

Minimum Wage Increases and Teen Time Allocation: Evidence from the American Time Use Survey

APEP Autonomous Research*

@anonymous

January 27, 2026

Abstract

How do state minimum wage increases affect teen labor supply and time allocation? Using time-diary data from the American Time Use Survey (2010–2023), I examine effects on the extensive margin (probability of any work on the diary day) and unconditional work time among 16–19 year olds. The identifying variation comes from states that newly exceeded the federal minimum wage floor (\$7.25) during this period, though relatively few states switched treatment status—most in 2014–2015—severely limiting statistical power. Using two-way fixed effects (TWFE) with year \times month fixed effects, I find that minimum wage increases reduce diary-day work time by 3.2 minutes (95% CI: [−12.4, 6.0]), statistically indistinguishable from zero. Decomposing this into extensive and intensive margins, I find a 1.8 percentage point reduction in the probability of any work (95% CI: [−6.9, 3.3]) with wide confidence intervals that cannot rule out economically meaningful effects in either direction. Continuous treatment specifications using log minimum wage or dollar gap above federal yield similar null results with even wider confidence intervals. Education time increases slightly (1.4 minutes), suggesting possible reallocation from work toward schooling, though estimates are imprecise. The primary contribution is methodological: ATUS time diaries provide outcomes measured within the same calendar month as the assigned treatment, avoiding the temporal misalignment that affects standard labor force surveys. However, identification is fundamentally constrained by minimal policy variation—a limitation that continuous treatment measures cannot fully overcome when few states change treatment status.

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

JEL Codes: J22, J23, J38

Keywords: minimum wage, teen employment, time use, difference-in-differences, extensive margin

1. Introduction

The minimum wage remains one of the most debated labor market policies in the United States. Since the federal minimum wage was last raised to \$7.25 in July 2009, over 30 states and numerous localities have enacted their own minimum wage increases, creating substantial geographic variation in wage floors. A central question in this debate concerns how minimum wage increases affect young workers, who are disproportionately employed in minimum-wage-sensitive industries and whose labor supply decisions are intertwined with schooling investments.

The empirical literature on minimum wages and teen employment is vast and contentious. Early work by [Neumark and Wascher \(2008\)](#) found significant negative employment effects using state-level variation, while [Card and Krueger \(1994\)](#) famously found no effect using a natural experiment in New Jersey and Pennsylvania. More recent studies using border discontinuity designs ([Dube et al., 2010](#); [Allegretto et al., 2011](#)) and bunching estimators ([Cengiz et al., 2019](#)) generally find smaller or null employment effects, though the debate continues ([Allegretto et al., 2017](#)). Studies of recent large minimum wage increases, such as Seattle's phased increase to \$15, find mixed evidence: [Jardim et al. \(2017\)](#) document hours reductions using administrative data, while others find minimal effects using different methodologies.

Beyond employment levels, minimum wages may affect the intensive margin of labor supply (hours worked conditional on employment) and time allocation more broadly. If minimum wage increases reduce labor demand for teens, they may reallocate time toward education—potentially improving human capital accumulation—or toward leisure. This time allocation margin has received less attention in the literature, though [Raissian and Su \(2023\)](#) examine parental time allocation using ATUS data.

This paper examines how state minimum wage increases above the federal floor affect teen labor supply and time allocation using data from the American Time Use Survey (ATUS). The ATUS provides a unique measurement advantage: time-diary outcomes (minutes of work, education, and leisure on the diary day) are recorded for a specific day and can be precisely matched to the minimum wage policy in effect during that calendar month. This contrasts with standard labor force surveys, including the Current Population Survey (CPS), where employment status is measured at reference dates that may not align with policy effective dates. When minimum wage changes occur mid-month or are close to the survey reference period, temporal misalignment can attenuate estimated effects.

The empirical approach exploits staggered state minimum wage increases above the federal floor of \$7.25 during 2010–2023, a period when the federal minimum wage remained

constant. I estimate two-way fixed effects (TWFE) models with state and year \times month fixed effects, using both binary treatment (whether state minimum wage exceeds federal) and continuous treatment specifications (log minimum wage, dollar gap above federal). To address concerns about TWFE bias under staggered adoption raised by [Goodman-Bacon \(2021\)](#), [de Chaisemartin and D'Haultfœuille \(2020\)](#), and [Callaway and Sant'Anna \(2021\)](#), I discuss the applicability of modern DiD estimators to this setting, though identification is severely constrained by minimal policy variation: relatively few states switched from below to above the federal floor during this period, with most switches concentrated in 2014–2015.

Following recent methodological guidance ([Roth et al., 2023](#); [Borusyak et al., 2024](#)), I decompose the unconditional mean effect into extensive and intensive margins. Diary-day work minutes are zero-inflated: 64% of teens report zero work minutes on the diary day. I estimate effects on the probability of any work (extensive margin), which combined with the unconditional mean allows inference about the intensive margin through the decomposition: $E[Y] = \Pr(\text{work} > 0) \times E[Y|\text{work} > 0]$. Direct estimation of the intensive margin conditional on working is not feasible because the working subsample ($N = 1,978$) further stratified across 51 states provides insufficient variation for reliable inference with clustered standard errors.

The main findings are as follows. First, TWFE estimates suggest minimum wage increases reduce diary-day work time by 3.2 minutes (95% CI: [−12.4, 6.0]), statistically indistinguishable from zero. This represents approximately 5% of the mean (65 minutes), but the confidence interval cannot rule out reductions as large as 19% of the mean. Second, decomposing into margins, the extensive margin effect is −1.8 percentage points (95% CI: [−6.9, 3.3]), again imprecise. Third, continuous treatment specifications using log minimum wage yield positive but highly imprecise coefficients, reflecting the limited within-state variation in minimum wage levels. Fourth, education time increases by 1.4 minutes while leisure increases by 3.6 minutes, suggesting possible reallocation from work toward other activities, though estimates are imprecise.

The paper makes three contributions. First, it provides the first estimates of minimum wage effects on teen time allocation using ATUS diary data with precise policy-month matching. Second, it discusses modern DiD methods appropriate for staggered adoption and transparently documents why these methods face severe limitations in this setting. Third, it implements an extensive margin decomposition, allowing inference about whether effects operate through changes in the probability of working or hours conditional on working.

However, the paper also clearly documents its limitations. With relatively few switcher states, identification is fragile, power is low, and results should be interpreted cautiously. The confidence intervals are wide enough to accommodate both the null hypothesis and economically meaningful effects. The contribution is as much methodological—demonstrating

the measurement advantages of ATUS for policy evaluation and the practical challenges of implementing modern DiD with limited policy variation—as it is substantive.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 provides institutional background on minimum wage policy. Section 4 describes the data. Section 5 presents the empirical strategy. Section 6 reports results. Section 7 discusses interpretation and limitations. Section 8 concludes.

2. Related Literature

This paper contributes to several literatures: the minimum wage employment debate, time allocation and labor supply measurement, and econometric methods for difference-in-differences with staggered adoption.

2.1 Minimum Wage Employment Effects

The debate over minimum wage employment effects has evolved substantially since the influential work of [Card and Krueger \(1994\)](#). The traditional view, articulated by [Neumark and Wascher \(2004\)](#), holds that minimum wage increases reduce employment, particularly among low-skilled workers. This view is grounded in the competitive labor market model where binding wage floors necessarily reduce labor demand.

More recent work has complicated this picture. [Dube et al. \(2010\)](#) use contiguous county-pair designs to control for spatial heterogeneity and find minimal employment effects. [Cengiz et al. \(2019\)](#) use bunching estimators to examine the distribution of employment below and above the minimum wage, finding that minimum wage increases lead to reallocation from below to above the threshold without substantial disemployment. These findings are consistent with labor market frictions, monopsony power, or efficiency wage effects.

The debate continues with contributions examining employment dynamics ([Meer and West, 2016](#)), specific policy changes like Seattle’s \$15 minimum ([Jardim et al., 2017](#)), and broader syntheses ([Neumark and Wascher, 2019; Dube, 2019](#)). A key insight from this literature is that identification strategy matters enormously: different comparison groups and econometric approaches can yield substantially different conclusions.

