reform. Moreover, effects gradually increase in magnitude over the years, resulting a decrease of about 2,000 employed people by 2018. Notably, the effects aimed at the average for all outcomes obtained from CF are similar to the findings from the standard event study approach based on the difference-in-differences framework.

CATE estimates exhibit levels of heterogeneity. Areas within the top 10% based on the magnitude of CATE estimates experienced reductions in uninsured rates ranging from -15 to -7 percentage points compared to the ATE estimate over the years 2014 and 2018. The variable importance matrix from Causal Forest suggest that the proportion of White votes for Obama in 2008 presidential election and uninsured rate in 2013 are the top two variables explaining heterogeneity in treatment effects. Areas most affected by the reform, as revealed by CATE estimates, saw increases in unemployment rates alongside decreases in total employment and labor force participation. These findings are consistent across analyses using labor market data from both the Bureau of Labor Statistics (BLS) and Quarterly Census of Employment and Wages (QCEW). Moreover, industry-level QCEW data highlight that labor market heterogeneity is concentrated within the food and accomodation and retailing sectors, with no statistically significant changes noted in the manufacturing and construction sectors.

In relation to the literature. A burgeoning body of literature evaluates the impact of ACA-Medicaid expansion on labor market outcomes. Most of these studies utilize state-level data under the difference-in-differences framework and find no significant effects of policy impact on labor market (Frisvold and Jung, 2018; Gooptu et al., 2016; Kaestner et al., 2017; Leung and Mas, 2018). Two drawbacks exist in these studies. First, the state level data can wash out the effects at the local level by aggregation. Second, comparision between the expansion versus non-expansion states systematically forms the treatment and control groups, such that the control group is heavily influenced by the southern states. Given spatial differences in labor market outcomes, such comparison may be problematic due to the potential of unobserved time varying changes affecting outcomes differently across the expansion versus non-expansion states. To account for both of these concerns, Peng et al. (2020) utilize county-level labor market data and employ identification strategy that focuses on bordering county-pairs by using variation in Medicaid expansion policy across contiguous states. The authors find that ACA-Medicaid expansion decreased total employment by 1.2 percent one year following the expansion. However, these effects are short-lived as employment returned to the pre-policy level two years after the policy implementation.

Separate to ACA-Medicaid expansion, some previous studies have also evaluated the impacts of gaining or losing Medicaid due to external shocks. Using the Oregon Health Insurance Experiment that assigned Medicaid coverage through lottery system, Baicker et al. (2014) find no impacts of the new-found Medicaid coverage on short-term employment. Conversely, Garthwaite et al. (2014) document sharp increases in labor supply following the abrupt disenrollment of Tennessee's Medicaid program (TennCare) in 2005. Dague et al. (2017) use a sudden enrollment cap in Wisconsin Medicaid program as the source of identification and find that employment reduced among individuals covered by Medicaid compared to those eligible but lacking coverage due to the imposed restriction. As it remains, the empirical evidence of Medicaid expansion on the labor market outcomes is mixed. The differences in findings may be explained due to differences in the types of policies implemented (Medicaid expansion through ACA versus expansion independent of ACA), geographical focus, macroeconomic conditions and the population of interest.

This study distinguishes itself from previous research in two significant ways. First and more importantly, unlike previous studies, the focus of this study is to examine the potential heterogeneous effects of ACA-Medicaid expansion on labor market outcomes. This can provide valuable insights to guide the policy and improve efficiency. For instance, by identifying which local areas tend to experience larger impacts, policymakers can design targeted initiatives to address any adverse effects on employment, if any, without drastically affecting areas that do not experience adverse employment effects. In a broader context, heterogeneous effects of the reform can also aid the accuracy of welfare impacts of the policy. The importance of heterogeneous treatment effect has grown substantially in medical research in the recent years, and this relevance extends to policymaking in healthcare sector. By incorporating heterogeneous treatment effect analysis, this study contributes to a more precise understanding of the ACA-Medicaid expansion's implications, which is crucial for informed and effective policy decisions.

Secondly, the study specifically focuses on the nuances of the American South. While this regional approach affects the external validity of the findings, as previously mentioned, several compelling reasons drive the decision to center the research on this particular region. Importantly, opposition against the Affordable Care Act (ACA) has been notably pronounced in the South, evidenced by the region's reluctance to fully implement the ACA's goals. For example, the majority of Southern states have opted not to expand Medicaid, thereby exacerbating the regional disparity in healthcare coverage.[6] Consequently, the Southern population may be more vulnerable to the impacts of changes in healthcare policies, making it a crucial area for in-depth analysis. This regional context offers a rich landscape for understanding how Medicaid expansion interacts with the unique institutional, socio-economic and healthcare challenges faced by Southern states, providing valuable insights into the broader implications of healthcare policy decisions.

2 Conceptual Framework

Although targeted as a federal policy, states had significant influence on the implementation of ACA. While the early provisions of ACA such as the dependent coverage and removal of copay on preventive cure went into effect in 2010, more comprehensive provisions such as ACA-related Medicaid expansion and the individual mandate were not fully implemented by states until 2014. Following the 2012 supreme court decision that allowed states to decide on Medicaid expansion, by the end of 2018, only 6 out of the 16 southern states implemented Medicaid expansion. Moreover,

[6] Compare map https://www.kff.org/affordable-care-act/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/ for systematic regional variation in Medicaid expansion. Note that Virginia and North Carolina implemented expansion in 2019 and 2023, respectively.

states had the option to set up their own exchange marketplace, partner with the federal government, or depend entirely on the federally operated exchange market. Other important avenues through which states exerted influence include but are not limited to the regulation of insurance plans, Medicaid waivers (Section 1115), and the establishment of outreach programs. In fact, Arkansas's Medicaid expansion is based on the Section 1115 waiver.

A significant body of research has documented the impact of the ACA-Medicaid expansion on reducing uninsured rates following its implementation (Frean et al., 2017; Courtemanche et al., 2017; Simon et al., 2017; Miller et al., 2021). Additionally, a strand of studies examine heterogeneous impacts of the expansion based on various observed factors such as race, age, gender, or pre-reform uninsured rates (Buchmueller et al., 2016; Simon et al., 2017; Wehby and Lyu, 2018; Hamilton, 2024). While these studies generally find that the ACA-Medicaid expansion reduced racial and ethnic disparities in insurance coverage and had greater impacts in areas with higher uninsured rates prior to the reform, the variables leading to heterogeneity in treatment effects are often handpicked and predetermined by researchers. This approach may overlook important variables that illustrate heterogeneity. In a more recent study, Shrestha (2023) suggests that the efficacy of the ACA is influenced by historical factors, notably the legacy of slavery, indicating that areas with a higher proportion of enslaved populations in 1860 benefit the least from the reform. While it is challenging to determine apriori which variables might mediate the differential effects of the ACA-Medicaid expansion, considering the politicized nature of the reform (Lanford and Quadagno, 2016; Grogan and Park, 2017; Michener, 2020), it is plausible that the reform's efficacy is greater in areas with strong Democratic support.

The standard economic theory typically predicts that an increase in public health insurance coverage, such as through the expansion of Medicaid, may suppress labor market activities. This works through the channel of Medicaid expansion potentially increasing employees' bargaining power resulting from higher reservation wages.[7] This phenomenon is particularly prevalent among individuals who are primarily employed to secure private health insurance such as employer-sponsored insurance (ESI) prior to the reform, allowing such individuals to break the "employment lock" following the reform. Furthermore, individuals who are slightly above the eligibility threshold for Medicaid may have an incentive to work less in order to qualify for Medicaid benefits. The phenomenon, known as the "income effect," suggests that as individuals receive more non-wage income — particularly relevant for Medicaid beneficiaries who typically pay no premiums and have low out-of-pocket expenditures — they may choose to work fewer hours or opt for part-time employment to maintain their eligibility for Medicaid benefits (known as the Medicaid notch). This income effect arises if Medicaid expansion effectively increases the overall financial resources available to individuals, thereby altering their incentives and preferences regarding labor market participation.

Specifically, individuals with high disutility of work are more suseptible to reducing their labor market activities following the expansion. The magnitude of the effect can vary across socio-

[7] A new-found insurance can reduce uncertainty, which can reduce an employee's willingness to work for lower wages.

economic background, accessibility to health insurance (prior to the reform), local labor market, macroeconomic conditions, as well as health condition. For instance, individuals with poor health conditions or those tied-up in local labor markets with substandard working conditions are more likely to reduce their participation in labor market. Additionally, local areas that experienced higher improvements in insurance outcomes following the expansion can be more susceptible to reduction in labor market activities. Furthermore, the reduction in labor market may be temporary as the safety-net through Medicaid can provide a platform for better employment prospects through job search and skill development. The potential for a temporary reduction in labor market activity followed by a return to employment highlights the dynamic nature of the impact of Medicaid expansion on individuals' labor market decisions and outcomes. Along these lines, Peng et al. (2020) provide evidence that Medicaid expansion reduced unemployment after a year of the expansion, with labor market activities returning to the pre-expansion period in two years following the reform. Moreover, the prospect of returning to the labor market may depend on macroeconomic conditions, further generating differential effects.

On the other hand, Medicaid coverage can have positive effects on labor market outcomes by improving an individual's health and reducing health-related emergencies. While several earlier studies evaluating the short-term impact on health have failed to detect any statistically significant effects of the reform on health outcomes, more recent studies provide evidence that Medicaid expansion has improved life expectancy outcomes (Chen, 2019; Borgschulte and Vogler, 2020; Miller et al., 2021). The "health effect" can increase the marginal product of labor, thus, increasing earnings and tax revenues. Given these opposing channels at play, how Medicaid coverage affects labor market outcomes and whether heterogeneity exists are questions that need empirical exploration.

3 Data

3.1 Insurance and Labor Market Outcomes

The outcome variables pertaning to uninsured rates and labor market outcomes are measured at the county level. The uninsured rates come from the Small Area Health Insurance Estimates (SAHIE) for years 2010 to 2018 pertaining to individuals between the age group of 19 and 64. SAHIE provides estimates for the overall population as well as for sub-groups defined by the following income categories in relation to the Federal Poverty Level (FPL): i) below the 138% of the FPL, ii) between 138% and 200%, iii) between 200% and 250%, iv) between 250% and 400%, and v) above 400% of FPL. Given that eligibility of ACA-Medicaid expansion is placed at 138% below the FPL, we specifically focus on uninsured rates among individuals below the 138% of FPL. Data for the federal Medicaid-CHIP transfers comes from the Bureau of Economic Analysis (BEA). BEA estimates the total value of the federal Medicaid-CHIP transfers allocated for each county. The federal Medicaid-CHIP transfer per person living under the federal poverty level is constructed from dividing the total transfer by the county-level population living in poverty in year 2010. Data for Medicaid-CHIP transfers are obtained from the Olvera et al. (2023) study replication package.

The labor market outcomes are extracted from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS) Program and Quarterly Census of Employment and Wages (QCEW) for years 2010 to 2018. The LAUS program collects and analyzes data from the Current Population Survey (CPS). Data is then used to estimate and report monthly and annual unemployment rates for different classification of geographical areas including states, Metropolitan Statistical Areas (MSA), counties and cities. We use three annual county-level variables to indicate labor market outcomes: i) unemployment rate (defined as the ratio of unemployed and labor force), ii) the total employment, and ii) the size of the labor force. The yearly estimates are used for the years 2010 to 2018.

Next, QCEW data includes quarterly count of employment and wages that are extracted from the state unemployment insurance programs. The database covers more than 95 percent of U.S. jobs and aggregates the labor market variables at the county, metropolitan statistical area (MSA), state, and national level. QCEW publishes employment level and wages by NAICS industry for more than 3,000 counties in America, which is valuable to inspect the impacts of the reform across industries that are more likely to contain individuals eligible for ACA Medicaid expansion.[8] Following Peng et al. (2020), we focus on the retail and trade (NAICS code 44-45), accomodation and food services (NAICS code 72), construction (NAICS code 23), and manufacturing (NAICS code 1013) sectors. One main limitation of both the LAUS and QCEW data is that the effects on labor market outcomes can only be evaluated at the extensive margin. One cannot evalute the effects on variables capturing the intensity measure at the intensive margin (say, the number of hours worked or the switch between full-time and part-time work). As is case with the previous studies, this limitation restricts the scope of a comprehensive analysis of the impacts of ACA-Medicaid expansion on labor market outcomes.

3.2 Additional county level variables

The information regarding Medicaid expansion is extracted from the Kaiser Family Foundation: *Status of State Medicaid Expansion Decisions* between years 2014 to 2018.[9] Five southern states expanded Medicaid in 2014, including Arkansas, Kentucky, West Virginia, Maryland and Delaware. Louisiana joined the expansion group in July 2016. Virginia and North Carolina have expanded Medicaid in more recent years but these expansions are not pertinent to the study as they are dated after 2018, the last year of the study.

Several additional pre-treatment county-level variables are incorporated into the analysis as features or covariates to enhance the depth of the study. These variables encompass a range of socio-demographic factors (county population, per capita income, percent with high school and college degree, poverty rate, and household median income all in 2010 and household income at age 35 in 2014 from the Opportunity Atlas), long-term climate data (average precipitation and

[8] More information regarding QCEW database can be obtained from `https://www.bls.gov/cew/about-data/data-files-guide.htm`. (Accessed on April 2, 2024).

[9] https://www.kff.org/affordable-care-act/issue-brief/status-of-state-medicaid-expansion-decisions-interactive-map/

temperature between year 1960 and 1990), geographic and environmental characteristics (longitude, latitude, land ruggedness, elevation, and PM 2.5 measure in 2010), a measure of cotton suitability, the malaria stability index, total Hill Burton expenses between 1947 and 1971, and contemporary political outcomes (an indicator for a democrat county, the proportion of White votes for Obama in the 2008 presidential election, and the proportion of Trump votes in the 2016 presidential election). Note that all of the covariates, except for the proportion of Trump votes during the 2016 election, are determined prior to the ACA-Medicaid expansion. A detailed documentation of these variables and their respective data sources is provided in the Appendix section 9.1 for further reference.

3.3 Descriptive results

The summary statistics for the outcome variables used in this study are presented in Table 1. The uninsured rate pertains to people below 138% of FPL. The table shows that uninsured rate drop dramatically by over 13 percentage points from 2013 to 2014. Although some provisions of ACA such as the dependent coverage mandate and prohibition of insurance coverage based on pre-existing condition went into effect in 2010, major components of ACA including the individual mandate, employer's mandate, and Medicaid expansion (through ACA) did not go into effect until 2014. All of these policies may explain the significant drop in uninsured population. It is important to highlight that over 20 percent of the population remained uninsured in 2014 and subsequent years. This persistent and alarmingly high uninsured rate, particularly among individuals falling below 138% of the Federal Poverty Level (FPL), underscores a critical issue: the prevalence of the "coverage gap" . People who typically fall in the "coverage gap" are those who do not qualify for subsidies on the exchange market but are ineligible for Medicaid coverage due lack of Medicaid expansion. Next, the labor market outcomes show improvements over the reported years as the economy moved away from the Great Recession. The unemployment rate gradually decreased from 10.3 percent in 2010 to 4.3 percent in 2018. This is consistent with the rise in the average number of employed individuals at the county level.

Figure 1 provides a more detailed breakdown of the statistics categorized by counties belonging to expansion and non-expansion states. Panel A shows that prior to 2014, uninsured rates (among those below 138% of the Federal Poverty Level) followed similar trends across both expansion and non-expansion units. However, a notable disparity emerged in 2014, with uninsured rates in expansion units declining by nearly 15 percentage points. In contrast, non-expansion units experienced a more gradual reduction in uninsured rates, albeit to a lesser extent. Panel B documents a reduction in unemployment rate across both expansion and non-expansion units. Panel C shows increases in the average number of employed people in counties over the years with such increments more pronounced among counties belonging to non-expansion states. Panel D shows the trend in labor force over time across the two groups. Both panels C and D exhibit a greater degree of labor market involvement in counties within the non-expansion states, which may be explained by a larger population size. On average, non-expansion states reports a significantly higher population size in 2010, with an average of 85,798 individuals per county, compared to 55,044 individuals per county in expansion states.

4 Method

The objective is to examine the varying impacts of ACA Medicaid expansion on insurance and labor market outcomes.

4.1 Event Study Specification

We first investigate the impacts of the reform on insurance and labor market outcomes in the American South using the standard event study framework. While this approach has been widely adopted in the realm of program evaluation literature, including evaluation of ACA, it serves as an arguably appropriate benchmark. This is particularly relevant as the causal forest approach modified for the context of this study (discussed in upcoming sections) shares assumption similar to the difference-in-differences framework. The event study specification is given as:

$$Y_{cst} = \alpha + \sum_{\underbrace{j=-4}_{j\neq -1}}^{5} \beta_j \times Expand_{cs} \times I(t-E=j) + \eta_c + \sigma_t + \epsilon_{cst} \quad (1)$$

where, Y_{cst} is the outcome variable pertaining to county c within state s in year t. $Expand_{cs}$ represents whether state s expanded Medicaid through ACA between the years 2010 and 2018 (sample period). The county and year fixed effects are represented by η_c and σ_t, respectively. The Medicaid expansion year is given by E and the indicator, $I(.)$, tracks the relative time from the expansion year. Note that the omitted category used is the year prior to the expansion year when the outcome is uninsured rate and two-years prior to the reform for labor market outcomes. A year prior to the reform is used as an adjustment period to allow for changes in labor market decisions in anticipation of ACA. This matter is discussed in more detail in section 4.5. The standard errors are clustered at the state level when estimating specification 1 to account for the correlation between the error terms at the state level.

The coefficient on β_j for $j = \{0,\ 1,\ ...,\ 5\}$ captures the dynamic effect of the expansion under the parallel trend assumption. Formally, using the potential outcome terminology, the parallel trend assumption can be formally written as:

$$E(Y^0_{ict} - Y^0_{icb}|Expand = 1,\ X_{ic}) = E(Y^0_{ict} - Y^0_{icb}|Expand = 0,\ X_{ic})\ \forall\, t \quad (2)$$

This illustrates that in absence of Medicaid expansion, the difference between average outcome in time t and base year b, would have been the same in the expansion and non-expansion areas. Equation 2 invokes counterfactual on states that expanded Medicaid (LHS), whereas actual outcomes are observed for non-expansion states (RHS). This only allows identification of the average treatment effect on the treated (ATT) rather than average treatment effect (ATE). As LHS of equation 2 is never observed, actual validation of the assumption is not possible. Typically the coefficient on β_j, $j < 0$, is used as a suggestive test for the parallel trend assumption, with coefficients close zero conventionally accepted as being supportive of the parallel trend.

4.2 Causal Forest

We use the causal forest (CF) methodology, developed within the Generalized Random Forest (grf) framework in Athey et al. (2019), to assess heterogeneous effects of Medicaid expansion. In essence, CF uses adaptive weights based on random forest to evaluate treatment effects, systematically adjusting the importance of observations to serve as comparisons to the target point (x) by employing the R-learner framework at the core. At a high level, in a causal tree observations falling in the same leaf as the target sample are the ones closest to the target point (Wager and Athey, 2018; Athey et al., 2019). Given that all features determining the treatment assignment are observed, variation in treatment across observations within a leaf is as good as random. This allows estimating treatment effect $\tau(x)$ for a test point x using treated and untreated observations in a neighborhood represented by a leaf.

The R-learner framework (Nie and Wager, 2021) is inspired by the decomposition originally proposed in Robinson (1988) to estimate parametric components in a partially linear models. Given the observed data $(X_i,\ Y_i,\ W_i)$ for $i = (1,\ 2,\ ...,\ n)$ where X_i is per-unit features or covariates, Y_i is the outcome, and $W_i = \{0,\ 1\}$ is the treatment status, the outcome can be expressed as a function of conditional means for both treated and untreated units as given below:

$$Y_i(0) = \mu_0(X_i) + \epsilon_i(0),\ for\ W = 0$$

$$Y_i(1) = \mu_0(X_i) + \tau(X_i) + \epsilon_i(1),\ for\ W = 1$$

$$Y_i = \tau(X_i) \times W + \mu(X_i)_0 + \epsilon_i \tag{3}$$

where, the first and the second lines pertain to the control and treated groups, respectively. The third line combines the first two lines together. Next, the R-learner (Nie and Wager, 2021) operationalizes the centered model, which is written as:

$$Y_i - m(X_i) = \tau(X_i)(W_i - e(X_i)) + \epsilon_i \tag{4}$$

Y_i is outcome of interest. The main effect, $m(X_i)$, is the conditional mean outcome and is defined as $m(x) = E[Y_i|X_i = x] = \mu_0(X_i) + e(X_i)\tau(X_i)$. Here, $\mu_0(X_i)$ is the baseline conditional expectation of the outcome without the treatment or the outcome regression. W_i as the treatment denotes whether a county belonged to the expansion state, and the propensity score is given as $e(x) = P(W_i = 1|X_i = x)$. In this context, $m(X_i)$ and $e(X_i)$ are the nuisance parameters that need to be estimated. $\tau(X_i)$ is allowed to vary with X_i or the features. This very aspect is the focal point for heterogeneous treatment effects. The linear specification given in 4 is estimated locally where the treatment effect is constant within a covariate space, determined by a causal tree.

$\tau(X_i = x)$ is identified as the conditional treatment effect (CATE) under the unconfoundedness assumption: $Y_i^{(0)}, Y_i^{(1)} \perp W_i|X_i$, where $Y_i^{(0)}$ and $Y_i^{(1)}$ are the potential outcomes. In other words, controling for covariates makes the treatment assignment as good as random. However, even if all covariates that determine the treatment assignment are available, trees may not be able to fully condition out all covariates in finite samples (Jacob, 2021). As such, a leaf may contain control and treatment observations that are systematically different. To alleviate such a concern, the observed confounding factors are directly orthogonalized from the outcome variable, resulting in $Y_i - m(X_i)$ on the LHS.

The estimated parameter of interest may not converge (in probability) to the true parameter due to regularization bias when estimating nuisance parameter $(m(X))$ as shown in Chernozhukov et al. (2018). The auxiliary prediction to estimate the conditional mean of X on W (propensity score) purges out the effect of covariates on treatment allocation, allowing for the residual-on-residual regression as given in equation 4. This approach of "double/debiased machine learning" not only more directly accounts for confounding variables but also has been shown to be effective in mitigating regularization bias. This bias typically arises from the use of ML procedures to estimate nuisance parameters (propensity score and outcome regression) for informative learning, balancing regularization bias to alleviate overfitting (Chernozhukov et al., 2018).

One challenge is that both $m(x)$ and $e(x)$ are not known in practice and need to be estimated. As discussed in Chernozhukov et al. (2018), directly plugging in estimates of $m(x)$ and $e(x)$ using ML approach can lead to over-fitting as the same sample is used for both training and estimation purposes. The authors suggest cross-fitting to overcome the problem of over-fitting using random sample splitting, where an auxiliary sample is used to estimate the nuisance parameters and the parameter of interest is estimated in the main sample. Additionally, in context of causal forest, using the same sample for the purposes of tree-building and estimating the treatment effects can result in inconsistent estimates due to overfitting. To address this issue, Athey et al. (2019) propose using *honest splitting*. Prior to building a tree structure, this approach randomly divides the sample into two halves. One segment of the sample is used to train the model (build trees) and the other is used to cast predictions (estimate treatment effects). Both $m(x)$ and $e(x)$ are estimated using random forests with *honest splitting* where predictions are carried out using the out-of-bag (OOB) sample. As such, the aspects of model building and estimation of treatment effects are independent from each other. By its nature, this approach integrates cross-fitting to overcome challenges in ML due to over-fitting.[10]

The important aspect of CF is the use of random forest as an adaptive neighborhood finder for the test sample x. While the standard random forest model for regression typically performs

[10] We also use boosted regression forest for estimating $m(x)$ and $e(x)$, where predictions are obtained from K fold sample splitting, $K = 10$, and replicate the entire findings of the study. The results are unchanged from the findings obtained from using random forest with *honest splitting*. These findings are available upon request.

a split to maximize the difference in means across two child nodes such as Breiman et al. (1984)'s Classification and Regression Trees (CART) algorithm, CF algorithm encodes the need to maximize the difference in treatment estimates when splitting a parental node. Theoretically, this means that for each potential axis aligned split that extends from the parent node, one would need to estimate treatment effects at two of the child nodes (τ_L and τ_R) and choose the split that maximizes the squared difference between child specific treatment effects. However, in practice this is highly computationally demanding and infeasible. The application of causal forest estimates τ_P at the parent node and uses the gradient based function to guide the split. At each (parent) node the treatment effect is estimated only once.[11]

Given the test point x, the goal is to provide higher weights to observations that are similar to x and lower weights to those that are not similar when estimating equation 4. The tree specific weight for a training observation i at the b^{th} tree is given as: $\alpha_{ib}(x) = \frac{1(X_i \in L_b(x))}{|L_b(x)|}$, where $L_b(x)$ is the leaf (neighborhood) that consists the test sample x. The forest is composed of B trees and the forest specific weight for an example i is given as: $\alpha_i(x) = \frac{1}{B}\sum_{b=1}^{B} \frac{1(X_i \in L_b(x))}{|L_b(x)|}$.[12] It measures the fraction of times an obsevation i falls on the same leaf as x in the course of the forest. Simply, it shows how similar X_i is to x.

