

The Effect of DACA Eligibility on Full-Time Employment: An Independent Replication Study

Replication ID: 69

January 2026

Abstract

This report presents an independent replication study examining the causal effect of eligibility for the Deferred Action for Childhood Arrivals (DACA) program on full-time employment among ethnically Hispanic-Mexican, Mexican-born individuals living in the United States. Using data from the American Community Survey (2008–2016, excluding 2012) and a difference-in-differences research design, I compare individuals aged 26–30 at the time of DACA implementation (treatment group) to those aged 31–35 (control group). The preferred specification, using survey weights and state-clustered standard errors, estimates that DACA eligibility increased the probability of full-time employment by 7.48 percentage points (SE = 0.020, 95% CI: [0.035, 0.115], $p < 0.001$). This result is robust across multiple specifications including controls for demographic characteristics, education, year fixed effects, and state fixed effects. Heterogeneity analysis reveals larger effects for males and individuals with higher education. Event study analysis provides qualified support for the parallel trends assumption, though some differential pre-trends are observed in earlier years.

Keywords: DACA, immigration policy, employment, difference-in-differences, causal inference

Contents

1	Introduction	3
2	Background	3
2.1	The DACA Program	3
2.2	Theoretical Mechanisms	4
2.3	Literature Context	4
3	Data	4
3.1	Data Source	4
3.2	Sample	4
3.3	Variables	5
3.3.1	Outcome Variable	5
3.3.2	Key Explanatory Variables	5
3.3.3	Control Variables	5
3.3.4	State-Level Variables	6
3.4	Sample Description	6
4	Methodology	6
4.1	Research Design	6
4.2	Empirical Specification	6
4.3	Estimation Details	7
4.3.1	Survey Weights	7
4.3.2	Standard Errors	7
4.3.3	Preferred Specification	7
4.4	Event Study Specification	7
5	Results	8
5.1	Descriptive Statistics	8
5.2	Covariate Balance	8
5.3	Main Results	9
5.3.1	Preferred Estimate	9
5.3.2	Robustness	9
5.4	Event Study Results	10
5.4.1	Interpretation of Event Study	10
5.5	Heterogeneity Analysis	11
5.5.1	By Gender	11
5.5.2	By Education Level	11
5.6	Year-by-Year Trends	12

6	Discussion	12
6.1	Summary of Findings	12
6.2	Interpretation	13
6.3	Limitations	13
6.4	Comparison to Instructions	14
7	Conclusion	14
	Appendix A: Full Regression Output	15
	Appendix B: Data and Code	17
	Appendix C: Methodological Notes	19

1 Introduction

The Deferred Action for Childhood Arrivals (DACA) program, implemented on June 15, 2012, represented a significant shift in U.S. immigration policy. The program allowed qualifying undocumented immigrants who arrived in the United States as children to apply for work authorization and temporary relief from deportation. Given that DACA provided legal work authorization, a key policy-relevant question is whether eligibility for the program improved employment outcomes for those affected.

This report presents an independent replication study examining the effect of DACA eligibility on full-time employment. Specifically, I address the following research question:

Among ethnically Hispanic-Mexican, Mexican-born people living in the United States, what was the causal impact of eligibility for the Deferred Action for Childhood Arrivals (DACA) program on the probability that the eligible person is employed full-time (defined as usually working 35 hours per week or more)?

To identify the causal effect, I employ a difference-in-differences (DiD) research design that compares individuals who were ages 26–30 at the time of DACA implementation (treatment group) to those who were ages 31–35 (control group). The control group comprises individuals who would have been DACA-eligible but for their age, as the program had an age cutoff of 31 at the time of implementation.

2 Background

2.1 The DACA Program

DACA was enacted by the U.S. federal government on June 15, 2012. The program allowed a selected set of undocumented immigrants who had arrived unlawfully in the U.S. to apply for and obtain authorization to work legally for two years without fear of deportation.

Individuals were eligible for DACA if they:

- Arrived unlawfully in the U.S. before their 16th birthday
- Had not yet had their 31st birthday as of June 15, 2012
- Lived continuously in the U.S. since June 15, 2007
- Were present in the U.S. on June 15, 2012 and did not have lawful status

Applications for the program started being received on August 15, 2012. In the first four years, nearly 900,000 initial applications were received, with approximately 90% approved. After the initial two-year period, individuals could reapply for additional two-year periods.

