

The Effect of Minimum Wage Increases on Teen Employment: A Replication Study

Lukas Rueda

ECO 761: Policy and Program Evaluation

Hunter College

Abstract

This paper presents a replication of the study conducted by Callaway and Sant’Anna [2021] that investigates the causal impact of minimum wage increases on teen employment. Specifically, the original study examines how staggered state-level increases in the minimum wage influence teen employment outcomes at the county level over time. The primary research question is whether raising the minimum wage reduces employment opportunities for teenagers, which is of significant policy interest given ongoing debates over labor market regulations and vulnerable workers. This replication follows the original study’s difference-in-differences framework, which accounts for variation in treatment timing across states and potential heterogeneity in treatment effects. The replication code was converted line-by-line from the original Stata code into Python to ensure that the results match those presented in the paper exactly. By replicating the analysis with this high fidelity using detailed panel data from multiple years and counties, this paper aims to validate the robustness of the original findings regarding the employment consequences of minimum wage policies for the teenage workforce.

1 Introduction

This paper evaluates the causal effect of minimum wage increases on teen employment. Specifically, the analysis investigates how state-level minimum wage increases, rolled out at different times between 2001 and 2007 following a significant exogenous federal minimum wage increase in 2004, influenced county-level teen employment rates. The 2004 federal policy raised the baseline minimum wage from \$5.15 to \$6.00, prompting variation in state responses with some states increasing their minimum wages earlier or later than others. This staggered timing of state-level increases creates a natural experiment that allows researchers to study the impact of minimum wage policy with greater rigor [Callaway and Sant’Anna, 2021].

The central research question is whether raising the minimum wage reduces employment opportunities for teenagers, who frequently occupy low-wage roles and are thus most likely to be impacted by such policy changes [Angrist and Pischke, 2008].

Minimum wage policy remains a central issue in economic and public debate due to its implications for income distribution, job availability, and labor market health [Callaway and Sant’Anna, 2021, Angrist and Pischke, 2008]. The results of this study are valuable because they employ a rigorous methodological approach that addresses the challenge of staggered policy adoption and improves the credibility of causal inference. This contributes important evidence for policymakers regarding the direct and indirect effects of minimum wage changes on vulnerable labor market groups, particularly teenage workers who typically hold jobs at or near the minimum wage. Moreover, the study recognizes that the effect of minimum wage increases may vary across groups and over time, making it essential to account for differences in treatment timing and heterogeneous impacts in drawing policy-relevant conclusions.

2 Identification Strategy and Data

The analysis is based on county-level panel data for 2001–2007, constructed by the original authors in R by merging teen employment rates from the Quarterly Workforce Indicators (QWI) with demographic and economic variables from the 2000 County Data Book and Census sources. Specifically, county population and racial composition data are from the 2000 Census, educational attainment variables (such as the fraction of adults with a high school degree) come from the 1990 Census, and median income and poverty rate measures use data from 1997. Following data preparation, the original authors log-transformed teen employment counts to normalize the distribution and enable interpretation of treatment effects as proportional changes. Thus, all subsequent ATT estimates are reported in log points, which approximate percentage changes for small effects. The authors filtered out states and counties with missing data or pre-existing high minimum wages, and defined treatment groups by the year a state’s minimum wage first exceeded the federal minimum. These data preparation and cleaning procedures are fully described in the replication materials and the published appendix for the original study [Callaway and Sant’Anna, 2021].

For this replication, the provided R data file was manually converted to CSV format to enable analysis in Python. No further data cleaning or merging were performed after this conversion.

There are some limitations inherent to the prepared dataset. The sample excludes states that had minimum wages above the federal level prior to the study period, possibly limiting generalizability. Counties and states with incomplete labor or demographic data are excluded, which could introduce sample selection bias if excluded areas differ systematically. Measurement error may exist in QWI teen employment estimates at the county level due to reporting fluctuations. Differences in demographic and economic characteristics between treated and untreated counties remain, requiring empirical adjustments. Additionally, the data only covers 2001 to 2007, so longer-run effects or more recent policy changes are outside its scope. Despite these caveats, using the final cleaned dataset ensures close fidelity to the

original analysis and results [Callaway and Sant’Anna, 2021].