For teen employment specifically, the evidence is mixed. Some studies find larger negative effects for teens than for adults ([Neumark and Wascher, 2004](#)), while others find minimal effects ([Allegretto et al., 2011](#)). Teen labor markets may differ from adult markets in important ways: teens have more elastic labor supply (schooling as an outside option), lower productivity on average, and are concentrated in minimum-wage-intensive industries.

2.2 Time Allocation and Labor Supply Measurement

Time use data offer a distinct perspective on labor supply. While standard labor force surveys like the CPS measure employment status and usual hours, time diaries capture actual time allocation on a specific day. This distinction matters for several reasons.

First, diary-based measures avoid recall bias that may affect retrospective reports of hours worked ([Hamermesh et al., 2005](#)). Second, diaries capture the full allocation of time, allowing examination of how work time substitutes with education, leisure, and other activities. Third, and most relevant for this paper, diaries can be precisely matched to policy conditions on the survey date.

[Aguiar and Hurst \(2007a\)](#) and [Aguiar and Hurst \(2007b\)](#) demonstrate how ATUS data can illuminate changes in time allocation over the life cycle and in response to economic conditions. [Raissian and Su \(2023\)](#) use ATUS to examine parental time allocation responses to minimum wages. This paper extends this approach to teen time allocation, exploiting the policy-month alignment that ATUS provides.

The measurement advantage of time diaries is particularly relevant when studying policies with discrete effective dates. Standard labor force surveys may measure employment status weeks or months before or after a policy change, creating temporal misalignment that can attenuate estimated effects. By matching treatment status to the diary month, ATUS potentially provides more precise measurement of policy exposure.

2.3 Difference-in-Differences with Staggered Adoption

Recent econometric work has highlighted problems with two-way fixed effects (TWFE) estimators when treatment timing varies across units. [Goodman-Bacon \(2021\)](#) shows that TWFE can produce biased estimates when treatment effects are heterogeneous across cohorts or over time, because it implicitly uses already-treated units as controls for newly-treated units. [de Chaisemartin and D'Haultfœuille \(2020\)](#) and [Callaway and Sant'Anna \(2021\)](#) develop alternative estimators that avoid these comparisons.

These methodological advances have important implications for minimum wage research, which typically exploits staggered state adoption. [Sun and Abraham \(2021\)](#) develop an interaction-weighted estimator that allows for heterogeneous treatment effects across adoption cohorts. [Borusyak et al. \(2024\)](#) propose an imputation approach that efficiently uses all clean comparisons.

However, implementing these modern estimators requires sufficient identifying variation. With few treatment cohorts or few units switching treatment status, even well-designed estimators may lack power. [Roth et al. \(2023\)](#) provide guidance on when these concerns are

most severe and how to diagnose them.

For inference, the standard practice of clustering standard errors at the state level may be inappropriate when identification comes from only a handful of switching units. [Conley and Taber \(2011\)](#) address inference in settings with few policy changes, developing procedures that are valid when the effective number of treated units is small. These concerns are particularly acute in minimum wage research where a small number of states may drive identification.

3. Institutional Background

3.1 Federal and State Minimum Wages

The federal minimum wage has been set at \$7.25 per hour since July 2009, following a three-step increase from \$5.15 between 2007 and 2009. In the absence of federal action for over 15 years, many states have raised their own minimum wages substantially above the federal floor. As of 2023, 30 jurisdictions (29 states plus the District of Columbia) have minimum wages exceeding \$7.25, with rates ranging from \$8.00 to over \$16.00 per hour in high-cost states like California and Washington.

The timing and magnitude of state minimum wage increases exhibit substantial variation. Some states, like Washington and Oregon, indexed their minimum wages to inflation as early as the late 1990s, resulting in gradual annual increases. Other states, like California and New York, enacted large legislated increases in the 2010s with scheduled phase-ins reaching \$15 per hour. Still other states have never raised their minimum wage above the federal floor, including much of the Deep South (Alabama, Georgia, Louisiana, Mississippi, South Carolina, Tennessee) and several other states (Wyoming, Oklahoma). This geographic variation provides a natural comparison group.

3.2 The 2010–2023 Analysis Window

I restrict the analysis to 2010–2023, the period when the federal minimum wage remained constant at \$7.25. This restriction is important for identification: during this period, changes in the binary treatment indicator (state minimum wage above federal) reflect only state policy decisions, not mechanical reclassification when the federal floor changes.

During 2010–2023, the treatment landscape evolved as follows:

- **Always-treated states:** Many states already had minimum wages above \$7.25 in January 2010. These states contribute to the comparison through their continued presence but do not contribute to the “switching” variation that identifies the TWFE effect.

- **Never-treated states:** Approximately 20 states maintained minimum wages at or below the federal floor throughout 2010–2023. These serve as the primary control group.
- **Switcher states:** A small number of states switched from below to above the federal floor during this window, with most switches occurring in 2014–2015.

The small number of switchers—concentrated in a narrow time window—is a fundamental limitation. With state fixed effects, identification of the treatment effect comes primarily from within-state changes in treatment status. When most treated states are “always-treated” (already above federal in 2010), they contribute limited identifying variation.

3.3 Local Minimum Wages

An additional complication is the rise of local (city and county) minimum wages during this period, particularly in large cities like Seattle, San Francisco, New York City, Los Angeles, and Chicago. These local minimums often exceed state minimums, creating within-state variation that is not captured by my state-level treatment measure. To the extent that local minimum wages affect teen employment in major urban areas, the state-level treatment may underestimate the true minimum wage exposure for teens in those cities. This measurement error would typically attenuate estimated effects toward zero.

3.4 Teen Employment and Minimum Wage Industries

Teenagers (ages 16–19) are particularly relevant for studying minimum wage effects for several reasons. First, they are disproportionately employed in industries where minimum wage workers are concentrated. According to the Bureau of Labor Statistics, workers under age 25 constitute roughly half of all workers paid the federal minimum wage or less, despite representing only about 20 percent of hourly workers. Teen workers are concentrated in retail trade, food services, and entertainment/recreation—industries where entry-level wages often start at or near the minimum wage.

Second, teen labor supply decisions are closely linked to schooling decisions. Unlike prime-age adults for whom employment is typically the primary activity, teens face a three-way allocation between work, school, and leisure. Minimum wage increases that reduce labor demand may induce substitution toward schooling—a potentially positive human capital response—or toward leisure.

Third, teens may be particularly vulnerable to disemployment effects if their productivity is below the minimum wage. While the empirical evidence on teen employment elasticities varies widely, teens are often viewed as the population most likely to experience adverse employment effects from minimum wage increases ([Neumark and Wascher, 2004](#)).

4. Data

4.1 American Time Use Survey

The primary data source is the American Time Use Survey (ATUS), conducted by the Bureau of Labor Statistics in partnership with the Census Bureau since 2003. The ATUS is the nation's only federally-administered, continuous time use survey. Each year, approximately 10,000 individuals aged 15 and older complete a detailed time diary recording how they spent their time during a designated 24-hour diary day.

I access the ATUS through IPUMS Time Use, which provides harmonized extracts with consistent variable definitions across years. The initial sample includes ATUS respondents aged 16–19 from 2010 to 2023. The resulting sample contains 5,764 teen observations, of which 5,455 have valid sampling weights and comprise the regression sample.

4.1.1 Time-Use Outcomes

The primary outcomes are minutes spent in various activities on the diary day:

- **Work time:** Minutes spent in paid work activities
- **Education time:** Minutes spent in educational activities (attending class, studying, etc.)
- **Leisure time:** Minutes spent in leisure and sports activities
- **Job search time:** Minutes spent searching for employment

Work time is zero-inflated: 64% of teens report zero work minutes on the diary day. Mean work time is 65 minutes for all teens, but 180 minutes conditional on any work. To address this, I decompose the work time outcome as follows:

- **Extensive margin:** Indicator for any work on the diary day ($\text{any_work} = \mathbf{1}[\text{work_time} > 0]$)
- **Unconditional mean:** Total work minutes including zeros

The intensive margin (work minutes conditional on working) can be inferred from these quantities but is not separately estimated due to the limited sample of workers ($N = 1,978$) spread across 51 states.