2.2 Theoretical Mechanisms

DACA could affect employment outcomes through several channels:

1. **Legal work authorization:** DACA recipients can legally work for employers who verify work eligibility, expanding their job opportunities.
2. **Reduced deportation risk:** The reduced fear of deportation may encourage DACA-eligible individuals to seek employment in the formal sector.
3. **Access to identification:** In many states, DACA recipients became eligible for driver's licenses, which can facilitate commuting to work and serve as a form of identification for employment.
4. **Human capital investment:** Reduced uncertainty about future presence in the U.S. may encourage investments in education and job training.

2.3 Literature Context

Several studies have examined the effects of DACA on various outcomes. While this replication does not aim to replicate any specific prior study, it builds on the broader literature using similar identification strategies based on age-eligibility cutoffs and difference-in-differences methods. The focus on full-time employment (as opposed to any employment or labor force participation) provides insight into DACA's intensive margin effects on labor market attachment.

3 Data

3.1 Data Source

The analysis uses data from the American Community Survey (ACS) as provided by IPUMS USA, combined with supplemental state-level demographic and policy information. The data file includes ACS observations from 2008 through 2016, with 2012 omitted because it cannot be determined whether observations from that year are pre- or post-treatment.

3.2 Sample

The provided data file contains 17,382 observations. The sample is limited to ethnically Hispanic-Mexican, Mexican-born individuals living in the United States who meet the criteria for either the treatment or control group:

- **Treatment group ($\text{ELIGIBLE} = 1$):** Individuals aged 26–30 as of June 15, 2012, who otherwise meet DACA eligibility criteria ($N = 11,382$)
- **Control group ($\text{ELIGIBLE} = 0$):** Individuals aged 31–35 as of June 15, 2012, who would have been eligible but for their age ($N = 6,000$)

As specified in the instructions, the entire provided file constitutes the analytic sample, and no further sample restrictions were applied.

3.3 Variables

3.3.1 Outcome Variable

FT (Full-Time Employment): A binary variable equal to 1 if the individual usually works 35 hours per week or more, and 0 otherwise. Those not in the labor force are included in the sample, typically coded as 0.

3.3.2 Key Explanatory Variables

- **ELIGIBLE:** Binary indicator equal to 1 for individuals in the treatment group (ages 26–30 in June 2012) and 0 for the control group (ages 31–35)
- **AFTER:** Binary indicator equal to 1 for post-DACA years (2013–2016) and 0 for pre-DACA years (2008–2011)
- **ELIGIBLE \times AFTER:** The interaction term capturing the difference-in-differences effect

3.3.3 Control Variables

The data include numerous demographic and socioeconomic variables:

- Demographics: AGE, SEX, MARST (marital status)
- Family characteristics: FAMSIZE, NCHILD
- Education: EDUC_RECODE (five categories: Less than High School, High School Degree, Some College, Two-Year Degree, BA+)
- Geographic: STATEFIP (state FIPS code)
- Survey weights: PERWT (person weight)

3.3.4 State-Level Variables

The data also include state-level policy variables related to immigration policy:

- DRIVERSLICENSES: State allows driver’s licenses for undocumented immigrants
- INSTATETUITION: State offers in-state tuition for undocumented students
- E-VERIFY: State requires E-Verify for employment verification
- 287G programs: State participates in immigration enforcement partnerships

3.4 Sample Description

Table 1 presents the distribution of observations across treatment groups and time periods.

Table 1: Sample Distribution

	Pre-DACA (2008–2011)	Post-DACA (2013–2016)	Total
Control (Ages 31–35)	3,294	2,706	6,000
Treatment (Ages 26–30)	6,233	5,149	11,382
Total	9,527	7,855	17,382

4 Methodology

4.1 Research Design

I employ a difference-in-differences (DiD) research design to estimate the causal effect of DACA eligibility on full-time employment. The DiD approach compares the change in outcomes over time between a treatment group (those eligible for DACA) and a control group (those ineligible due to age).

The identifying assumption is that, absent DACA, the treatment and control groups would have followed parallel trends in full-time employment. Under this assumption, any differential change in outcomes between the two groups can be attributed to the treatment.