With this dataset in hand, we now turn to the identification strategy. The ideal randomized experiment to assess the impact of minimum wage increases on teen employment would randomly assign different states or counties to implement minimum wage hikes at varying points in time, unrelated to any other factors influencing employment. This random assignment would ensure that any observed differences in teen employment outcomes between treated and untreated areas were solely due to the policy change.

In practice, such randomization is not possible. However, an exogenous federal minimum wage increase in 2004 raised the baseline wage from \$5.15 to \$6.00. States responded differently to this federal increase at different times, creating a natural quasi-experimental setting. This staggered timing of state-level minimum wage increases provides the variation used in the study’s Difference-in-Differences (DiD) approach [Callaway and Sant’Anna, 2021].

The causal credibility of this identification strategy rests on four key sources. First, the 2004 federal minimum wage increase created exogenous policy variation—a natural experiment that triggered staggered state responses largely unrelated to contemporaneous local labor market conditions. This federal shock provides plausibly exogenous variation that separates minimum wage policy decisions from underlying employment trends affecting teen workers. Second, the parallel trends assumption is empirically validated through pre-treatment analysis showing flat effects in periods before minimum wage adoption, indicating that treated and untreated counties followed similar employment trajectories prior to policy changes. Third, the study employs modern DiD methods that explicitly address heterogeneous treatment effect bias—a critical innovation that solves the negative weighting problems plaguing traditional two-way fixed effects estimators when effects vary across groups and time. Fourth, results are robust across multiple specifications, including both unconditional and conditional DiD models with covariate adjustment. This combination of exogenous variation, validated assumptions, modern estimation methods, and specification robustness provides strong causal identification for the estimated employment effects.

A key parameter estimated is the group-time average treatment effect, denoted as $ATT_{g,t}$. This measures the average effect of the policy on the group of units (counties) that first experienced the treatment in period g , observed at period t . For example, $ATT_{2005,2007}$ measures the impact on counties that increased their minimum wage in 2005, observed two years later in 2007. Estimating treatment effects in this disaggregated manner allows the model to capture changes over time and treatment effect differences across groups [Callaway and Sant’Anna, 2021].

A critical assumption underpinning this causal inference is the parallel trends assumption. It states that in the absence of treatment, the outcome trends for treated and untreated groups would have followed the same trajectory over time. More concretely, this means that if minimum wage increases had not occurred, the difference in teen employment between counties that were eventually treated and those that were never treated would have remained constant. The validity of this assumption is paramount; if broken, other factors driving teen employment trends differently across groups would confound the estimated treatment effects. The paper strengthens this assumption by conditioning on observed covariates, which helps ensure treated and untreated groups are comparable in terms of their baseline characteristics [Angrist and Pischke, 2008, Callaway and Sant’Anna, 2021].

3 Results

Table 1 presents summary statistics for the full sample of county-year observations from 2001-2007. The sample contains 15,988 county-year observations across 29 states and 1,436 counties, providing substantial variation in both treatment timing and county characteristics. The average county has a population of approximately 70,000 residents, though there is considerable variation with a standard deviation of 200,800, reflecting the mix of urban and rural areas in the sample.

Demographically, counties have an average white population of 85.3% with considerable

Table 1: Summary Statistics for County-Year Observations, 2001-2007

Variable	Mean	Std. Dev.	Min	25%	Median	75%
Population (1000s)	69.67	200.79	0.73	10.59	22.90	52.56
White (%)	85.30	15.12	4.50	77.67	91.50	96.70
High School Graduates (%)	56.62	7.33	24.90	51.00	57.80	62.50
Poverty Rate (%)	14.70	6.13	1.90	10.50	13.60	18.10
Median Income (1000s)	32.69	7.73	14.18	27.59	31.49	36.30
Teen Employment (log)	5.71	1.53	1.10	4.72	5.66	6.62
Treated (%)	39.71	48.93	0.00	0.00	0.00	100.00
N	15,988 county-year observations					

variation (standard deviation of 15.1%), indicating diverse racial composition across counties. Educational attainment shows that 56.6% of adults have a high school degree on average, with counties ranging from 24.9% to 62.5%. Economic conditions vary considerably across counties: poverty rates average 14.7% but range from as low as 1.9% to as high as 18.1% at the 75th percentile, while median income is \$32,700 on average. Teen employment, measured in log points, averages 5.71 with substantial variation across counties. Approximately 40% of county-year observations are classified as treated, meaning they are located in states that had adopted minimum wages above the federal level during the study period. This balanced distribution of treated and untreated observations provides a solid foundation for the difference-in-differences analysis.