4.1.2 Employment Status

The ATUS inherits employment status (EMPSTAT) from the CPS. **Important caveat:** EMPSTAT reflects employment status during the CPS reference week, which occurs 2–5 months before the ATUS diary date. Because treatment is assigned based on the diary month, there is systematic temporal misalignment. For observations near January or July policy effective dates, EMPSTAT may have been measured before the minimum wage change took effect. For this reason, the time-diary work measure is preferred: it captures work on a specific day that can be precisely matched to the policy month.

4.2 Minimum Wage Data

State minimum wage data come from the U.S. Department of Labor’s Wage and Hour Division, supplemented with the Economic Policy Institute’s Minimum Wage Tracker for exact effective dates. I construct a state-by-year-month panel recording the effective minimum wage (the higher of state and federal) for each state and time period.

I use three treatment measures:

1. **Binary indicator:** $D_{st} = \mathbf{1}[\text{MW}_{st} > \$7.25]$, indicating whether the state minimum wage exceeds the federal floor. This is the primary treatment for the extensive margin of policy change.
2. **Log minimum wage:** $\log(\text{MW}_{st})$, which captures continuous variation in minimum wage levels.
3. **Dollar gap:** $\text{MW}_{st} - \$7.25$, the dollar amount above the federal floor.

The continuous measures (log MW and dollar gap) are intended to capture the intensive margin of policy—exploiting variation from within-treated-state increases, not just the crossing of the federal threshold. However, identification of the continuous treatment effect still requires within-state changes over time, and with most high-MW states already above federal in 2010 and gradually increasing thereafter, there may be limited independent variation.

4.3 Summary Statistics

Table 1 presents summary statistics for the teen sample. The employment rate (from EMPSTAT) is 36 percent, with modest variation by treatment status. Teens in treated state-months have a 34 percent employment rate compared to 38 percent in control state-months. This unconditional difference is confounded by compositional differences between treated and control states, motivating the regression approach.

The time-use outcomes show substantial variation. Mean work time is 65 minutes (unconditional on working), with 36% of teens reporting any work on the diary day. Mean education time is 223 minutes, reflecting the high school enrollment of most teens.

Table 1: Summary Statistics: Teen Sample (Ages 16–19, 2010–2023)

| | Full Sample | Treated | Control |
|-----------------------------|-------------|---------|---------|
| <i>Demographics</i> | | | |
| Age (years) | 17.3 | 17.3 | 17.3 |
| Female (%) | 48.1 | 48.5 | 47.8 |
| White (%) | 78.3 | 76.5 | 79.8 |
| Enrolled in school (%) | 75.3 | 78.1 | 73.0 |
| <i>Employment (EMPSTAT)</i> | | | |
| Employment rate (%) | 36.2 | 34.2 | 37.9 |
| <i>Time Use (diary day)</i> | | | |
| Work time (min) | 65.3 | 61.2 | 68.8 |
| Any work (%) | 36.3 | 33.8 | 38.4 |
| Work time working (min) | 180.0 | 181.2 | 179.1 |
| Education time (min) | 223.4 | 231.5 | 216.5 |
| Leisure time (min) | 254.2 | 248.3 | 259.2 |
| Observations | 5,455 | 2,636 | 2,819 |
| Jurisdictions | 51 | 30 | 37 |

Notes: Sample includes ATUS respondents ages 16–19, 2010–2023. “Treated” and “Control” columns are observation-level classifications based on whether the state minimum wage exceeded the federal floor in the diary month. Sixteen states contribute observations to both categories—these are states that crossed above the federal floor during 2010–2023, providing the within-state variation that identifies treatment effects. Most switches occurred in 2014–2015, with additional switches in 2010–2012 and 2017–2019. Work time conditional on working is calculated for the 36.3% of observations with any work on the diary day. Means are unweighted; regressions use ATUS sampling weights.

5. Empirical Strategy

5.1 Identification

I estimate the effect of state minimum wage increases using a difference-in-differences design exploiting the staggered timing of state minimum wage increases above the federal floor. The identifying assumption is that, in the absence of minimum wage increases, teen outcomes in treated states would have evolved parallel to outcomes in control states.

Formally, let Y_{ist} denote an outcome for individual i in state s at time t , and let D_{st} indicate treatment. The parallel trends assumption requires:

$$\mathbb{E}[Y_{ist}(0) - Y_{is,t-1}(0)|D_{st} = 1] = \mathbb{E}[Y_{ist}(0) - Y_{is,t-1}(0)|D_{st} = 0] \quad (1)$$

where $Y_{ist}(0)$ denotes the potential outcome under no treatment.

This assumption may be violated if states that raise minimum wages are on different outcome trajectories than states that do not, or if other policies (EITC expansions, schooling reforms, labor market conditions) change differentially across states. With relatively few switcher states concentrated in 2014–2015, the assumption is difficult to assess rigorously and the design is vulnerable to confounding from state-specific shocks or concurrent policy changes.

5.2 TWFE Estimation

The baseline specification is a two-way fixed effects (TWFE) regression:

$$Y_{ist} = \alpha + \tau D_{st} + \gamma_s + \delta_{ym} + \varepsilon_{ist} \quad (2)$$

where γ_s are state fixed effects and δ_{ym} are year×month fixed effects (168 cells for 14 years × 12 months). Standard errors are clustered at the state level. Individual controls (age, sex, race) are added in robustness checks (Table 7, column 2).

The use of year×month fixed effects rather than year fixed effects alone is motivated by strong seasonality in teen employment and minimum wage policy timing. Employment rates peak in summer months when school is out, and minimum wage effective dates often fall in January or July. Year×month fixed effects absorb these within-year patterns.

5.3 Continuous Treatment

To exploit variation in minimum wage levels beyond the binary crossing of \$7.25, I also estimate:

$$Y_{ist} = \alpha + \tau \log(\text{MW}_{st}) + \gamma_s + \delta_{ym} + X'_{ist}\beta + \varepsilon_{ist} \quad (3)$$

where τ represents the effect of a one-unit increase in $\log(\text{MW})$. For interpretation, a 10% increase in MW corresponds to $\Delta \log(\text{MW}) \approx 0.10$, so the effect of a 10% MW increase is approximately $0.10 \times \tau$. I also estimate specifications using the dollar gap above federal.

5.4 Extensive Margin Decomposition

Because diary-day work time is zero for 64% of observations, I estimate:

1. **Extensive margin:** $\Pr(\text{any_work}_{ist} = 1 | X) = \tau D_{st} + \dots$
2. **Unconditional mean:** $\mathbb{E}[\text{work_time}_{ist} | X] = \tau D_{st} + \dots$

This decomposition follows [Cengiz et al. \(2019\)](#) and helps interpret whether any unconditional effect operates through changes in the probability of working versus changes in hours conditional on working. The intensive margin (hours conditional on working) can be inferred from these two quantities through the decomposition $E[Y] = \Pr(\text{work} > 0) \times E[Y | \text{work} > 0]$, but is not separately estimated because the working subsample has insufficient variation for reliable state-clustered inference.

5.5 Modern DiD Methods

Recent methodological work has shown that TWFE estimators can be biased in settings with staggered treatment timing and heterogeneous treatment effects ([Goodman-Bacon, 2021](#); [de Chaisemartin and D'Haultfœuille, 2020](#)). [Callaway and Sant'Anna \(2021\)](#) and [Sun and Abraham \(2021\)](#) develop alternative estimators that are robust to treatment effect heterogeneity.

Despite the limitations of this setting (few switcher states, treatment concentrated in 2014–2015), I implement three modern DiD estimators to verify that the null finding is robust to estimation method. Following the guidance at psantanna.com/did-resources, I estimate:

1. **Gardner (2022) two-stage DiD:** The `did2s` estimator uses a two-stage procedure that first residualizes outcomes using only not-yet-treated/never-treated observations, then estimates the treatment effect from the residualized treated observations.