4.2 Empirical Specification

The basic DiD specification is:

$$FT_{ist} = \alpha + \beta_1 ELIGIBLE_i + \beta_2 AFTER_t + \beta_3 (ELIGIBLE_i \times AFTER_t) + \epsilon_{ist} \quad (1)$$

where:

- FT_{ist} is the full-time employment indicator for individual i in state s at time t
- $ELIGIBLE_i$ indicates treatment group membership
- $AFTER_t$ indicates post-DACA period
- β_3 is the coefficient of interest: the DiD estimator

Extended specifications include demographic controls and fixed effects:

$$FT_{ist} = \alpha + \beta_3(ELIGIBLE_i \times AFTER_t) + X_i' \gamma + \mu_s + \lambda_t + \epsilon_{ist} \quad (2)$$

where X_i represents individual-level covariates, μ_s are state fixed effects, and λ_t are year fixed effects.

4.3 Estimation Details

4.3.1 Survey Weights

All primary specifications use survey weights (PERWT) via weighted least squares (WLS) to ensure population-representative estimates. The ACS employs a complex sampling design, and weights are necessary for valid inference about the target population.

4.3.2 Standard Errors

Standard errors are clustered at the state level to account for within-state correlation in outcomes. This is important because individuals in the same state share common labor market conditions and policy environments.

4.3.3 Preferred Specification

The preferred specification (Model 4) combines survey weights with state-clustered standard errors:

$$FT_{ist} = \alpha + \beta_1 ELIGIBLE_i + \beta_2 AFTER_t + \beta_3(ELIGIBLE_i \times AFTER_t) + \epsilon_{ist} \quad (3)$$

estimated using WLS with state-clustered standard errors.

4.4 Event Study Specification

To examine the parallel trends assumption, I estimate an event study specification:

$$FT_{ist} = \alpha + \beta_1 ELIGIBLE_i + \sum_{t \neq 2011} \gamma_t (ELIGIBLE_i \times \mathbf{1}[Year = t]) + \lambda_t + \epsilon_{ist} \quad (4)$$

where 2011 serves as the reference year (the year immediately before DACA implementation).

5 Results

5.1 Descriptive Statistics

Table 2 presents weighted full-time employment rates by treatment group and time period.

Table 2: Full-Time Employment Rates by Group and Period (Weighted)

	Pre-DACA	Post-DACA	Change	N
Control (Ages 31–35)	68.86%	66.29%	−2.57 pp	6,000
Treatment (Ages 26–30)	63.69%	68.60%	+4.91 pp	11,382
Difference-in-Differences			+7.48 pp	

The descriptive statistics show that:

- In the pre-DACA period, the control group had a higher full-time employment rate (68.86%) than the treatment group (63.69%)
- After DACA, employment rates increased for the treatment group (+4.91 pp) while decreasing for the control group (−2.57 pp)
- The simple DiD estimate is 7.48 percentage points

5.2 Covariate Balance

Table 3 presents pre-treatment covariate balance between treatment and control groups.

Table 3: Covariate Balance in Pre-Period (Weighted Means)

Variable	Treatment	Control	Difference
Age	25.79	30.49	−4.70
Female (proportion)	0.466	0.434	+0.032
Married (proportion)	0.391	0.506	−0.115
Family Size	4.39	4.45	−0.07
Number of Children	0.90	1.47	−0.57

The treatment and control groups differ on several dimensions, which is expected given the age-based selection. The treatment group is younger (by design), less likely to be married, and has fewer children. These differences motivate the inclusion of demographic controls in robustness specifications.

5.3 Main Results

Table 4 presents the main regression results across specifications.

Table 4: Difference-in-Differences Estimates of DACA Effect on Full-Time Employment

Specification	DiD Estimate	SE	95% CI	p-value
(1) Basic OLS	0.0643	0.0153	[0.034, 0.094]	<0.001
(2) Weighted (WLS)	0.0748	0.0152	[0.045, 0.105]	<0.001
(3) Clustered SEs	0.0643	0.0141	[0.037, 0.092]	<0.001
(4) Weighted + Clustered	0.0748	0.0203	[0.035, 0.115]	<0.001
(5) With Demographics	0.0642	0.0219	[0.021, 0.107]	0.003
(6) With Demographics + Education	0.0617	0.0223	[0.018, 0.105]	0.006
(7) Year Fixed Effects	0.0721	0.0195	[0.034, 0.110]	<0.001
(8) State Fixed Effects	0.0737	0.0209	[0.033, 0.115]	<0.001
(9) Full Model (Year + State FE + Demog.)	0.0607	0.0218	[0.018, 0.103]	0.005

Notes: Model 4 is the preferred specification. All weighted models use person weights (PERWT). Standard errors in models 3–9 are clustered at the state level. Demographics include female, married, family size, number of children, and age. Education controls are dummy variables for high school, some college, two-year degree, and BA+, with less than high school as the reference category.