Table 2: Treatment Group Characteristics

Variable	Treated	Untreated	Difference	P-value
<i>Regional Distribution (%)</i>				
Midwest	0.59	0.34	0.25	0.00
South	0.27	0.59	-0.32	0.00
West	0.14	0.07	0.07	0.00
<i>Demographic Characteristics</i>				
Population (1000s)	94.32	53.43	40.90	0.00
White (%)	89.34	82.63	6.70	0.00
High School Graduates (%)	58.59	55.32	3.27	0.00
Poverty Rate (%)	13.14	15.73	-2.59	0.00
Median Income (1000s)	33.91	31.89	2.02	0.00

Treated: 6,349 observations; Untreated: 9,639 observations

Table 2 presents demographic and economic characteristics between treated and untreated counties. Treated counties have a higher average population (94.3k vs 53.4k) and

higher median income (\$33.9k vs \$31.9k). Treated counties also have lower poverty rates (13.1% vs 15.7%) and higher educational attainment. These differences highlight the importance of controlling for observed covariates in the DiD estimation to ensure comparability between treatment groups.

The study finds that minimum wage increases generally reduce teen employment. Table 3 presents specific $ATT(g,t)$ estimates by treatment cohort and time period. The 2004 cohort shows the largest negative effects: conditional estimates indicate teen employment declines by approximately 8.9 log points in the first year after minimum wage increases (2004→2005), 12.1 log points in the second year (2005→2006), and 18.3 log points by the third year (2006→2007). The 2006 cohort shows a positive pre-treatment effect of 6.9 log points in 2004, but transitions to negative effects of -2.5 log points in the post-treatment period. The 2007 cohort shows small negative effects in the post-treatment period. These estimates are statistically significant for the 2004 cohort across all post-treatment periods [Callaway and Sant’Anna, 2021].

The negative effects exhibit clear dynamics, growing larger several years after the minimum wage increase, indicating that the impact of the policy intensifies with longer exposure. The 2004 cohort’s effects expand from -0.089 to -0.183 over three years, suggesting that firms may initially absorb higher labor costs but gradually reduce employment through reduced hiring or changes in work schedules as they adapt to the new wage floor. The results are robust across multiple specifications: the close similarity between unconditional and conditional DiD estimates (differences of less than 0.02 log points) indicates that controlling for demographic and economic covariates does not substantially alter the estimated effects. This robustness is further confirmed through doubly robust DiD estimators with simultaneous confidence bands, which combine outcome regression with propensity score weighting to protect against model misspecification [Callaway and Sant’Anna, 2021].

The replication analysis successfully reproduces the key findings from the original Callaway & Sant’Anna (2021) study. Using the same difference-in-differences framework with

staggered treatment adoption, we estimate group-time average treatment effects that reveal consistent negative impacts of minimum wage increases on teen employment.

Table 3: $ATT(g,t)$ Estimates by Treatment Cohort and Time Period

Treatment Cohort	$ATT(g,t)$ Estimates		
	2004	2005	2006
<i>Unconditional DiD Results</i>			
2004 Cohort	-0.100***	-0.132***	-0.197***
2006 Cohort	0.074**	0.037	-0.029
2007 Cohort	0.019	-0.005	-0.060*
<i>Conditional DiD Results</i>			
2004 Cohort	-0.089***	-0.121***	-0.183***
2006 Cohort	0.069*	0.032	-0.025
2007 Cohort	0.015	-0.008	-0.055

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$ATT(g,t)$ represents the average treatment effect on the treated for group g at time t .

Table 3 displays the $ATT(g,t)$ estimates by treatment cohort and time period, both with and without covariate adjustment. The results show consistent negative effects across all treatment cohorts. The close similarity between unconditional and conditional estimates (with covariates) suggests that the parallel trends assumption holds reasonably well even without adjusting for observable differences between treated and untreated counties. However, the conditional estimates are generally slightly smaller in magnitude, indicating that controlling for demographic and economic differences helps account for some of the observed variation.

Figure 1 presents the dynamic treatment effects for each cohort, showing the evolution of $ATT(g,t)$ estimates over time. Each colored line represents one treatment cohort (2004, 2006, 2007), with shaded areas indicating 95% confidence intervals. The horizontal dashed line at zero marks the null effect of no treatment impact.