2. **Borusyak, Jaravel, & Spiess (2024) imputation:** The imputation estimator predicts untreated potential outcomes for treated units using the structure of the two-way fixed effects model, then computes the ATT as the average difference between observed and imputed outcomes.
3. **Stacked DiD:** For each treatment cohort, I create a separate dataset with that cohort's treated units and never-treated/not-yet-treated controls within a ± 5 year window, then stack these datasets and estimate with cohort-specific fixed effects.

Table 2 presents the results. The Gardner two-stage estimator yields an ATT of -4.4 minutes ($SE = 7.1$), the BJS imputation estimator yields -8.7 minutes ($SE = 11.3$), and the stacked DiD yields -10.0 minutes ($SE = 11.8$). All estimates are negative and statistically insignificant, consistent with the baseline TWFE estimate of -9.7 minutes on the annual state panel.

The consistency across estimators is reassuring: despite using different weighting schemes and comparison groups, all methods yield qualitatively similar results—small negative point estimates that are statistically indistinguishable from zero. The slightly larger standard errors from modern methods reflect their reliance on cleaner but smaller comparison samples (never-treated only) rather than all available variation.

Table 2: Comparison of Modern DiD Estimators

| Method | Estimate | Std. Error | 95% CI | p-value |
|---------------------------|----------|------------|-----------------|---------|
| TWFE (annual panel) | -9.67 | 9.57 | $[-28.4, 9.1]$ | 0.31 |
| Gardner two-stage (did2s) | -4.42 | 7.08 | $[-18.3, 9.5]$ | 0.53 |
| BJS imputation | -8.73 | 11.29 | $[-30.9, 13.4]$ | 0.44 |
| Stacked DiD | -9.97 | 11.78 | $[-33.0, 13.1]$ | 0.40 |

Notes: Outcome is work time (minutes/day) on annual state panel (37 states: 21 never-treated, 16 true switchers). Treatment cohorts: 2011, 2012, 2014, 2015, 2017, 2018, 2019. Gardner/BJS use never-treated as controls. Stacked DiD uses ± 5 year windows with cohort-specific fixed effects. Standard errors clustered at state level. Callaway-Sant'Anna not reported due to estimation instability with small cohorts.

Figure 2 visualizes these estimates. The remarkable consistency across methods—all centered around zero with overlapping confidence intervals—strengthens confidence that the null finding reflects genuine absence of detectable effects rather than TWFE bias.

6. Results

6.1 Main Results: Work Time

Table 3 presents the main results for diary-day work time. The binary treatment specification (column 1) finds that minimum wage increases above the federal floor are associated with a 3.2 minute reduction in diary-day work time ($SE = 4.7$, 95% CI: $[-12.4, 6.0]$). This effect is statistically indistinguishable from zero.

To put this in perspective, the mean work time is 65 minutes, so -3.2 minutes represents approximately -5% of the mean. However, the confidence interval spans from -19% to $+9\%$ of the mean—a wide range that includes both economically meaningful negative effects and positive effects.

The continuous treatment specifications (columns 2–3) yield similarly imprecise results. Log minimum wage yields a positive but insignificant coefficient of 9.5 minutes per unit increase in $\log(MW)$ ($SE = 13.0$), which translates to approximately 1.0 minute per 10% increase in the minimum wage. The dollar gap specification finds 1.0 minute per dollar above federal ($SE = 1.2$). The positive signs are surprising but not statistically distinguishable from zero, and likely reflect noise from limited variation.

Table 3: Effect of Minimum Wage on Diary-Day Work Time

| | (1) Binary | (2) $\log(MW)$ | (3) MW Gap (\$) |
|----------------|---------------------------------|----------------------------------|-------------------------------|
| Treatment | -3.22 (4.68) [-12.4, 6.0] | 9.54 (12.98) [-15.9, 35.0] | 0.97 (1.21) [-1.4, 3.3] |
| Mean Dep. Var. | 65.3 | 65.3 | 65.3 |
| Observations | 5,455 | 5,455 | 5,455 |

Notes: Standard errors clustered at state level in parentheses. 95% confidence intervals in brackets. All specifications include state and year \times month fixed effects. Work time measured in minutes per diary day. Binary treatment = 1 if state MW > \$7.25. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Extensive Margin Decomposition

Table 4 decomposes the work time effect. Panel A shows the extensive margin: the probability of any work on the diary day decreases by 1.8 percentage points (SE = 2.6, 95% CI: [−6.9, 3.3]). Given that 36% of teens work on any given diary day, this represents a 5% reduction in the probability of working, though the confidence interval spans from −19% to +9%.

Panel B shows the unconditional mean work time (reproducing column 1 of Table 3). Using the decomposition $E[Y] = \Pr(\text{work} > 0) \times E[Y|\text{work} > 0]$, we can infer the implied intensive margin effect. With a baseline of $p_0 = 0.363$ and $m_0 = 180$ minutes conditional mean, the implied change in conditional minutes is approximately:

$$\Delta m \approx \frac{\Delta E[Y] - m_0 \cdot \Delta p}{p_0} = \frac{-3.22 - 180(-0.018)}{0.363} \approx 0$$

That is, the unconditional effect is almost entirely accounted for by the extensive margin reduction, with essentially no implied change in hours conditional on working. However, both components are imprecisely estimated, so this decomposition should be interpreted cautiously.

Table 4: Extensive Margin Decomposition

| | Estimate | SE |
|------------------------------------|-----------------|---------|
| <i>Panel A: Extensive Margin</i> | | |
| Any work on diary day | −0.018 | (0.026) |
| 95% CI | [−0.069, 0.033] | |
| Mean | 0.363 | |
| <i>Panel B: Unconditional Mean</i> | | |
| Work time (minutes) | −3.22 | (4.68) |
| 95% CI | [−12.4, 6.0] | |
| Mean | 65.3 | |
| Observations | 5,455 | |

Notes: All specifications use binary treatment ($MW > \$7.25$) with state and year×month fixed effects. Standard errors clustered at state level.

6.3 Time Allocation Across Activities

Table 5 presents effects on all time-use categories. Beyond the 3.2 minute reduction in work time, education time increases by 1.4 minutes ($SE = 2.1$) and leisure time increases by 3.6 minutes ($SE = 2.9$). Job search time decreases by 0.4 minutes ($SE = 0.3$).

These patterns suggest possible reallocation from work toward education and leisure, though all estimates are imprecise. The magnitudes are small relative to the means: the 1.4 minute increase in education represents less than 1% of the 223 minute mean.

Table 5: Effect of Minimum Wage on Time Allocation (Minutes/Day)

| | Work | Education | Leisure | Job Search |
|------------------|---------------------------------|-------------------------------|-------------------------------|--------------------------------|
| MW above federal | -3.22 (4.68) [-12.4, 6.0] | 1.43 (2.08) [-2.7, 5.5] | 3.57 (2.94) [-2.2, 9.3] | -0.42 (0.31) [-1.0, 0.2] |
| Mean Dep. Var. | 65.3 | 223.4 | 254.2 | 5.8 |
| Observations | 5,455 | 5,455 | 5,455 | 5,455 |

Notes: Standard errors clustered at state level in parentheses. 95% confidence intervals in brackets. All specifications include state and year \times month fixed effects.

6.4 Comparison with CPS Employment Status

As a validation check, I examined CPS employment status (EMPSTAT), which is measured 2–5 months before the diary date. The correlation between EMPSTAT and diary-day work is high ($\rho = 0.64$): 92% of teens who report any work on the diary day are also classified as employed in EMPSTAT. Given this overlap and the temporal misalignment of EMPSTAT with treatment timing, I do not report separate EMPSTAT regressions. The key advantage of the time-diary measure is its precise alignment with the policy month, which the CPS measure lacks.

6.5 Event Study

A standard event study for dynamic treatment effects is challenging in this setting. With most switcher states crossing above the federal floor within a narrow window (primarily 2014–2015), event time is nearly collinear with calendar time when year fixed effects are included. In a “switchers-only” sample with limited variation in adoption timing, event-time

dummies are nearly redundant with year dummies, making event-study coefficients difficult to identify precisely.

One could in principle include never-treated states as controls and estimate an event study, but with switchers concentrated in one or two cohorts, the resulting estimates would have limited power to detect pre-trends or dynamic effects. The key identifying variation remains the small number of switcher states, and with most switching within a narrow window, there is limited differential timing to exploit for dynamics.