Once the adaptive weights pertaining to the test point x are determined using random forests, a weighted least square (WLS) is performed on equation 4 using the weights. Note that for a new-point x_{new}, weights are re-estimated before estimating equation 4. This process is carried out using the GRF package in R.[13] As the treatment occurs at the state level, cluster random sampling is applied at the state level to generate a bootstrap sample for each tree.[14]

4.3 A Modified Causal Forest Approach

The causal forest (CF) approach poses two primary challenges within the scope of this study. Firstly, it is tailored for cross-sectional settings, thus lacking the capacity to accommodate the panel structure inherent in the data. Secondly, the assumption of unconfoundedness may prove to be overly stringent within observational settings. To address these challenges, we employ a slightly adapted version of the causal forest method, aligning it more closely with the framework of the difference-in-differences model. This modification serves to alleviate concerns regarding the identification assumption governing CF, shifting the basis of identification from unconfoundedness to the parallel trend assumption. As such, the modified causal forest approach builds on the parallel trend assumption of the difference-in-differences framework.

The main modification arises from the first differencing of the outcome variable in relation to the base year. The *no anticipation* assumption, which states that individuals are unable to

[11] Providing details of how gradient based functions are used to create pseudo outcomes at the parental nodes is beyond the scope of the discussion and the readers are directed to the Athey et al. (2019) study for details.

[12] The weights sum up to 1.

[13] The comprehensive discussion of the package can be found in `https://grf-labs.github.io/grf/`.

[14] For this process, the whole state (cluster) is selected randomly and random sampling is performed within the cluster.

anticipate the implementation of Medicaid expansion and thus repond accordingly prior to the implementation, allows using a year prior to the expansion as the base year. The validity of the *no anticipation* assumption is uncertain as individuals might influence treatment assignment by adjusting their labor market behaviors in anticipation of policy changes, such as Medicaid expansion eligibility criteria, in the years leading up to the expansion. To capture the possibility of *anticipation effect* of the policy we assume that individuals may anticipate treatment one year before the expansion. As such, we use the year prior to the policy as an adjustment period when individuals may change labor market activities in anticipation of the expansion. Thus, two years prior to the expansion year is used as the base year when examining the labor market outcomes. For instance, if the ACA Medicaid expansion was implemented in 2014, the year 2012 is used as the base year. Since insurance through Medicaid coverage is unlikely to change prior to the implementation of Medicaid expansion, a year prior to the reform is used as the base year for insurance outcome.

We modify the outcome variables to measure the difference in outcome in county i between year $t = (2010, ..., 2018;\ s.t.\ t \neq 2012)$ and 2012 (2013 for insurance outcome). ΔY_i for county i represents five outcomes: i) the change in uninsured rate among people below 138% of FPL, ii) the change in Medicaid transfers per capita, iii) the change in employment rate, iv) the change in total employed individuals, and v) the change in labor force. The first differencing differences out the time invarant unobserved heterogeneity in outcomes across counties by construction similar to the difference-in-differences framework.[15] Following this modification, the identification of causal forest relies on the parallel trend assumption. Explicitly this states that the trend in treatment units would mimic that of control units within a given leaf in absence of the treatment. A regular causal forest is estimated using the modified outcome variables for each sample year separately. This is similar to estimating a causal forest in the full sample where the first split is forced on time periods, before considering other splits.

The modified causal forest approach proposed here is similar to the "dynamic causal forest (DCF)" method discussed in Gavrilova et al. (2023). The authors show that the DFC approach identifies consistent estimate for the treatment effect under the assumptions of parallel trend and overlap. The parallel trend assumption disscussed previously allows estimation of the conditional average treatment effect on the treated (CATT) analogous to the average treatment effect on treated (ATT) under the difference in differences framework. This limits estimation of treatment effect only among units in the expansion states. However, more direct policy relevant question may ask how labor market activities would change if areas without the expansion (control units) were to implement the expansion policy. To investigate this we invoke an additional assumption that trends in outcomes among control units would evolve similar to the treatment units had they received treatment. This allows estimation of CATE. Next, the overlap assumption states that for each X_i the probability of treatment is between 0 and 1 ($0 < P(W_i|X_i) < 1$). Besides the parallel trend assumption all other assumptions used by Gavrilova et al. (2023) are similar to the Athey

[15] For example, time invarant regional disparity in labor market outcomes will be differenced out using this differencing approach.

et al. (2019) study.

4.4 Obtaining ATE estimate from CATE

Rather that simply averaging the CATE estimates, the estimate for ATE is obtained by summarizing CATE estimates using the Augmented Inverse Probability Weighted (AIPW) estimator given as:

$$\begin{aligned} \tau_{AIPW} &= \frac{1}{n}\sum_{i=1}^{n}(\mu(X_i,\ 1) - \mu(X_i,\ 0) + W_i.\frac{Y_i - \mu(X_i,\ 1)}{e(X_i)} - (1 - W_i).\frac{Y_i - \mu(X_i,\ 1)}{1 - e(X_i))} \\ &= \frac{1}{n}\sum_{i=1}^{n}(\tau(X_i) + W_i.\frac{Y_i - \mu(X_i,\ 1)}{e(X_i)} - (1 - W_i).\frac{Y_i - \mu(X_i,\ 1)}{1 - e(X_i))} \\ &= \frac{1}{n}\Gamma_i \end{aligned} \tag{5}$$

where, $\mu(X_i,\ W_i)$ represents the conditional means at each treatment arm, $\mu(X_i,\ W_i) = E(Y_i|X_i = x, W_i = w)$ for $W_i = \{0,\ 1\}$, and $e(X_i)$ denote propensity scores. A well known propery of AIPW approach is its double robustness in the sense that $\hat{\tau}_{AIPW}$ is consistent if either $\hat{\mu}(X_i,\ W)$ or $\hat{e}(X_i)$ is consistent. The unit specific CATE estimate given by $\tau(X_i)$ can be noisy point estimates. AIPW approach alleviates biases by applying Inverse Probability Weighting (IPW) to the residuals, $Y_i - \mu(X_i,\ W_i)\ for\ W \in \{0,\ 1\}$.

A plug-in approach is used to obtain an estimate for τ_{AIPW}, with $\hat{\tau}(x)$ estimated using CF. The nuisance parameters $e(X_i)$ and $\mu(X_i,\ W_i)$ are both estimated using separate random forests regression with *honest splitting*. $\hat{\tau}_{AIPW}$ is the average of the estimated scores, $\hat{\Gamma}_i$. The linear projection of covariates is a regression of $\hat{\Gamma}_i$ on covariates, which depicts the suggestive influence of the covariates on CATE.

To first inspect the role of propensity score adjustment in improving comparision between the treatment and control units, Figure I1 exhibit absolute versus the adjusted difference in means using the standardized version of X_is. The absolute difference represents the simple difference in mean between treatment and control groups, whereas the adjusted differences use the inverse of propensity scores to adjust for differences in treatment and control units ($\frac{1}{N_T}\sum\frac{X_i \times I(W_i=1)}{e(X_i)} - \frac{1}{N_C}\sum\frac{X_i \times I(W_i=0)}{1-e(X_i)}$). As shown in the figure, differences in covariates shrink dramatically when units are adjusted using the inverse propensity scores.

4.5 Investigating heterogeneity using Rank-Weighted Average Treatment Effects (RATE)

Our interest of the study is to determine whether the conditional treatment effect estimates (CATE), denoted by $\hat{\tau}(x)$, demonstrate a significant magnitude of heterogeneity. Specifically, we examine the effectiveness of ranking based on CATE estimates in detecting heterogeneity by using the

RATE metrics as developed in Yadlowsky et al. (2021). The motivation behind RATE is to fit a heterogeneous treatment effect model based on a score measure to provide "prioritization rule" that can distinguish units with the most treatment benefit. The priority score, $S(.)$, is provided by the user and can include CATE estimates (learned separately), baseline risk measure, or other baseline characteristics. The study uses CATE estimates, estimated using the modified CF framework, as the priority score.

The utilization of RATE to assess heterogeneity in ACA-Medicaid expansion reform aligns well with the motivational viewpoint as disscussed in Yadlowsky et al. (2021). Decision-makers frequently weigh intervention benefit against costs when designing final policies. For instance, prioritizing individuals with greater treatment benefits, such as drugs for medical treatment, could conceptually streamline policy design. In the context of this study, the objective is to identify whether the benefits on insurance outcomes and potential adverse labor market effects are concentrated in certain areas. This nuanced analysis offers policymakers invaluable insights into tailoring the reform's efficacy according to county-level characteristics. For example, by identifying areas with heightened labor market distortions, policymakers can tailor targeted interventions such as job training programs to mitigate adverse impacts and enhance overall economic outcomes.

In sum, RATE metric uses the Targeting Operator Characteristics (TOC) and area under the TOC (AUTOC) to characterize heterogeneity. TOC is defined as:

$$TOC(q) = E[Y_i(1) - Y_i(0)|S(X_i) > F^{-1}_{S(X_i)}(1-q)] - E[Y_i(1) - Y_i(0)] \tag{6}$$

In other words, $TOC(q)$ for $0 \leq q \leq 1$ is defined as the difference in ATE among units above the q^{th} percentile of $S(X_i)$ and the overall ATE. ATEs are estimated using the scores, Γ_is, estimated using the AIPW approach as shown in equation 5. In presence of significant heterogeneity across the priority score, $TOC(q)$ is significantly higher in magnitude for the lower values of q, while approaching towards 0 as q gets larger. $TOC(q = 1)$ is simply 0. One measure to summarize TOC curve is to calculate area under the curve. Formally, RATE is defined as the area under the TOC curve (AUTOC).

In this study $S(X_i)$ needs to be learned as $S(X_i) = \hat{\tau}(X_c)$ (CATE). The training set is used for the estimation of $\hat{\tau}(X_i)$ and RATE evaluation is performed in the evaluation set. As such, the same data is not used for the estimation as well as evaluation to prevent over-fitting.

5 Results

5.1 Effects on the average

We start with the event study results on insurance and labor market outcomes estimated using specification 1. Subfigures A and B in Figure 2 present the dynamic effects of Medicaid expansion on uninsured rate (subfigure A) and per capita Medicaid funds (subfigure B). Data for the former comes from SAHIE, while the latter variable is extracted from BEA. The omitted category represents one period prior to the Medicaid expansion. In subfigure A, the coefficients reveal a notable decrease in the uninsured rate among individuals below the 138% FPL in expansion states compared to non-expansion states following the implementation of Medicaid expansion. On average, uninsured rates decreased by 12 percentage points during the expansion year among those individuals below 138% of FPL residing in expansion states, with slightly higher magnitudes observed in subsequent years. Consistent to such findings, subfigure B shows a rise in Medicaid transfer funds in counties belonging to the expansion states after the policy, with per capita Medicaid amount on average increasing by over $0.5. Moreover, the estimates prior to the policy year in the top two panels are close to zero and statistically insignificant at the conventional levels. This provides suggestive evidence in support of the parallel trend assumption used for identification in the difference-in-differences framework.

Subfigures C and D present findings for labor market outcomes. Following the discussion provided earlier, the omitted period used in labor market cases is two years prior to the policy implementation year to account for potential pre-policy adjustments in labor market activities a year before the policy. In subfigure C, examining the unemployment rate fails to reveal any distinct pattern in coefficients post-expansion year. Notably, these coefficients represent the average treatment effect on the treated (ATT) estimates, which may obscure nuances offered by the estimation of heterogeneous effects. Subfigure D delves into the total number of employed individuals and the size of the labor force. The findings indicate a relative reduction in both outcomes post-Medicaid expansion, with these impacts intensifying over subsequent years. Specifically, during the policy year, the size of the labor force decreased by an average of 800 individuals. Remarkably, this impact more than doubled in magnitude four years after the policy, leading to a reduction of little over 2,000 individuals in the labor force. Such a pattern is suggestive of dynamic and increasing labor market distortions following Medicaid expansion. While the findings presented here are not directly comparable to those from Peng et al. (2020) due to differences in methodology and specific area of investigation, the results document that the employment distortions are not transient but rather become more pronounced over time.

Next, we turn to the results summarized by the average treatment effect estimates using the augumented inverse probability weighting (AIPW) approach following the CF estimation with 10,000 trees. The findings are presented in Figure 3. Panels A and B use changes in uninsured and Medicaid transfers per capita (in relation to the year 2013), Panel C uses percentage points change in unemployment rate, and Panel D uses changes in the total employed and labor force participation (in relation to the year 2012). Panel A shows reductions in uninsured rates following the expansion, with counties within the expansion states on average experiencing 12 percentage points drop in the uninsured rate among people below the 138% of the FPL in the year 2014. These effects hover in between 12 and 15 percentage points for years the 2015 to 2018. The confidence intervals illustrated by the error bars show that the ATE estimates are statistically significant at least the 5% level. Moreover, the ATE estimates pertaining to the years prior to ACA are precisely estimated at zero, showing absence of any pre-existing differences in insurance outcome between the treatment and control groups prior to the expansion conditional upon the covariates. The findings in Panel B is consistent with Panel A, indicating increases in per capita Medicaid transfer funds following the expansion year. For the most part, the ATEs obtained from the AIPW approach using CATE estimates, obtained from CF, are noticably similar to the event study estimates shown in Figure 2.

Turning to the analysis of unemployment rates, panel C reveals that there are no statistically significant changes in unemployment rate following the reform. The ATE estimates are very close to zero. Moving on to panel D, the figure shows relative reductions in labor force participation and total number of employed people in counties belonging to the expansion states following the policy implementation. On average the reductions in labor force and total employment ranges from 500-2,000 individuals between 2014 and 2018. Most of the estimates are significant at the 5% level. It is important to note that the estimates pertaining to labor market variables should be perceived as the Intent-to-Treat (ITT) estimates. This is because the measured outcome reflects the overall population, including those unaffected by the Medicaid expansion, rather than solely those directly impacted.

Although structured in similar ways, the interpretation of event study estimates in Figure 2 and results summarized using the CF approach in Figure 3 differ in terms of interpretation given the underlying identification assumptions. First, the event study estimates resemble the average treatment effect on the treated given the assumption that outcomes within the treatment group would have evolved similar to the control units in absence of the expansion. As discussed in section 4, an additional assumption involving the counterfactual for the control units such that outcomes in control units would trend similar to the treatment units in presence of treatment allows estimating the average treatment effect (ATE). As such, findings in Figure 3 resemble the ATE estimates.

5.2 Accessing heterogeneity by important features

In the context of causal forests (CF), an important facet is the ability to assess the importance of the features used for splitting the trees. This measure helps trace the significance of the variables in maximizing heterogeneity in treatment effects. By identifying these key variables, researchers can gain suggestive insights into the underlying mechanisms or channels through which interventions, such as Medicaid expansion, impact response variables.

In Figure I4, subfigures display the variable importance plots for the respective response variables as indexed in the y-axis and referenced below the figure. The vertical line marks the 5 percent importance threshold. Variables with importance measures exceeding the 5 percent are listed in Table 2 for the corresponding response variable. The infant mortality rate from 2010 to 2013 and longitude of the county's centriod explains 6 percent of the split for changes in uninsured rate. Uninsured rate in 2013 accounts for 9 percent of heterogeneity. Notably, the proportion of Whites who voted for Obama in the 2008 presidential election accounts for 9 percent of the treatment heterogeneity, which aligns with the highly politicized nature of the ACA and the predominant opposition from Republicans. Interestingly, the proportion of Whites' votes for Obama is also responsible for explaining 5 percent of heterogeneity when unemployment rate is used as the dependent variable.

The total count of employed people in 2010 account for 9 and 8 percent of treatment heterogeneity for the total employed and labor force.

To assess heterogeneous effects of Medicaid expansion we conduct event study exercises by dividing counties into groups above and below the median of three pertinent features explaining heterogeneity: *i*) the proportion of White votes for Obama in the 2008 election (results shown in Figure 4), *ii*) longitude of the county's centriod (Figure 5), and *iii*) the count of employed people in 2010 (Figure 6). Subfigure A in Figure 4 shows that while uninsured rate decreased following the reform for both groups, the reduction in uninsured rate is about 2 percentage points higher for counties with above-median White votes for Obama in the year of the expansion and the subsequent year. This gap gradually subsides and nearly disappears by the fourth year of the reform. However, differences across high versus low voting groups are not statistically significant, as the confidence intervals overlap. Moreover, the pre-reform estimates are close to zero and statistically insignificant. Subfigure B shows no discernible differences in unemployment rate between counties with high and low White votes for Obama. Subfigures C and D show that the relative decrease in total employment and the size of the labor force is significantly higher in counties where Whites favored Obama. For example, the relative decline in total employed is sightly over 2,500 during the expansion year for counties with high White votes for Obama, compared to less than 500 in counties with low votes for Obama. Moreover, the magnitude in decline grows over the years for both groups, but the decrease is much steeper for counties with Whites favoring Obama. Subfigure D reflects similar patterns as observed in subfigure C. These findings should be interpreted with caution since the event study estimates prior to the reform show some pre-trends and are statistically different from zero.

Next, subfigures in Figure 5 show similar effects in outcomes across counties with above-median versus below-median longitude measure except for unemployment rate, as illustrated in subfigure B. While counties below median longitude value show no effects on unemployment rate following the reform, the unemployment rate increases in counties with above-median longitude value. Moreover, the estimates pertaining to periods prior to reform are close to zero.

Figure 6 presents assessment of treatment heterogeneity by above-median versus below-median groupings based on the total count of employed people in 2010. The reduction in uninsured rate following the reform is consistently higher for the counties in below-median employment group. This may be driven due to lower baseline uninsured rate among counties with above-median total employment. Subfigure B shows that unemployment increased gradually in counties with above-median total employment. Both subfigures C and D show a relative supression in labor market outcomes concentrated in counties with above-median total employment in 2010, as shown by the relative decline in both total employment as well as labor force. Notably, none of the pre-reform estimates are significantly different from zero, providing suggestive evidence in favor of the identification assumption.

As previously discussed, CF approach allows prediction of treatment effects for each county based on their covariates, which are in fact the conditional average treatment effect (CATE) estimates. The histogram of raw CATE estimates are presented in the Appendix figures I5, I6, I7, and I8 for the outcomes defined as the first difference in uninsured rate for people below 138% of FPL, unemployment rate, total employed, and labor force participation. The CATE estimates for the uninsured rate ranges between -17 and -7 in year 2014 and distributions of the estimates are discernibly bimodal for years 2014, 2015, and 2016. The histogram bars for CATE estimates pertaining to unemployment rate in I6 predominantly fall on the positive segment of the real line, except for the year 2016, where the distribution is centered around zero. Next, the distributions of CATE estimates for both the count of total employed and labor force are left tailed and predominantly fall on the negative segment of the real line as shown in figures I7 and I8, suggesting that ACA-Medicaid expansion suppressed labor market outcomes in the American South.

5.3 Assessing heterogeneity based on the ranking of CATE estimates

Figure 7 evaluates heterogenous effects based on CATE estimates for uninsured rate using the Rank Weighted Average Treatment Effect (RATE) approach as previously discussed in section 4.5. All subfigures pertaining to years 2014-2018 demonstrate some levels of heterogeneity in treatment. For example, panel A suggests that counties ranking in the top 10% of the distribution of CATE estimates (in terms of magnitude) exhibit 13 percentage points more of reductions in uninsured rate in 2014 compared to the ATE estimate. This indicates that given the ATE estimate of -12 percentage points reduction (panel A, Figure 3, year 2014), counties that benefitted most from the treatment experienced a drop in uninsured rate by as much as 25 percentage points in 2014. The incidence of heterogeneity is prevalent across the reported years, although the level is notably higher for the years 2014, 2015, and 2016. The AUTOC estimates summarizing the TOC, displayed at the bottom of subfigures alongside the standard errors, indicate that the presence of heterogeneity is statistically significant at the 10 percent level for most of the reported years. This suggests that there are notable variations in the effects of Medicaid expansion in providing access to insurance across different subgroups. Similarly, the subfigures Figure I9 depict heterogeneity in Medicaid transfers per capita between the reported years and the year 2013. The TOC curves are downward sloping suggesting that some areas received higher Medicaid transfers (per capita) compared to others. However, the confidence intervals are slightly over the zero line across all the reported years.

Figure 8 presents findings from the RANK approach when the response variable is the change in the unemployment rate over the years (compared to the year 2012), obtained from BLS. Subfigure A, for the year 2014, shows that unemployment rate increased by as much as 2 percentage points among the most affected counties (i.e., those in the top 10% bracket of CATE estimates). The AUTOC measure summarizes an increase in the unemployment rate of almost 1 percentage point, with the estimate statistically significant at the 5 percent level. Notably, the estimates pertaining to the average, shown in subfigure C of Figures 2 and 3, display no effect on unemployment rate during the year of the reform. This observation underlines the importance of the nuanced analysis provided by the CATE and AUTOC approaches, which uncover differing effects that might not be immediately apparent in the overall estimates referring to the average. More importantly, similar patterns in the TOC curves are observed for the subsequent years, as shown in subfigures B-E, with the AUTOC measure gradually increasing in magnitude. This indicates that the labor market distortions may have been long-lasting.

Figures 9 and 10 display TOC curves, focusing on changes in the total employed individuals and labor force participation as the respective outcome variables. The TOC curves, shown in Figure 9, are upward sloping. This documents differing effects on labor market across counties. For instance, the employment numbers dropped by around 6,000 individuals in 2014 among counties in the top 10% of the distribution of CATE estimates. The sub-figures for years 2015-2018 show similar pattern as to 2014. However, it is essential to consider the precision of the CATE estimates and that provided by AUTOC method. For all of the years examined (2014-2018), the confidence intervals include zero and the AUTOC estimates appear to be statistically insignificant at the 10 percent level. The results showcased in Figure 9 are further reinforced by examining the change in labor force as the outcome variable in Figure 10, where the pattern in the TOC curves for labor force participation closely resemble those observed for total employment. While the TOC curves are imprecise, the underlying patterns are quite revealing. When considered collectively, Figures 9 and 10 show that the decline in the number of employed individuals is primarily attributed to exits from the labor market.

Next, I turn to results derived from analyses utilizing labor market variables from QWEC dataset. As previously mentioned, QWEC draws its data directly from state unemployment insurance tax records. In contrast, the labor market variables provided by the BLS are based on surveys, specifically the Current Population Survey (CPS). As such, QWEC data may offer enhanced reliability compared to BLS, as the latter is suseptible to sampling error.

