NBER WORKING PAPER SERIES

RESTAURANT EMPLOYMENT, MINIMUM WAGES, AND BORDER DISCONTINUITIES

Arindrajit Dube
Michael Reich
Akash Bhatt
Denis Sosinskiy

Working Paper 32902
http://www.nber.org/papers/w32902

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2024, Revised May 2026

We are grateful to the Institute for Research on Labor and Employment at UC Berkeley for research support, to Ben Zipperer for his discussion of the JNR paper at ASSA 2024, and to Attila Lindner for his valuable comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Restaurant Employment, Minimum Wages, and Border Discontinuities
Arindrajit Dube, Michael Reich, Akash Bhatt, and Denis Sosinskiy
NBER Working Paper No. 32902
September 2024, Revised May 2026
JEL No. J20, J39, J88

ABSTRACT

Dube, Lester and Reich (2010), using cross-state border county pairs (BCPs), did not detect restaurant employment losses from higher minimum wages. Jha, Neumark and Rodriguez- Lopez (2024, JNR), using multi-state commuting zones (MSCZs), do find disemployment effects. However, JNR's results are confounded by 1990s-era parallel trends violations amplified by the two-way-fixed-effects (TWFE) model, which is also unreliable for detecting such violations. With post-2000 data, JNR's specification does not indicate substantial disemployment. BCPs are much less affected by such pre-trends than MSCZs. Importantly, modern event study designs (using either MSCZs or BCPs) show no local pre-tends, and find no meaningful employment decline.

Arindrajit Dube
University of Massachusetts, Amherst
Department of Economics
and NBER
adube@econs.umass.edu

Michael Reich
University of California, Berkeley
mreich@econ.berkeley.edu

Akash Bhatt
University of Massachusetts, Amherst
akashbhatt@umass.edu

Denis Sosinskiy
University of California, Berkeley
dsosinskiy@berkeley.edu

1 Introduction

Dube et al. (2010) (DLR) contributed to the minimum wage literature by bridging the local-area case study approach of Card and Krueger (1994, 2000) with the two-way fixed effects (TWFE) panel regression proposed by Neumark and Wascher (1992). DLR argued that using nearby areas as controls, which likely experience similar underlying shocks, provides a more credible counterfactual than the standard TWFE using all state-by-state comparisons. Using the Quarterly Census of Employment and Wages (QCEW), DLR showed that a standard TWFE model implied negative employment effects. Yet the same TWFE model restricted to variation within border-county pairs (BCPs) found restaurant employment largely unchanged while earnings rose substantially.

Jha et al. (2024) (JNR) revisit DLR. Substituting multi-state commuting zones (MSCZs) for BCPs but otherwise estimating a similar two-way fixed-effects (TWFE) specification as DLR, they find negative effects on restaurant employment, both within DLR's 1990-2006 period (with QCEW data) and in a longer 1990-2016 period (with County Business Patterns data).

This paper reconciles the contrasting conclusions of DLR and JNR. We make four main points: 1. JNR's negative employment estimates stem from parallel trends violations within MSCZs in the 1990s that the TWFE estimator magnifies throughout the sample. 2. JNR's TWFE-based test of pre-trends is inadequate. 3. BCP specifications do not exhibit these parallel trends violations in the 1990s, thereby providing a more credible counterfactual. 4. Modern difference-in-differences (DiD) event-study estimators avoid the TWFE biases. Applied to both MSCZ and BCP comparisons, they reveal no adverse employment effects, regardless of the inclusion of the 1990s. Updating DLR's border-county approach with modern tools thus corroborates DLR's original TWFE-based findings, because—unlike MSCZs—BCPs do not seem to suffer from significant violations of parallel trends, even over a long time horizon.

To motivate our focus on violations of the parallel trends assumption, Figure 1, adapted from Dube and Lindner (2024), contrasts state-level employment trends for states that raised their minimum wage between 1990 and 2019 ("ever-treated") with those that did not. During the

1990s, per-capita restaurant employment declined in ever-treated states relative to other states, despite negligible minimum wage differences, suggesting sizeable departures from parallel trends.[1] Nonetheless, restaurant employment trends are largely stable in the 2000s, including in the period with a large number of state minimum wage events (mostly in two waves around 2002-2006, and from 2014-2019). This pattern appears inconsistent with large losses in restaurant employment resulting from state minimum wage changes since 2000.

These cross-state divergences underscore that a successful research design needs control groups that experience these shocks similarly to the treated areas. Moreover, state-level employment changes during the 1990s were much more closely correlated with MSCZ-level changes than with changes in border counties, indicating a greater risk of bias for the MSCZ design.[2] In theory, border discontinuity designs should mitigate these state-level concerns. In practice, however, MSCZs and BCPs differ in how effectively they control for pre-trends. A map illustrating the boundaries of MSCZs and border counties (Figure A1) sheds some light on this discrepancy. The map reveals that some MSCZs are nearly as large as entire states, making them potentially more vulnerable to confounding state-level shocks (as shown in Figure 1) compared to border counties.

We first show that TWFE estimates based on within-MSCZ variation are sensitive to including 1990s data, whereas BCP-based estimates are not. Restricting the sample to post-2000—when most minimum-wage variation occurs—yields uniformly small or positive TWFE estimates of restaurant employment, regardless of the dataset or geographic design (MSCZ or BCP).

Does the sensitivity to 1990s data imply that the MSCZ design fails to account for shocks specific to that decade? JNR use a distributed-lag TWFE model to conduct an event study, and argue that they find parallel pre-treatment trends. However, in a TWFE setting, the sign or size of the leads is an unreliable diagnostic: the estimator's weighting scheme can hide substantial departures from parallel trends.[3] With staggered policy adoption, violations at any point in the

[1] For overall employment, Cengiz et al. (2019) show that such differences largely faded by 1993. However, as Figure 1 shows, restaurant employment continued to diverge into the late 1990s.

[2] The correlation between state-level and MSCZ restaurant employment changes (per capita) between 1990 and 2000 was 0.52, compared to 0.35 with border counties.

[3] This point also applies to the original DLR paper, which used distributed lag TWFE before the design's limitations were well understood. This underscores the value of our updating DLR's event study analysis with modern DiD tools.

sample can contaminate all distributed-lag coefficients.[4]

We therefore adopt modern DiD methods with clearly defined control units and event windows to evaluate design validity and produce more credible estimates. When applied to either MSCZ or BCP strategies, these methods indicate negligible employment effects, whether the analysis begins in 1990 or 2000. QCEW-based own-wage elasticity estimates range from -0.38 to $+0.49$, none statistically different from zero at the five-percent level. All confidence intervals exclude JNR's preferred value of -1.5.

Reassuringly, the DiD event studies do not exhibit pre-trends in the three years leading up to treatment, whether looking within BCPs or MSCZs. However, this short window cannot rule out the possibility of longer-run violations—especially in the 1990s—that would bias TWFE estimates using earlier data. To address this issue, we extend the event study to include pre-treatment comparisons back to 1990. In this extended period, we find clear evidence of pre-trends within MSCZs, across a host of specifications, datasets and starting years. By contrast, extended-period pre-trends are far smaller (and not statistically significant) in within-BCP comparisons.

These extended-period pre-trends within MSCZs breach the core parallel trends assumption required by JNR's TWFE model. Section 5 illustrates the consequences of this violation using simulations that combine the actual pattern of minimum-wage changes in MSCZs with placebo outcomes where the true treatment effect is zero. When we introduce parallel trends violations in the 1990s (calibrated to match empirical magnitudes), the MSCZ-based TWFE model produces large, spurious negative estimates. In contrast, the same exercise with BCPs yields estimates correctly centered around zero. In sum, JNR's framework is vulnerable to early-period parallel trends violations within MSCZs. DLR's estimator is less exposed because BCPs account for those shocks more effectively. Crucially, when we re-estimate the effects with modern DiD methods with well-defined event windows and "clean" controls, both the BCP or MSCZ designs perform well. Applying these updated event-study tools to the border-county setting—one of this paper's main contributions—reproduces DLR's original results.

[4] Appendix C demonstrates this point, using a simulation adapted from Dube and Lindner (2024).

2 The fragility of TWFE-based results

JNR estimate a TWFE-style regression in which time effects vary by MSCZ, rather than by BCP as in DLR. The key independent variable in JNR's preferred specification is the log of the minimum wage. We can write the general estimating equation as follows:

$$\ln y_{it} = \beta \ln(MW_{it}) + X_{it}\Lambda + \gamma_i + \delta_{jt} + \nu_{it} \tag{1}$$

Here i is the local area, which may be a state, a county, or a commuting-zone-by-state group, and γ_i is the unit fixed effect. Additionally, δ_{jt} is a j-specific time fixed effect, where j may be (1) a commuting zone pair (TWFE-logMW-CZ) or (2) a contiguous county pair (TWFE-logMW-BCP). The outcome $\ln y$ is either log restaurant employment or log of average restaurant earnings. When estimating the TWFE models, we also follow JNR in using a vector of controls, X_{it}, which includes the log of working-age population and the log of non-restaurant employment or earnings.[5] Finally, β is the minimum wage elasticity.[6]

The first three columns of Table 1 display JNR's core findings, using their County Business Patterns replication data.[7] The last three columns show analogous estimates, but focusing on the 2000-2016 period, when most of the minimum wage variation in the data occurred. We report both the minimum wage elasticity (MWE) and the own-wage elasticity (OWE) of employment.[8]

JNR's preferred specification (TWFE-logMW-CZ, without weights) yields a MWE of -0.242 (s.e. 0.126) for the 1990-2016 sample (column 2), statistically significant at the 10 percent level (and an OWE of -1.517 (s.e. 0.945)). But when we narrow the focus to the 2000-2016 period, the

[5]To avoid challenges arising when including time-varying covariates, our preferred event study estimation in Section 3 uses the log of per-capita employment as the outcome, and does not include any covariates (Dube and Lindner, 2024). However, this specification choice makes little difference, as we show in Appendix Table D4.