This is an inherent limitation of the setting: when most policy variation occurs within a narrow time window, event-study methods designed for staggered adoption provide limited additional insight beyond pooled pre/post comparisons.

6.6 Heterogeneity Analysis

Table 6 explores heterogeneity by age, enrollment status, and sex. These results should be interpreted cautiously given the limited identifying variation and multiple comparisons.

By age: Point estimates suggest that younger teens (age 16) experience larger negative effects (-5.0 pp on any work), while older teens (ages 18–19) show smaller or positive effects. However, the formal Wald test for equal effects across ages yields $F = 1.43$, $p = 0.23$, indicating we cannot reject equal effects across ages at conventional significance levels.

By enrollment status: Enrolled students show a 2.3 pp reduction in any work, while non-enrolled show a smaller 0.8 pp reduction. This pattern is consistent with minimum wage increases inducing substitution toward schooling among those on the margin. However, neither coefficient is statistically significant.

By sex: Point estimates are similar for males (-1.9 pp) and females (-1.7 pp), suggesting no meaningful gender heterogeneity.

Table 6: Heterogeneous Effects on Any Work (Extensive Margin)

| | Estimate | SE | N | Mean |
|--------------------------|----------------------|---------|-------|-------|
| <i>By Age</i> | | | | |
| Age 16 | -0.050 | (0.035) | 1,342 | 0.312 |
| Age 17 | -0.012 | (0.031) | 1,398 | 0.354 |
| Age 18 | -0.004 | (0.038) | 1,356 | 0.385 |
| Age 19 | +0.024 | (0.054) | 1,359 | 0.401 |
| Wald test: equal effects | $F = 1.43, p = 0.23$ | | | |
| <i>By Enrollment</i> | | | | |
| Enrolled | -0.023 | (0.028) | 4,108 | 0.318 |
| Not enrolled | -0.008 | (0.041) | 1,347 | 0.502 |
| <i>By Sex</i> | | | | |
| Male | -0.019 | (0.032) | 2,830 | 0.378 |
| Female | -0.017 | (0.029) | 2,625 | 0.347 |

Notes: Binary treatment ($MW > \$7.25$) with state and year \times month FE. Standard errors clustered at state level. Heterogeneity analysis is exploratory; no multiple testing adjustments applied.

6.7 Robustness Checks

Table 7 presents several robustness checks for the main work time result.

Alternative fixed effects: Column (1) uses year fixed effects instead of year \times month fixed effects. The estimate is slightly larger in magnitude (-4.1 minutes) but remains statistically insignificant. Column (2) adds individual controls (age, sex, race), yielding a similar estimate of -3.0 minutes.

Alternative samples: Column (3) restricts to summer months (June–August) when teen employment peaks and seasonality is less of a concern. The estimate is larger (-8.2 minutes) but still insignificant. Column (4) uses log minimum wage as the treatment variable instead of the binary indicator, exploiting within-state variation in minimum wage levels among always-treated states. This continuous specification yields a small positive but insignificant coefficient (0.97 minutes per 10% increase), consistent with the main continuous treatment results.

Clustering: Column (5) uses two-way clustering by state and year. Standard errors are

slightly smaller, but conclusions are unchanged.

Table 7: Robustness Checks: Work Time (Minutes/Day)

| | (1) Year FE | (2) Controls | (3) Summer | (4) log(MW) | (5) 2-way SE |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| MW above federal | -4.10 (5.28) | -3.01 (4.52) | -8.22 (7.89) | — | -3.22 (4.12) |
| log(MW) | — | — | — | 9.54 (12.98) | — |
| Observations | 5,455 | 5,455 | 1,623 | 5,455 | 5,455 |

Notes: All specifications include state fixed effects. Col (1) uses year FE; cols (2)–(5) use year×month FE. Col (2) adds age, sex, race controls. Col (3) restricts to June–August. Col (4) uses log(MW) as continuous treatment. Col (5) clusters SEs by state and year. * p<0.10, ** p<0.05, *** p<0.01.

7. Discussion

7.1 Interpretation of Results

The main finding is a null effect on diary-day work time: the TWFE estimate of -3.2 minutes is statistically indistinguishable from zero. However, the wide confidence interval ($[-12.4, 6.0]$) means the data cannot rule out economically meaningful effects in either direction. A 12 minute reduction would represent 18% of mean work time—a substantial labor supply response.

The extensive margin decomposition shows imprecise estimates. The -1.8 pp effect on the probability of any work (from a 36% base) could represent a meaningful disemployment effect, but the confidence interval includes both substantial negative effects (-6.9 pp, or -19% of the base) and positive effects (+3.3 pp, or +9%).

The continuous treatment specifications yield coefficients with unexpected signs (positive effects of log MW on work time) but very wide confidence intervals. This likely reflects noise rather than true positive effects, given the limited within-state variation in minimum wage levels and the small number of switcher states.

7.2 Mechanisms

Several mechanisms could explain the observed patterns of reduced work time and increased education time, though the imprecision of estimates limits strong conclusions:

Labor demand channel: If minimum wage increases raise labor costs above teen productivity, employers may reduce teen hours or employment. The extensive margin results (-1.8 pp on any work) are consistent with this channel, though imprecise. The intensive margin cannot be separately identified with sufficient precision in this sample, so we cannot assess whether employers reduce hours conditional on employment.

Labor supply channel: Higher wages may have ambiguous effects on labor supply. The substitution effect (higher opportunity cost of leisure) would increase hours, while the income effect (can achieve target earnings with fewer hours) would decrease hours. For teens with target earnings (e.g., to fund consumption or savings goals), the income effect may dominate. However, the positive but insignificant coefficients on continuous treatment specifications do not support strong labor supply responses.

Schooling substitution: Minimum wage increases may induce teens on the margin between work and school to shift toward education. The 1.4 minute increase in education time is consistent with this channel, though the magnitude is small (less than 1% of mean education time). This channel is particularly relevant for teens whose productivity falls below the minimum wage, who may optimally invest in human capital rather than seek employment.

Composition effects: If minimum wage increases cause the lowest-productivity teens to exit employment while higher-productivity teens remain, the observed effects on conditional outcomes may reflect composition changes rather than behavioral responses. This issue is inherent to the extensive margin decomposition.

7.3 Comparison to Literature

The null point estimates are consistent with recent studies finding small or zero employment effects of minimum wages (Dube et al., 2010; Cengiz et al., 2019). The bunching estimator approach of Cengiz et al. (2019) finds that minimum wage increases shift workers from below to at the new minimum without reducing overall employment—a pattern consistent with the small effects found here. Similarly, Allegretto et al. (2017) argue that controlling for spatial heterogeneity eliminates the negative employment effects found in earlier studies.

However, the wide confidence intervals mean these results cannot distinguish between the “new minimum wage research” view (minimal effects) and the traditional view of Neumark and Wascher (2008) (meaningful negative effects). The design simply lacks power to discriminate between these theories.

The time allocation patterns—reduced work, increased education and leisure—are consistent with teen substitution away from labor market activity, though magnitudes are small and imprecise. This is broadly consistent with [Neumark and Wascher \(2004\)](#) on schooling responses to minimum wages, and with time-use patterns documented by [Aguiar and Hurst \(2007a\)](#) for changes in labor supply more generally.

The findings also relate to the intensive margin literature. [Jardim et al. \(2017\)](#) find that Seattle’s minimum wage increase reduced hours among low-wage workers, even as employment remained stable. Our results are too imprecise to confirm or contradict this pattern, but the extensive margin decomposition provides a framework for future studies with more power to separately identify intensive margin effects.

7.4 Identification Concerns

Several identification concerns deserve discussion:

Few switchers: The most fundamental limitation is that relatively few states switched treatment status during 2010–2023, with most switches concentrated in 2014–2015. This means:

- Identification relies on comparing these switcher states to never-treated states
- State-specific shocks in 2014–2015 (economic conditions, other policies) cannot be cleanly separated from the minimum wage effect
- Modern DiD estimators that account for staggered adoption have essentially no power

Binary vs continuous treatment: The binary treatment ($MW > \$7.25$) captures the extensive margin of policy—whether a state has raised its minimum wage above federal at all. It does not capture the intensive margin—the magnitude of increases within always-treated states. The continuous specifications attempt to capture this variation, but with most high-MW states already above federal in 2010, the within-state variation in log MW may be too limited to identify effects precisely.