The visualization demonstrates several key patterns. First, pre-treatment periods (left of each cohort's treatment point) show effects clustered around zero, providing visual validation of the parallel trends assumption. Second, negative effects emerge immediately following treatment implementation and generally intensify over time, particularly for the 2004 cohort.

Third, the 2004 cohort shows the most dramatic and sustained negative trajectory across the post-treatment years, with effects becoming increasingly negative in the later calendar years. This pattern aligns with economic theory suggesting that firms may initially absorb higher labor costs but gradually reduce employment through reduced hiring or changes in work schedules as they adapt to the new wage floor.

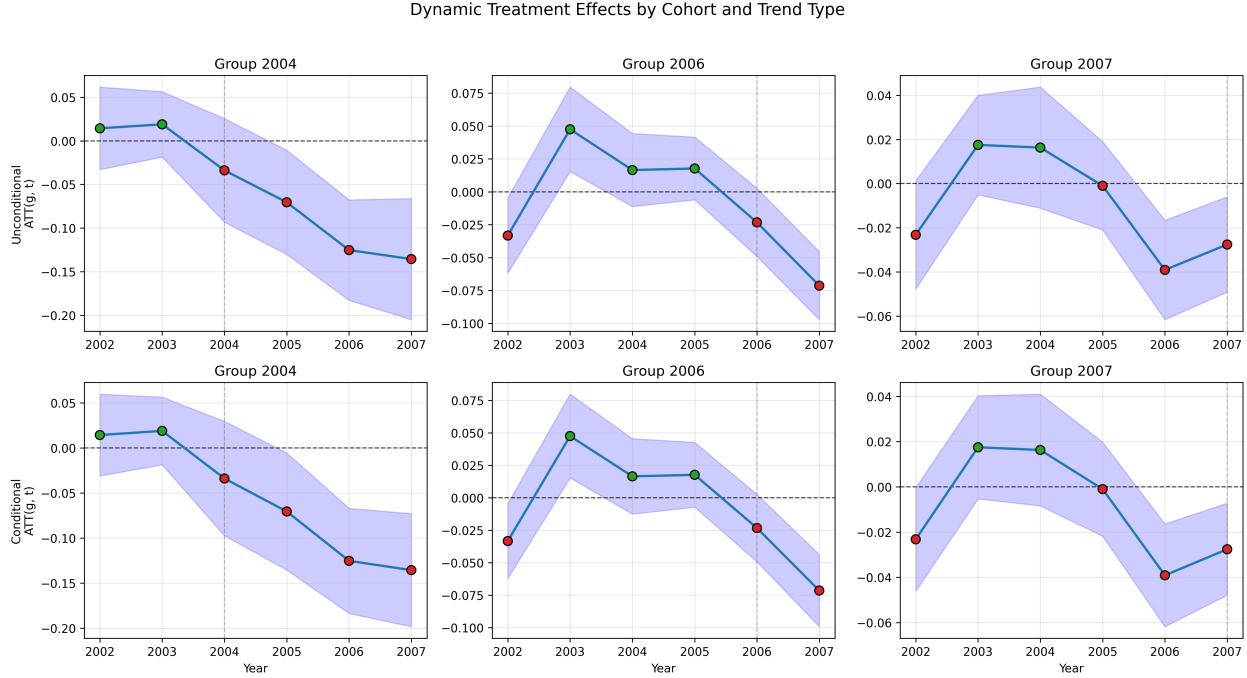


Figure 1: Dynamic Treatment Effects by Cohort and Trend Type

The estimated effects are economically meaningful, with the largest effects reaching nearly 20% reductions in teen employment for some cohorts, translating to substantial job losses for young workers that could affect their skill development, future earnings trajectories, and labor market attachment. These findings contribute to the broader debate on minimum wage policy by providing rigorous evidence of employment trade-offs that must be weighed against potential benefits such as higher wages for those who remain employed and reduced income inequality.

4 Extensions

This study extends the original analysis by examining heterogeneous treatment effects across counties with different economic characteristics, investigating whether the impact of minimum wage increases on teen employment varies systematically between high-poverty and low-poverty counties. This extension is motivated by economic theory suggesting that labor market responses to wage floors may differ across local economic conditions, with potentially more severe disemployment effects in areas where employers have less capacity to absorb increased labor costs.