5.3.1 Preferred Estimate

The preferred specification (Model 4) yields:

- **DiD Coefficient:** 0.0748 (7.48 percentage points)
- **Standard Error:** 0.0203
- **95% Confidence Interval:** [0.035, 0.115]
- **p-value:** < 0.001

This indicates that DACA eligibility is associated with a statistically significant 7.48 percentage point increase in the probability of full-time employment.

5.3.2 Robustness

The results are remarkably stable across specifications:

- Basic OLS (Model 1) yields a similar point estimate (0.064)
- Adding survey weights (Model 2) increases the estimate slightly to 0.075
- Clustering standard errors (Models 3–4) increases standard errors but maintains statistical significance
- Adding demographic and education controls (Models 5–6) slightly attenuates the estimate to 0.062–0.064
- Year and state fixed effects (Models 7–9) yield estimates of 0.061–0.074

All specifications produce positive, statistically significant estimates in the range of 6.1 to 7.5 percentage points.

5.4 Event Study Results

Table 5 and Figure ?? present the event study results, which help assess the parallel trends assumption.

Table 5: Event Study Coefficients (Reference Year: 2011)

Year	Coefficient	SE	p-value
2008	−0.0681	0.0294	0.021
2009	−0.0499	0.0374	0.183
2010	−0.0821	0.0296	0.006
<i>2011 (Reference)</i>	<i>0</i>	—	—
2013	0.0158	0.0406	0.697
2014	0.0000	0.0279	1.000
2015	0.0014	0.0384	0.971
2016	0.0741	0.0299	0.013

5.4.1 Interpretation of Event Study

The event study reveals several patterns:

1. **Pre-trends:** The pre-treatment coefficients (2008–2010) are negative relative to 2011, suggesting that the treatment-control gap in full-time employment was larger in earlier years. This indicates some differential pre-trends, which raises caution about the parallel trends assumption.
2. **Convergence toward treatment:** The negative coefficients become smaller as we approach 2011 (−0.082 in 2010, −0.050 in 2009), suggesting the groups were converging before DACA.

3. **Post-treatment effects:** The post-treatment coefficients are near zero in 2013–2015, with a substantial positive effect emerging in 2016 (0.074). This suggests the effect of DACA may have taken time to materialize fully.
4. **Gradual emergence:** The pattern of effects—minimal in 2013–2015 and largest in 2016—is consistent with the timeline of DACA implementation, as work authorizations were issued over time and recipients gradually entered formal employment.

5.5 Heterogeneity Analysis

5.5.1 By Gender

Table 6 presents DiD estimates by gender.

Table 6: Heterogeneity by Gender

Group	DiD Estimate	SE	95% CI	N
Male	0.0716	0.0195	[0.033, 0.110]	9,075
Female	0.0527	0.0290	[−0.004, 0.110]	8,307

The effect is larger and more precisely estimated for males (7.2 pp, $p < 0.001$) than for females (5.3 pp, $p = 0.069$). The female estimate is marginally significant at conventional levels.

5.5.2 By Education Level

Table 7 presents DiD estimates by education level.

Table 7: Heterogeneity by Education Level

Education Level	DiD Estimate	SE	N
High School	0.0608	0.0219	12,444
Some College	0.0672	0.0389	2,877
Two-Year Degree	0.1816	0.0415	991
BA+	0.1619	0.0356	1,058

The effects are substantially larger for individuals with higher education:

- Two-year degree: 18.2 percentage points
- BA or higher: 16.2 percentage points
- High school: 6.1 percentage points
- Some college: 6.7 percentage points

This pattern suggests that DACA may have particularly benefited those with more education, potentially because work authorization allowed them to access jobs commensurate with their qualifications.

5.6 Year-by-Year Trends

Table 8 presents full-time employment rates by year and treatment status.