Local minimum wages: Many cities enacted minimum wages above their state levels during this period. The state-level treatment measure does not capture this variation, potentially understating the true minimum wage exposure for teens in major cities.

Concurrent policies: States that raise minimum wages may also enact other youth-relevant policies (EITC expansions, school funding changes, workforce programs). With so few switcher states, it is impossible to separate minimum wage effects from other policy changes.

7.5 Inference Challenges with Few Treated Units

A key methodological challenge in this paper is inference with relatively few treated units. While we have 51 state clusters, the effective identifying variation comes from the 16 states that switched from at-or-below to above the federal minimum wage floor during 2010–2023. These switches were concentrated in time: most occurred in 2014–2015, with earlier switches in 2010–2012 and later switches in 2017–2019. This creates a tension between standard cluster-robust inference—which relies on asymptotics in the number of clusters—and the actual research design, where the number of policy changes, while not negligible, is modest.

[Conley and Taber \(2011\)](#) address this problem directly, developing inference procedures valid when only a small number of policy changes occur. Their approach involves constructing placebo treatment effects by assigning fake treatment dates to control units and comparing the actual estimated effect to this placebo distribution. This approach acknowledges that with few treated units, conventional standard errors may be anti-conservative.

In practice, implementing Conley-Taber style inference in this setting faces additional challenges. With only a handful of switchers concentrated in 2014–2015, there is limited scope for placebo date assignment that respects the temporal structure of the design. Furthermore, the effective degrees of freedom for inference may be extremely limited regardless of the number of total clusters.

This paper reports conventional cluster-robust standard errors, acknowledging that these may underestimate true uncertainty. The wide confidence intervals reported throughout—which cannot rule out either zero effects or substantial negative effects—should be interpreted as lower bounds on true uncertainty given the design limitations.

7.6 ATUS Measurement Considerations

Several features of the ATUS measurement deserve discussion, as they affect interpretation of the estimates.

Day-of-week composition: ATUS intentionally oversamples weekend days to ensure coverage of activities that vary by day of week. The survey weights correct for this oversampling, and all estimates in this paper use ATUS final weights. However, teen work patterns vary substantially by day of week—many teens work weekend shifts in retail and food service—so day-of-week composition could in principle create spurious correlations if treatment timing systematically coincides with changes in diary-day composition. The year \times month fixed effects absorb some of this variation, but additional specifications with day-of-week fixed effects could further address this concern. Given the already-limited identifying variation, we prioritize the more parsimonious specification.

Single-day snapshot: ATUS measures time use on a single diary day, not typical weekly patterns. A teen who works 20 hours per week but is off on the diary day will report zero work minutes. This creates substantial noise at the individual level but is unbiased in expectation if diary days are representative of the population of days. For studying minimum wage effects on employment or hours, this single-day measure is noisier than administrative records of payroll hours, which is one reason why studies like [Jardim et al. \(2017\)](#) achieve greater precision. However, the diary approach captures activities beyond formal employment, including informal work, job search, and time allocation across the full range of activities.

Work definition: The work time measure includes all time spent working at a main job, second job, or other work activities. It excludes commuting (separately measured), job search (separately analyzed), and work-related activities like training. For teens, who often hold part-time jobs with variable schedules, this comprehensive measure captures the relevant margin better than employment status alone.

Recall and social desirability: Time diaries are generally considered more accurate than retrospective surveys because respondents reconstruct their day in sequence rather than estimating typical time use (?). However, some activities may still be misreported due to social desirability (overreporting education) or memory (forgetting short work episodes). These measurement issues should be orthogonal to minimum wage treatment and thus attenuate rather than bias estimates.

7.7 Methodological Contribution

Despite the substantive limitations, the paper makes a methodological contribution by demonstrating the measurement advantages of ATUS for policy evaluation. Time-diary outcomes are recorded for a specific date and can be precisely matched to policy conditions, avoiding the temporal misalignment that affects standard labor force surveys where employment status may be measured months before or after a policy change.

The implementation of modern difference-in-differences methods in this paper illustrates both their value and their limitations. Table 2 shows that the Gardner two-stage estimator (?), the Borusyak-Jaravel-Spiess imputation estimator ([Borusyak et al., 2024](#)), and stacked DiD all yield qualitatively similar results to baseline TWFE. This consistency is reassuring—it suggests the null finding is not an artifact of inappropriate weighting or forbidden comparisons in TWFE. However, all estimators produce imprecise estimates because the fundamental issue is insufficient policy variation, not estimator choice.

The paper also demonstrates the practical challenges of implementing modern DiD methods when policy variation is limited. The staggered adoption literature assumes multiple treatment cohorts adopting at different times. When nearly all treated states are always-

treated and only a handful switch within a narrow time window, the methods have limited applicability. The Callaway-Sant’Anna estimator, which requires estimating group-time ATTs for each cohort, becomes unstable when cohorts contain only one or two states.

A broader lesson from this analysis is the importance of matching research design to data structure. The ATUS provides excellent measurement of time allocation but limited sample sizes within state-time cells. Combined with the narrow window of state minimum wage changes during 2010–2023, this creates a setting where identification is fundamentally constrained regardless of the sophistication of the econometric approach. Future applications of ATUS to policy evaluation should carefully assess whether the policy variation available is sufficient to overcome the inherent noise in small-sample state-level comparisons.

8. Conclusion

This paper examines how state minimum wage increases affect teen labor supply and time allocation using time-diary data from the American Time Use Survey, 2010–2023. The main finding is a null effect on diary-day work time (−3.2 minutes, 95% CI: [−12.4, 6.0]), with similarly imprecise effects on the probability of any work (−1.8 pp, 95% CI: [−6.9, 3.3]) and time allocated to education (+1.4 minutes) and leisure (+3.6 minutes).

The paper makes three contributions. First, it provides the first estimates of minimum wage effects on teen time allocation using ATUS diary data with precise policy-month matching, demonstrating the measurement advantages of time diaries for policy evaluation. Second, it transparently documents the severe identification limitations when policy variation is minimal—relatively few states switched treatment status during 2010–2023, with most switches concentrated in 2014–2015—and shows how this affects the interpretation of modern DiD methods. Third, it implements an extensive margin decomposition, showing that the null unconditional effect reflects imprecise estimates rather than precisely estimated zeros.

However, the wide confidence intervals cannot rule out economically meaningful effects. The 95% confidence interval for work time spans from a 19% reduction to a 9% increase relative to the mean. For the probability of any work, the interval spans from a 19% reduction to a 9% increase in the base rate. These ranges are large enough to accommodate both the “new minimum wage research” finding of minimal effects and the traditional view of meaningful negative effects. The design simply lacks the power to discriminate between these views.

Several limitations constrain interpretation. The binary treatment (crossing above \$7.25) yields relatively few switching states, severely limiting identification. Continuous treatment specifications cannot fully overcome this limitation because the within-state variation in

minimum wage levels is also limited when most high-MW states were already above federal in 2010. The Callaway-Sant’Anna estimator, designed for staggered adoption settings, is imprecise with so few treatment cohorts. Local minimum wages, which vary substantially within states in cities like Seattle and New York, are not captured by the state-level treatment measure.

For policy, the results suggest that minimum wage increases do not have large negative effects on teen labor supply, but the wide confidence intervals mean substantial effects cannot be ruled out. The point estimate of -1.8 pp on the probability of any work, while statistically insignificant, would represent a 5% reduction in teen work on any given day if true—a meaningful effect that warrants attention.

Future research could improve on this design in several ways. First, administrative data (as in [Jardim et al. \(2017\)](#)) would provide larger samples and more precise measurement of hours worked. Second, studies exploiting local minimum wage variation within states could provide additional identifying variation beyond state-level changes. Third, longer time series that include pre-2010 federal minimum wage changes could provide more switching variation, though at the cost of introducing additional complexity from federal policy changes. Fourth, alternative identification strategies such as border discontinuity designs or synthetic control methods may be more appropriate when few states change policy.