[6]Throughout our paper, we present both unweighted estimates and estimates weighted by population. There is no reason to favor either (Solon et al., 2015). The unweighted estimates reflect the expected effect for a randomly chosen unit (a county or commuting zone), whereas the population-weighted estimates capture the expected effect for a randomly chosen individual.

[7]JNR cluster on both state and border segment, following the approach in DLR. It has become clear since DLR that clustering at just the state level is more appropriate, because states are the unit at which treatment occurs and state-clustering accounts for any duplication of areas due to stacking (Wing et al. (2024)). Therefore, we cluster only at the state level. In practice, double-clustering slightly reduces the standard errors in JNR's preferred specification.

[8]The OWE scales the employment effect by the wage effect, providing a more economically meaningful measure of the welfare effects of the policy (see Dube and Lindner (2024)). However, the OWE is only meaningful when there is a clear wage effect (i.e., a "first stage" in an instrumental variables interpretation of the OWE).

MWE estimate drops by more than half, to -0.114 (s.e. 0.123),[9] and it is no longer statistically distinguishable from zero (column 4).[10] In contrast, the TWFE-logMW-BCP estimates (columns 3 and 6, unweighted) tend to be more stable and close to zero in both samples, with MWEs ranging between -0.081 (s.e. 0.065) and 0.044 (s.e. 0.059), and OWEs ranging between -0.492 (s.e. 0.430) and 0.609 (s.e. 0.381).

Columns 1-9 of Table 2 display a complete set of estimates with the same specification as JNR but using QCEW data. For JNR's preferred sample period (1990-2016) and specification (within-CZ, unweighted), the QCEW data also produces a statistically significant and sizable MWE estimate of -0.154, (s.e. 0.077); the OWE estimate is -0.731 (0.450). Our discussion below, which leans on QCEW data, therefore does not reflect an inability to find a substantial negative effect using QCEW data with JNR's preferred specification and sample period.

Figure 2 plots the MWE estimates using the TWFE-logMW-CZ specification, but with rolling start years between 1990 and 2005. The figure also shows estimates using the QCEW, separately for samples that end in 2016 (as in JNR) and samples that extend through 2019. Figure 2 exhibits three key results: 1) Estimates from both the CBP and QCEW are more negative when the sample includes the 1990s, while results from 2000 and later suggest a minimum wage employment elasticity much smaller in magnitude and not statistically distinguishable from zero. 2) Extending JNR's sample through 2019 (which is feasible with QCEW data) leads to a minimum wage employment elasticity estimate that is substantially smaller and statistically indistinguishable from zero.[11] 3) The CBP results are more negative, as are unweighted estimates. These results arise from including earlier years. Estimates starting in 2002 or later are smaller than -0.1 in magnitude and statistically indistinguishable from zero, when we use weights—and whether we use the QCEW or the CBP.[12]

[9] The OWE for the 2000-2016 sample is -1.375, but is highly imprecise because of a weak first-stage wage effect.

[10] With population weights, the estimate declines in magnitude from -0.141 (s.e. 0.077) for the 1990-2016 sample to -0.057 (s.e. 0.047) and not statistically significant for the sample starting in 2000.

[11] For comparability with JNR's annual CBP data, we use annual QCEW data, which also reduces confidentiality-related suppression. Appendix Table A2 shows results using quarterly QCEW data are similar, although the TWFE-logMW-CZ results move closer to zero in the smaller quarterly sample.

[12] QCEW data yield more precise employment estimates: the standard errors in JNR's preferred specification are much larger with CBP data. Moreover, although QCEW restaurant data cover fewer areas due to confidentiality-related suppression, the suppressed areas represent a small share of the population (see Appendix Table B1). Additionally, the CBP data only extend through 2016, limiting its utility for future research and preventing out-of-sample assessment of

3 Difference-in-differences event study

The preceding sensitivity checks reveal that *within*-MSCZ TWFE estimates are sensitive to including 1990s data, whereas *within*-BCP estimates are not. Figure 1 suggests a possible explanation: restaurant employment in ever-treated parts of MSCZs declined relative to controls during the 1990s, well before most minimum wage increases, indicating a violation of parallel trends. To assess this possibility, we first explain why JNR's TWFE-based approach fails as a test of pre-trends, and then present our DiD event study framework.

3.1 TWFE with distributed lags cannot reliably detect pre-trends

JNR present evidence on pre-trends using a distributed lag TWFE-logMW-CZ model, mirroring DLR.[13] They then conduct an event study analysis using binary minimum wage changes, but otherwise keeping to a distributed lag TWFE analysis ("TWFE-binary"), implemented using `xtevent` in Stata. However, a TWFE distributed lags model is ill-suited for evaluating pre-trends, for the same reason that it may produce unreliable estimates of the effects of treatment (either continuous or binary). With staggered treatment adoption, heterogeneity in the true distributed lag coefficients—reflecting *either* treatment effects or pre-trends—can produce highly spurious dynamic estimates (Sun and Abraham, 2021). We present an "existence proof" of such bias using simulations in Appendix C. In a stylized setting with absorbing, binary treatment, we confirm that sizable (but heterogeneous) pre-trend departures can go undetected in distributed lags TWFE models, including with JNR's estimator (`xtevent`), even as they produce biased treatment effect estimates.

3.2 DiD specifications

We therefore adopt a pooled DiD event-study design with explicit control groups and event windows, following Dube and Lindner (2024). This Local Projections-DiD (LP-DiD) model provides the

JNR's findings using later variation.

[13]They find a positive trend (between three years and one year prior to treatment) for BCPs and a negative trend for MSCZs. They flag the former as troubling but dismiss the latter as anticipation. Yet Table A1 shows an even larger MSCZ decline between four and two years before treatment—too early to plausibly reflect anticipation. JNR's interpretation is thus questionable, even leaving aside problems with distributed lag TWFE estimates (see Appendix A.2 for more).

same estimate as the stacked approach of Cengiz et al. (2019), aggregating individual case studies, but in the original panel rather than by stacking (Dube et al., 2025).[14] An *event* is a state minimum wage increase of at least five per cent and \$0.25; multiple phase-ins are combined into a single *combined event.* These rules yield 45 combined events between 1990 and 2019 (reflecting 135 underlying increases), plotted in Appendix Figure D1. Of these, 34 can be evaluated in balanced MSCZ panels and 42 in balanced BCP panels. Average post-treatment windows span 4.2 years; mean minimum wage growth is virtually identical for MSCZs (24.3%) and BCPs (24.6%).

Treatment areas are compared to *clean controls*: areas with no state minimum wage change in the three years before an event and none in the entire post-treatment window. Years with a federal increase are removed from the post-treatment window. Based on these events, we estimate:

$$\overline{\Delta \ln y_{it}} = \beta \times D_{it} + \delta_{jt} + \nu_{it} \tag{2}$$

where $\overline{\Delta \ln y_{it}} = \frac{1}{6}\sum_{k=0}^{5}\left(\ln y_{i,t+k} - \ln y_{i,t-1}\right)$ is the average difference in the log outcome between the post- and pre-treatment periods over the post-treatment window.[15] The indicator $D_{it} = 1$ when unit i (a county or a CZ-state cell) is newly treated at date t. The sample contains only newly treated units and "clean" controls. Time effects δ_{jt} vary by MSCZ or BCP. Critically, β is a convex, non-negative weighted average of event-specific DiD effects.[16] Unlike TWFE, however, this aggregation avoids negative weighting as well as biases due to parallel trends violations that occurred outside the event window (see Appendix C).

To recover dynamic responses, we also use the same sample to estimate separate regressions for each k with long-differenced outcomes—a regression analog to Callaway and Sant'Anna (2021):

$$\ln y_{i,t+k} - \ln y_{i,t-1} = \beta_k \times D_{it} + \delta_{jt} + \nu_{it} \tag{3}$$

where the estimation sample consists of observations that are either newly treated or clean control units. These regressions aggregate individual event-specific estimates. Unlike TWFE, however,

[14]The details of event construction are provided in Appendix D, Section D.1, while our DiD event study methodology is more fully described in Section D.2.

[15]The post-period has a *maximum* length of 6 years, but is truncated for certain events due to federal increases or end-of-sample considerations. For details, see Appendix D.1.

[16]This is a variance-weighted ATT; see Dube et al. (2025). Re-weighting yields an equal-weighted ATT identical to the Callaway and Sant'Anna (2021) estimator.

this aggregation avoids biases due to parallel trends violations that occurred outside the event window.

3.3 Difference-in-differences event study results

Figure 3 and Table 3 present estimates from the event study analyses using QCEW data. Since JNR's preferred specification does not use weights, we show event study figures without weights for comparability. The tables report both weighted and unweighted estimates.

Panels A and B of Figure 3 show the impact of the events on the minimum wage differential between treated and control units; these lead to relative increases in restaurant earnings in treated areas over the post-period (see Panels C and D). Importantly, Panels E and F do not reveal parallel trends violations for estimates of the log employment rate in the three years prior to treatment (we revisit this issue below using longer pre-treatment windows). The effect in the post-period remains flat up to six years after the events, when we look within MSCZs and within BCPs.

Table 3 shows that all unweighted specifications for the full period (Columns 1 and 2, Panel A) exhibit strong earnings effects,[17] and employment effects that are statistically indistinguishable from zero. The associated MWE is -0.001 (s.e. 0.047) for MSCZs and -0.028 (s.e. 0.039) for BCPs. The implied OWEs are small: -0.004 (s.e 0.335) for MSCZs, and -0.235 (s.e. 0.331) for BCPs, similar to OWEs in the minimum wage research literature and restaurant event studies comparing across all states (Dube and Lindner (2024)).

These results hold whether we weight our regressions by population (Columns 1 and 2, Panel B) or focus on events post-2000 (Columns 3 and 4).[18] This consistency contrasts with the sensitivity of the TWFE estimates to the sample period. These results also differ substantially from the TWFE-logMW-MSCZ estimates (unweighted, QCEW 1990-2016 sample), where the MWE is -0.154 and the OWE is -0.731. Appendix D.3 provides a battery of robustness tests for our event study design. Table D3 finds that the results hold across a variety of event definitions.[19] Table D4

[17] To compare with TWFE wage elasticities, the event study earnings estimates have to be divided by the percent change in minimum wages for events (0.24 (0.25) on average for MSCZs (BCPs)).