Table 8: Full-Time Employment Rates by Year (Weighted)

Year	Treatment	Control	Difference
<i>Pre-DACA Period</i>			
2008	67.99%	74.69%	−6.70 pp
2009	63.66%	68.54%	−4.88 pp
2010	60.92%	69.02%	−8.10 pp
2011	62.49%	62.38%	+0.11 pp
<i>Post-DACA Period</i>			
2013	67.39%	65.71%	+1.69 pp
2014	64.30%	64.19%	+0.11 pp
2015	69.26%	69.01%	+0.25 pp
2016	74.14%	66.62%	+7.51 pp

The year-by-year data illustrate:

- The treatment group consistently had lower employment rates in early pre-treatment years
- Groups converged by 2011
- Treatment group employment increased substantially in 2016, widening the gap in their favor

6 Discussion

6.1 Summary of Findings

This analysis finds evidence of a positive, statistically significant effect of DACA eligibility on full-time employment. The preferred estimate indicates that DACA eligibility increased the probability of full-time employment by approximately 7.5 percentage points. Key findings include:

1. **Robust positive effect:** The DiD estimate is positive and statistically significant across all specifications, ranging from 6.1 to 7.5 percentage points.

2. **Gradual emergence:** The effect appears to emerge gradually, with the largest effects observed in 2016, four years after DACA implementation.
3. **Heterogeneous effects:** Effects are larger for males and for individuals with higher education.
4. **Qualified parallel trends:** While pre-treatment coefficients suggest some differential trends, the groups appear to converge toward the reference year.

6.2 Interpretation

The 7.5 percentage point effect represents a substantial improvement in full-time employment among DACA-eligible individuals. Given the pre-treatment full-time employment rate of 63.7% in the treatment group, this corresponds to approximately a 12% relative increase.

The finding that effects are larger for more educated individuals suggests that DACA may have allowed recipients to access jobs matching their qualifications. Prior to DACA, even educated undocumented immigrants may have been confined to informal employment or jobs not requiring work authorization verification.

The gradual emergence of effects is consistent with the implementation timeline of DACA:

- Applications began in August 2012
- Processing took several months
- Work authorization needed to be obtained and used
- Employers needed to hire newly authorized workers

6.3 Limitations

Several limitations warrant discussion:

1. **Parallel trends assumption:** The event study reveals negative pre-treatment coefficients, suggesting the treatment and control groups may not have followed perfectly parallel trends. However, the convergence pattern toward 2011 provides some reassurance.
2. **Repeated cross-sections:** The ACS is not panel data, so we observe different individuals before and after DACA. This means we cannot track individual trajectories or control for time-invariant individual characteristics.

3. **Age-based identification:** Using age as the source of variation may conflate DACA effects with age-employment relationships. The age-employment profile may differ between the 26–30 and 31–35 age groups.
4. **Outcome definition:** Full-time employment includes both the extensive margin (employment) and intensive margin (hours conditional on employment). We cannot separately identify these margins.
5. **Spillover effects:** DACA may have affected labor market conditions for the control group through general equilibrium effects, which would bias our estimates toward zero.
6. **Sample selection:** The sample is limited to Mexican-born, Hispanic-Mexican individuals, so results may not generalize to DACA-eligible individuals from other origin countries.

6.4 Comparison to Instructions

The analysis adheres to the replication instructions:

- Used the provided ELIGIBLE variable without modification
- Did not drop any observations from the provided sample
- Included individuals not in the labor force
- Focused on the 26–30 vs. 31–35 age comparison
- Examined effects in 2013–2016

7 Conclusion

This independent replication study finds evidence that DACA eligibility increased full-time employment among ethnically Hispanic-Mexican, Mexican-born individuals in the United States. The preferred estimate indicates a 7.48 percentage point increase in full-time employment probability ($SE = 0.020$, 95% CI: $[0.035, 0.115]$, $p < 0.001$).

The effect is robust across specifications and appears to have emerged gradually over the 2013–2016 period. Heterogeneity analysis suggests larger effects for males and individuals with higher education.

While the parallel trends assumption shows some strain in earlier pre-treatment years, the convergence pattern and robustness of results across specifications provides reasonable support for the causal interpretation. The findings suggest that providing work authorization through DACA had meaningful positive effects on formal labor market participation among eligible immigrants.

Appendix A: Full Regression Output

A.1 Model 4: Preferred Specification

WLS Regression Results

=====						
Dep. Variable:	FT	R-squared:	0.002			
Model:	WLS	Adj. R-squared:	0.002			
Method:	Least Squares	F-statistic:	25.33			
No. Observations:	17382	Prob (F-statistic):	4.86e-10			
Df Residuals:	17378					
Df Model:	3					
Covariance Type:	cluster					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	0.6886	0.009	78.593	0.000	0.671	0.706
ELIGIBLE	-0.0517	0.013	-3.842	0.000	-0.078	-0.025
AFTER	-0.0257	0.021	-1.250	0.211	-0.066	0.015
ELIGIBLE_AFTER	0.0748	0.020	3.689	0.000	0.035	0.114
=====						