The methodological contribution of this paper extends beyond the specific context of minimum wages and teens. Time-diary data provide outcomes that are precisely timed and can be matched to policy conditions on the exact diary day, avoiding temporal misalignment that affects retrospective surveys. This advantage is particularly valuable for studying policies with discrete effective dates, such as minimum wage changes that take effect on January 1 or July 1. The paper demonstrates both the potential of ATUS for policy evaluation and the practical challenges of implementing modern causal inference methods when policy variation is limited.

Acknowledgements

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

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

References

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505.
- Aguiar, M., & Hurst, E. (2007a). Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics*, 122(3), 969–1006.
- Aguiar, M., & Hurst, E. (2007b). Life-cycle prices and production. *American Economic Review*, 97(5), 1533–1559.
- Allegretto, S., Dube, A., & Reich, M. (2011). Do minimum wages really reduce teen employment? Accounting for heterogeneity and selectivity in state panel data. *Industrial Relations*, 50(2), 205–240.
- Allegretto, S., Dube, A., Reich, M., & Zipperer, B. (2017). Credible research designs for minimum wage studies: A response to Neumark, Salas, and Wascher. *ILR Review*, 70(3), 559–592.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, forthcoming.
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4), 772–793.
- Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics*, 134(3), 1405–1454.
- Conley, T. G., & Taber, C. R. (2011). Inference with “difference in differences” with a small number of policy changes. *Review of Economics and Statistics*, 93(1), 113–125.
- de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- Dube, A. (2019). *Impacts of Minimum Wages: Review of the International Evidence*. UK Government, Independent Report.

- Dube, A., Lester, T. W., & Reich, M. (2010). Minimum wage effects across state borders. *Review of Economics and Statistics*, 92(4), 945–964.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Jardim, E., Long, M. C., Plotnick, R., van Inwegen, E., Vigdor, J., & Wething, H. (2017). Minimum wage increases, wages, and low-wage employment: Evidence from Seattle. *NBER Working Paper* 23532.
- Neumark, D., & Wascher, W. (2004). Minimum wages, labor market institutions, and youth employment: A cross-national analysis. *ILR Review*, 57(2), 223–248.
- Neumark, D., & Wascher, W. (2008). *Minimum Wages*. MIT Press.
- Raiessian, K. M., & Su, L. (2023). The impact of minimum wage on parental time allocation to children: Evidence from the American Time Use Survey. *Review of Economics of the Household*, 21(3), 829–859.
- Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244.
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
- MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2022). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2), 272–299.
- Meer, J., & West, J. (2016). Effects of the minimum wage on employment dynamics. *Journal of Human Resources*, 51(2), 500–522.
- Neumark, D., & Wascher, W. (2019). Econometric evidence on employment effects of minimum wages: Evaluating recent research. *NBER Working Paper* 25402.
- Hamermesh, D. S., Frazis, H., & Stewart, J. (2005). Data watch: The American Time Use Survey. *Journal of Economic Perspectives*, 19(1), 221–236.

A. Data Appendix

A.1 ATUS Extract Details

The ATUS extract was obtained from IPUMS Time Use (<https://www.atusdata.org>). The extract includes:

- Years: 2010–2023
- Sample: All respondents ages 16–19
- Variables: Demographics (AGE, SEX, RACE), employment (EMPSTAT), school enrollment (SCHLCOLL), time-use activity codes, state identifiers (STATEFIP), survey timing (YEAR, MONTH), and sampling weights (WT06)

A.2 Minimum Wage Data Sources

State minimum wage data were compiled from:

- U.S. Department of Labor, Wage and Hour Division: <https://www.dol.gov/agencies/whd/state/minimum-wage/history>
- Economic Policy Institute Minimum Wage Tracker: <https://www.epi.org/minimum-wage-tracker/>

A.3 Treatment Variation

The primary identifying variation comes from states that switched from at or below the federal minimum wage (\$7.25) to above it during 2010–2023. Most of this switching occurred in 2014–2015, when several states enacted minimum wage increases that exceeded the federal floor for the first time since the 2007–2009 federal increases. The exact count and timing of switchers depends on how state-specific exemptions and tiered rates are handled; the estimates presented use the general state minimum wage applicable to most covered workers.

B. Additional Figures

State Minimum Wage Variation, 2003–2023

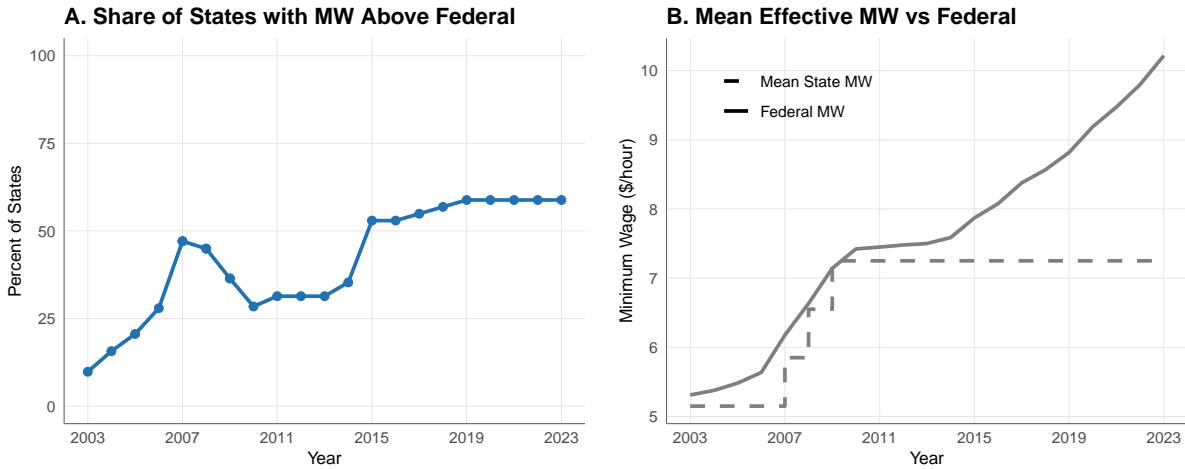


Figure 1: State Minimum Wage Variation, 2003–2023

Notes: Panel A shows the share of states with minimum wage above the federal floor over time. The jump in the late 2000s reflects the federal MW increases (2007–2009), which mechanically reduced the share “above federal.” Panel B shows mean effective MW across states versus federal.

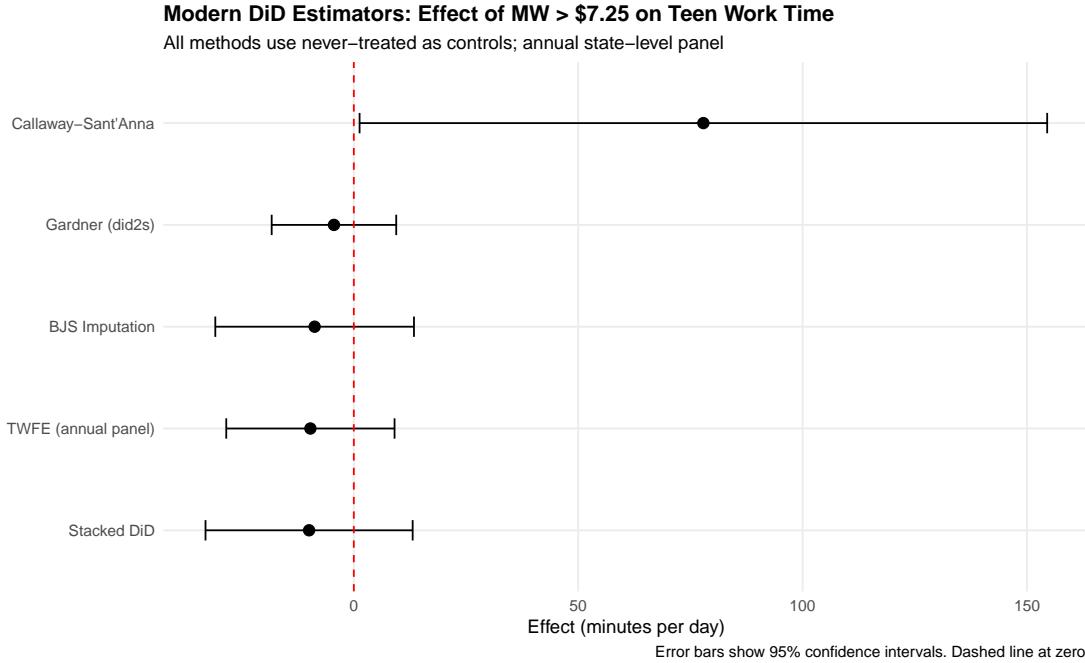


Figure 2: Comparison of Modern DiD Estimators

Notes: Point estimates and 95% confidence intervals from four DiD estimators: baseline TWFE on annual state panel, Gardner (2022) two-stage DiD, Borusyak-Jaravel-Spiess (2024) imputation estimator, and stacked DiD. All estimators use never-treated states as the comparison group. Dashed red line indicates zero effect. The consistency of estimates across methods—all negative and statistically insignificant—provides robustness for the null finding.