[18] The lone exception is a positive and marginally significant MSCZ OWE of 0.487.

[19] This includes use of individual (not combined) events and alternative approaches to accounting for federal minimum

shows the estimates are robust to: 1) considering two or four year pre-treatment periods to define clean controls and 2) JNR's strategy of using log employment as the outcome while controlling for log population and log of non-restaurant outcomes as covariates.[20]

To summarize, the event study estimates are less sensitive to the choice of period and the choice of counterfactual, whether within CZs or BCPs. The consistency of the results and the stability of pre-treatment trends suggest that the difference-in-differences design obtains credible and robust estimates, locally within the event window and for the average event, whether within BCPs as in DLR, or within MSCZs as JNR propose. The employment estimates for these events, obtained using a clear difference-in-differences design, contrast sharply with the TWFE estimates.

4 Parallel trends violations reconcile TWFE and DiD estimates

Why are within-MSCZ TWFE estimates sensitive to including 1990s data, while modern DiD event-study estimates remain stable? And why are within-BCP estimates more robust to the choice of period under both methods? This section considers the most plausible explanation: within-MSCZ comparisons are compromised by 1990s parallel trends violations, biasing TWFE estimates that include that decade; in contrast, parallel trends within BCPs hold during the entire period.

4.1 Checking for parallel trends violation in extended pre-treatment periods

The DiD event study specification described above allows us to check for pre-trends, but only within a short window (e.g., three years) before treatment. Since most of the events we study occur after 2000, we investigate whether shocks from the 1990s bias the TWFE estimates. To do so, we extend the pre-treatment window back to 1990 and modify equation 3 to additionally include long differences going back to years $\tilde{t} \in 1990, ..., 1998$:

$$\ln y_{i,\tilde{t}} - \ln y_{i,t-1} = \beta_{\tilde{t}} \times D_{it} + \delta_{jt} + \nu_{it} \tag{4}$$

wages, including estimating the effects of federal increases (in addition to state increases).

[20] We also show that the event study estimates do not result from varying post-treatment window lengths for different event cohorts; Tables D5 and D6 present estimates for each event time from 0 to 5 holding event composition constant, and find similar patterns.

The dependent variable is the outcome difference between (event) year $t-1$ (the year before the treatment) and (calendar) year $\tilde{t}$.[21] Figure 4 presents the event study estimates augmented with these extended pre-period comparisons. Besides reporting β_k coefficients from equation 3 in dark blue (k is event time stretching from -3 to +5), we also report the extended pre-period coefficients $\beta_{\tilde{t}}$ from equation 4 in light blue. These $\beta_{\tilde{t}}$ coefficients measure the outcome gaps between treated and control units in the 1990s, averaged over all events occurring since 2000. Outcomes include log minimum wage, log earnings and log employment rate—for BCP and MSCZ designs and QCEW data.

Panels A and B of Figure 4 plot these event study estimates with log minimum wage as the outcome. Reassuringly, there is little differential trend in minimum wages between the 1990s and the year before treatment: the coefficients are close to zero and not statistically significant. Panels C and D show similar findings for log earnings. We see no significant difference in long-run pre-trends between treated and control units, even when the pre-treatment window extends into the 1990s. Employment trends we might observe are not caused by different minimum wage trajectories.

Panels E and F use log employment rate as the outcome. Panel E shows results within MSCZs. While no trend is visible in the three years before treatment, the extended pre-period reveals clear violations. The 1990 coefficient is $\beta_{1990} = 0.09$, and statistically significant. The 1998 coefficient is about 0.05, also significant. On average, log employment rates in future-treated areas decline relative to controls during the 1990s. In contrast, Panel F shows that when using BCPs, the earlier-year coefficients are roughly one-third as large and not statistically distinguishable from zero. Thus, we reject the strong parallel trends assumption within MSCZs, but not within BCPs.

[21] Since the available pre-period length differs depending on when the treatment occurs, we use calendar time rather than event time to measure earlier trends. As the dependent variable is the difference between event date -1 and particular years in the 1990s outside the baseline 3-year pre-treatment window, we must restrict attention to the 34 post-2000 events (of the total of 45) for which we can plot the extended calendar-time effects that occur *prior to* event date -3. So, usable events are those where the largest $\tilde{t} = 1998$ corresponds to $\leq t-3$. For example, for a 2001 event, the first coefficient measures the differential change between the treatment and control areas between 1990 and 2000.

4.2 Additional evidence of extended pre-trends in a DiD event study design

Appendix Table 4 provides further evidence of parallel trends violations across a range of specifications, time periods and data sources (QCEW and CBP). We find significant evidence of pre-trends for MSCZ pairs (column 1), consistent with the visual evidence in Figure 4. These pre-trends are similar whether we use $(t-1)$ or $(t-3)$ as the terminal period (rows 1 and 2, respectively); this evidence is inconsistent with anticipation effects. Below, we vary specifications, datasets, and samples to assess robustness of the pre-trends found in column 1.

First, columns 6 and 7 show that these pre-trends are robust to using 1990 or 1997 as the starting period instead of 1993 (the starting year in other columns, including column 1). This result suggests that for restaurant employment, parallel trends violation extends beyond 1993.

Column 3 examines log employment (instead of log employment rate) and adds log population and log non-restaurant employment as controls, mirroring the specification in JNR. Column 5 uses the same specification as Column 1, but weights by population.

Columns 2 and 4 use CBP data instead of QCEW data. In Column 2, the CBP estimates are similar to those with the QCEW, but the standard errors are 2 to 2.5 times larger.[22] The pre-trends in column 4, which mirrors column 3 in using JNR's log employment specification with controls, are significant at the 10 percent level. Each of these results suggests the presence of significant pre-trends in the extended pre-treatment window.

Finally, Columns 8-10 show results for pairs of counties instead of MSCZ pairs (note the change in sample sizes). Column 8 consists of all border county pairs (like Figure 4) and does not exhibit any significant pre-trends. Columns 9 and 10 examine the robustness of our estimates to dividing cross-state county pairs into JNR's sub-samples. Column 9, which corresponds to JNR's sub-sample 1, consists of all the contiguous BCPs that belong to different CZs. Column 10, which corresponds to JNR's subsample 3, consists of all pairs formed by counties within the same CZ—whether contiguous or not. JNR find a small employment effect when using subsample 1 and a large negative effect with subsample 3. They claim that the small employment effect in

[22] If we double cluster standard errors, as in JNR, these estimates are significant at the 10 percent level.

subsample 1 is less reliable than estimates based on all county pairs within CZs (as in subsample 3). However, Table 4 shows that subsample 3 also exhibits pre-trends, while subsample 1 does not. In other words, the JNR subsample that yields a large negative effect exhibits larger parallel trend violations.

5 Monte Carlo simulation: 1990s parallel trends violations bias TWFE estimates within CZs but not BCPs

Next, we use Monte Carlo simulations to precisely demonstrate that these early-period violations bias the static and dynamic TWFE estimators when using within-MSCZ variation. Our minimum wage simulation (1) incorporates the actual minimum wages within CZs or BCPs between 1990-2016, (2) generates a placebo outcome under the assumption of no causal effect of minimum wages, and (3) introduces negative employment trends in the 1990s for treated units, calibrated to match the magnitude of the extended-period pre-trends observed in actual data (as documented in Section 4.1), separately for CZs and BCPs. This setup allows us to assess whether the different TWFE estimators can recover the true (null) effect in the presence of realistic 1990s-era violations of the parallel trends assumption.

While the shortcomings of the TWFE framework are now well known (Sun and Abraham, 2021; Dube and Lindner, 2024), it is difficult to pin down analytically how those shortcomings distort estimates from a TWFE-logMW-MSCZ specification. With staggered adoption of a binary, absorbing treatment and perfect parallel trends, the source of bias is transparent: some treated observations carry negative weights. Yet Appendix C shows that even in this simpler setting, once parallel trends are violated, those violations map into bias in individual distributed-lag coefficients in complicated ways (for more on event windows and strong-form parallel trends assumptions, see Roth et al. (2023)). The challenge is far greater for TWFE-logMW models, where treatment is continuous, recurring, and non-absorbing. Here, units cannot be cleanly labeled "treated" or "control," so we cannot identify which observations receive the most problematic weights, nor can we easily trace analytically how early-sample shocks propagate into bias across the lag structure.

This difficulty highlights the value of the simulation presented in this section. Under minimal assumptions—a true minimum-wage effect of zero, 1990s-era parallel-trend violations calibrated to the data, and the actual pattern of minimum-wage changes within MSCZs—the TWFE-logMW-CZ estimator delivers a markedly negative bias.

5.1 Data generating process for the 1990–2016 simulation

We draw on data for commuting-zone pairs (or border county pairs) from 1990 to 2016. We populate the data with the *actual* timing and size of all state and federal minimum wage changes (in logs). We then simulate an outcome variable as follows:

- The causal impact of any minimum wage change is set to zero.
- Parallel trends are violated in the early 1990s (circa 1990–1997) for those states (or counties) that enacted state-level minimum wages after 2000. The magnitude of this early-sample violation is set to be equal to the average event-study estimates (with extended pre-treatment periods) in our real data (see Figure 4 of the main text), separately for the MSCZ and BCP samples.
- We specify an outcome model of the form:

$$\text{Outcome}_{it} = \text{Trend}_{it} + \text{Trend\_error}_{it} + \text{timeFE}_t + \text{unitFE}_i + \varepsilon_{it},$$

  - Trend_{it} is a fixed additive term for each year in 1990–1997 in units that adopt a state minimum wage after 2000 (calibrated to the observed pre-trend magnitudes in Figure 4 for the log per-capita restaurant employment outcome).[23]
  - Trend_error_{it} is a small random draw with positive or negative sign, but applied only to the early-sample period for those same pairs. The standard deviation of the error is calibrated to 0.015 to approximately match the sampling error of the event study pre-treatment leads.