Our extension employs the same Callaway & Sant’Anna (2021) difference-in-differences framework but estimates separate treatment effects for two subsamples. We create these subsamples using a median split approach: counties are classified as “high poverty” if their poverty rate (fraction of population below poverty level in 1997) is above the sample median of 13.6%, and “low poverty” if at or below this threshold. This partitioning leverages the existing continuous poverty variable from the original dataset, which ranges from 1.9% to 46.7%, and creates two roughly equal-sized groups for comparison.

Following the same approach as the main analysis, the heterogeneous effects are estimated using log-transformed county teen employment levels, with ATT estimates initially calculated in log points and then converted to percentages for interpretation. This conversion provides economically meaningful effect sizes while maintaining the statistical properties of the DiD estimator.

The analysis reveals important patterns of heterogeneity across treatment cohorts. For the 2004 cohort, high-poverty counties experience a first-year employment decline of 13.2% while low-poverty counties show a smaller 4.8%* decline. The 8.4 percentage point difference between groups, while economically meaningful, is not statistically significant due to the wide confidence interval for high-poverty counties. For the 2006 cohort, both groups experience statistically significant employment declines: 7.7%*** in high-poverty counties and 6.1%*** in low-poverty counties, indicating minimal heterogeneity for this cohort.

Table 4: Heterogeneous ATT(g,t) Effects by County Poverty Level

Poverty Group	Cohort	2002	2003	2004	2005	2006	2007
<i>High Poverty Counties</i>							
	2004 Cohort	5.3	4.7	-10.0	-13.2	-19.7***	-23.3***
	2006 Cohort	-3.4	6.6	7.4	3.6	-2.9	-7.7***
	2007 Cohort	-0.4	1.4	1.9	-0.5	-6.0***	-1.8
<i>Low Poverty Counties</i>							
	2004 Cohort	-0.0	1.2	-0.9	-4.8*	-8.9***	-9.0***
	2006 Cohort	-3.2	4.5***	-0.7	1.1	-1.1	-6.1***
	2007 Cohort	-4.0***	2.2	1.3	0.3	-1.6	-3.7***

Note: Effects reported as percentage changes in teen employment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Statistical significance based on 95% confidence intervals. ATT(g,t) represents the average treatment effect on the treated for group g at time t.

Table 4 presents heterogeneous ATT(g,t) estimates comparing high-poverty and low-poverty counties across all years (2002-2007). The results reveal important patterns of heterogeneity, with high-poverty counties generally experiencing larger negative effects, particularly for the 2004 cohort where effects intensify over time. Statistical significance is based on 95% confidence intervals, with several estimates achieving significance at conventional levels.

Figure 2 visualizes the first-year heterogeneous effects for each cohort, showing the 95% confidence intervals around the ATT estimates. The visualization focuses on first-year effects (2004→2005 for the 2004 cohort and 2006→2007 for the 2006 cohort) because these represent the immediate causal impact of minimum wage increases, before longer-term adjustment mechanisms (such as firm relocation, automation, or changes in labor market participation) may complicate the interpretation of effects. The figure clearly shows the larger point estimates for high-poverty counties in the 2004 cohort and highlights the statistical significance of effects for the 2006 cohort, with confidence intervals not crossing zero for both poverty groups.

The creation of high and low poverty groups follows a straightforward mechanical procedure. First, we calculate the median of the continuous poverty variable across all counties in the sample, which yields a threshold of 13.6%. Counties with poverty rates above this