Figure 11 displays findings pertaining to the total employment following the reform. Similar to the results obtained from BLS data, the TOC curves lie on the negative quadrant and are upward sloping. The AUTOC estimates indicate increasing levels of heterogeneity documented by the drop in the total employment across the reported years. For comparison, counties falling within the top 10% of CATE distribution experienced a drop in employment count by close to 6,000 in 2014, while the number rose to over 20,000 in 2018. Additionally, it should be noted that the results obtained from QWEC dataset exhibit similar AUTOC measure across the reported years but with greater precision compared to the findings from BLS dataset. As discussed earlier, this discrepancy may be the result of sampling error in the BLS data.

QWEC data allows investigation at the industry level, providing opportunities for more nuanced analysis. Figure I3 in the Appendix illustrates the average effects obtained by summarizing the CATE estimates from causal forest using the AIPW approach for the retail-trade, food and accommodation, construction, and manufacturing sectors in subfigures A-D. The estimates shown in subfigures A and B reveal a pattern of reduced total count of employed people in retail and food sectors following the reform. Conversely, such labor market distortions are not apparent in construction and manufacturing sectors in subfigures C and D.

Figures 12, 13, 14, and 15 document heterogeneous impacts on retail trade, accomodation and food services, construction, and manufacturing sectors, respectively. The results in Figures and 12 and 13 document that employment reduced in food/accomodation and retail/trade sectors. The sub-figures in Figure 12 show that the employment counts decreased among the most affected counties (in the top 10% of CATE distribution) in accomodation and food industry by nearly 1,500 jobs in 2014. This reduction steadily progressed, culminating in a decrease of over 4,000 jobs by 2018. Moreover, the AUTOC estimates are negative and precisely estimated at least at the 10 percent level of significance. However, no significant impacts are exhibited for construction and manufacturing sectors, as shown in Figures 14 and 15. Overall, these findings align with the results that the retail/trade, as well as food services sectors, harbor the largest percentage of low-wage workers. In contrast, the construction sector consists of the smallest proportion of low-wage workers.[16]

6 Additional Results

To evaluate the sensitivity of CF method to changes in an approach used for estimating the nuisance parameters, we re-analyze the CF model using the boosted regression to estimate both $m(x)$ and $e(x)$. Figure I10 replicates the balance plot, using the absolute versus adjusted differences in means of covariates across the treatment and control units. The figure clearly shows that the adjusted differences in means shrink to zero once weighted by the inverse propensity score obtained from the boosted regression forest. This indicates an effective balancing of covariates, similar to what was achieved with the initial random forest approach.

Furthermore, the results of the auxiliary analysis, presented in Figures I11 to I15, reinforce our primary findings. These figures show consistent patterns in the estimated Average Treatment Effects (ATE) and Conditional Average Treatment Effects (CATE) when using boosted regression as opposed to random forests. The robustness of the results across different methods of estimating nuisance parameters underscores the reliability of the CF method in capturing the heterogeneous impacts of the ACA-Medicaid expansion. Specifically, the figures illustrate that the reduction in uninsured rates, the observed increases in unemployment rates, and the decreases in total employment and labor force participation are consistent regardless of the method used for estimating $m(X_i)$ and $e(X_i)$. This consistency strengthens the credibility of our findings and suggests that the CF method is not overly sensitive to the choice of technique for estimating nuisance parameters.

7 Conclusion

One primary argument against the expansion of Medicaid program has been its perceived status as an entitlement program and the potential disincentives to work as predicted by the standard economic theory. This has prompted policymakers to consider work requirement policies as a

[16] According to 2022 estimates published by BLS, 7.1 percent of the total workers at or below the minimum wage falls in the construction sector, whereas 15.7 falls in wholesale and retail trade sector. Compare https://www.bls.gov/opub/reports/minimum-wage/2022/home.htm (Accessed on April 4, 2024).

means to tie coverage to employment criteria. While the implementation of such work requirement proposals have faced legal challenges and blocks, the matter is far from closed and is likely to be influenced by changes in political landscape.

Past studies evaluating the average impacts of ACA Medicaid program on labor market outcomes have generally found little to no effect (Frisvold and Jung, 2018; Gooptu et al., 2016; Kaestner et al., 2017; Leung and Mas, 2018), although some suggest transient adverse impacts (Peng et al., 2020). As the conversation continues, it is evident that the balance between access to healthcare and employment incentives will remain a pivotal part of ACA Medicaid expansion. Thus, gaining nuanced understanding of Medicaid expansion on labor market outcomes is crucial for informed policy decisions moving forward.

In this study, we move one-step further than the "average impact" and assess the heterogenous impacts of the ACA Medicaid expansion on insurance coverage and labor market outcomes using the Causal Forest approach to derive conditional average treatment effect (CATE) estimates. The study's specific focus in the American South is driven by several key reasons. Firstly, a striking fact is that eight out of ten states that are yet to expand Medicaid are situated in the South. The regional disparity becomes even more ironic after considering that Southerners generally have lower rate of insurance coverage and reduced life expectancy compared to people residing in other regions. Secondly, southern states including Arkansas and Kentucky have been among the first to consider the implementation of work requirements. Given this context and the role of historical institutions in policymaking, any potential roll-out of such policies, if implemented, would likely originate from the South.

The findings from this study reveal substantial heterogeneity in insurance and labor market impacts following the ACA Medicaid expansion. In particular, counties that were most affected, specially those within the top 10 percent of the distribution of Conditional Average Treatment Effects (CATE) estimates, saw a notable reduction in uninsured rate by approximately 12-15 percentage points more than the average effect in the years 2014 and 2015. Although the level of heterogeneity seems to have reduced over the years, the measure of heterogeneity as summarized by the AUTOC, remains statistically significant. This implies that variability in impacts persisted, even if some what reduced. Moreover, it is interesting to note that the machine learning approach identified the proportion of White votes for Obama during the 2008 presidential election as one of the primary variables mediating heterogeneity. The empirical results are consistent with the fact that ACA was highly politicized.

The Medicaid expansion reform significantly distorted labor market outcomes in the American South. The counties in the top 10% of the distribution of effects, which were the most adversely impacted, exhibited a decline of 6,000 jobs in 2014 compared to the average impact. This negative trend intensified over the following years, with these counties experiencing a substantial employment decrease of 13,000 jobs by 2018 relative to the average impact. The reductions in labor market activities were particularly concentrated in the food and accommodation and retail trade sectors, which have a disproportionately higher presence of low-wage workers. In contrast, no significant impacts were observed in the manufacturing and construction sectors.

While direct comparisons with past studies in the literature are challenging due to differing study objectives, broadly the findings reveal that the Medicaid expansion suppressed labor market outcomes consistent to the findings in Peng et al. (2020). However, employment effects are not transient, in contrast to what was previously shown in Peng et al. (2020). We note that such a comparison should be viewed with high caution as Peng et al. (2020) focus on the whole of U.S. and evaluate average impacts, whereas the findings presented in this study are specific to the American South and depict heterogeneous impacts.

The implications of these nuanced findings may be profound. The ACA Medicaid expansions have significantly reduced uninsured rates, particularly among people below 138% of the FPL. Moreover, recent studies show improvements in health outcomes (Miller et al., 2021; Borgschulte and Vogler, 2020). Given these advancements, implementing strict measures to restrict ACA Medicaid expansion, such as work requirement policies intended to counteract labor market distortions, could have severe consequences by stripping insurance from those who benefit most from the policy. Instead, more proactive policies aimed at enhancing local labor market outcomes—such as job training programs and place-based job policies—focused in areas where labor market outcomes are most affected by the reform can help mitigate labor market distortions without negatively impacting insurance outcomes. This balanced approach may preserve health benefits of Medicaid expansion while addressing labor market challenges.

References

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

Arias, E., Bastian, B., Xu, J., and Tejada-Vera, B. (2021). National Vital Statistics Reports. *National Vital Statistics Reports*, 70(1).

Athey, S., Tibsharani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Baicker, K., Finkelstein, A., Song, J., and Taubman, S. (2014). The impact of Medicaid on labor market activity and program participation: evidence from the Oregon Health Insurance Experiment. *American Economic Review*, 104(5):322–328.

Baker, R. S. (2022). The historical racial regime and racial inequality in poverty in the American south. *American Journal of Sociology*, 127(6):1721–1781.

Borgschulte, M. and Vogler, J. (2020). Did the ACA Medicaid expansion save lives? *Journal of Health Economics*, 72:102333.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). Classification and regression trees. *CA: Wadsworth International Group*.

Buchmueller, T. C., Levinson, Z. M., Levy, H. G., and Wolfe, B. L. (2016). Effect of the affordable care act on racial and ethnic disparities in health insurance coverage. *American journal of public health*, 106(8):1416–1421.

Chen, J. (2019). Does medicaid save lives? *Available at SSRN 3432701*.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal*, 21(1).

Courtemanche, C., Marton, J., Ukert, B., Yelowitz, A., and Zapata, D. (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.

Currie, J., Voorheis, J., and Walker, R. (2023). What caused racial disparities in particulate exposure to fall? New evidence from the clean air act and satellite-based measures of air quality. *American Economic Review*, 113(1):71–97.

Dague, L., DeLeire, T., and Leininger, L. (2017). The effect of public insurance coverage for childless adults on labor supply. *American Economic Journal: Economic Policy*, 9(2):124–154.

Esposito, E. (2019). The side effects of immunity: Malaria and African slavery in the United States. *American Economic Journal: Applied Economics*.

Foundation, K. F. (2023). Tough tradeoffs under republican work requirement plan: Some people lose medicaid or states could pay to maintain coverage. Technical report.

Frean, M., Gruber, J., and Sommers, B. D. (2017). Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act. *Journal of Health Economics*, 53:72–86.

Frisvold, D. E. and Jung, Y. (2018). The impact of expanding Medicaid on health insurance coverage and labor market outcomes. *International journal of health economics and management*, 18:99–121.

Garthwaite, C., Gross, T., and Notowidigdo, M. J. (2014). Public health insurance, labor supply, and employment lock. *The Quarterly Journal of Economics*, 129(2):653–696.

Gavrilova, E., Langørgen, A., and Zoutman, F. (2023). Dynamic causal forests, with an application to payroll tax incidence in norway. *CESifo Working Paper*.

Gooptu, A., Moriya, A. S., Simon, K. I., and Sommers, B. D. (2016). Medicaid expansion did not result in significant employment changes or job reductions in 2014. *Health affairs*, 35(1):111–118.

Greibrok, M. (2023). Universal work requirements for welfare programs are a win for all involved. https://thefga.org/research/universal-work-requirements/.

Grogan, C. M. and Park, S. (2017). The racial divide in state Medicaid expansions. *Journal of Health Politics, Policy and Law*, 42(3):539–572.

Guth, M. and Musumeci, M. (2022). An overview of Medicaid work requirements: What happened under the Trump and Biden administrations. *Kaiser Family Foundation, May*, 3.

Hamilton, C. (2024). The impact of the 2014 medicaid expansion on the health, health care access, and financial well-being of low-income young adults. *Health Economics*.

Harris, E. and Mok, S. (2015). *How CBO estimates the effects of the Affordable Care Act on the labor market*. Congressional Budget Office.

Hornbeck, R. and Naidu, S. (2014). When the levee breaks: Black migration and economic development in the American South. *American Economic Review*, 104(3):963–90.

Jacob, D. (2021). Cate meets ml: Conditional average treatment effect and machine learning. *Digital Finance*, 3(2):99–148.

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

Lanford, D. and Quadagno, J. (2016). Implementing Obamacare: the politics of Medicaid expansion under the Affordable Care Act of 2010. *Sociological Perspectives*, 59(3):619–639.

Leung, P. and Mas, A. (2018). Employment effects of the Affordable Care act Medicaid expansions. *Industrial Relations: A Journal of Economy and Society*, 57(2):206–234.

Michener, J. (2020). Race, politics, and the affordable care act. *Journal of Health Politics, Policy and Law*, 45(4):547–566.

Miller, S., Johnson, N., and Wherry, L. R. (2021). Medicaid and mortality: new evidence from linked survey and administrative data. *The Quarterly Journal of Economics*.

Nie, X. and Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319.

Olvera, J. G., Smith, C. W., 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.

Peng, L., Guo, X., and Meyerhoefer, C. D. (2020). The effects of Medicaid expansion on labor market outcomes: evidence from border counties. *Health economics*, 29(3):245–260.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, pages 931–954.

RWJF (2020). 2020 county health rankings key findings. Technical report, University of Wisconsin Population Health Institute and Robert Wood Johnson Foundation.

Shrestha, V. (2023). The institution of american slavery, current-day preference to repeal the affordable care act, and efficacy of the reform. Technical report.

Simon, K., Soni, A., and Cawley, J. (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.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wehby, G. L. and Lyu, W. (2018). The impact of the aca medicaid expansions on health insurance coverage through 2015 and coverage disparities by age, race/ethnicity, and gender. *Health services research*, 53(2):1248–1271.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., and Wager, S. (2021). Evaluating treatment prioritization rules via rank-weighted average treatment effects. *arXiv preprint arXiv:2111.07966*.

8 Figures and Tables

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.
uninsured rate 2010	1,418	45.329	7.847
uninsured rate 2011	1,418	44.596	7.544
uninsured rate 2012	1,418	43.631	7.231
uninsured rate 2013	1,418	42.807	7.533
uninsured rate 2014	1,418	35.798	9.108
uninsured rate 2015	1,418	31.235	10.150
uninsured rate 2016	1,418	29.060	10.134
uninsured rate 2017	1,418	28.899	10.398
uninsured rate 2018	1,418	28.574	10.256
unemployment rate 2010	1,418	10.294	2.733
unemployment rate 2011	1,418	9.706	2.620
unemployment rate 2012	1,418	8.602	2.466
unemployment rate 2013	1,418	8.051	2.374
unemployment rate 2014	1,418	6.900	2.038
unemployment rate 2015	1,418	6.129	1.853
unemployment rate 2016	1,418	5.699	1.794
unemployment rate 2017	1,418	4.908	1.468
unemployment rate 2018	1,418	4.326	1.308
employed (QCEW) 2010	1,418	30,706.400	99,420.230
employed (QCEW) 2011	1,418	31,106.440	101,305.500
employed (QCEW) 2012	1,418	31,664.180	103,896.000
employed (QCEW) 2013	1,418	32,219.530	106,519.300
employed (QCEW) 2014	1,418	32,921.010	109,649.100
employed (QCEW) 2015	1,418	33,676.430	112,534.200
employed (QCEW) 2016	1,418	34,267.350	114,367.600
employed (QCEW) 2017	1,418	34,767.740	116,123.400
employed (QCEW) 2018	1,418	35,365.180	118,244.600
employed (BLS) 2010	1,418	35,013.470	97,849.750
employed (BLS) 2011	1,418	35,621.090	100,344.600
employed (BLS) 2012	1,418	36,224.890	103,116.300
employed (BLS) 2013	1,418	36,607.260	105,187.300
employed (BLS) 2014	1,418	37,176.640	107,661.400

employed (BLS) 2015	1,418	37,688.690	109,056.800
employed (BLS) 2016	1,418	38,450.710	111,193.000
employed (BLS) 2017	1,418	39,277.610	113,818.200
employed (BLS) 2018	1,418	39,902.130	115,719.400
labor force (BLS) 2010	1,418	38,670.350	107,452.800
labor force (BLS) 2011	1,418	39,077.370	109,414.200
labor force (BLS) 2012	1,418	39,269.610	111,078.400
labor force (BLS) 2013	1,418	39,384.610	112,448.000
labor force (BLS) 2014	1,418	39,563.970	113,909.700
labor force (BLS) 2015	1,418	39,786.770	114,528.600
labor force (BLS) 2016	1,418	40,413.040	116,585.600
labor force (BLS) 2017	1,418	41,039.590	118,868.900
labor force (BLS) 2018	1,418	41,472.860	120,191.600
Medicaid tranfers per capita 2010	1,418	1.421	0.655
Medicaid tranfers per capita 2011	1,418	1.471	0.678
Medicaid tranfers per capita 2012	1,418	1.492	0.687
Medicaid tranfers per capita 2013	1,418	1.565	0.729
Medicaid tranfers per capita 2014	1,418	1.708	0.845
Medicaid tranfers per capita 2015	1,418	1.822	0.927
Medicaid tranfers per capita 2016	1,418	1.878	0.959
Medicaid tranfers per capita 2017	1,418	1.923	0.981
Medicaid tranfers per capita 2018	1,418	1.986	1.019

Note: Summary statistics for outcomes used in the study. The uninsured rate pertains to people below 138% of the Federal Poverty Level (FPL) and comes from the Small Area Health Insurance Estimates (SAHIE). The unemployment rate, total employment count, and labor force data comes from the Bureau of Labor Statistics (BLS). The count of total employed comes from the Quarterly Census of Employment and Wages (QCEW). Data for Medicaid-CHIP transfer funds is obtained from Olvera et al. (2023) replication package and is originally extracted from the Bureau of Economic Analysis (BEA).

Table 2: Variable Importance Matrix from Causal Forest

	response variable	features	importance measure
1	01. uninsured rate	percent high school (2010)	0.05
2	01. uninsured rate	mortality rate (2010-2013) per 1,000 population	0.06
3	01. uninsured rate	longitude	0.07
4	01. uninsured rate	uninsured rate 2013 (income $<$ 138% FPL)	0.09
5	01. uninsured rate	proportion of White votes for Obama (2008)	0.09
6	02. unemp. rate	PM 2.5 measure (2010)	0.05
7	02. unemp. rate	proportion of White votes for Obama (2008)	0.05
8	02. unemp. rate	unemployment rate 2010	0.06
9	02. unemp. rate	mortality rate (2010-2013) per 1,000 population	0.08
10	02. unemp. rate	unemployment rate 2012	0.10
11	02. unemp. rate	longitude	0.13
12	03. employed	Hill Burton Funds expenses	0.06
13	03. employed	labor force 2010	0.08
14	03. employed	labor force 2012	0.08
15	03. employed	county population (2010)	0.08
16	03. employed	employed 2012	0.09
17	03. employed	employed 2010	0.09
18	04. labor force	per capita income (2000) in 2010 dollars	0.06
19	04. labor force	labor force 2012	0.06
20	04. labor force	employed 2012	0.07
21	04. labor force	labor force 2010	0.07
22	04. labor force	county population (2010)	0.07
23	04. labor force	Hill Burton Funds expenses	0.07
24	04. labor force	employed 2010	0.08

Note: Variable importance matrix obtained from Causal Forest with 10,000 trees for the reported response variables.

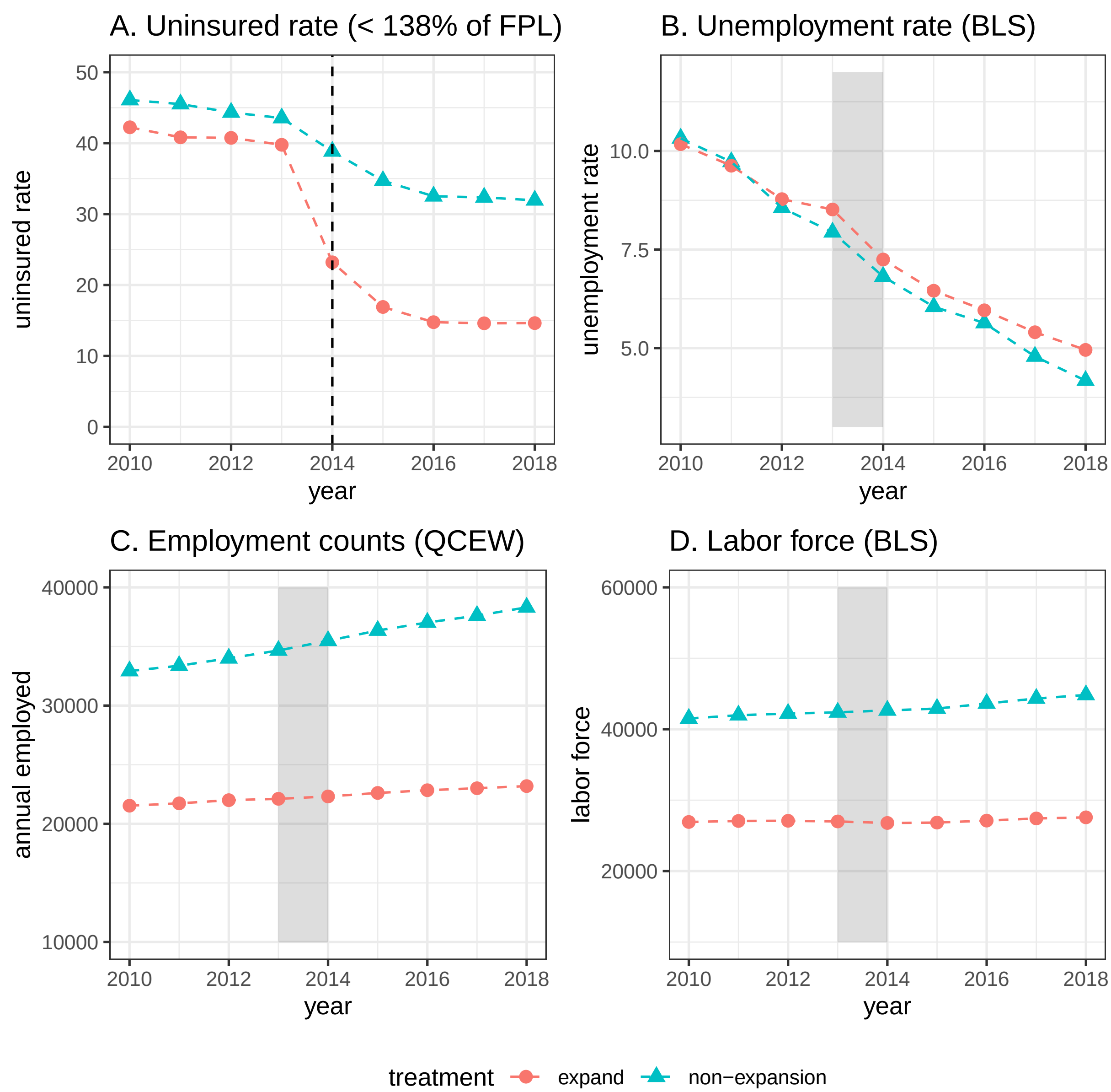

Figure 1: Descriptive figures

Note: The subfigures depict average outcomes by year and Medicaid expansion status in the American South. The vertical line in figure A denotes the expansion year for all the southern states that expanded Medicaid except Louisiana. Louisiana expanded Medicaid in 2016. The rectangular grey blocks between 2013-2014 illustrated in panels B-D show the adjustment period used for the purpose of labor market analyses, where individuals are allowed to adjust their behavior in anticipation of the policy. Data for panel A comes from SAHIE, panels B and D comes from BLS, and panel C comes from QCEW.