[23] This calibration assigns the same average Trend_{it} terms (trend violations) to all states adopting minimum wages post-2000. Allowing further variation in these trends within these states could help match the empirical estimates even more closely, including for pre-trends, along the lines of Appendix C.

- timeFE_t and unitFE_i are drawn from normal distributions and assigned at the year and pair levels, respectively.
- ε_{it} is drawn from a normal distribution whose variance is calibrated to the residual variance of the actual log per-capita restaurant employment outcome (after partialling out unit and pair-time effects).

Crucially, $\log(\text{MW}_{it})$—which is *actual* policy variation for each pair i in each year t—enters the simulation as a regressor with a *true coefficient of zero*. Thus, any estimated effect for $\log(\text{MW}_{it})$ arises strictly from spurious correlation with the early-sample trend violation interacting with staggered adoption.

5.2 Simulation results

We then estimate the following models in each of the 500 simulation replications:

1. **TWFE-logMW-CZ** and **TWFE-logMW-BCP** specifications
   - A "static" model, regressing the outcome on $\log(\text{MW}_{it})$ with two-way fixed effects.
   - A "distributed lag" version, including 3 leads and 3 lags of $\log(\text{MW}_{it})$.
2. **LP-DiD event study model** based on the same *combined events* with sizable state MW changes used in the main paper.
3. **XTEVENT**, which uses all binary minimum wage changes and estimates a distributed lag TWFE-binary model. Importantly, this does not restrict to "clean controls" and bins up distant leads/lags, just as in JNR.
4. We conduct these simulations separately for **MSCZ-based** pairs—where pre-trend issues in the early 1990s are large and statistically significant—and **BCP-based** pairs, where such violations are smaller.

5.2.1 MSCZ sample TWFE results

Figure 5 presents the results. In the MSCZ sample (left panel), where we know from real data that the early 1990s parallel trends violation was substantial, the static TWFE-logMW-CZ estimate of log(MW) is around -0.13 ($SD = 0.03$) for the 1990-2016 sample. This result strongly resembles actual estimates from the QCEW-based TWFE-logMW-CZ regressions in the paper. In contrast, the estimate for 2000-2016 is 0.00 ($SD = 0.03$), again resembling estimates using real data (not shown in the figure). The distributed lag version reveals positive early leads and increasingly negative effects in the more distant post-treatment lags. These patterns closely match analogs using actual data, as in Appendix Figure A3 (panel A).

5.2.2 BCP sample TWFE results

For the border-county pair sample (right panel), where the measured pre-trends are small (statistically indistinguishable from zero), the static TWFE-logMW-BCP estimate centers near zero (-0.02 with SD = 0.02). (Restriction to post-2000 sample, not shown in the figure, makes little difference.) The distributed lag coefficients also hover around zero, showing little sign of bias, contrary to the MSCZ results. These simulation results also closely align with results using actual data in the paper, and with earlier findings in Dube et al. (2010).

5.2.3 DiD event study estimates

Finally, we replicate our (LP-DiD) event-study estimates in the main paper with this simulated data, using both within-MSCZ and within-BCP variation. Both approaches yield estimates close to zero, which (1) closely resemble estimates using actual data, and (2) are close to the true zero effect under the DGP used in the simulation. The MSCZ event study results are slightly (spuriously) negative owing to the 1990s trend violations, but unlike the TWFE-based XTEVENT estimates, this bias is quite small because the comparisons are more localized around the time of the events, and because most events are in the 2000s.

Furthermore, extending the pre-period back to 1990 in these LP-DiD event studies (similar to

Figure 4 of the paper using actual data) correctly recovers the early-period violations in the 1990 sample, which are large for the MSCZ sample but small for the BCP sample.

Hence, the key takeaway: The 1990s parallel trends violation within CZs causes the TWFE-logMW-CZ to produce spurious negative estimates, but not for the TWFE-logMW-BCP estimate because the violations within BCPs are small. Just imposing the actual (average) parallel trend violations based on our event studies, and assuming a true zero causal effect, can rationalize a sizable, negative minimum wage elasticity from the TWFE-logMW-CZ estimate and QCEW data. And we can replicate the findings that modern DiD estimators, restricted to narrower event windows (around state minimum wage events), avoid this contamination from the early sample and consistently find no disemployment effect.

Simply put, our simulations show that violations of parallel trends in the 1990s can rationalize the divergence in findings using MSCZ- and BCP-based strategies when using the TWFE estimator, and the similarity when using modern DiD estimators.

6 Conclusion

Border discontinuity designs mitigate bias when treatment assignment correlates with state-level potential outcomes. Contiguous counties across state borders often share similar shocks, making them effective counterfactuals, as noted by Dube et al. (2016). In principle, JNR's MSCZ design could play a similar role. In practice, the parallel trends assumption does not hold within MSCZs during the 1990s, which leads to a bias in the within-MSCZ TWFE estimates. The disemployment results from JNR's TWFE-logMW-CZ specification largely disappear once 1990s data are excluded, regardless of the dataset. In contrast, BCP-based estimates remain robust across time periods and datasets. Modern DiD event studies perform well with either geography, but for long samples BCPs better guard against earlier shocks. Most importantly, we demonstrate how shocks during the 1990s confound the TWFE estimates using a within-MSCZ design, but do not bias the within-BCP estimates.

Two broader methodological lessons follow. First, the TWFE framework effectively imposes a “strong-form” parallel trends requirement: violations that occur well outside the window when most minimum wage changes happen can still contaminate estimates through the TWFE weighting scheme, including in distributed-lag implementations commonly used to diagnose dynamics. In that setting, small or statistically insignificant lead coefficients do not constitute persuasive evidence of parallel trends. Second, modern DiD designs with explicit event windows and clean controls provide a clearer link between identifying assumptions and estimates, and they deliver stable results across samples and counterfactual geographies. Applied to both MSCZs and BCPs, these estimators yield negligible employment effects and implied own-wage elasticities that are far smaller in magnitude than JNR’s preferred values.

Substantively, our evidence reconciles DLR and JNR without requiring a change in the underlying economic relationship between minimum wages and restaurant employment. The divergence arises from the interaction of early-period pre-trends within MSCZs and the TWFE estimator, not from a robust negative employment response to minimum wage increases. Our extended pre-period diagnostics show that future-treated areas within MSCZs experienced relative declines in restaurant employment during the 1990s even when minimum wage differences were small, while comparable violations are much weaker within border-county pairs. Our Monte Carlo simulations then show that injecting violations of the magnitude observed in the data is sufficient to reproduce the negative TWFE estimates within MSCZs under a true zero effect, while leaving BCP-based TWFE estimates centered near zero and leaving modern event-study estimates largely unaffected.

Crucially, our update of DLR using within-BCP variation along with modern DiD methods reaffirms the original DLR conclusion: to date, minimum wage increases have not caused substantial losses in restaurant jobs. These findings are also consistent with recent research using matched employer-employee data sources and recent minimum wage variation (Rao and Risch, 2024).

References

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

Card, David and Alan B. Krueger (1994) "Minimum Wages and Employment: A Case Study of the New Jersey and Pennsylvania Fast Food Industries," *American Economic Review*, 84 (4), 772–793.

Card, David and Alan B Krueger (2000) "Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply," *American Economic Review*, 90 (5), 1397–1420.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 134 (3), 1405–1454.

Dube, Arindrajit, Daniele Girardi, Oscar Jorda, and Alan M Taylor (2025) "A Local Projections Approach to Difference-in-Differences," *Journal of Applied Econometrics*.

Dube, Arindrajit, T. William Lester, and Michael Reich (2010) "Minimum wage effects across state borders: estimates using contiguous counties," *The Review of Economics and Statistics*, 92 (4), 945–964.

Dube, Arindrajit, T William Lester, and Michael Reich (2016) "Minimum wage shocks, employment flows, and labor market frictions," *Journal of Labor Economics*, 34 (3), 663–704.

Dube, Arindrajit and Attila Lindner (2024) "Minimum Wages in the 21st Century," in *Handbook of Labor Economics, volume 5-6*.

Jha, Priyaranjan, David Neumark, and Antonio Rodriguez-Lopez (2024) "What's across the border? Re-evaluating the cross-border evidence on minimum wage effects," *Journal of Political Economy Microeconomics*, Forthcoming.

Neumark, David and William Wascher (1992) “Employment effects of minimum and subminimum wages: panel data on state minimum wage laws,” *ILR Review*, 46 (1), 55–81.

Rao, Nirupama and Max Risch (2024) “Who’s Afraid of the Minimum Wage? Measuring the Impacts on Independent Businesses Using Matched US Tax Returns.”

Roth, Jonathan, Pedro HC Sant’Anna, Alyssa Bilinski, and John Poe (2023) “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature,” *Journal of Econometrics*, 235 (2), 2218–2244.

Solon, Gary, Steven J Haider, and Jeffrey M Wooldridge (2015) “What are we weighting for?” *Journal of Human resources*, 50 (2), 301–316.

Sun, Liyang and Sarah Abraham (2021) “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects,” *Journal of Econometrics*, 225 (2), 175–199.