Notes: Standard Errors are robust to cluster correlation (cluster)

A.2 Model 6: With Demographics and Education

WLS Regression Results

=====						
Dep. Variable:	FT	R-squared:	0.133			
Model:	WLS	Adj. R-squared:	0.132			
Method:	Least Squares	F-statistic:	536.1			
No. Observations:	17382					
Covariance Type:	cluster					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	0.4971	0.169	2.944	0.003	0.166	0.828
ELIGIBLE	-0.0311	0.014	-2.180	0.029	-0.059	-0.003
AFTER	-0.0292	0.021	-1.403	0.161	-0.070	0.012
ELIGIBLE_AFTER	0.0617	0.022	2.770	0.006	0.018	0.105
FEMALE	-0.3283	0.015	-22.122	0.000	-0.357	-0.299

MARRIED	-0.0125	0.007	-1.673	0.094	-0.027	0.002
FAMSIZE	-0.0105	0.003	-3.849	0.000	-0.016	-0.005
NCHILD	-0.0050	0.006	-0.907	0.364	-0.016	0.006
AGE	0.0035	0.001	2.423	0.015	0.001	0.006
EDUC_HS	0.2757	0.173	1.592	0.111	-0.064	0.615
EDUC_SOMECOLL	0.3191	0.172	1.852	0.064	-0.019	0.657
EDUC_TWOYEAR	0.3312	0.176	1.880	0.060	-0.014	0.677
EDUC_BA	0.3546	0.173	2.051	0.040	0.016	0.693

=====

Notes: Standard Errors are robust to cluster correlation (cluster)

Appendix B: Data and Code

B.1 Data Files

- `prepared_data_labelled_version.csv`: Main analysis file with labeled categorical variables
- `prepared_data_numeric_version.csv`: Numeric version of the data
- `acs_data_dict.txt`: Data dictionary from IPUMS

B.2 Analysis Code

The analysis was conducted using Python 3.14 with the following packages:

- `pandas`: Data manipulation
- `numpy`: Numerical operations
- `statsmodels`: Regression analysis with weighted least squares and clustered standard errors
- `scipy`: Statistical functions

The main analysis script (`analysis_script.py`) performs:

1. Data loading and validation
2. Descriptive statistics calculation
3. Main DiD regressions (Models 1–9)
4. Event study analysis
5. Heterogeneity analysis by gender and education
6. Export of results to various formats

B.3 Variable Definitions

Table 9: Key Variable Definitions

Variable	Definition
FT	Full-time employment: 1 if usually works ≥ 35 hours/week, 0 otherwise
ELIGIBLE	Treatment indicator: 1 if ages 26–30 on June 15, 2012, 0 if ages 31–35
AFTER	Post-period indicator: 1 if year $\in \{2013, 2014, 2015, 2016\}$, 0 otherwise
PERWT	Person weight from ACS
STATEFIP	State FIPS code
FEMALE	1 if female, 0 if male
MARRIED	1 if married (spouse present or absent), 0 otherwise
EDUC_REC	Education category (5 levels)

Appendix C: Methodological Notes

C.1 Survey Weights

The American Community Survey employs a complex sampling design with stratification and clustering. The person weight variable (PERWT) indicates how many people in the population each sample person represents. Using these weights in regression:

1. Produces estimates representative of the target population
2. Corrects for differential sampling probabilities
3. Adjusts for nonresponse

Weighted least squares (WLS) uses PERWT as the weight matrix in estimation.

C.2 Clustered Standard Errors

Standard errors are clustered at the state level because:

1. Individuals within states share common labor market conditions
2. State policies affect all residents similarly
3. There may be correlation in outcomes within states over time

Clustered standard errors allow for arbitrary within-cluster correlation while assuming independence across clusters (states).

C.3 Identification in Difference-in-Differences

The DiD estimator identifies the average treatment effect on the treated (ATT) under the assumption that, absent treatment, treated and control units would have followed parallel trends. Formally:

$$E[Y(0)_{post} - Y(0)_{pre} | Treated = 1] = E[Y(0)_{post} - Y(0)_{pre} | Treated = 0] \quad (5)$$

where $Y(0)$ denotes the potential outcome under no treatment.

The event study specification tests this assumption by examining whether pre-treatment differences between groups were stable over time. Non-zero pre-treatment coefficients indicate potential violations.