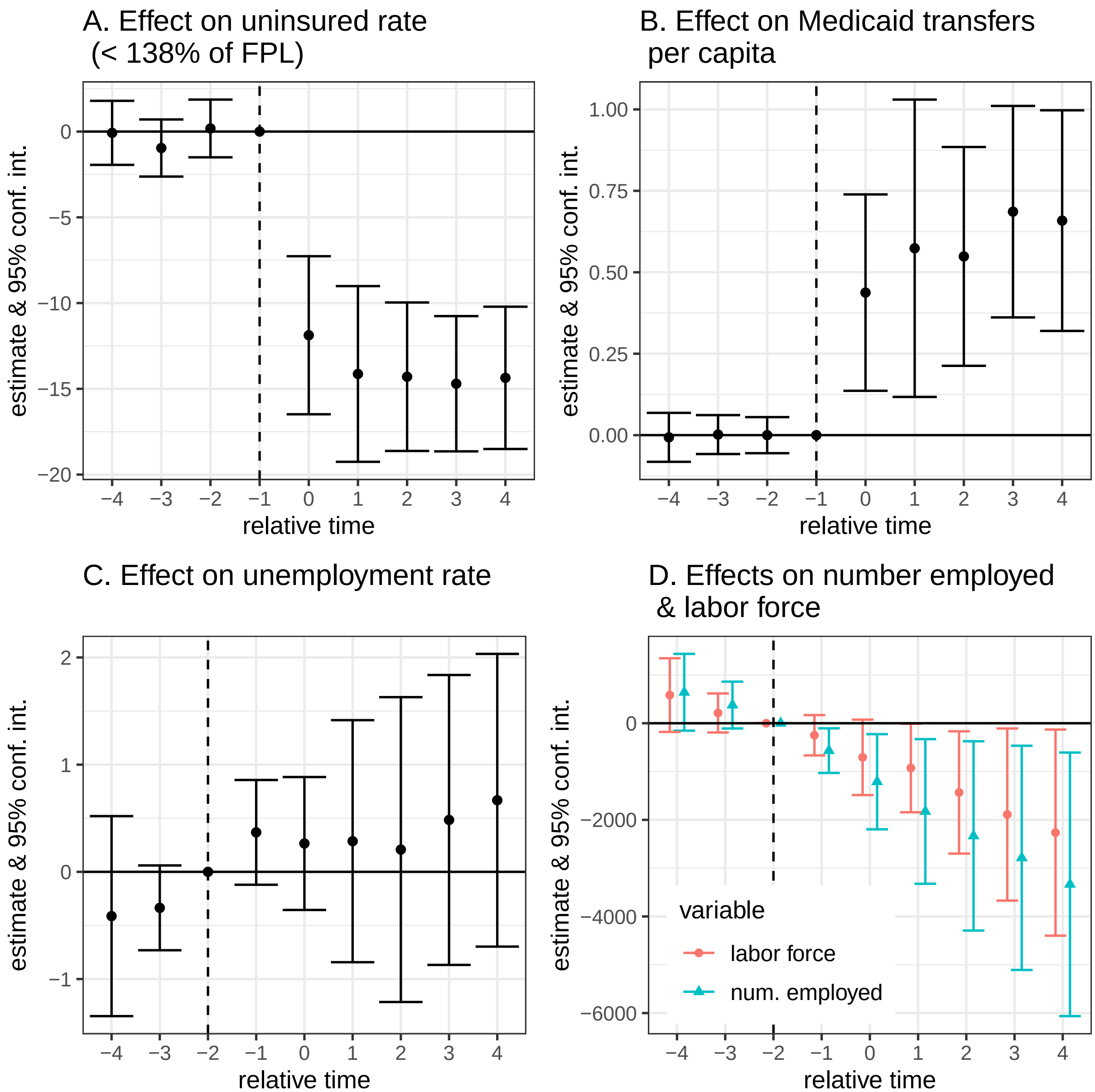

FIGURE 2: Event study estimates

Note: The subfigures depict event study estimates obtained from estimating the specification illustrated in equation 1 along with the 95% confidence interval bars obtained from the standard errors clustered at the state level. The omitted relative year is one year prior to the implementation year when evaluating uninsured rates, whereas the omitted category is two years prior to the implementation year for the labor market outcomes to allow for adjustments in labor market decisions a year prior to the reform due to anticipation.

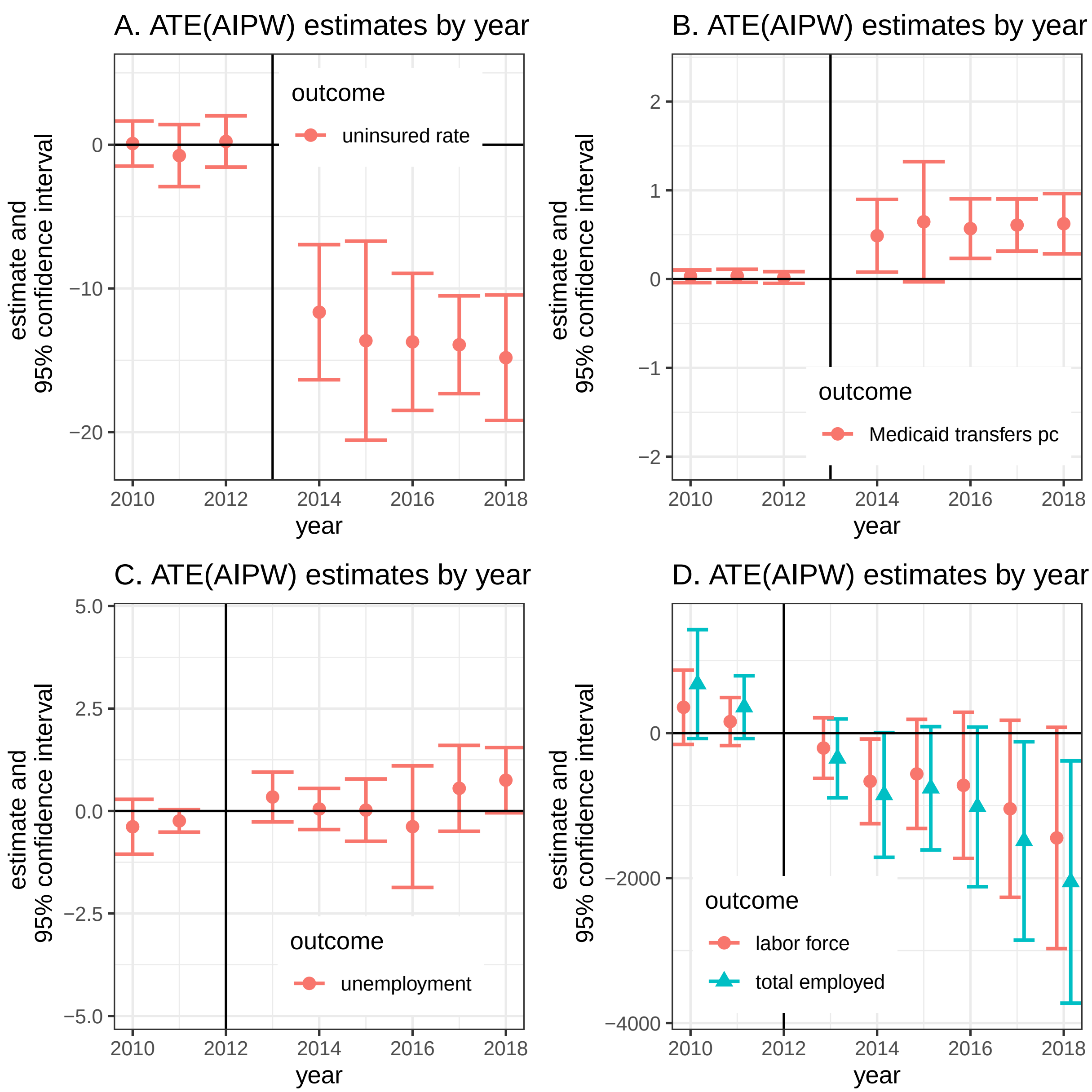

Figure 3: ATE Estimates from Augumented Inverse Probability Weighting (AIPW)

Note: The subfigures depict ATE estimates obtained using the AIPW approach discussed in 4.4 for the reported response variables along with the 95% confidence interval bars. The base year to create the first difference outcome variable is 2013 for uninsured rate and Medicaid funds, whereas the year 2012 is used for labor market outcomes.

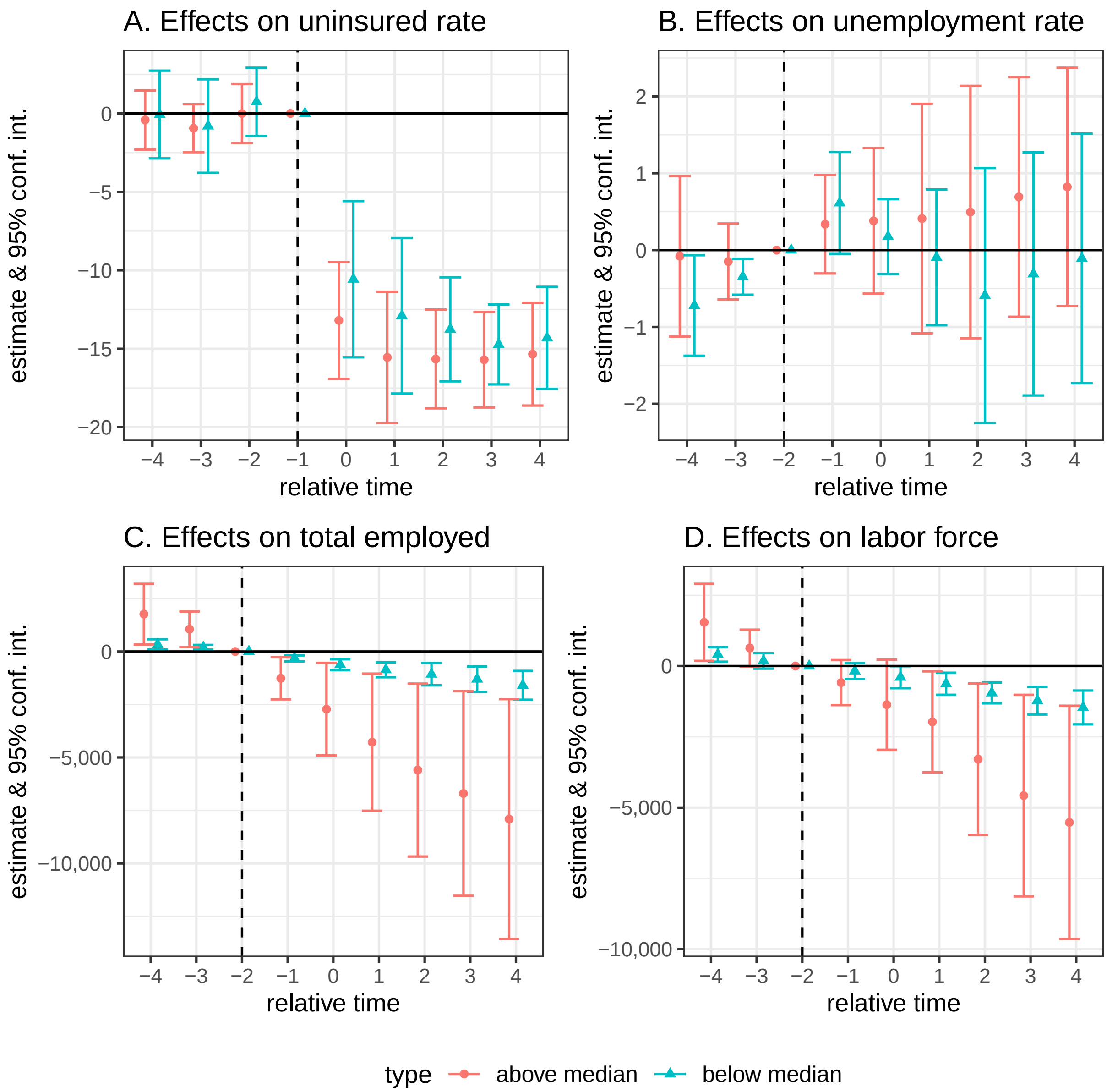

FIGURE 4: Event study estimates by the classification of the proportion of White votes for Obama in the 2008 presidential election

Note: The subfigures depict event study estimates obtained using the specification illustrated in equation 1 along with the 95% confidence interval bars obtained from the standard errors clustered at the state level. The omitted relative year is one year prior to the implementation year when evaluating uninsured rates, whereas the omitted category is two years prior to the implementation year for the labor market outcomes to allow for adjustments in labor market decisions a year prior to the reform due to anticipation. The event study estimations are carried out for counties above and below the median proportion of White votes for Obama in the 2008 presidential election. All the estimation are weighted by the inverse of the propensity scores obtained from the regression forest.

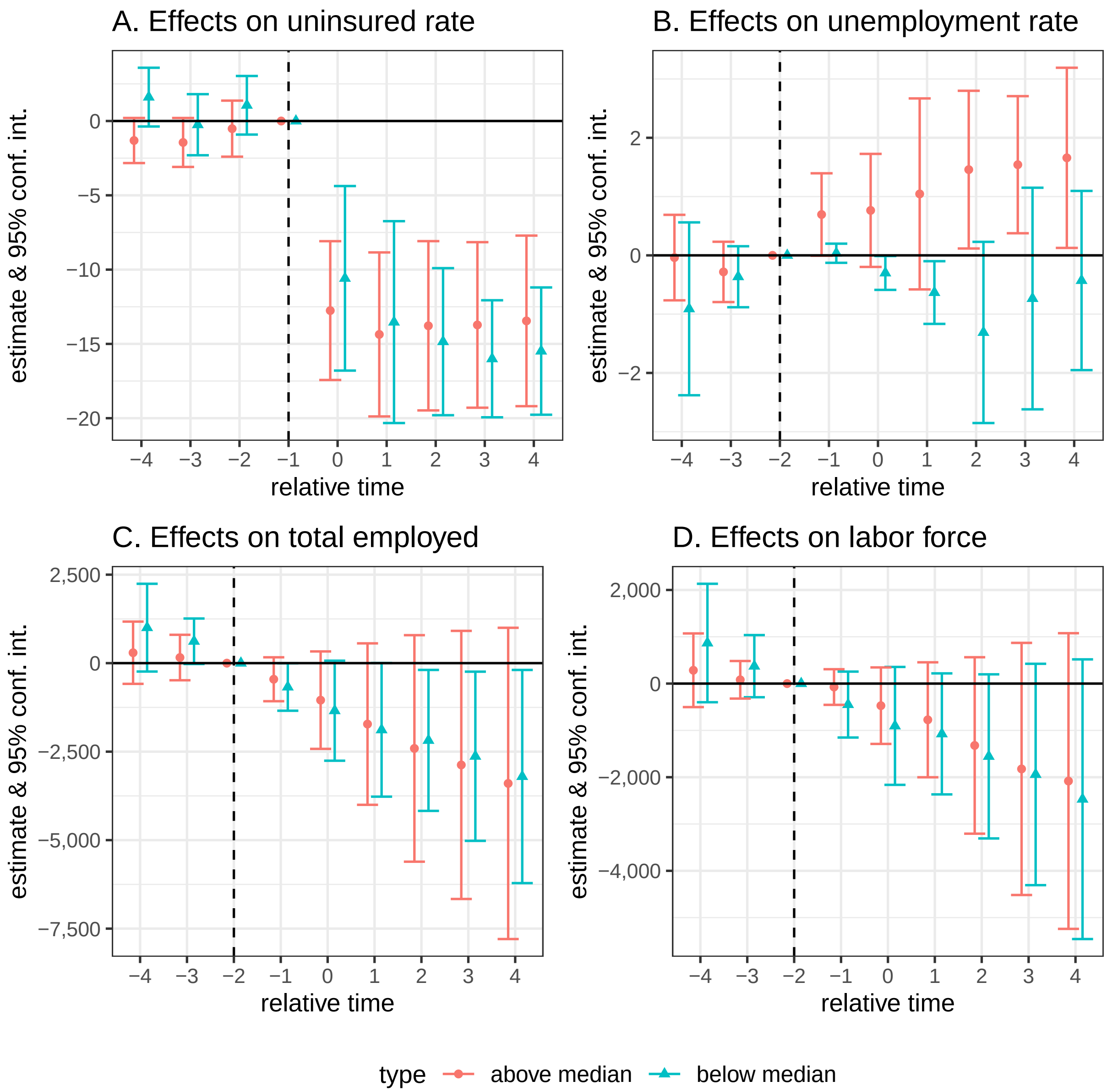

FIGURE 5: Event study estimates by longitude

Note: The subfigures depict event study estimates obtained using the specification illustrated in equation 1 along with the 95% confidence interval bars obtained from the standard errors clustered at the state level. The omitted relative year is one year prior to the implementation year when evaluating uninsured rates, whereas the omitted category is two years prior to the implementation year for the labor market outcomes to allow for adjustments in labor market decisions a year prior to the reform due to anticipation. The event study estimations are carried out for counties above and below the median longitude value. All the estimation are weighted by the inverse of the propensity scores obtained from the regression forest.

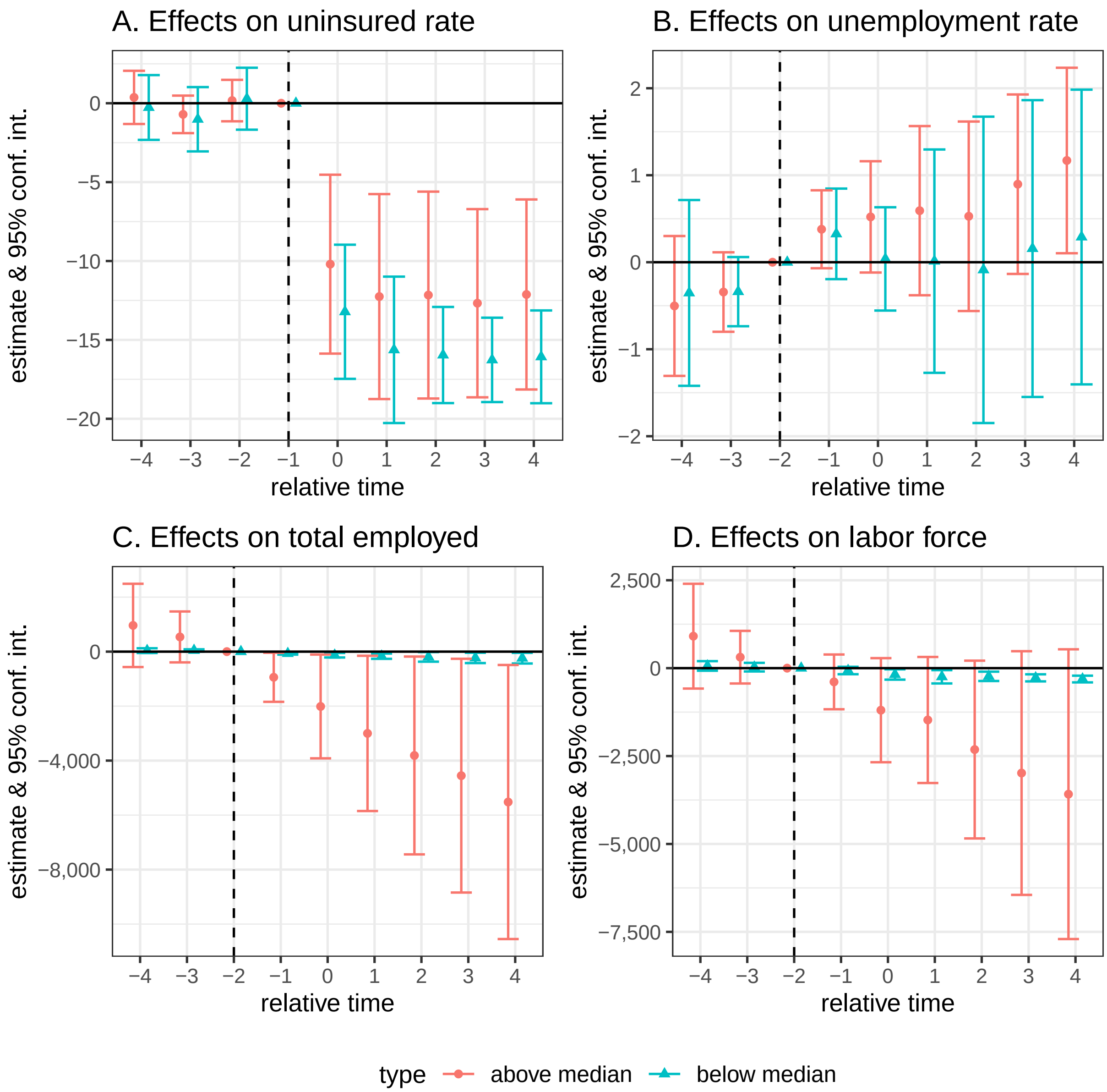

FIGURE 6: Event study estimates by the classification based on the total employed people in 2010

Note: The subfigures depict event study estimates obtained using the specification illustrated in equation 1 along with the 95% confidence interval bars obtained from the standard errors clustered at the state level. The omitted relative year is one year prior to the implementation year when evaluating uninsured rates, whereas the omitted category is two years prior to the implementation year for the labor market outcomes to allow for adjustments in labor market decisions a year prior to the reform due to anticipation. The event study estimations are carried out for counties above and below the median employment value in 2010. All the estimation are weighted by the inverse of the propensity scores obtained from the regression forest.

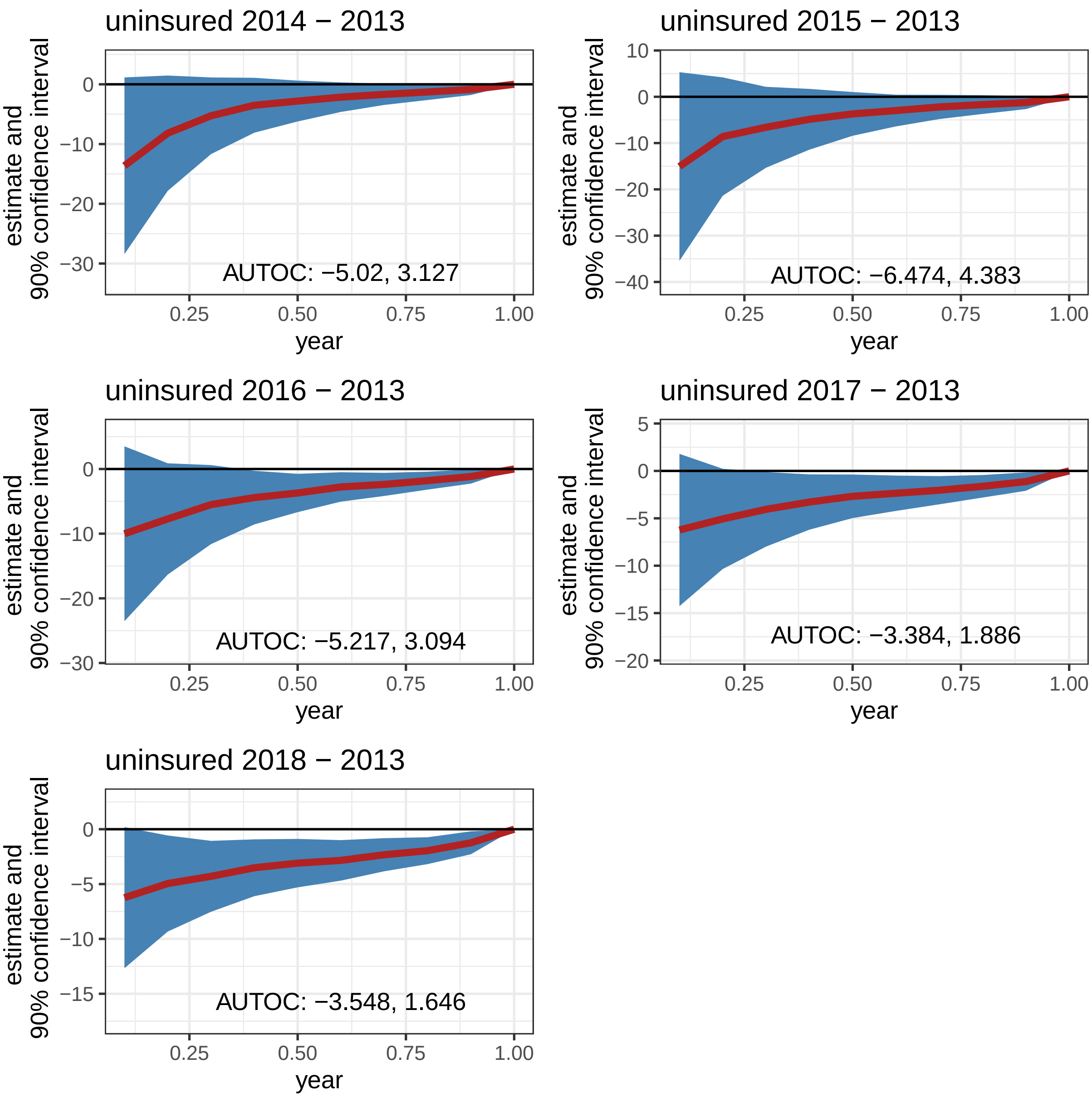

FIGURE 7: TOC curve from RATE: Uninsured Rate for people below 138% of FPL

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in uninsured rate for people living under 138% of FPL between the given year and 2013 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for the outcome variable area extracted from SAHIE. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