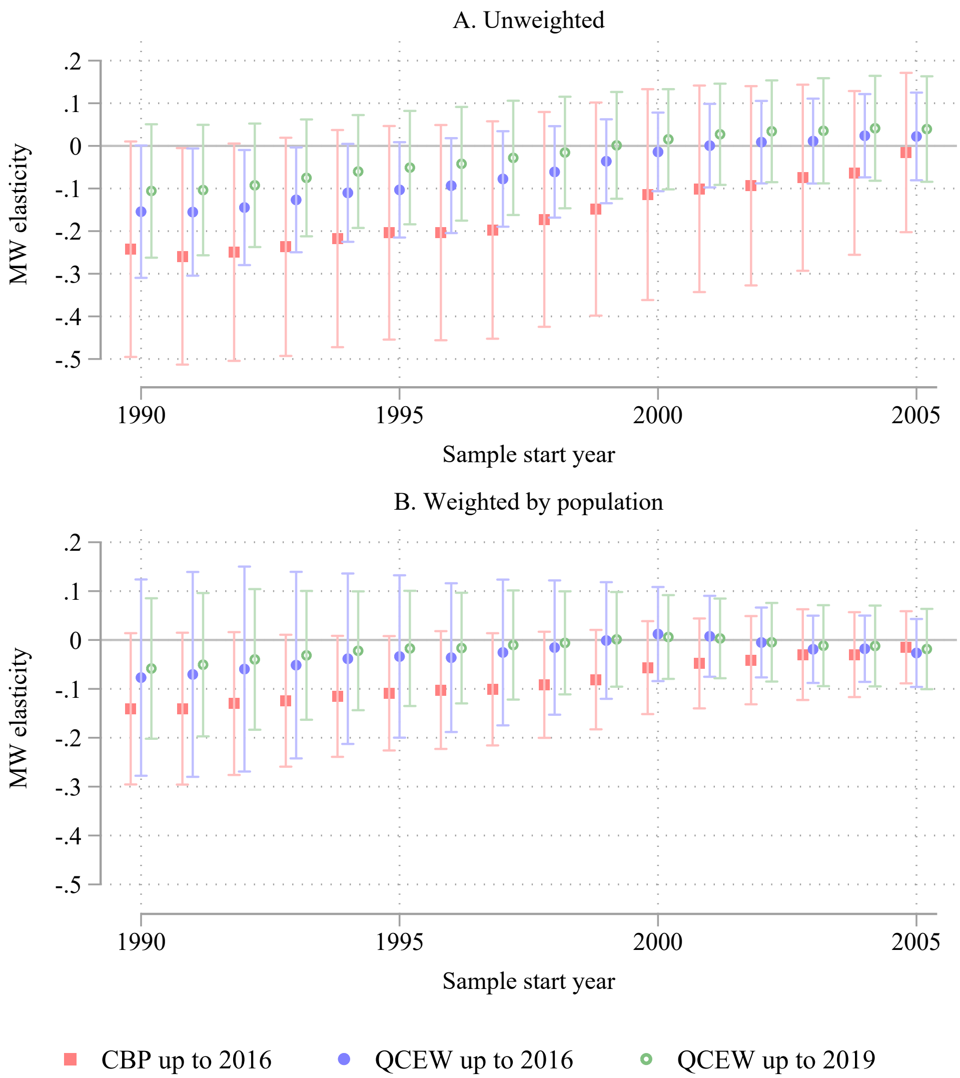
Wing, Coady, Seth M Freedman, and Alex Hollingsworth (2024) “Stacked difference-in-differences,”Technical report, National Bureau of Economic Research.

Tables and Figures

Figure 1: Difference in minimum wages and restaurant employment between "ever-treated" and "never-treated" states

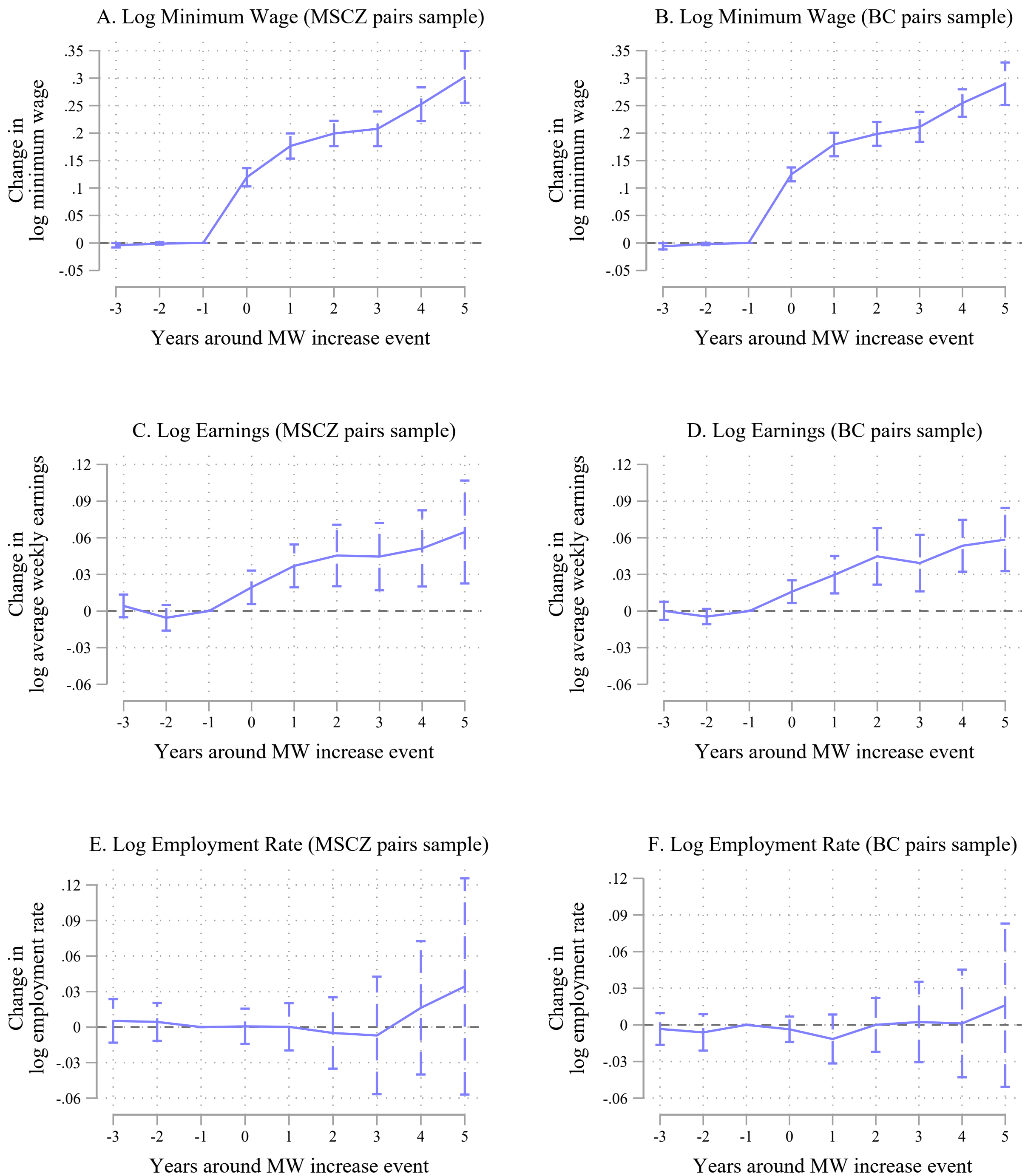
Notes: This figure, adapted from Dube and Lindner (2024), plots the difference in restaurant log per-capita employment (left axis) and the difference in log minimum wages (right axis) between 34 ever-treated and 16 never-treated states. Ever-treated states had at least one state minimum wage increase, over and above any federal increases between 1990 and 2019. Never-treated states did not have any such increases.

Figure 2: Minimum wage employment elasticity, TWFE-logMW-CZ specification: different start and end years using CBP and QCEW



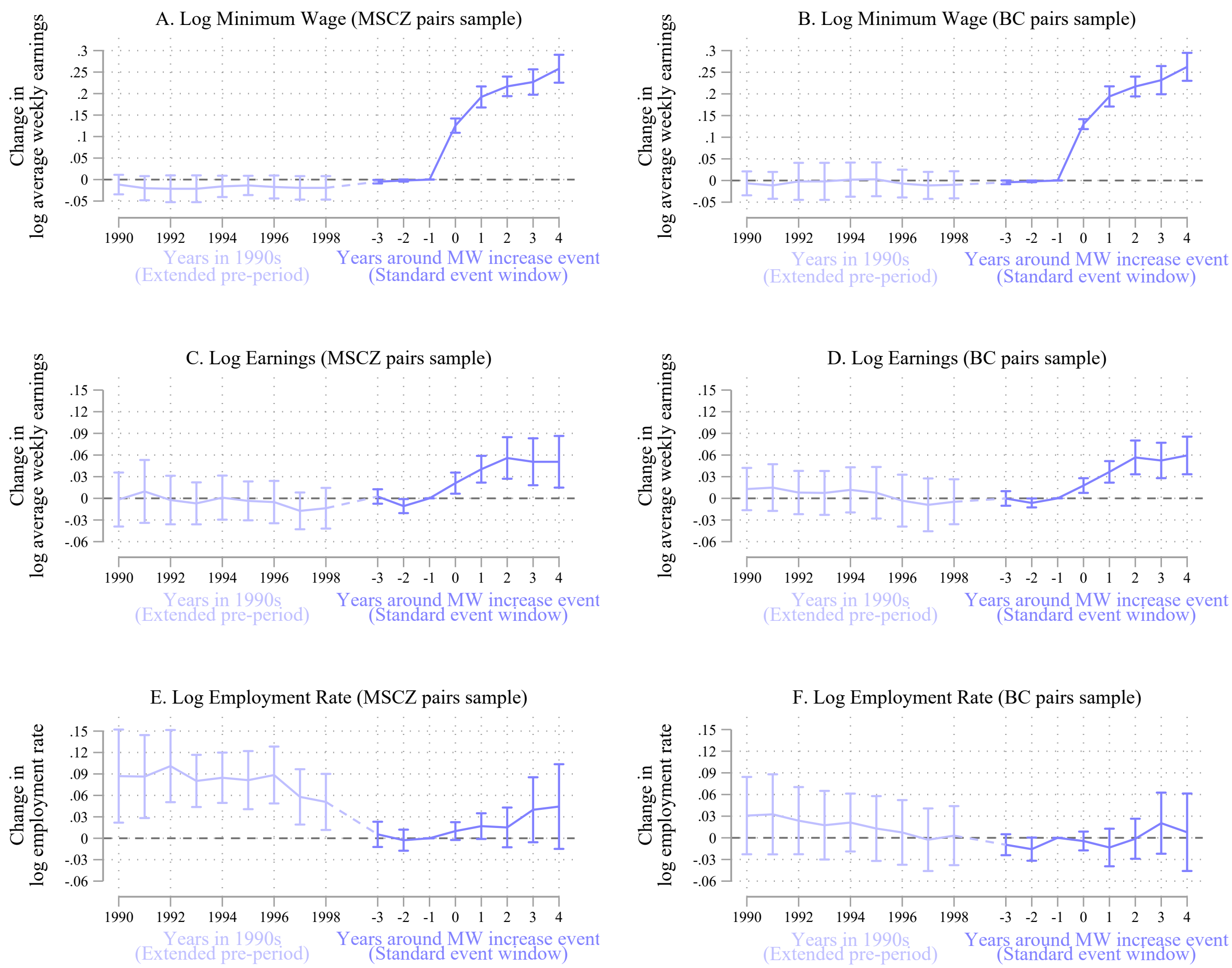
Notes: Estimated using the TWFE-logMW-CZ specification described in equation (1). Each point corresponds to an estimate from a regression using a particular dataset, starting in the year shown on the x-axis and ending in either 2016 or 2019. The horizontal axis reports the first year in the sample. The vertical axis depicts the effect of log minimum wage on log employment. Each color-shape combination refers to a particular dataset used and the final year in the sample, as described by the legend. Lines show 95 percent confidence intervals. All estimates use MSCZ state pairs and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. Estimates in Panel A are unweighted, Panel B uses working-age population weights. All specifications include log employment outside of the restaurant sector and log working-age population as control variables. Standard errors are clustered at the state level.