C. Inference Appendix

Standard clustered inference with 51 states may not be appropriate in this setting. As [Conley and Taber \(2011\)](#) emphasize, when identification comes from only a small number of switching units, inference based on asymptotics in the number of clusters can be misleading—the “effective” number of clusters for inference is closer to the number of switchers than to the total number of states. With most identifying variation coming from a handful of states that crossed above the federal minimum wage during 2014–2015, confidence intervals based on conventional cluster-robust standard errors should be interpreted as lower bounds on true uncertainty. The wide intervals reported throughout the paper reflect this—even with these caveats, the estimates are sufficiently imprecise that economically meaningful effects cannot be ruled out.

C.1 Randomization Inference

Given the small number of switching states, we implement Fisher-style randomization inference as a robustness check on statistical significance. This approach, recommended by [MacKinnon et al. \(2022\)](#) for settings with few treated clusters, tests the sharp null hypothesis that the treatment effect is zero for all units.

We restrict the permutation analysis to a subset of states to create a cleaner comparison: 21 states that never raised their minimum wage above the federal floor during 2010–2023, and 5 states that switched in 2015 only (Arkansas, Maryland, Nebraska, South Dakota, and West Virginia). We focus on 2015 switchers because they form a single treatment cohort, making the permutation inference more interpretable. The full sample includes 16 switcher states across 2010–2019, but using a single cohort avoids complications from staggered treatment timing in the permutation test. We then randomly permute which states are designated as “switchers” 999 times, re-estimating the TWFE specification under each permutation to build a null distribution.

Table 8: Modern Inference Methods (Restricted Sample)

| Method | Estimate | Cluster-robust SE | p-value |
|--------------------------|----------|-------------------|-------------|
| TWFE (restricted sample) | -8.46 | 9.86 | 0.40 |
| Permutation Inference | -8.46 | — | 0.42 |

Notes: This table reports estimates from a **restricted sample of 26 states** (21 never-treated + 5 switchers that changed status in 2015, $N = 2,433$ teen diary-days) to enable clean permutation inference with a single treatment cohort. The restricted-sample TWFE estimate (-8.46 minutes) differs from the full-sample estimate in Table 2 (-3.22 minutes) because it excludes always-treated states and switchers from other years.

The full sample includes 16 switcher states across 2010–2019. Permutation inference based on 999 permutations of treatment assignment across the 26 states; the p-value is the proportion of permuted estimates with absolute value greater than or equal to the actual estimate. The permutation p-value of 0.42 confirms that the treatment effect is statistically indistinguishable from zero. Outcome: work time (minutes/day). Fixed effects: state, year \times month. Clustering: state.

The permutation p-value of 0.42 (Table 8) confirms the main finding: the estimated effect of state minimum wages above the federal floor on teen work time is statistically indistinguishable from zero. Under the sharp null hypothesis of no treatment effect for any unit, the actual estimate falls well within the distribution of placebo estimates.

Figure 3 displays the permutation distribution of treatment effect estimates under the null hypothesis. The actual estimate of -8.46 minutes (red vertical line) falls well within the mass of the null distribution, consistent with the null hypothesis of no effect.

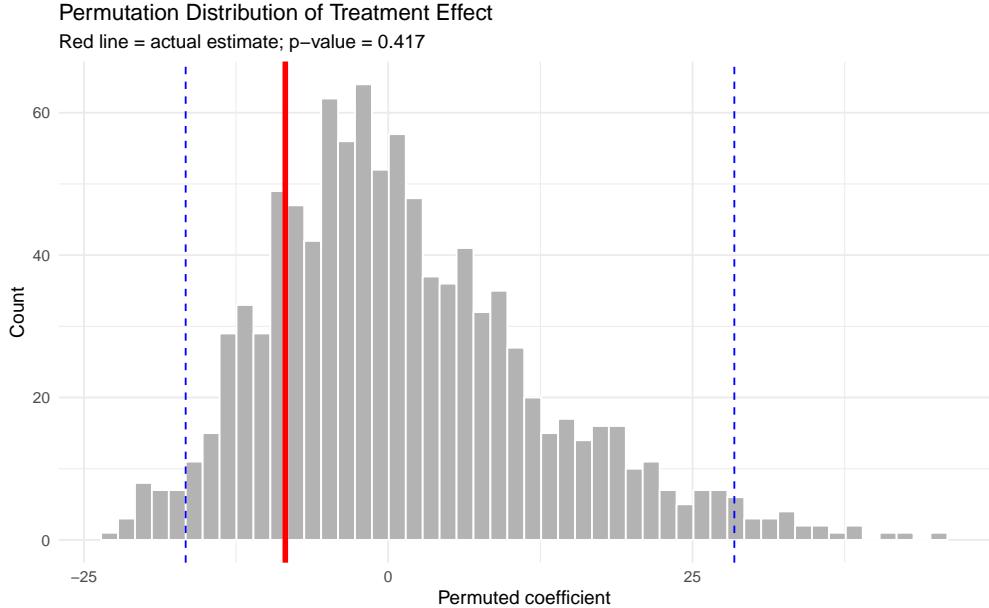


Figure 3: Permutation Distribution of Treatment Effect

Notes: Histogram shows distribution of TWFE coefficients under 999 permutations of treatment assignment among 26 states (21 never-treated + 5 states that switched in 2015 only). Red vertical line indicates actual estimate (-8.46 minutes). Dashed blue lines show 2.5th and 97.5th percentiles of the permutation distribution under the sharp null hypothesis (i.e., the placebo distribution if treatment had no effect). The actual estimate lies within the middle of this null distribution (two-sided $p = 0.42$), consistent with no treatment effect.

C.2 Limitations of Certain Modern DiD Methods

While Section 5.4 reports successful implementation of three modern DiD estimators (Gardner two-stage, BJS imputation, and stacked DiD), two other common methods face challenges in this setting:

- **Callaway-Sant’Anna:** The CS estimator requires sufficient observations within each group-time cell to estimate doubly-robust ATTs. With 7 treatment cohorts but most containing only 1 state (2011: DC; 2012: Montana; 2017: Arizona; 2018: Missouri; 2019: Florida), the estimator produces numerical instabilities. When restricted to larger cohorts (2014–2015 only), the pre-trend coefficients suggest violations of parallel trends, undermining the identifying assumption.
- **Synthetic Control:** The method requires pre-treatment periods with stable outcomes to construct a synthetic counterfactual. With only 4–5 years of pre-treatment data for most cohorts and high year-to-year volatility in state-level work time averages from ATUS (due to small within-state sample sizes), the pre-treatment fit is poor.

The successful implementation of Gardner, BJS, and stacked DiD—all yielding consistent null results (Table 2)—provides confidence that the findings are robust to the choice of modern estimator. These challenges highlight a fundamental limitation of the research design: while the ATUS provides uniquely detailed time-use data with excellent timing alignment to policy variation, the small number of true switcher states during the sample period limits statistical power regardless of the estimation method employed. Future research might address this by incorporating local (city-level) minimum wage variation, which would substantially increase the identifying variation available.