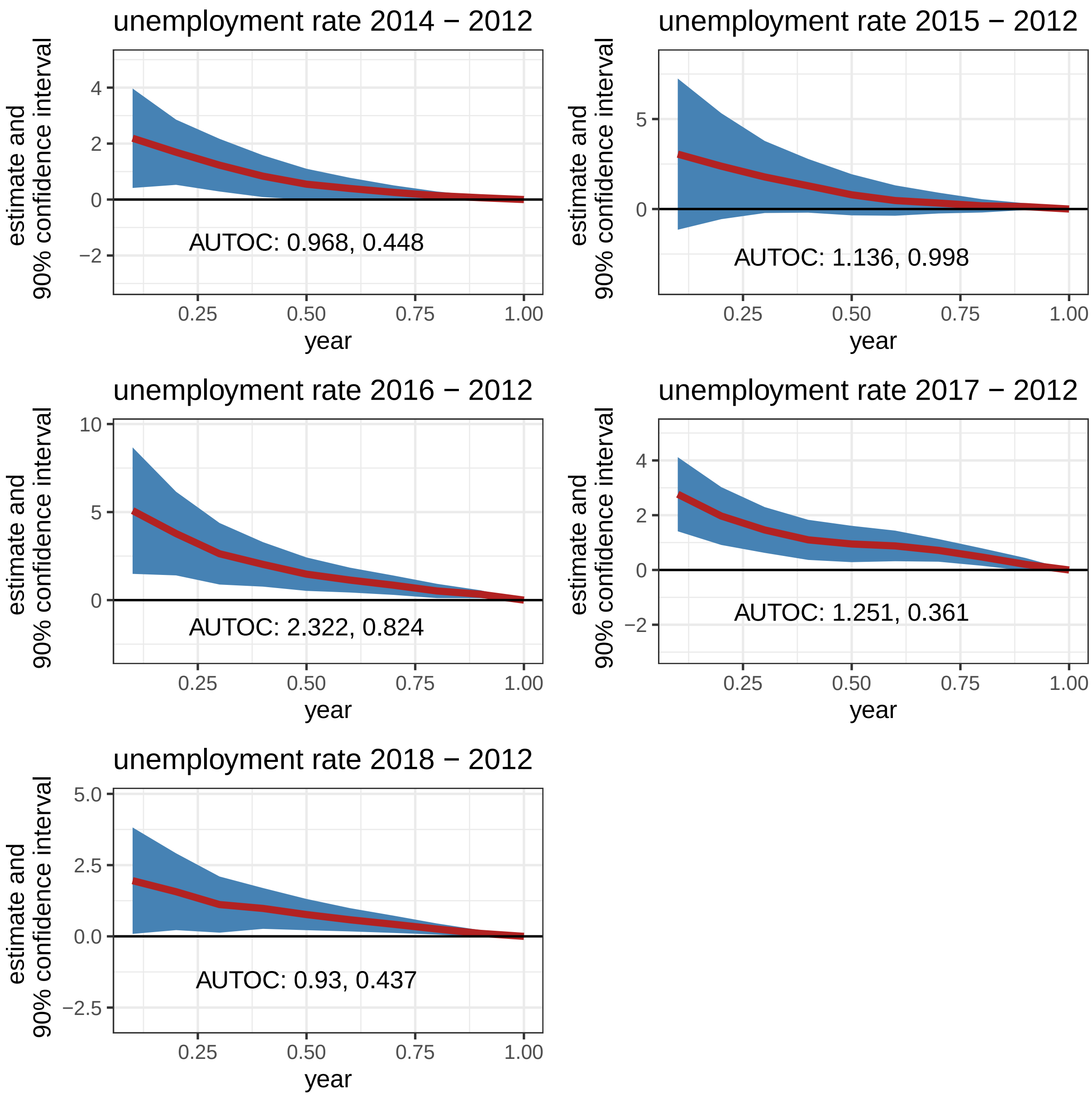

FIGURE 8: TOC curve from RATE: Unemployment Rate (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in unemployment rate between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for unemployment rate comes from BLS. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

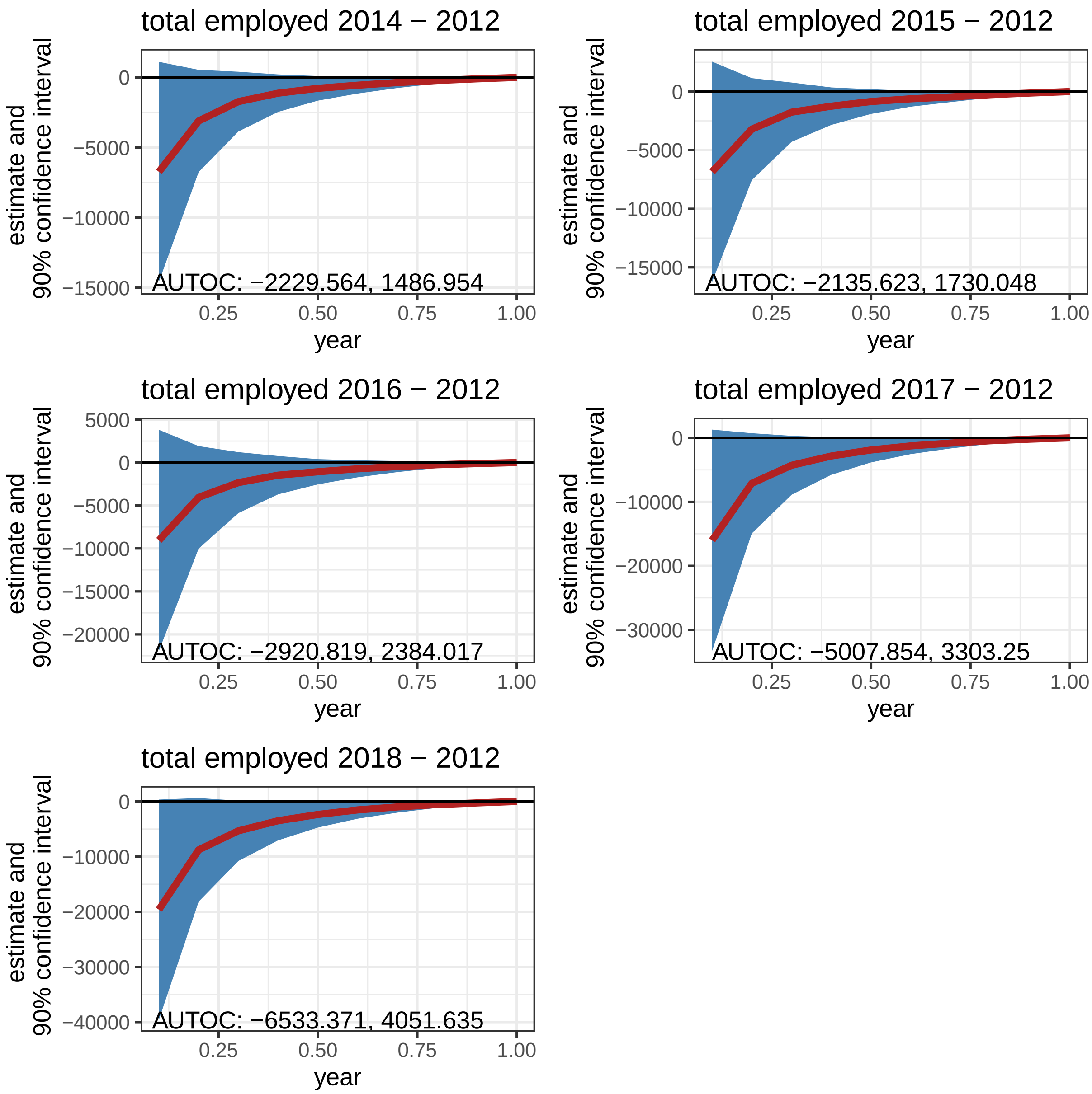

FIGURE 9: TOC curve from RATE: Total Employed (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for total employed comes from BLS. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the top.

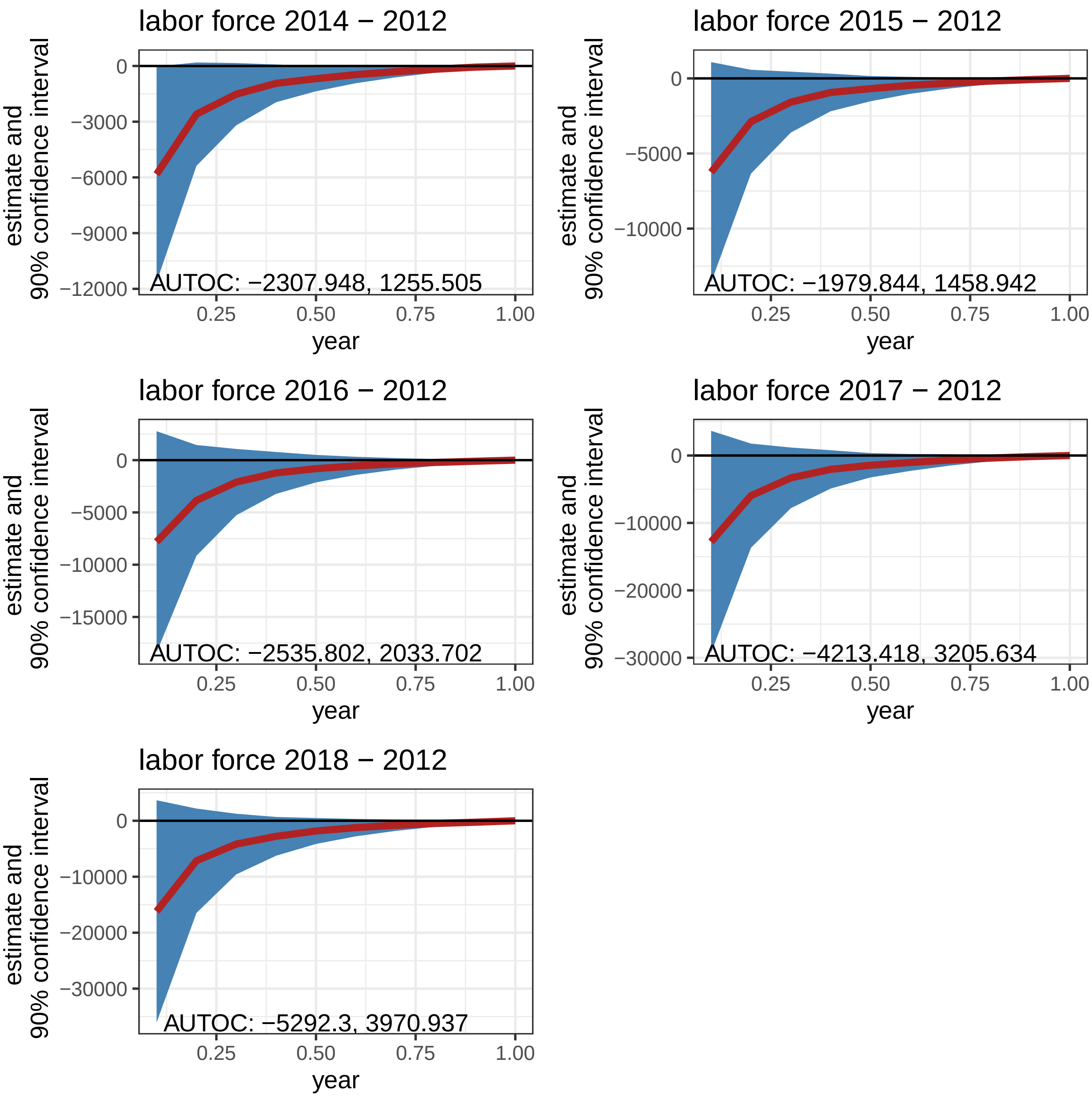

FIGURE 10: TOC curve from RATE: Labor Force (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in labor force between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for labor force participation comes from BLS. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the top.

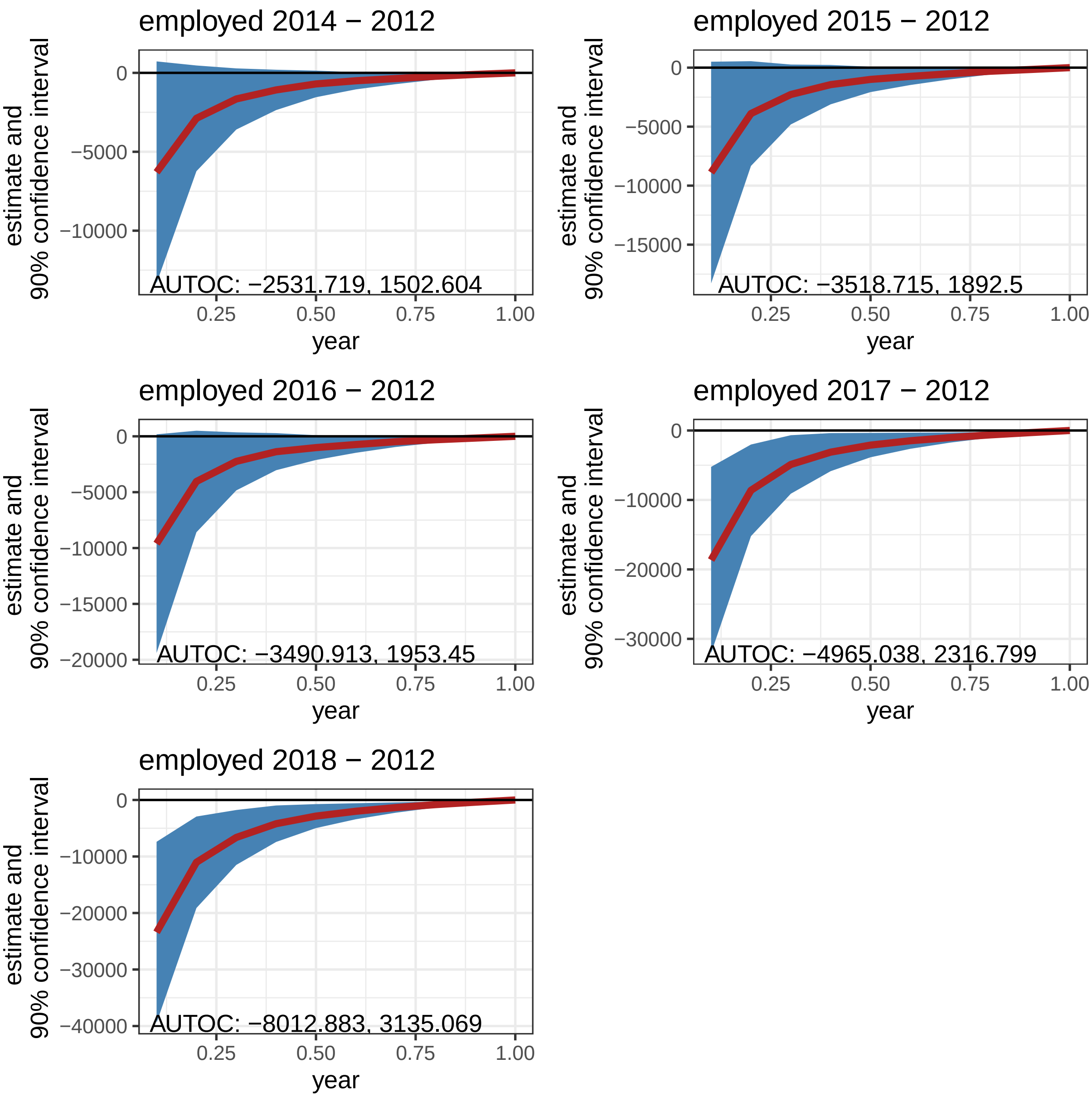

Figure 11: TOC curve from RATE: Total Employed (QCEW)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for total employment count comes from QCEW. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

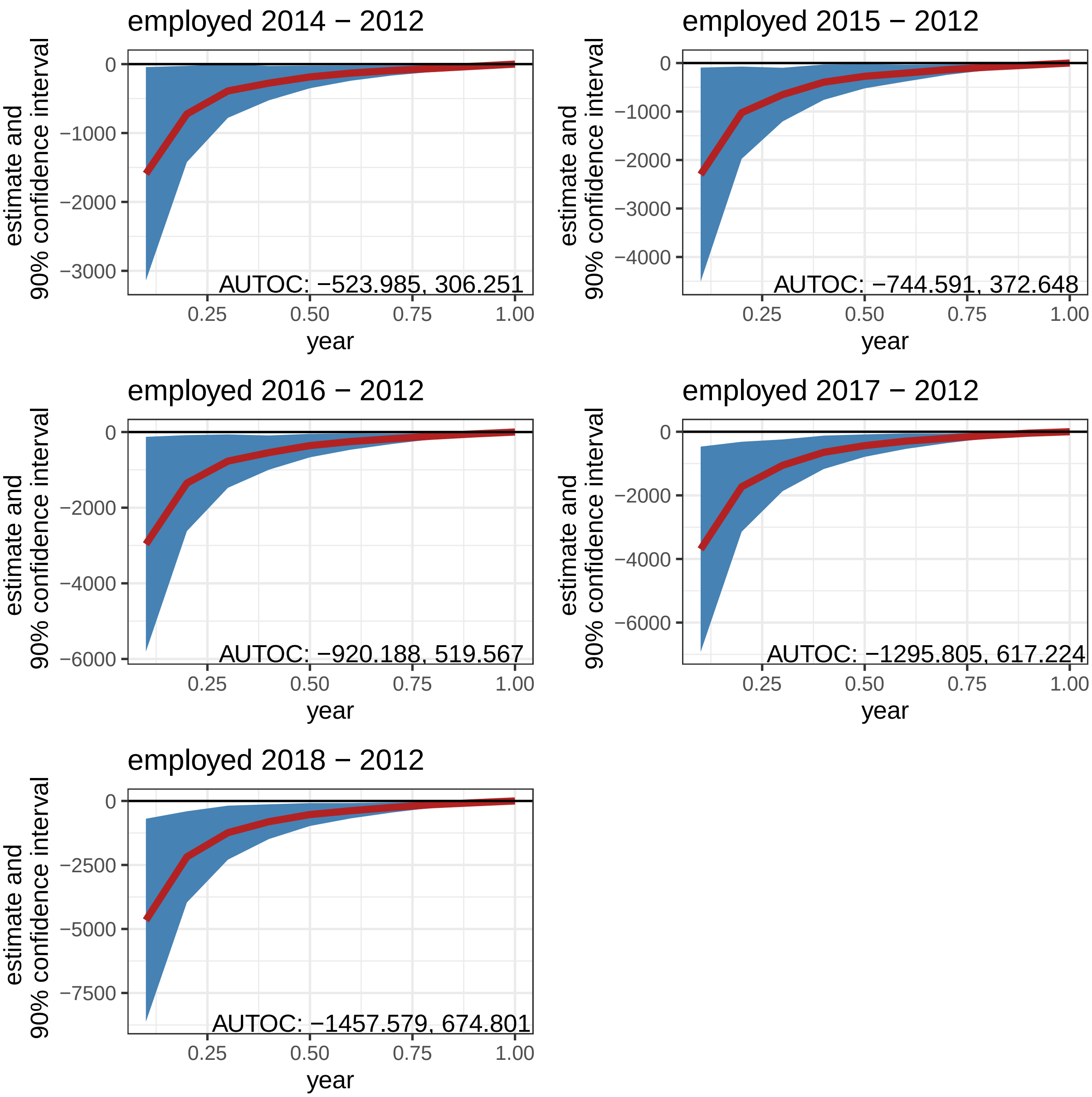

FIGURE 12: TOC curve from RATE: Total Employed in Food Sector (QCEW)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed in food and accomodation sector between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

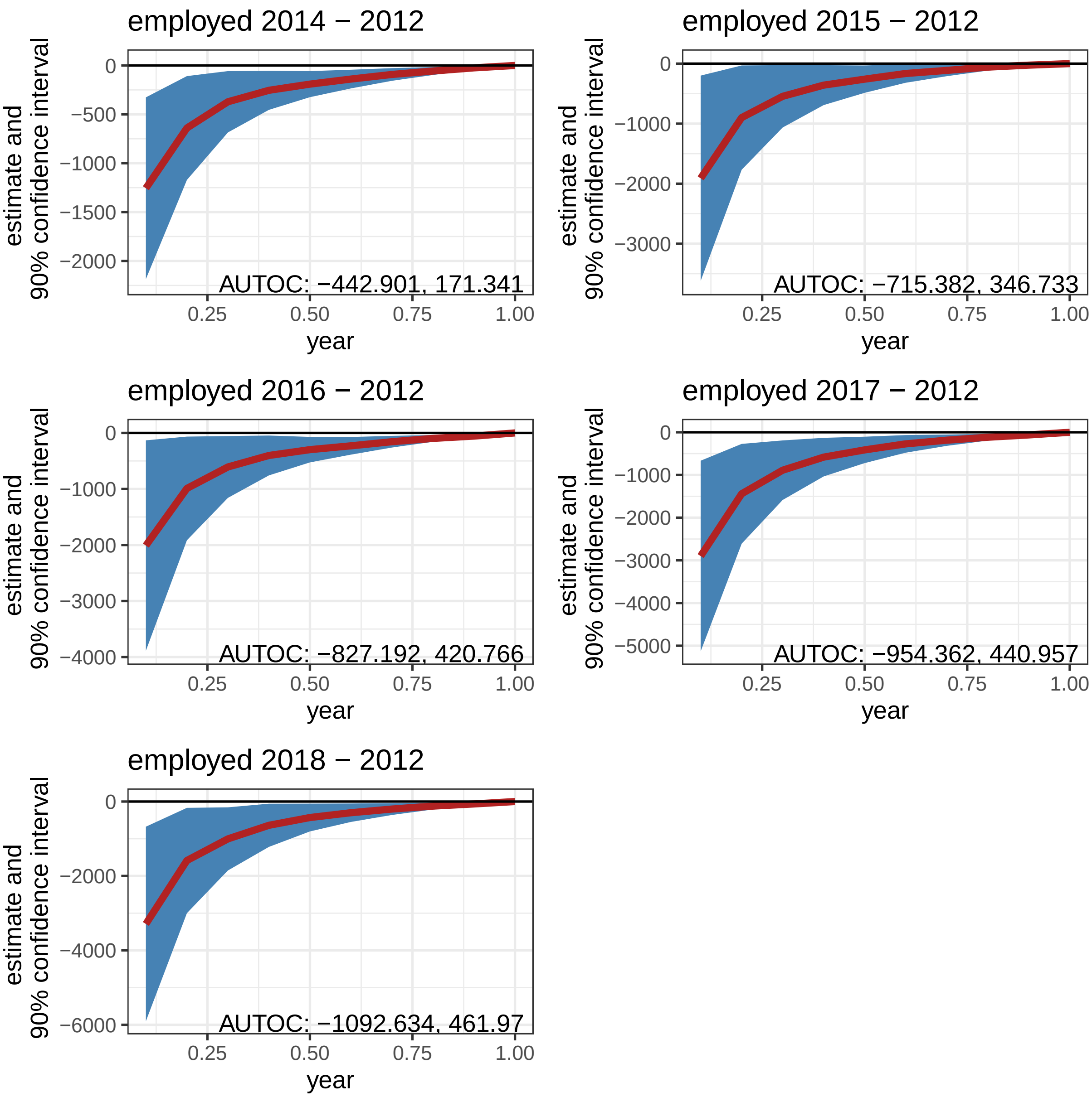

FIGURE 13: TOC curve from RATE: Total Employed in Retail Sector (QCEW)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed in retail/trade sector between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