Figure 3: Event study estimates: 1990-2019



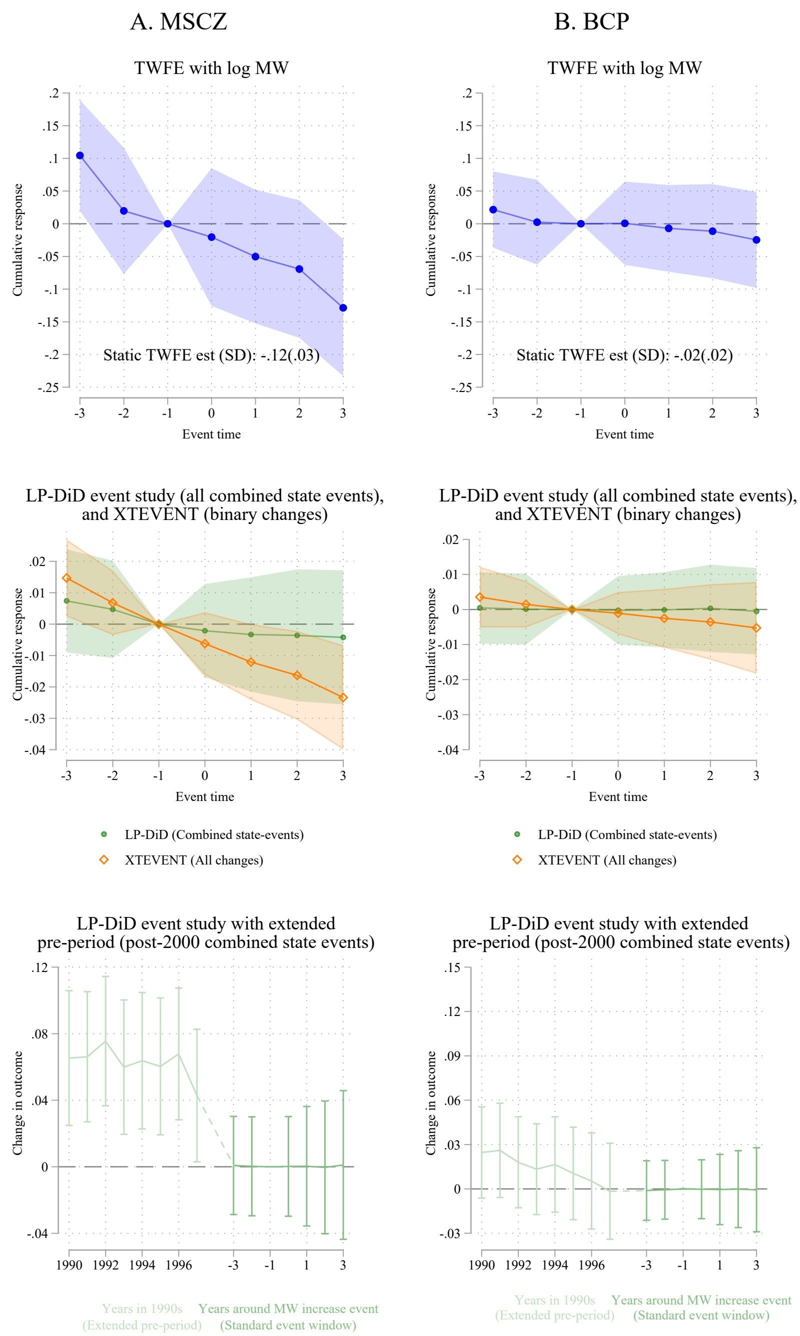
Notes: The plotted points represent coefficients from regressions described in equation (3) using QCEW data for 1990-2019. The left column uses MSCZ state pairs and includes pair-year fixed effects. The right column uses BCPs and also includes pair-year fixed effects. The top row plots effects on log minimum wage. The middle row shows results for log of weekly earnings; while the bottom row shows results for log of employment rate. All estimates are unweighted. Standard errors are clustered at the state level. Lines indicate 95 percent confidence intervals.

Figure 4: Event study estimates and 1990s pre-trends for events post-2000



Notes: The plotted points on the left (lighter blue) are coefficients from regressions following equation 4. The plotted points on the right (darker blue) represent coefficients from regressions described in equation (3) using data from the QCEW for post-2000 events. The left column uses MSCZ state pairs and includes pair-year fixed effects. The right column uses a sample of border county pairs and also includes pair-year fixed effects. The top row plots effects on log minimum wage. The middle row shows results for log weekly earnings; the bottom row shows results for log employment rate: All estimates are unweighted. Standard errors are clustered at the state level. Lines indicate 95 percent confidence intervals.

Figure 5: Monte Carlo simulation results using actual minimum wages, placebo outcomes, and 1990s pre-trends



Notes: This figure shows results from the simulation (with 500 replications) using actual 1990-2016 minimum wages, generated placebo outcomes assuming a zero treatment effect, and 1990s pre-trends calibrated to empirical results from Figure 4. The figures on the left use MSCZ pairs and on the right use BPC pairs. The top row reports estimates using a TWFE-distributed lag model. The second row shows Local Projections DiD event study, and a TWFE-binary event study suing `xtevent`. The LPDiD uses "combined events" as defined in the paper, while `xtevent` uses all minimum wage changes as events as in JNR. In both the first and second rows, the shaded regions represent 2xSD intervals based on the estimated treatment effects. The bottom row, which is similar to Figure 4 in the main text, calculates differences from *t - 1* not only for years in the event-period, but also for years in the 1990s, and only uses events occurring after 2000.

Table 1: MW effects on log average earnings and log employment in restaurants: TWFE estimates for different samples and specifications using CBP data

	1990-2016			2000-2016		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Unweighted						
Log wages	0.215***	0.163***	0.156***	0.100***	0.086	0.109*
	(0.037)	(0.050)	(0.044)	(0.029)	(0.071)	(0.064)
Log employment	-0.338***	-0.242*	-0.081	-0.077	-0.114	0.044
	(0.089)	(0.126)	(0.065)	(0.063)	(0.123)	(0.059)
OWE	-1.531***	-1.517	-0.492	-0.723	-1.375	0.385
	(0.371)	(0.945)	(0.430)	(0.611)	(2.276)	(0.588)
B. Weighted by population						
Log wages	0.192***	0.141***	0.174***	0.120***	0.098*	0.131***
	(0.021)	(0.048)	(0.019)	(0.020)	(0.058)	(0.039)
Log employment	-0.054	-0.141*	0.029	0.061	-0.057	0.079
	(0.072)	(0.077)	(0.090)	(0.056)	(0.047)	(0.060)
OWE	-0.340	-1.013*	0.170	0.518	-0.580	0.609
	(0.319)	(0.536)	(0.473)	(0.507)	(0.461)	(0.381)
N	23,361	8,134	62,228	14,703	5,116	39,146
Fixed effects						
MSCZ pair-year		Y			Y	
BC pair-year			Y			Y
Sample						
All CZ-state	Y			Y		
MSCZ pairs		Y			Y	
Border county pairs			Y			Y

Notes: Estimated using specification described in equation 1 and County Business Patterns data. Columns (1)-(3) use the years 1990-2016, and (4)-(6) use 2000-2016. Columns (1) and (4) use the sample of all commuting-zones-by-state, and include CZ-state and year fixed effects. Columns (2) and (5) use multi-state commuting-zone-by-state pairs sample and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. Columns (3) and (6) similarly use the sample of border county pairs and include fixed effects for border counties (each time a county is in a pair) and pair-year fixed effects. Outcomes include log weekly earnings, log employment, and own wage elasticity (OWE). OWE is the IV estimate from regressing log employment on log earnings instrumented by log minimum wage. † indicates OWE estimates with a weak first-stage. Estimates in Panel A are unweighted, while estimates in Panel B are weighted using the working-age population. All specifications include log of earnings/log of employment outside the restaurant sector, and working-age population as control variables. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: MW effects on log average earnings and log employment in restaurants: TWFE estimates for different samples and specifications using QCEW data