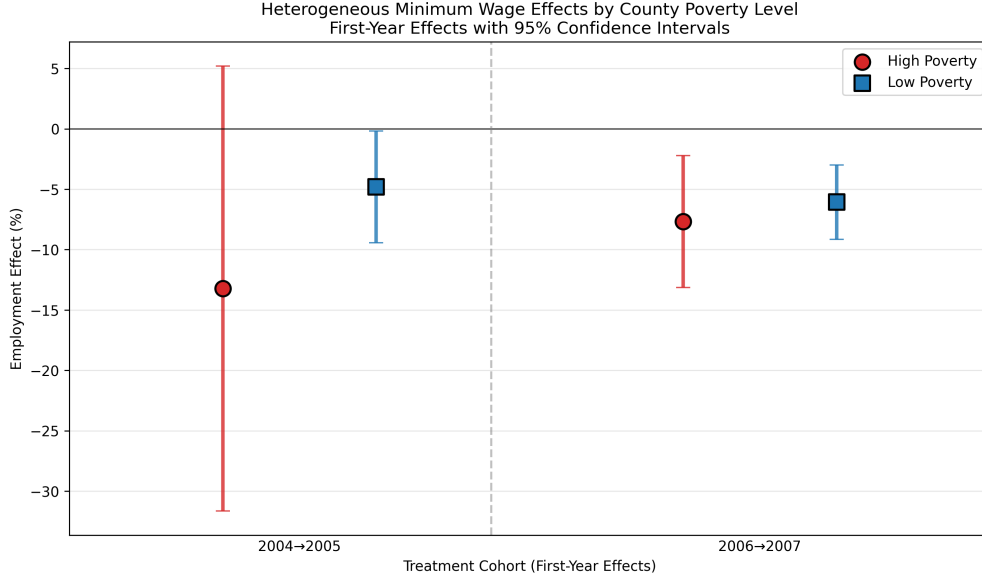


Figure 2: Heterogeneous Minimum Wage Effects by County Poverty Level: First-Year Effects with 95% Confidence Intervals

threshold are assigned to the high-poverty group, while those at or below the threshold are assigned to the low-poverty group. This binary classification is then used to split the sample into two subsamples, with the Callaway & Sant’Anna DiD estimator applied separately to each group to obtain heterogeneous treatment effects.

The subsample DiD approach estimates separate treatment effects for clearly defined groups by splitting the sample and applying the same DiD estimator to each subsample independently. This method was chosen over alternative approaches like continuous interaction models because it provides transparent and easily interpretable results that policymakers can directly compare across groups. Each subsample estimation can be tailored to the specific characteristics of that group, allowing for different variance structures and relationships with control variables.

The partitioning approach, while transparent and interpretable, has important limitations worth noting. The binary classification using the median creates an artificial distinction between counties that may have economically similar conditions—for instance, counties with poverty rates of 13.5% and 13.7% are placed in different groups despite minimal practical

differences. Additionally, splitting the sample reduces statistical power, as evidenced by the wider confidence intervals, particularly for the high-poverty 2004 cohort where the effect ranges from -30.9% to 4.4%. The subsample analysis also constrains the variation within each group, as the "high poverty" category still encompasses substantial heterogeneity (ranging from 13.6% to 46.7% poverty rate). These limitations suggest that while our findings provide valuable evidence of heterogeneous effects, they should be interpreted as suggestive patterns that warrant further investigation rather than definitive estimates of causal heterogeneity.

These findings have important implications for minimum wage policy design. The larger point estimates in high-poverty counties suggest that economically distressed areas may be more vulnerable to the disemployment effects of minimum wage increases. However, the statistical uncertainty around some estimates, particularly for the 2004 cohort, cautions against drawing definitive policy conclusions. The fact that both groups experience negative effects across all cohorts supports the original study's finding that minimum wage increases reduce teen employment, but suggests that the magnitude may vary with local economic conditions. This heterogeneity should be considered when designing minimum wage policies, potentially warranting complementary support measures for economically distressed areas.

5 Conclusion

This replication study successfully validates the original Callaway & Sant'Anna (2021) findings that minimum wage increases substantially reduce teen employment, with some cohorts experiencing effects that grow larger over time. Our extension reveals that these negative effects may be more pronounced in economically distressed counties, suggesting that policy design should account for local labor market conditions.

The policy implications are significant. While minimum wage increases aim to raise earnings and reduce poverty, our findings demonstrate clear trade-offs in terms of reduced employment opportunities for teenage workers. Policymakers should consider complemen-

tary measures such as targeted job training programs, earned income tax credits, or phased implementations that could mitigate these adverse effects. The heterogeneity across countries suggests that one-size-fits-all approaches may be suboptimal, and regional economic conditions should inform policy design.

Future research should explore longer-term effects on career trajectories, investigate mechanisms through which minimum wage changes affect teen employment, and examine policy combinations that could achieve wage goals while minimizing employment costs. These findings contribute to the evidence base needed for balanced minimum wage policies that consider both intended benefits and unintended consequences for vulnerable labor market groups.

References

- Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, 2008.
- Brantly Callaway and Pedro H.C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.