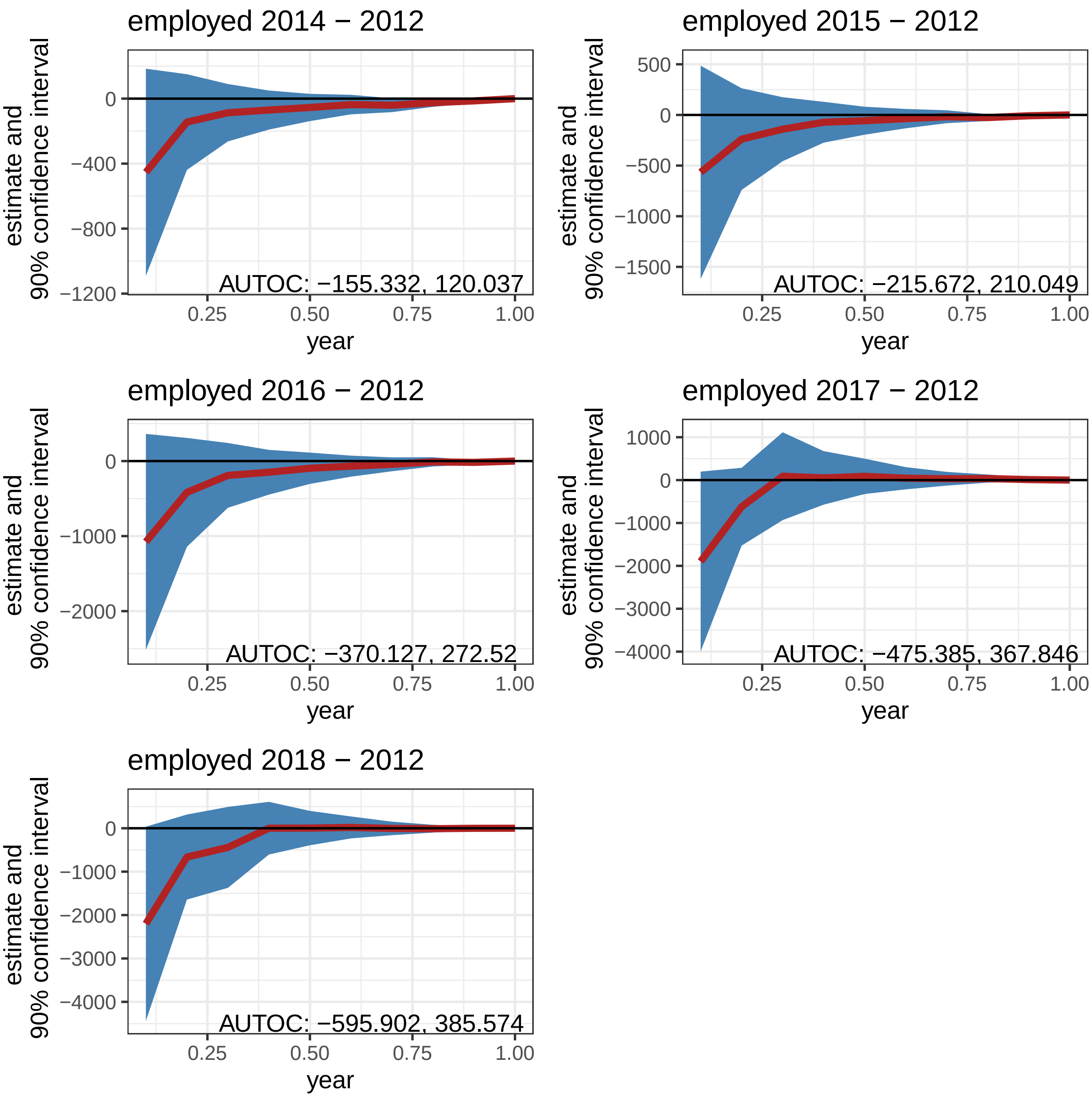

Figure 14: TOC curve from RATE: Total Employed in Construction Sector (QCEW)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed in construction sector between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

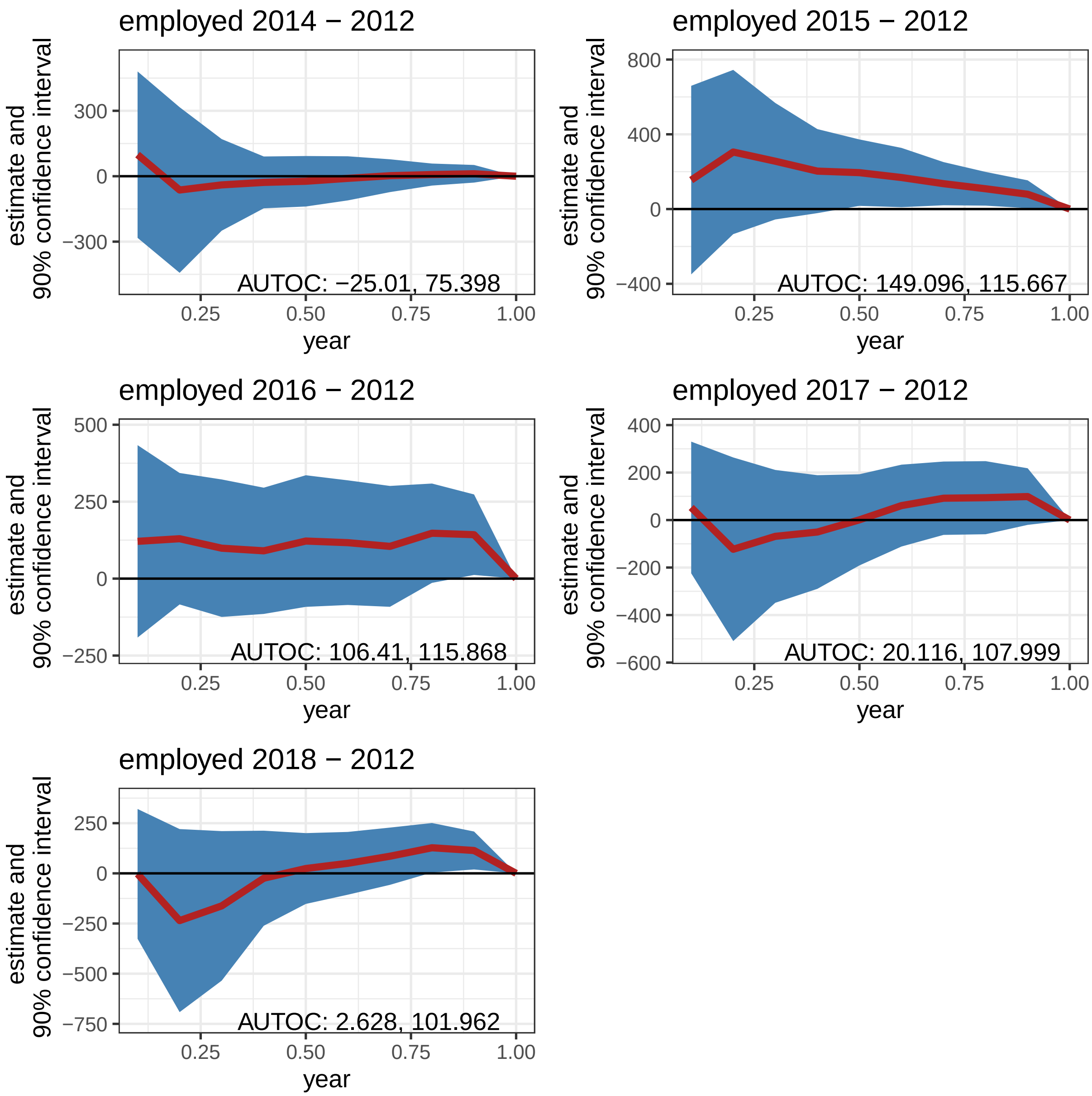

FIGURE 15: TOC curve from RATE: Total Employed in Manufacturing Sector (QCEW)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed in manufacturing sector between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. CATE estimates obtained from the trained CF model using an independent split of the sample are used as the priority scores, S(.). The RANK evaluation is carried out in another spilt of the sample as discussed in section 4.5. This assures that priority scores and RANK evaluation are not based on the same sample, which otherwise might increase over fitting. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

9 Appendix

9.1 Data Source of the County Level Variables Used

1) ABS = Acharya et al. (2016); 2) Olvera 2023 = Olvera et al. (2023); 3) Hornbeck-Naidu = Hornbeck and Naidu (2014); 4) Currie 2023 = Currie et al. (2023); 5) SAIPE: Small Area Income and Poverty Estimates; 6) USDA: U.S. Department of Agriculture; 7) CRU: Climate Research Unit; 8) FAO: UN Food and Agriculture Organization.

Variable	Main Source	Obtained
proportion Blacks in 2010	U.S. Census	
proportion Whites in 2010	U.S. Census	
total population in 2010	U.S. Census	
per capita income in 2000 (2010 dollars)	SAIPE	
state minimum wage		Olvera 2023 rep. files
percent high school (2010)	USDA	
percent college (2010)	USDA	
poverty rate (2010)	USDA	
hh. median income (2010)	SAIPE	
hh. income Whites in 2014 (age 35)	Opportunity Atlas	
Decomcrat county		Olvera 2023 rep. files
estimated proportion of White votes for Obama		ABS rep. files
proportion Trump votes in 2016	ICPSR	
PM 2.5 measure in 2010		Currie 2023 rep. files
county area	2000 Census Shapefiles	ABS
land ruggedness	Hornbeck-Naidu	ABS
latitude and longitude	2000 Census Shapefiles	ABS
elevation		ABS
average precipitation (1961-1990)	CRU	
average temperature (1961-1990)	CRU	
cotton suitability measure	FAO	ABS rep. files
malaria stability index	author's calculation	
	following Esposito (2019)	
Hill-Burton expenses (1947 – 1971)		Heidi L. Williams
FSSNAP Benefit for 3 person family		Olvera 2023 rep. files
State EITC rate		Olvera 2023 rep. files
Refundable state EITC (Yes)		Olvera 2023 rep. files

9.2 Additional Results

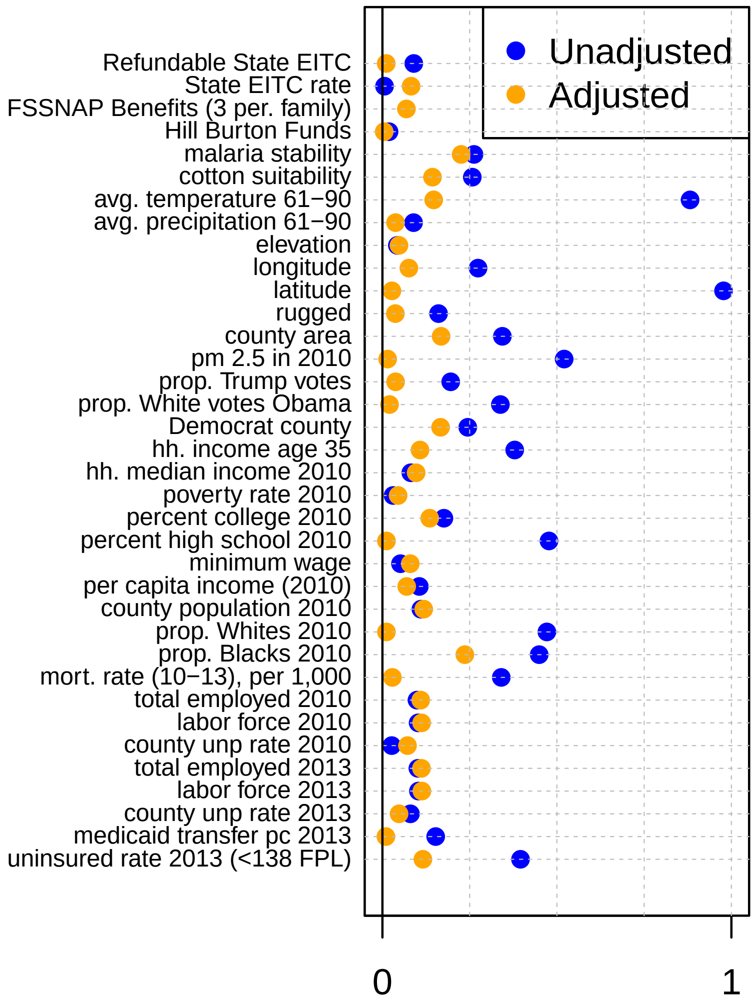

Figure I1

Note: The figure shows the absolute value of the standardized difference in mean between the treatment (expansion) and control (non-expansion) units. The unadjusted difference is the simple difference in mean between the treated and control units. The adjusted mean difference plots the absolute difference in means once the covariates are adjusted using inverse propensity weights.

Table I1: Summary Statistics

Statistic	N	Mean	St. Dev.
Medicaid transfers per capita in 2013	1,418	1.565	0.729
mortality rate (between 2010-2013 per 1,000 population)	1,410	9.624	1.303
proportion Blacks 2010	1,416	0.167	0.180
proportion Whites 2010	1,416	0.815	0.182
county population 2010	1,418	79,812.140	211,359.100
per capita income (2010 dollars)	1,395	28,056.190	6,309.248
state minimum wage	1,418	7.015	0.721
percent high school 2010	1,418	35.688	6.526
percent college 2010	1,418	17.052	8.104
poverty rate 2010	1,418	19.589	6.383
household median income 2010	1,418	39,897.480	10,554.690
household income Whites (age 35)	1,413	45,089.740	6,381.488
Democrat county	1,418	0.159	0.366
proportion White votes for Obama (2008)	1,275	0.268	0.129
proportion Trump vote 2016	1,417	0.660	0.153
PM 2.5 measure 2010	1,418	10.200	1.620
county area	1,417	611.568	404.294
ruggedness	1,338	43.047	47.492
latitude	1,417	34.433	3.111
longitude	1,417	−87.652	6.953
elevation	1,418	0.240	0.248
average precipitation (1961-1990)	1,418	96.863	22.406
average temperature (1961-1990)	1,418	15.845	3.004
cotton suitability measure	1,345	0.448	0.169
malaria stability	1,418	0.145	0.365
Hill Burton expenses	1,126	19,553,743.000	97,288,286.000
FSSNAP Benefitfor 3 person family	1,418	497.000	0.000
State EITC Rate	1,418	0.019	0.059
Refundable State EITC (Yes)	1,418	0.061	0.239

Summary statistics for features used in the study.

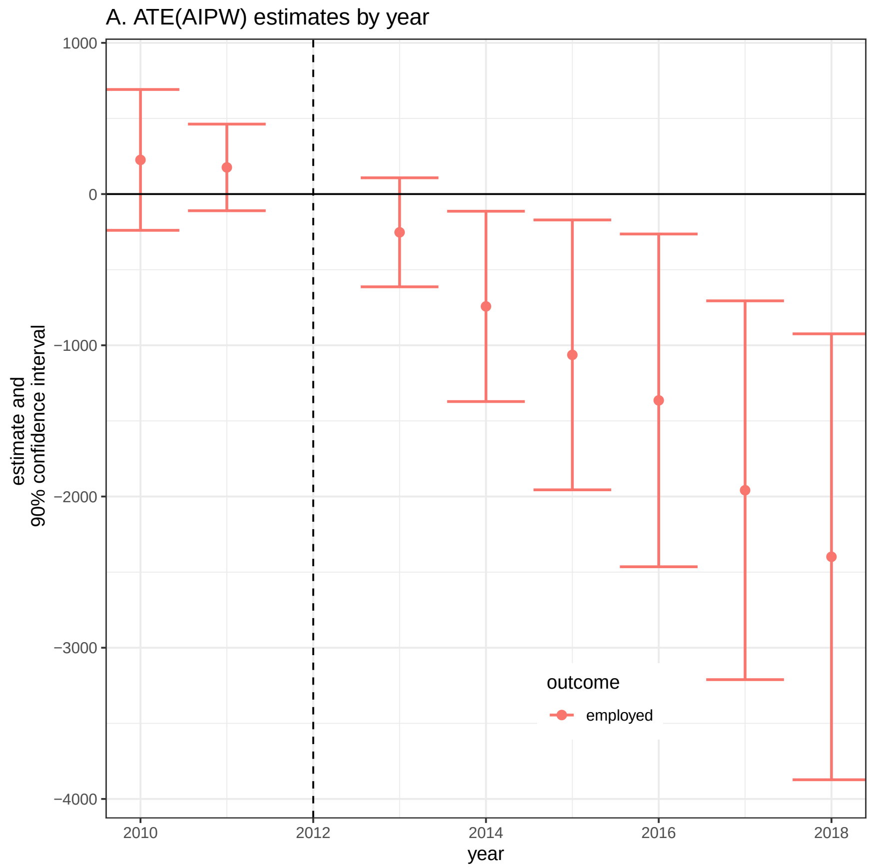

Figure I2: Event study estimates for the number employed (QCEW)

Note: The figure depicts ATE estimates obtained using the AIPW approach discussed in 4.4 for the reported response variable, the counts of total employed from QCEW, along with the 95% confidence interval bars. The year 2012 is used as the base year used to create the first difference outcome.

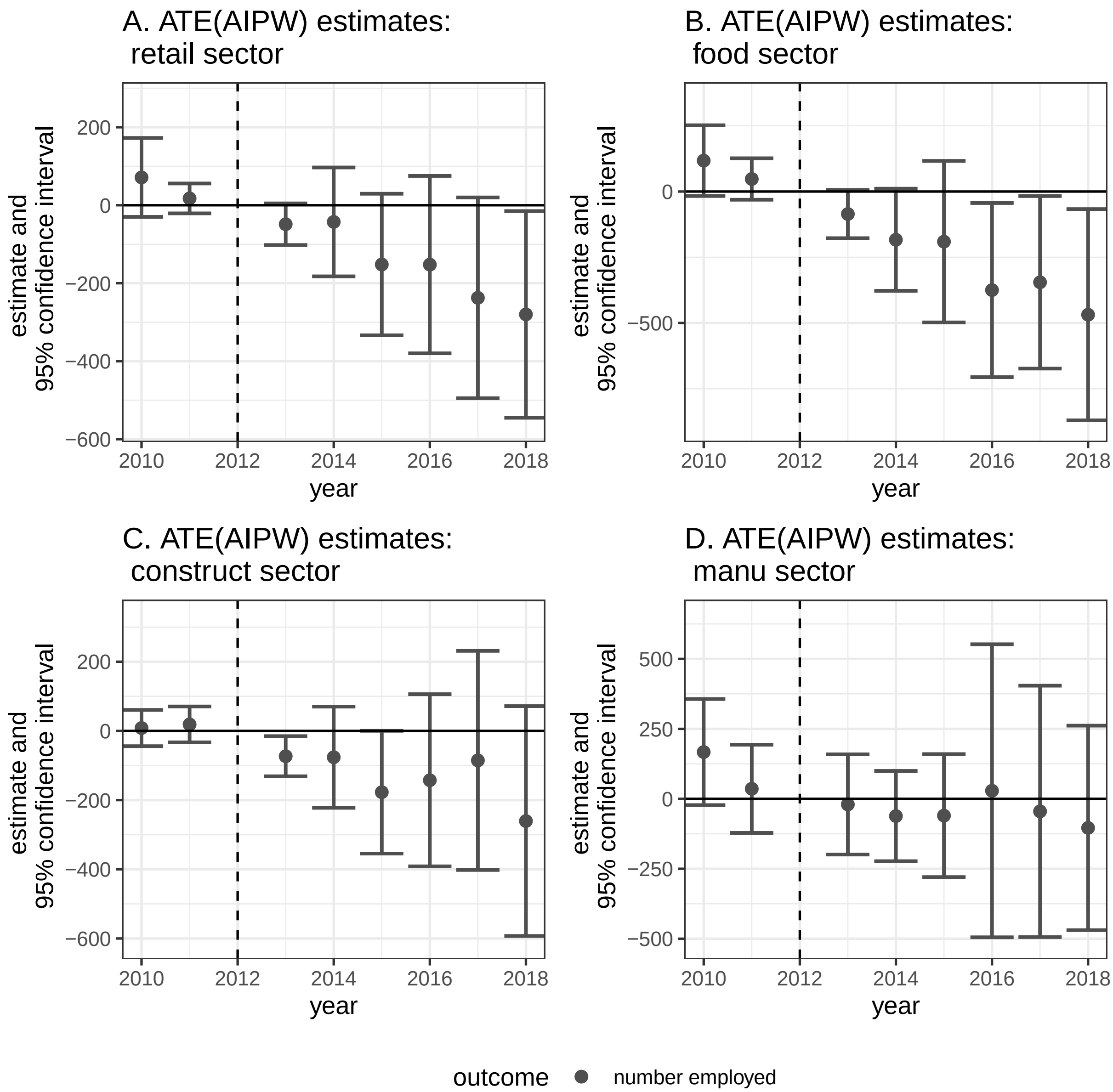

FIGURE I3: Event study estimates for the number employed by the industry sector (QCEW)

Note: The figure depicts ATE estimates obtained using the AIPW approach discussed in 4.4 for the reported response variable, the counts of total employed from QCEW for the given industry, along with the 95% confidence interval bars.

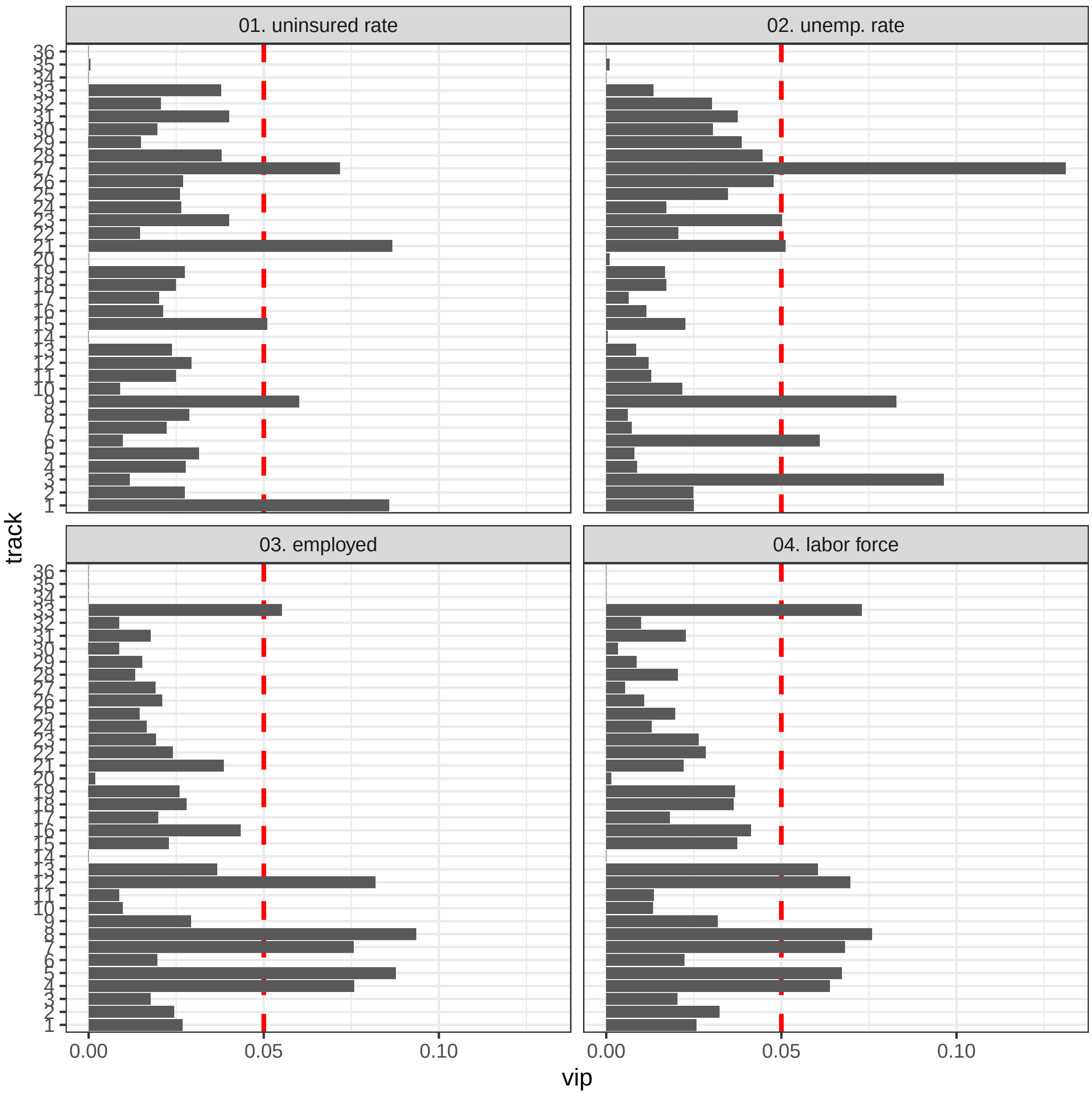

Figure I4

Note: The subfigures plot variable importance plot (VIP) from Causal Forest with 10,000 trees. 1) uninsured rate 2013 (income < 138% FPL), 2) Medicaid transfers per capita (2013), 3) unemployment rate 2012, 4) labor force 2012, 5) employed 2012, 6) unemployment rate 2010, 7) labor force 2010, 8) employed 2010, 9) mortality rate (2010-2013) per 1,000 population, 10) proportion Blacks (2010), 11) proportion Whites (2010), 12) county population (2010), 13) per capita income (2000) in 2010 dollars, 14) State minimum wage, 15) percent high school (2010), 16) percent college (2010), 17) poverty rate (2010), 18) household median income (2010), 19) household income age 35 (Whites), 20) Democrat county, 21) proportion of White votes for Obama (2008), 22) proportion of Trump votes (2016), 23) PM 2.5 measure (2010), 24) county area, 25) ruggedness, 26) latitude, 27) longitude, 28) elevation, 29) average precipitation (1961-1990), 30) average temperature (1961-1990), 31) cotton suitability measure, 32) malaria stability index, 33) Hill Burton Funds expenses, 34) FSSNAP benefits (3 person family), 35) State EITC Rate, 36) Refundable State EITC.