	1990-2019			1990-2016			2000-2016		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Unweighted									
Log wages	0.254***	0.228***	0.242***	0.223***	0.196***	0.216***	0.211***	0.220***	0.225***
	(0.028)	(0.056)	(0.032)	(0.029)	(0.061)	(0.039)	(0.027)	(0.029)	(0.021)
Log employment	-0.265***	-0.106	0.028	-0.308***	-0.154*	0.021	-0.077	-0.014	0.106**
	(0.088)	(0.078)	(0.056)	(0.084)	(0.077)	(0.055)	(0.053)	(0.046)	(0.048)
OWE	-1.050***	-0.436	0.115	-1.365***	-0.731	0.092	-0.363	-0.061	0.474**
	(0.301)	(0.359)	(0.228)	(0.318)	(0.450)	(0.247)	(0.236)	(0.196)	(0.226)
B. Weighted by population									
Log wages	0.214***	0.229***	0.200***	0.196***	0.215***	0.169***	0.144***	0.162***	0.179***
	(0.024)	(0.038)	(0.030)	(0.025)	(0.043)	(0.024)	(0.022)	(0.048)	(0.028)
Log employment	0.025	-0.058	0.077	-0.035	-0.077	0.087	0.066	0.012	0.188**
	(0.078)	(0.071)	(0.076)	(0.068)	(0.100)	(0.080)	(0.060)	(0.048)	(0.071)
OWE	0.078	-0.265	0.348	-0.221	-0.355	0.463	0.462	0.057	1.051**
	(0.365)	(0.291)	(0.344)	(0.309)	(0.441)	(0.426)	(0.451)	(0.253)	(0.467)
N	21,570	5,700	25,800	19,413	5,130	23,220	12,223	3,230	14,620
Fixed effects									
MSCZ pair-year		Y			Y			Y	
BC pair-year			Y			Y			Y
Sample									
All CZ-state	Y			Y			Y		
MSCZ pairs		Y			Y			Y	
Border county pairs			Y			Y			Y

Notes: Estimated using specification described in equation 1 and Quarterly Census of Employment and Wages data. Columns (1)-(3) use the years 1990-2019, (4)-(6) use 1990-2016, (7)-(9) use 2000-2016. Columns (1), (4), and (7) use the sample of all commuting-zones-by-state, and include CZ-state and year fixed effects. Columns (2), (5), and (8) use multi-state commuting-zone-by-state pairs sample and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. Columns (3), (6), and (9) similarly use the sample of border county pairs and include fixed effects for border counties (each time a county is in a pair) and pair-year fixed effects. Outcomes include log weekly earnings, log employment, and own wage elasticity (OWE). OWE is the IV estimate from regressing log employment on log earnings instrumented by log minimum wage. OWE estimates are only reported when there is a strong first-stage effect. Estimates in Panel A are unweighted, while estimates in Panel B are weighted using the working age population. All specifications include log earnings/log employment outside of the restaurant sector, and working-age population as control variables. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: sym* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Effects of increased MW - event study estimates with QCEW data

	1990-2019		2000-2019	
	(1)	(2)	(3)	(4)
A. Unweighted				
Log wages	0.031***	0.027***	0.033***	0.032***
	(0.008)	(0.006)	(0.008)	(0.006)
Log employment rate	-0.000	-0.006	0.016	-0.005
	(0.010)	(0.009)	(0.009)	(0.011)
MWE	-0.001	-0.028	0.067	-0.019
	(0.047)	(0.039)	(0.040)	(0.044)
OWE	-0.004	-0.235	0.487*	-0.154
	(0.335)	(0.331)	(0.258)	(0.347)
B. Weighted by population				
Log wages	0.020*	0.019*	0.024**	0.026***
	(0.011)	(0.010)	(0.010)	(0.010)
Log employment rate	-0.007	0.000	-0.002	0.003
	(0.006)	(0.005)	(0.006)	(0.006)
MWE	-0.030	0.001	-0.006	0.009
	(0.024)	(0.018)	(0.024)	(0.023)
OWE	-0.381	0.009	-0.063	0.097
	(0.298)	(0.243)	(0.264)	(0.240)
N	2,132	10,414	1,220	6,170
Events evaluated	34	42	26	32
Fixed effects				
MSCZ pair-year		Y		Y
BC pair-year				
Sample				
MSCZ pairs		Y		Y
Border county pairs				

Notes: Estimated using specification described in equation 2 using Quarterly Census of Employment and Wages data. Columns (1) and (2) use the years 1990-2019, and (3) and (4) use 2000-2019. Columns (1) and (3) use multi-state commuting-zone-by-state pairs sample and include pair-year fixed effects. Columns (2) and (4) similarly use the sample of border county pairs and also include pair-year fixed effects. Outcomes include log weekly earnings, log employment rate, minimum wage elasticity (MWE) and own wage elasticity (OWE). MWE is the IV estimate from regressing log employment rate on the log minimum wage instrumented by the event dummy. OWE is the IV estimate from regressing log employment rate on log earnings instrumented by the event dummy. Estimates in Panel A are unweighted, while estimates in Panel B are weighted using the working-age population. The estimation sample for all years includes all units that are either newly treated or are clean controls for that year. However, the number of events evaluated is calculated only from pair-years that exhibit treatment variation - i.e one side of the pair has an event start, while the other side is a clean control as defined in the text. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Pre-trends check for post-2000 events

	MSCZ pairs							County pairs		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) All BCP	(9) Non-MSCZ BCP (JNR sub-sample 1)	(10) MSCZ county pairs (JNR sub-sample 3)
Trend until t[-1]	0.080***	0.072	0.074***	0.069*	0.058**	0.087**	0.058***	0.017	-0.027	0.059**
	(0.018)	(0.045)	(0.017)	(0.041)	(0.022)	(0.032)	(0.019)	(0.024)	(0.026)	(0.022)
Trend until t[-3]	0.075***	0.075	0.072***	0.072*	0.075***	0.082**	0.052***	0.027	-0.012	0.065***
	(0.017)	(0.044)	(0.016)	(0.041)	(0.017)	(0.032)	(0.018)	(0.020)	(0.022)	(0.018)
N	1,220	1,956	1,220	1,956	1,220	1,220	1,220	6,170	4,640	5,098
CBP dataset		Y		Y						
Outcome is log emp			Y	Y						
Pop and non-rest emp controls			Y	Y						
Weighted by population					Y					
Start year 1990						Y				
Start year 1997							Y			
Events evaluated	26	22	26	22	26	26	26	32	32	26

This table reports coefficients from the specification described in Equation 4, with the outcome being log of employment rate in all columns apart from Columns 3 and 4. The top row is the exact specification as the equation, while the bottom row takes the trend until t[-3] in event-time instead of t[-1]. Unless indicated otherwise (only in Columns 6 and 7), the long difference is taken from 1993. The dataset is QCEW in all columns apart from Column 2 and 4. All estimates, with the exception of Column 5 are unweighted. Columns 1 - 7 use data from MSCZ pairs, while Columns 8 - 10 use pairs of counties. Column 8 uses all border county pairs, while Columns 9 and 10 correspond to sub-samples 1 and 3 as identified by JNR. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix A Descriptive figures, and additional results using TWFE estimators

A.1 TWFE with log population weights

Our TWFE estimates in Figure 2 and Tables 1 and 2 use working age population as weights in the weighted specifications. JNR also use log of working age population as weights, which arbitrarily reduces the representation of the most populous commuting zones. Use of log population weights is unusual, and to our knowledge has never been used in the minimum wage literature. Nevertheless, in Appendix Figure A2 we recreate Panel B of Figure 2 (which plots coefficients from a TWFE-logMW-CZ specification with different start years and different datasets) but with JNR's preferred log population weights. Consistent with our findings, the coefficients from both QCEW and CBP data become insignificant as data from the 1990s is removed. The results also accord with our result that extending the sample to 2019 (with QCEW data) makes the employment coefficient less negative (more positive) across all start years.

A.2 Trends and anticipation effects in TWFE distributed lag models

JNR analyze pre-trends using the distributed lag TWFE framework, following the approach developed by DLR. They observe a positive trend for BCP between three years and one year before treatment and a negative trend for MSCZ during the same period. While they consider the positive trends problematic, they attribute the negative trends to possible anticipation effects.

To assess this claim, we re-estimate distributed lag models by extending the analysis to include leads up to four years before treatment. Specifically, we compare trends between four years and two years prior to treatment instead of between three years and one year. Table A1 presents these re-estimated results. The negative trends for MSCZ become more pronounced and remain statistically significant when comparing four years and two years before treatment. In contrast, the positive trends for BCP remain similar in magnitude, but are smaller than those for MSCZ, and they lose statistical significance.

This finding suggests that the negative trends observed for MSCZ do not result from anticipation effects, contrary to JNR's claim. Moreover, a two-year anticipation effect is inconsistent with most of the extant minimum wage literature. Importantly, as we note in the paper, recent research has raised concerns about the reliability of distributed lag results, which we also show using simulation results in Appendix C. For this reason, we implement pre-trends tests using our DiD event study design. However, the findings from Table A1 suggest that even considering the TWFE-logMW-CZ distributed lag model favored by JNR, their interpretation of the results is questionable.

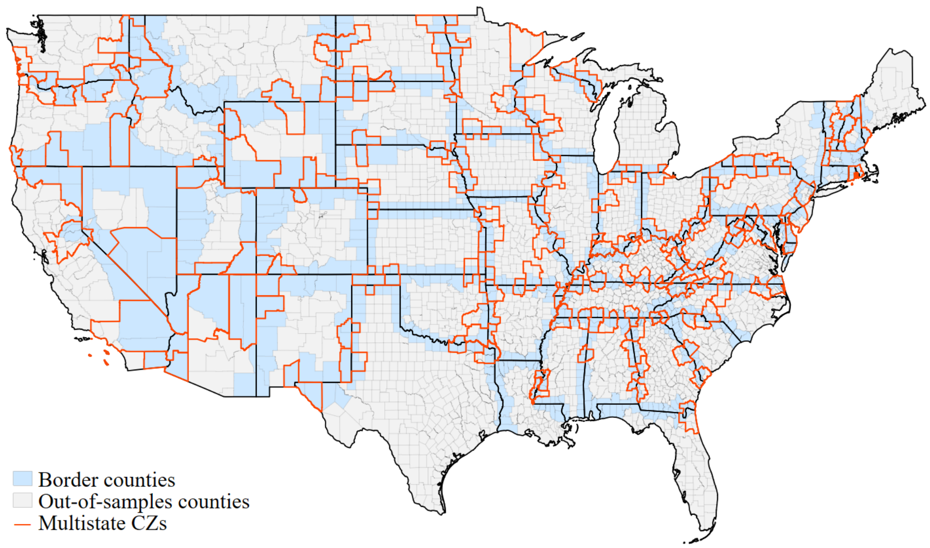
A.3 TWFE results with quarterly QCEW data

For comparability with JNR's annual CBP data, we use annual QCEW data, which also reduces confidentiality-related suppression. However, in Appendix Table A2 we show that results using quarterly QCEW data are similar, although the TWFE-logMW-CZ results move modestly closer to zero in the quarterly sample. In principle, this change in coefficient could be driven either by (1) change in the sample, since there are fewer balanced county panels with quarterly than annual data (due to confidentiality reasons), (2) difference in the frequency of observation within the same sample.