Table I2: Variable Importance Matrix obtained from Causal Forest (year prior to reform)

	response variable	features	importance measure
1	01. uninsured rate	county population (2010)	0.05
2	01. uninsured rate	household median income (2010)	0.06
3	01. uninsured rate	average precipitation (1961-1990)	0.06
4	01. uninsured rate	ruggedness	0.08
5	02. unemp. rate	ruggedness	0.05
6	02. unemp. rate	unemployment rate 2012	0.07
7	02. unemp. rate	Medicaid transfers per capita (2013)	0.07
8	02. unemp. rate	cotton suitability measure	0.08
9	02. unemp. rate	mortality rate (2010-2013) per 1,000 population	0.10
10	02. unemp. rate	PM 2.5 measure (2010)	0.11
11	03. employed	percent college (2010)	0.05
12	03. employed	labor force 2010	0.05
13	03. employed	labor force 2012	0.05
14	03. employed	employed 2012	0.06
15	03. employed	employed 2010	0.06
16	04. labor force	percent college (2010)	0.05
17	04. labor force	household median income (2010)	0.05
18	04. labor force	labor force 2012	0.05
19	04. labor force	per capita income (2000) in 2010 dollars	0.05
20	04. labor force	labor force 2010	0.06
21	04. labor force	Hill Burton Funds expenses	0.06
22	04. labor force	county population (2010)	0.07

Note: Variable importance matrix from Causal Forest for the reported response variables. The respose variable used is the first difference between the variables between the years 2012 and 2011 for all cases.

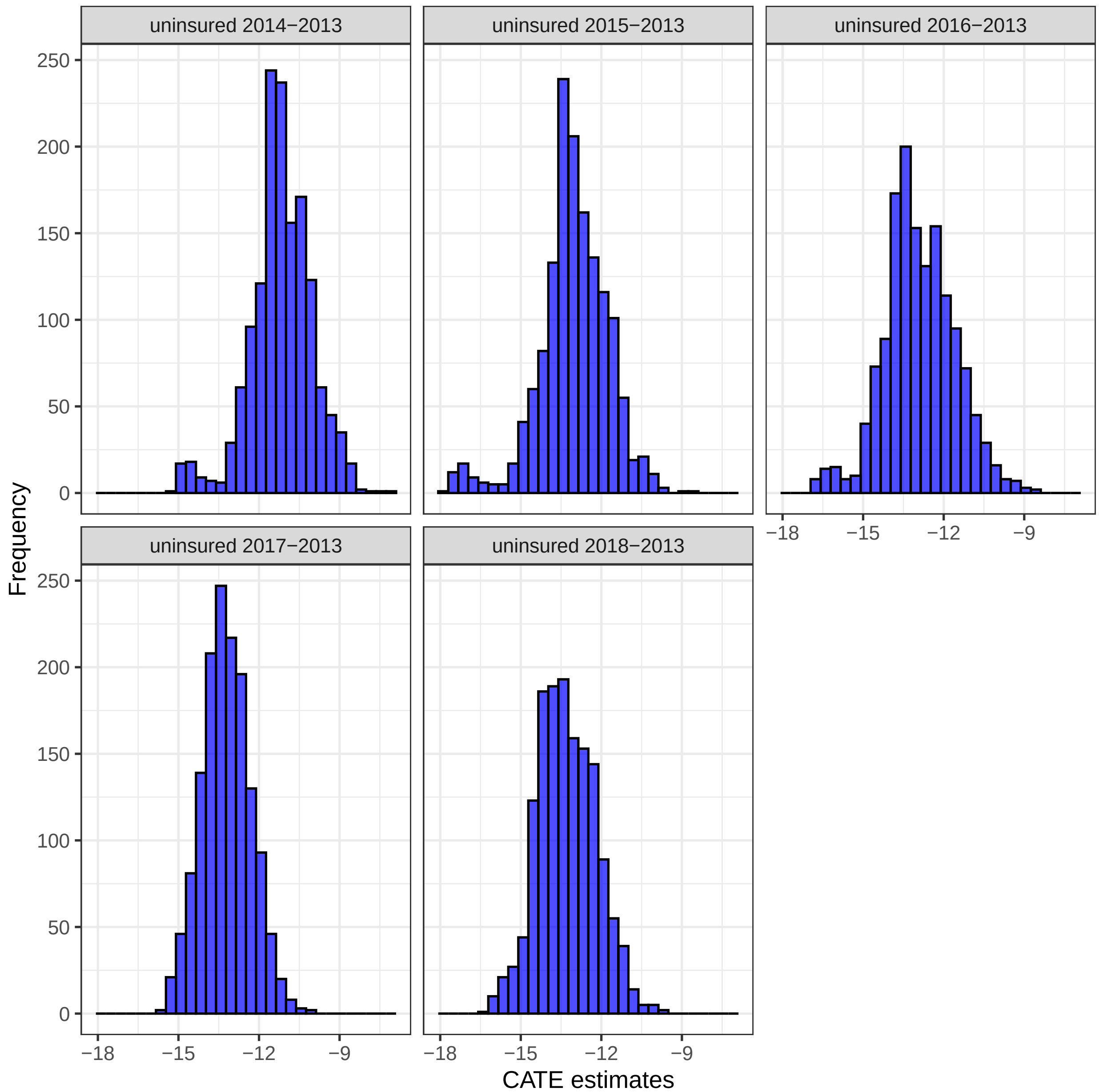

FIGURE I5: Histogram of CATE estimates: uninsured rate (138% below FPL)

Note: The figure plots the the histrogram of CATE estimates for each year when the outcome variable is the first difference in uninsured rate among people below 138% of FPL between years 2014 and 2013.

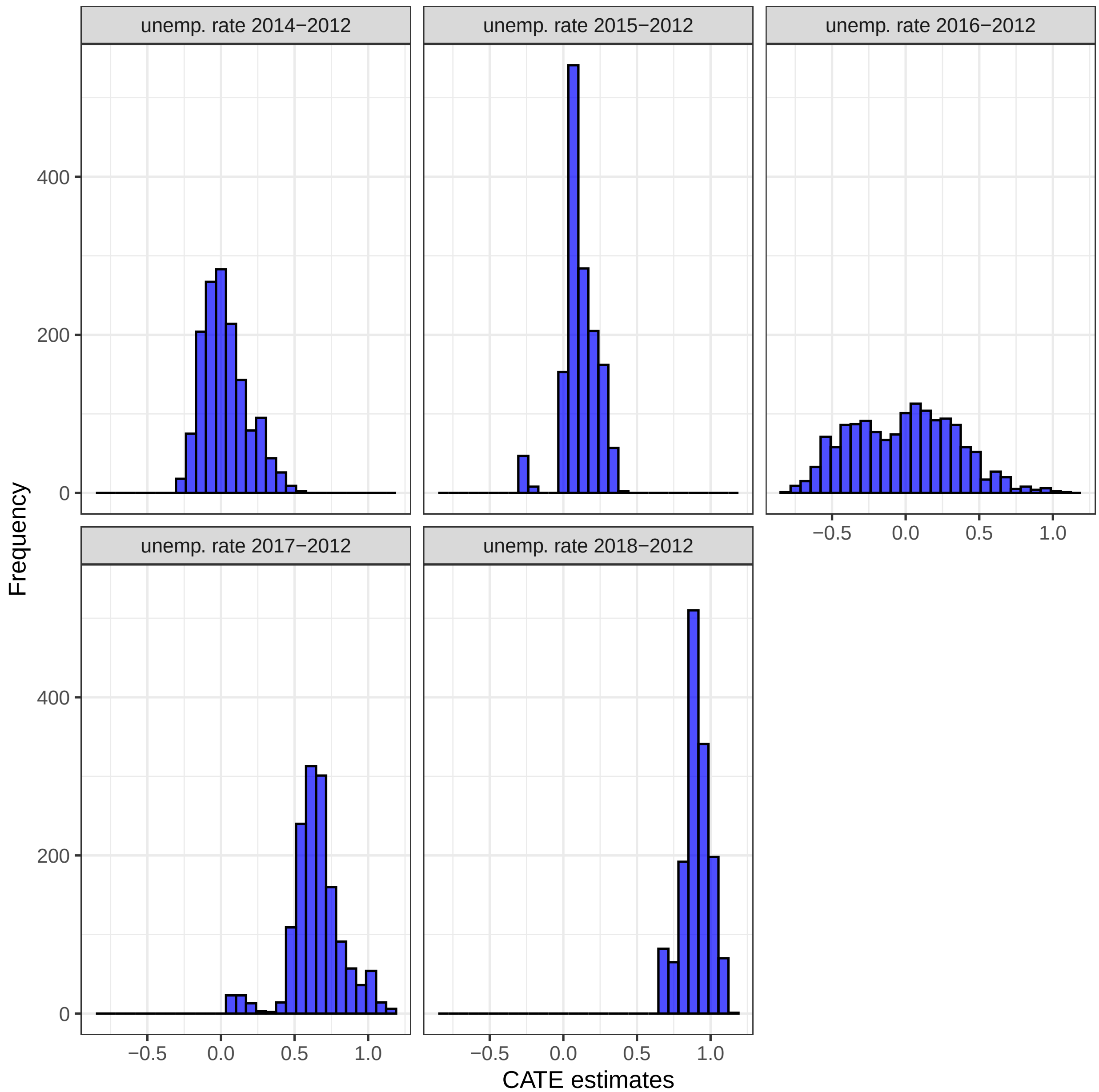

Figure I6: Histogram of CATE estimates: unemployment rate

Note: The figure plots the the histrogram of CATE estimates for each year when the outcome variable is the first difference in unemployment rate between years 2014 and 2013.

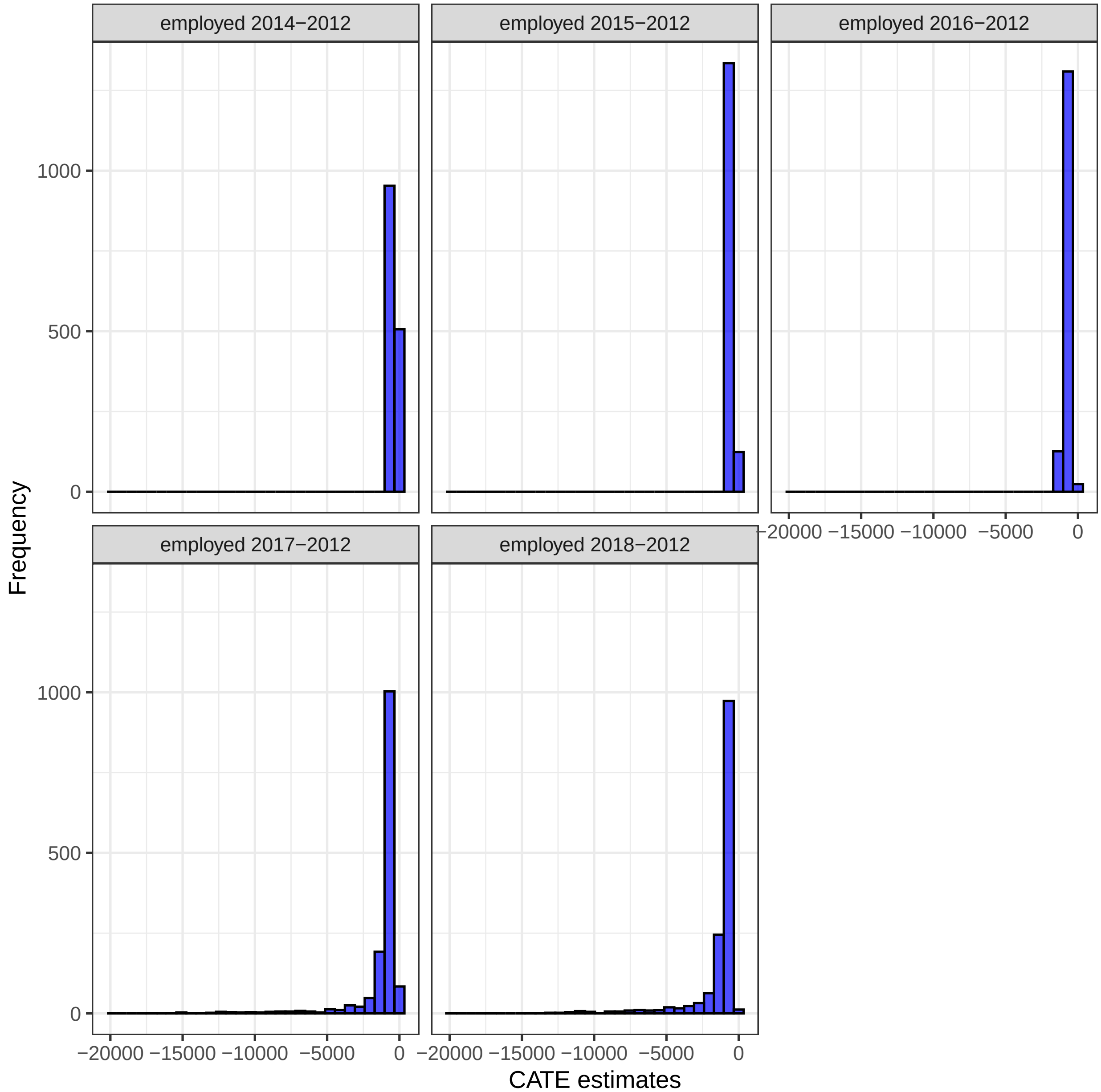

Figure I7: Histogram of CATE estimates: total employed

Note: The figure plots the the histrogram of CATE estimates for each year when the outcome variable is the first difference in total employed between years 2014 and 2013.

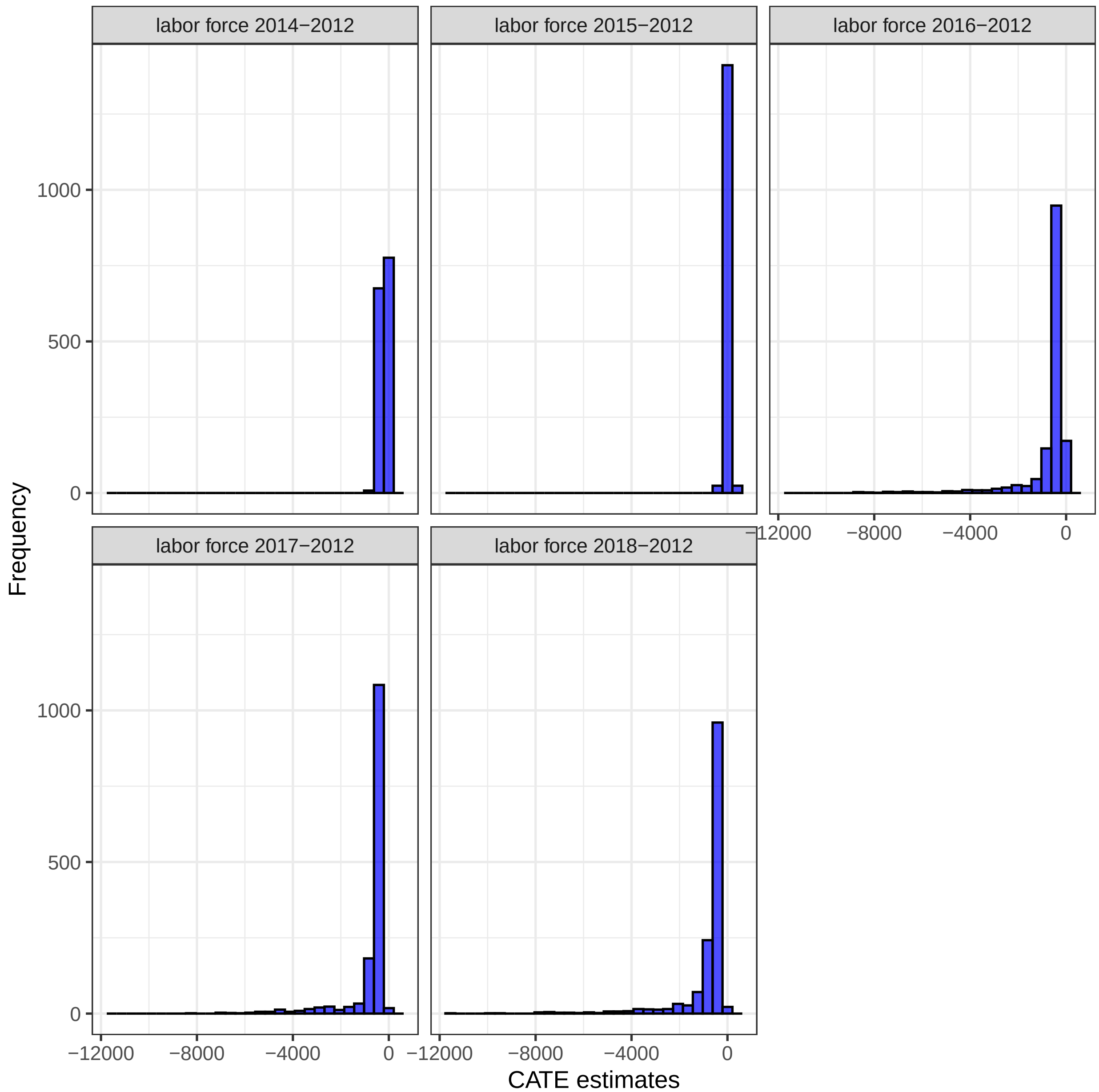

FIGURE I8: Histogram of CATE estimates: labor force

Note: The figure plots the the histrogram of CATE estimates for each year when the outcome variable is the first difference in labor force participation between years 2014 and 2013.

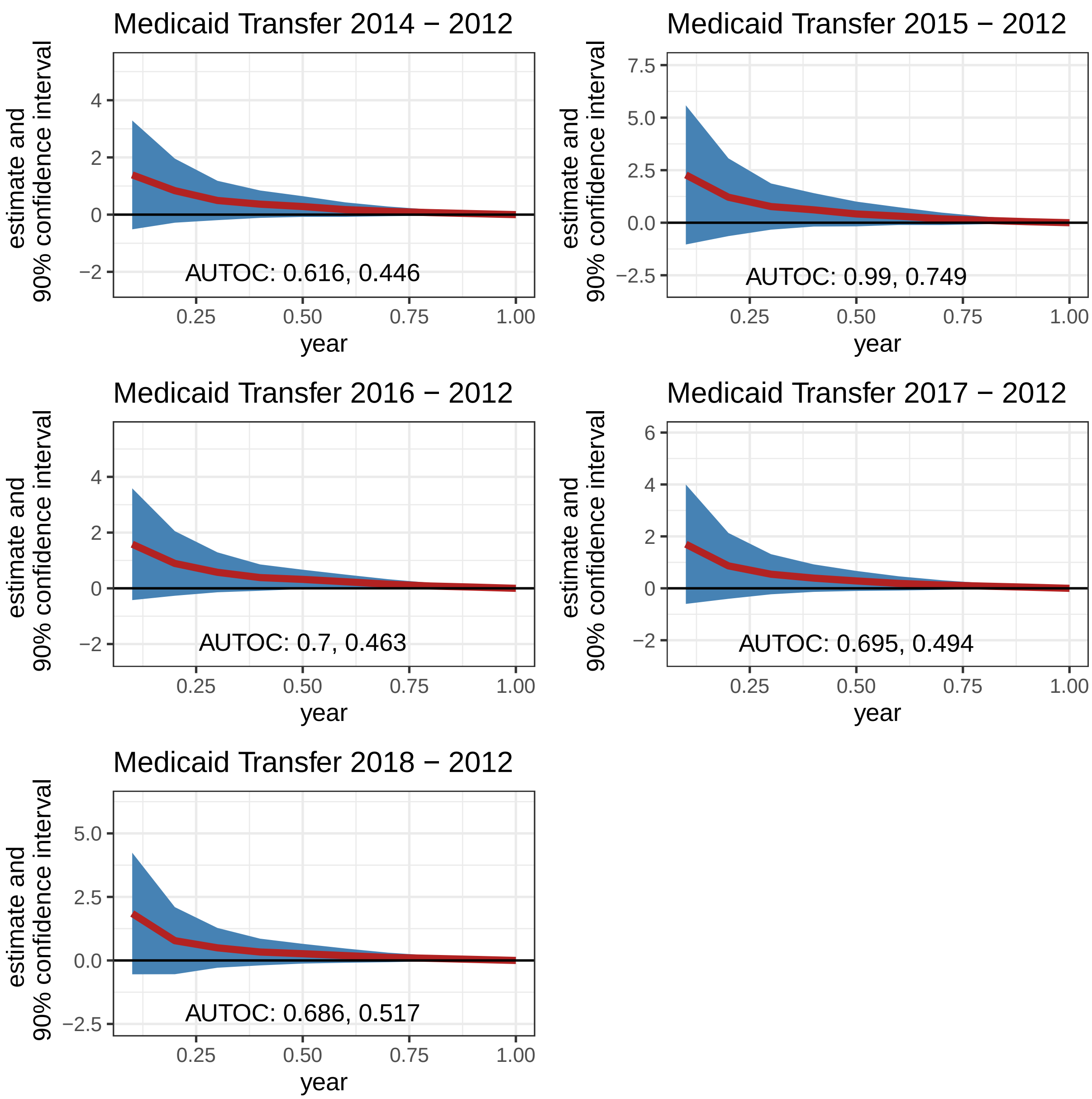

FIGURE I9: Histogram of CATE estimates: Medicaid transfers per capita

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in Medicaid transfers per capita between the given year and 2013 alongside the corresponding 90% confidence intervals. CATE estimates are used as the priority scores, S(.). The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the top.

9.3 Results from CF with boosted regression forest

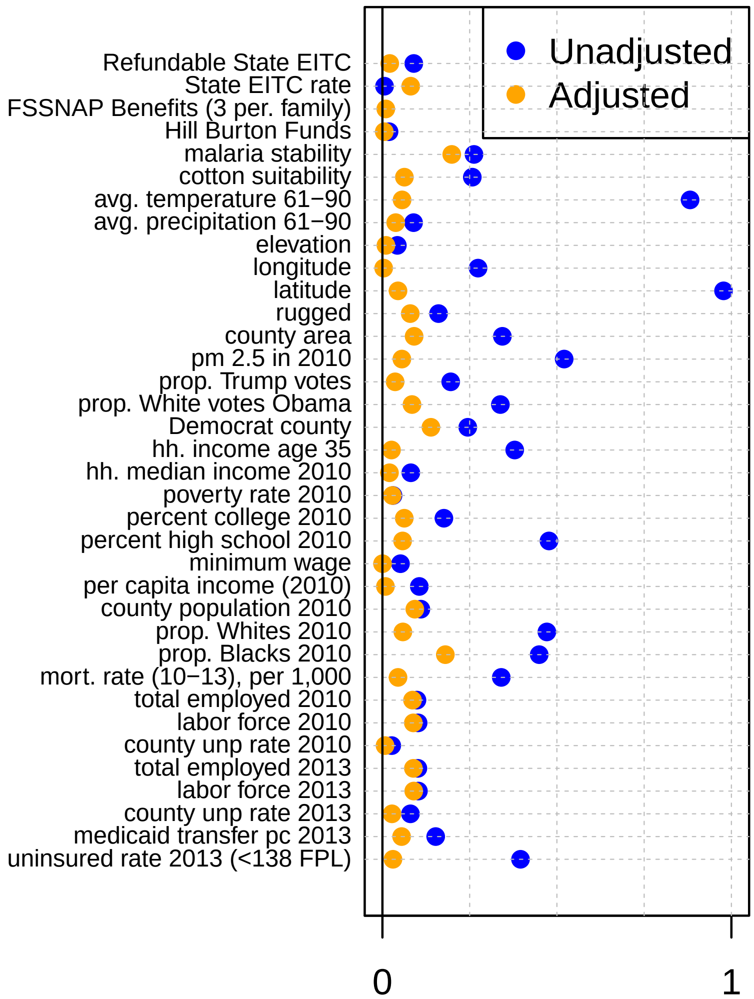

Figure I10

Note: The figure shows the absolute value of the standardized difference in mean between the treatment (expansion) and control (non-expansion) units. The unadjusted difference is the simple difference in mean between the treated and control units, where the estimates for $\hat{e}(x)$ are based on the boosted regression forest. The adjusted mean difference plots the absolute difference in means once the covariates are adjusted using inverse propensity weights.

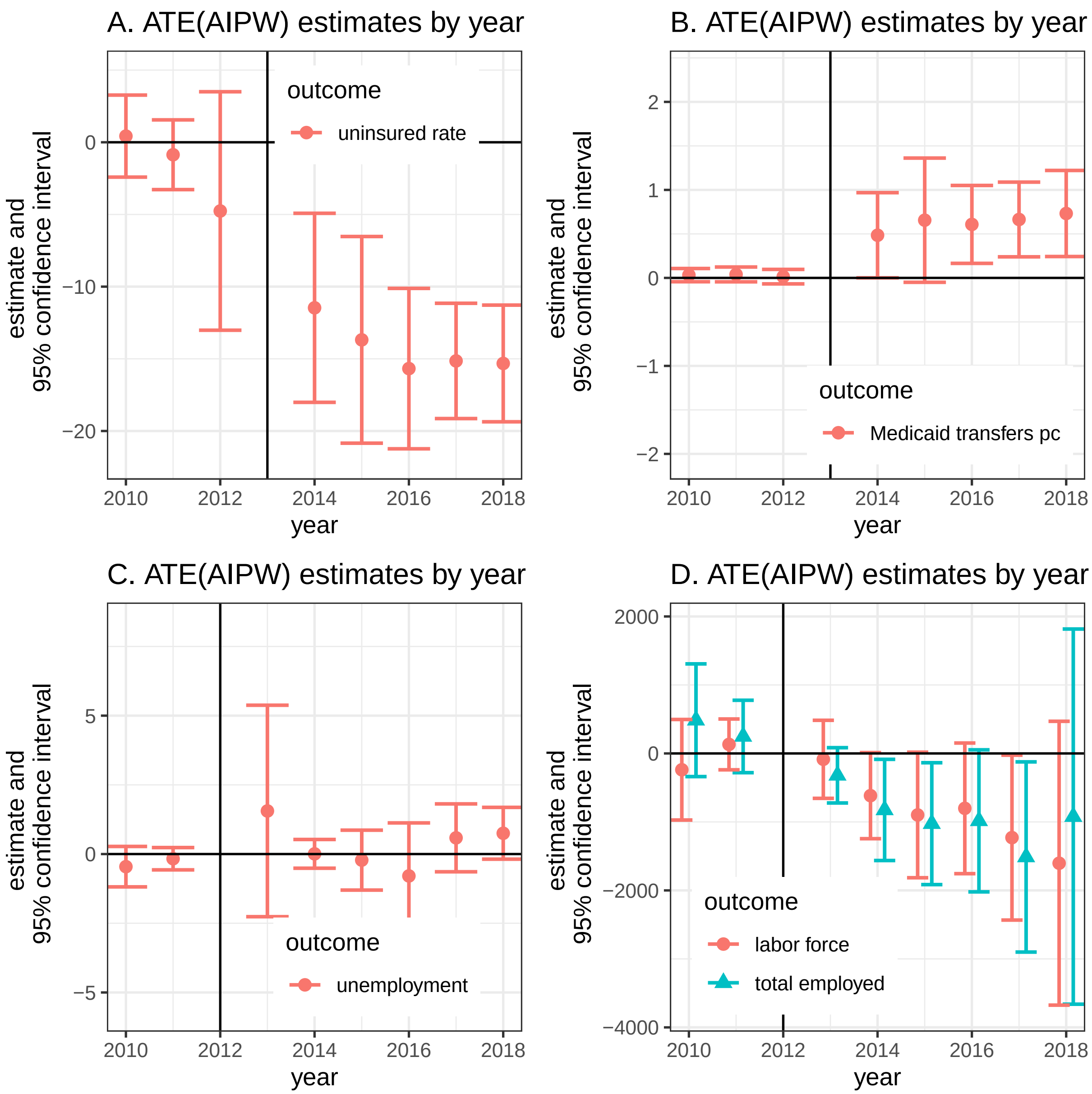

FIGURE I11: ATE Estimates from Augumented Inverse Probability Weighting (AIPW)

Note: The subfigures depict ATE estimates obtained using the AIPW approach discussed in 4.4 for the reported response variables along with the 95% confidence interval bars. The estimations of $\hat{e}(x)$ and $\hat{m}(x)$ are based on the boosted regression forest. The Causal Forest is ran on the centered and first differenced outcome. The year 2013 is used as the base year for uninsured rate, whereas the year 2012 is used for the labor market variables.

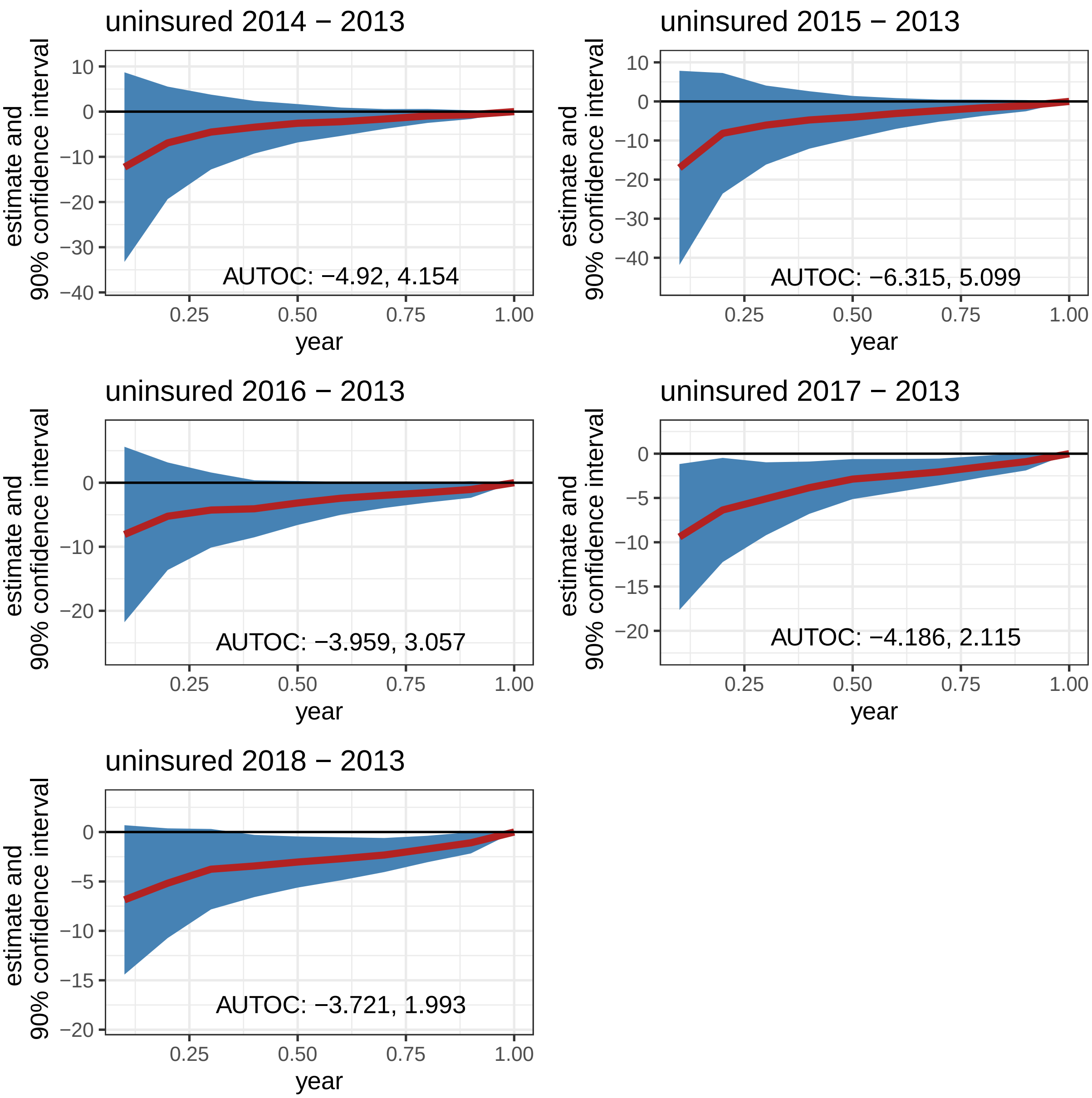

FIGURE I12: TOC curve from RATE: Uninsured Rate for people below 138% of FPL

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in uninsured rate for people living under 138% of FPL between the given year and 2013 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for the outcome variable area extracted from SAHIE. CATE estimates obtained from the trained CF model are used as the priority scores, S(.), when evaluating the RANK approach as discussed in section 4.5. CATE estimates are based on the model trained using a split of the sample that is independent from the subsample used for the purpose of RANK evaluation. Note that the estimations of $\hat{e}(x)$ and $\hat{m}(x)$ are based on the boosted regression forest. The Causal Forest is ran on the centered and first differenced outcome. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

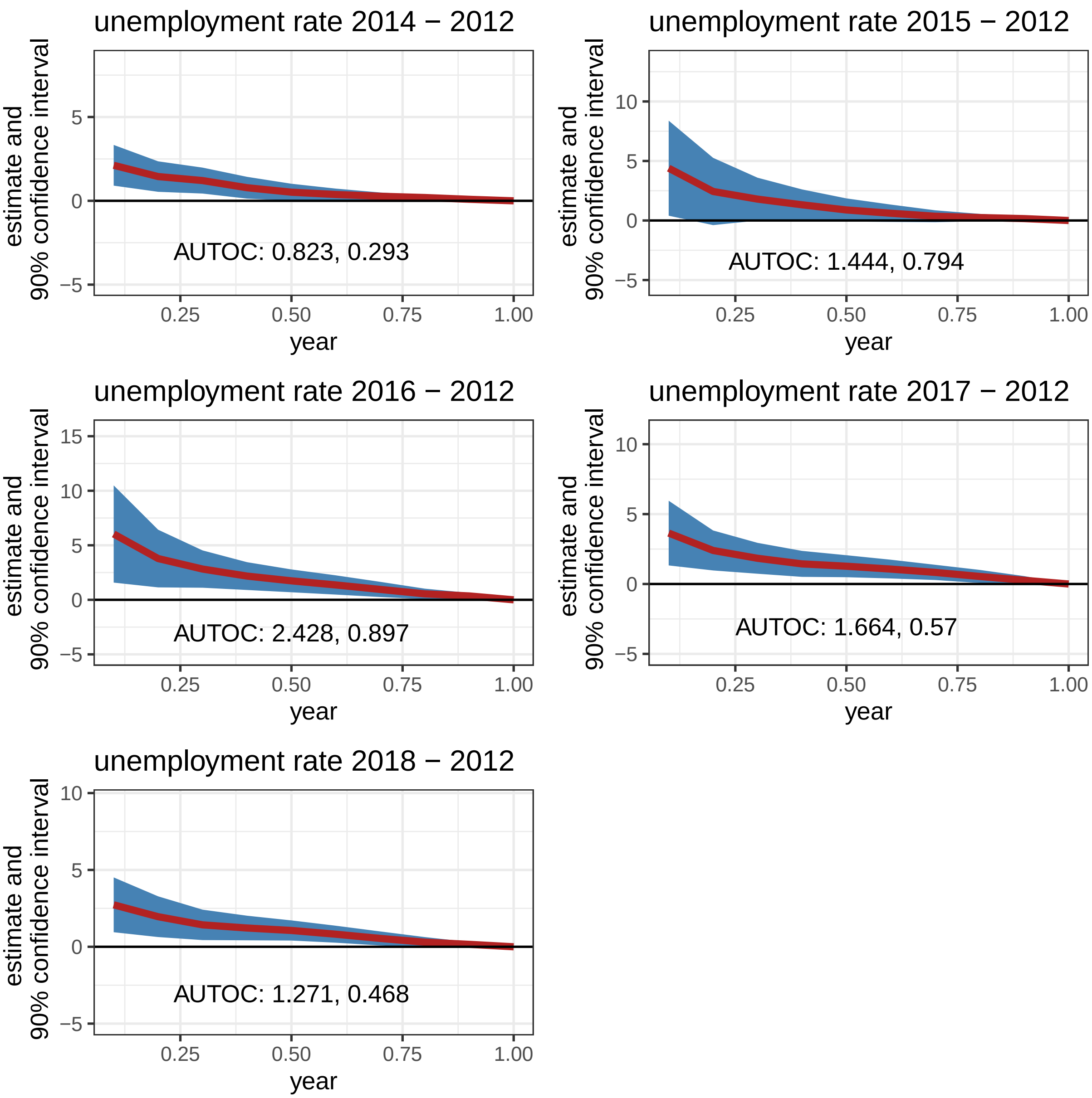

FIGURE I13: TOC curve from RATE: Unemployment Rate (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in unemployment rate between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for unemployment rate comes from BLS. CATE estimates obtained from the trained CF model are used as the priority scores, S(.), when evaluating the RANK approach as discussed in section 4.5. CATE estimates are based on the model trained using a split of the sample that is independent from the subsample used for the purpose of RANK evaluation. Note that the estimations of $\hat{e}(x)$ and $\hat{m}(x)$ are based on the boosted regression forest. The Causal Forest is ran on the centered and first differenced outcome. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the bottom.

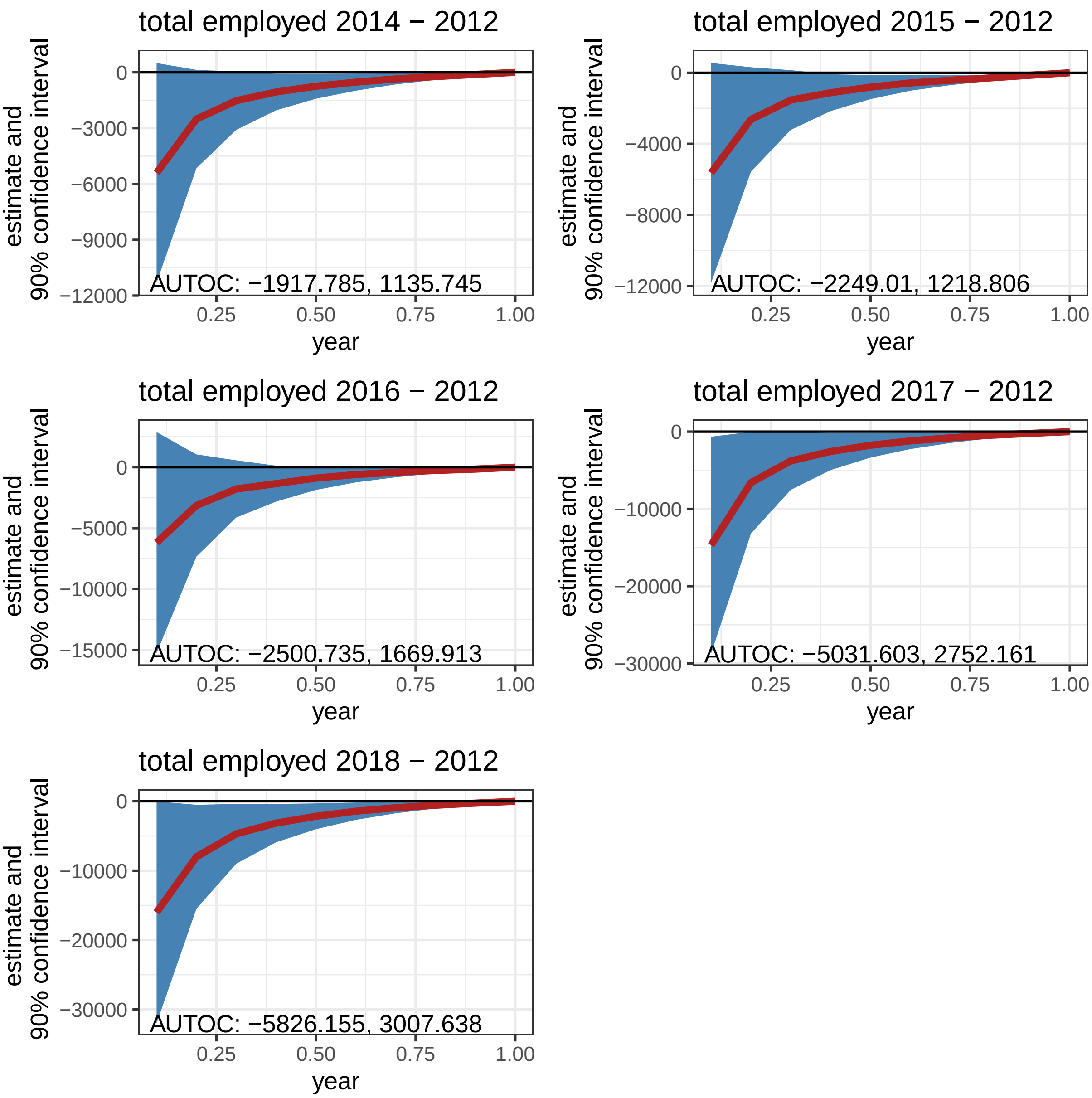

FIGURE I14: TOC curve from RATE: Total Employed (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in total employed between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for total employed comes from BLS. CATE estimates obtained from the trained CF model are used as the priority scores, S(.). CATE estimates are based on the model trained using a split of the sample that is independent from the subsample used for the purpose of RANK evaluation. Note that the estimations of $\hat{e}(x)$ and $\hat{m}(x)$ are based on the boosted regression forest. The Causal Forest is ran on the centered and first differenced outcome. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the top.

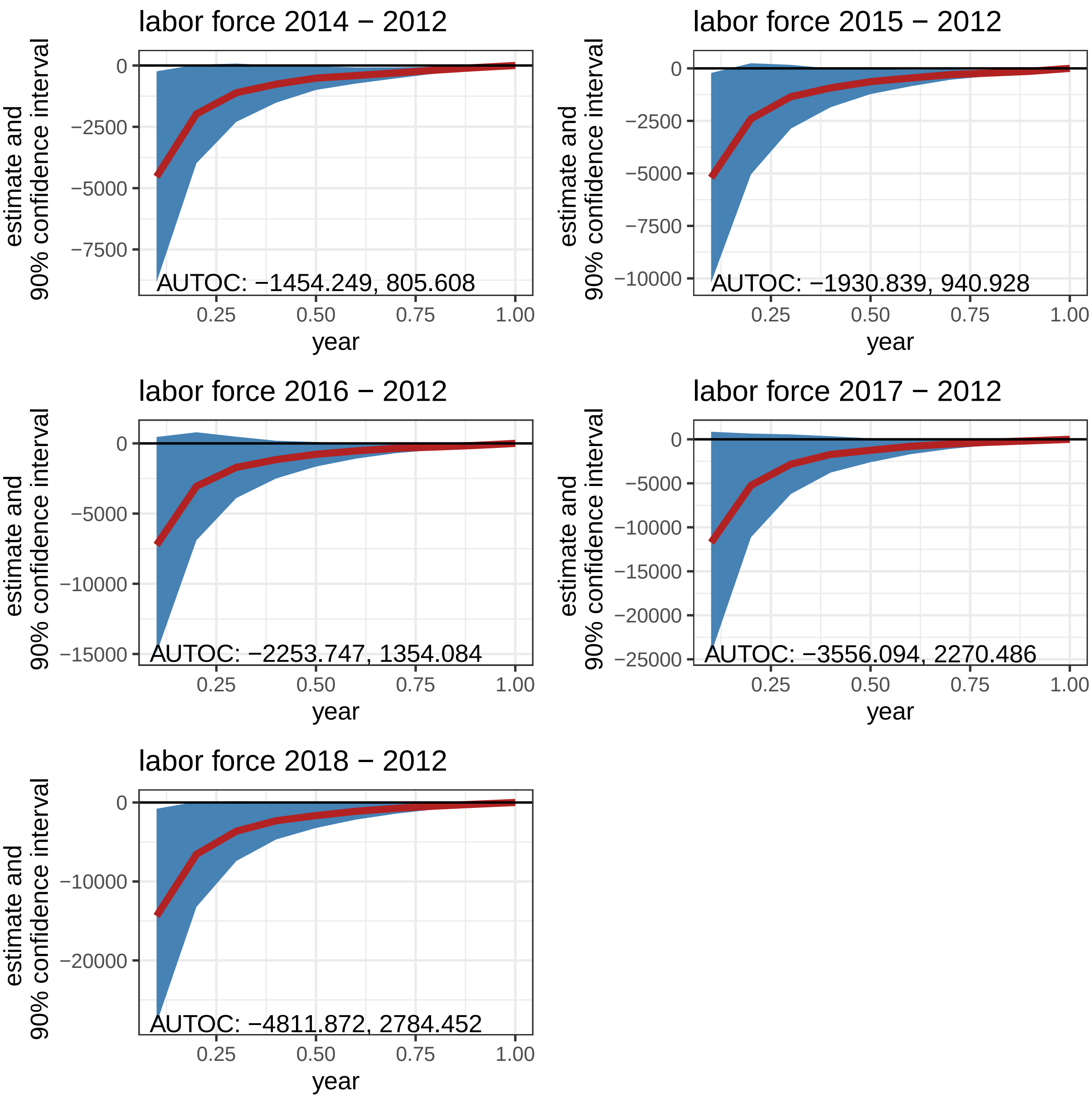

FIGURE I15: TOC curve from RATE: Labor Force (BLS)

Note: The subfigures depict Targetting Operator Characteristics (TOC) curves for the change in labor force between the given year and 2012 alongside the corresponding 90% confidence intervals obtained from bootstrap replications. Data for labor force participation comes from BLS. CATE estimates obtained from the trained CF model are used as the priority scores, S(.). CATE estimates are based on the model trained using a split of the sample that is independent from the subsample used for the purpose of RANK evaluation. Note that the estimations of $\hat{e}(x)$ and $\hat{m}(x)$ are based on the boosted regression forest. The Causal Forest is ran on the centered and first differenced outcome. The TOC curve is estimated using the framework given in equation 6. The AUTOC estimate and standard error are reported on the top.