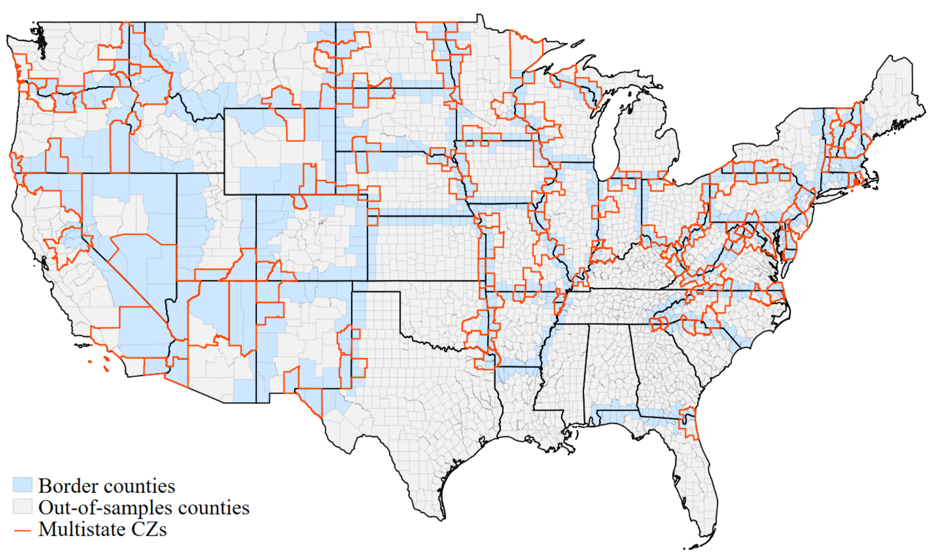
To assess these two possibilities, Appendix Table A3 additionally reports annual-level estimates but only using the quarterly sample. These results are nearly the same as the quarterly-level estimates from the quarterly sample in Appendix Table A2. Therefore, the modest change in coefficients from using quarterly data largely arises from the change in sample. This also underscores why we use annual data as our primary sample, in order to correspond as closely as possible to the CBP sample used by JNR.

Figure A1: Geographical comparison of border counties and commuting zones

A. All BCPs and MSCZ



B. BCPs and MSCZs with different MW in any year between 1990-2016

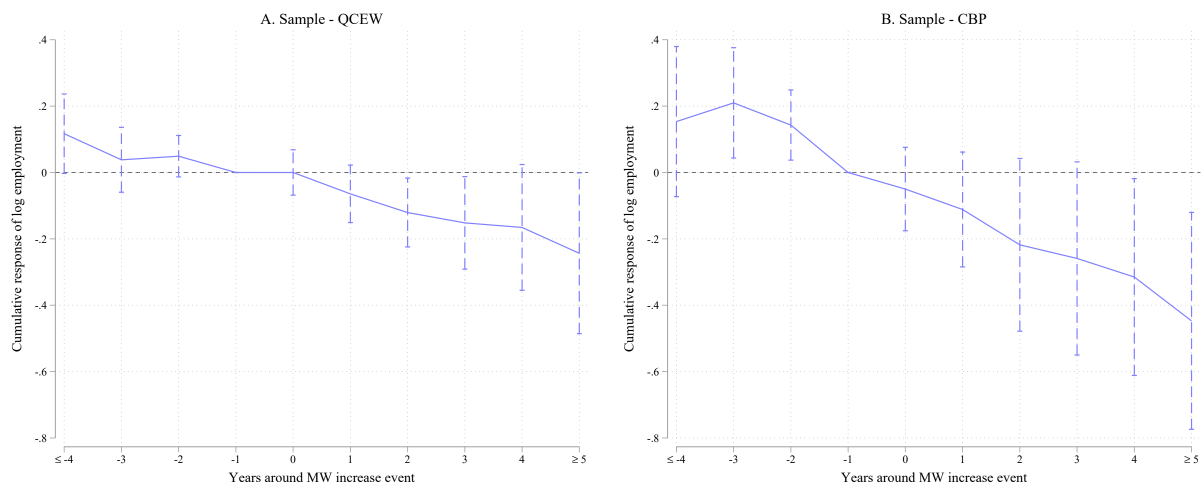


Notes: Each panel shows a map of the U.S. mainland. State borders are represented by thick black lines, while gray lines depict county limits. Border counties are shaded blue, while non-border counties are shaded light gray. Borders of MSCZs are in red. Panel A shows all BCPs and MSCZs. Panel B shows the subset of BCPs and MSCZs that experienced a minimum wage difference in any year between 1990 and 2016.

Figure A2: Minimum wage employment elasticity, TWFE-logMW-CZ specification weighted by log population: different start years, CBP and QCEW

Notes: Estimated using the TWFE-logMW-CZ specification described in equation (1). All regressions are weighted by log of working-age population. Each point corresponds to an estimate from a regression using a particular dataset, starting in the year shown on the x-axis and ending in either 2016 or 2019. The horizontal axis reports the first year in the sample. The vertical axis depicts the effect of log minimum wage on log employment. Each color-shape combination refers to a particular dataset used and the final year in the sample, as described by the legend. Lines show 95 percent confidence intervals. All estimates use MSCZ state pairs and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. All specifications include log employment outside of the restaurant sector and log working-age population as control variables. Standard errors are clustered at the state level.

Figure A3: Cumulative response of log employment to a log point increase in minimum wage; TWFE model with distributed lags (1990-2016, unweighted, MSCZ pairs samples)



Notes: This figure plots the cumulative response of log employment to a log-point increase in the minimum wage using a distributed lag model with 3 leads and 5 lags of log minimum wage. The cumulative responses are normalized relative to event date (-1). 95 percent confidence intervals are based on standard errors clustered by state, and all regressions use state population weights. Panel A uses data from QCEW (from 1990-2016) and Panel B uses data from CBP (again from 1990-2016). Unlike in Table A1 we use the full data from the CBP - by augmenting CBP data with minimum wage data in out-of-sample periods, so that any in-sample period is not dropped due to missing lags and leads. Both panels use multi-state commuting-zone-by-state pairs data and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. All estimates are unweighted. All specifications include log employment outside of the restaurant sector and working-age population as control variables.

Table A1: Modification of JNR table 7

	(1) MSCZ pairs	(2) BC pairs
A. Trend between t: -3 and t: -1		
t: -3	-0.074	0.023
	(0.060)	(0.064)
t: -1	-0.166**	0.104
	(0.069)	(0.078)
t: 0	-0.420***	-0.042
	(0.122)	(0.095)
Trend (t[-1] - t[-3])	-0.092**	0.080*
	(0.046)	(0.044)
B. Trend between t: -4 and t: -2		
t: -4	0.033	0.002
	(0.072)	(0.066)
t: -2	-0.099	0.051
	(0.093)	(0.083)
t: 0	-0.279*	-0.054
	(0.147)	(0.089)
Trend (t[-2] - t[-4])	-0.133**	0.049
	(0.052)	(0.042)

Notes: Panel A estimated using the specification in Equation 7 in JNR. Panel B is a modification of the equation with the third and first lag replaced with the fourth and second, respectively. Both columns use data from County Business Patterns. JNR do not augment their data with out-of-sample minimum wage observations, so some years get dropped because of missing leads of log minimum wage. Thus, their results contain observations up till 2013. To ensure replicability of Panel A, we also restrict to observations until 2013. For Column B, as we have an additional lead, we augment with minimum wage data from 2017, so that we can still use data up to 2013. The distributed lag figures in Figure A3, on the other hand, use our preferred method of using all available data until 2016. In this table, Column 1 uses multi-state commuting-zone-by-state pairs sample and column 2 uses the sample of border county pairs. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: sym* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: MW effects on log average earnings and log employment in restaurants: TWFE estimates for different samples and specifications using quarterly QCEW data

	1990-2019			1990-2016			2000-2016		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Unweighted									
Log wages	0.271***	0.252***	0.252***	0.239***	0.236***	0.226***	0.228***	0.200***	0.220***
	(0.028)	(0.039)	(0.032)	(0.027)	(0.042)	(0.037)	(0.026)	(0.037)	(0.024)
Log employment	-0.262***	-0.085	0.034	-0.290***	-0.118	0.025	-0.092*	-0.022	0.129**
	(0.090)	(0.076)	(0.060)	(0.087)	(0.079)	(0.056)	(0.050)	(0.039)	(0.058)
OWE	-0.961***	-0.319	0.134	-1.176***	-0.472	0.107	-0.373*	-0.098	0.562**
	(0.291)	(0.278)	(0.236)	(0.322)	(0.304)	(0.238)	(0.192)	(0.181)	(0.238)
B. Weighted by population									
Log wages	0.230***	0.236***	0.212***	0.204***	0.220***	0.174***	0.149***	0.151***	0.177***
	(0.023)	(0.036)	(0.030)	(0.024)	(0.045)	(0.025)	(0.022)	(0.051)	(0.026)
Log employment	0.007	-0.051	0.057	-0.049	-0.070	0.042	0.024	0.010	0.112**
	(0.065)	(0.076)	(0.058)	(0.050)	(0.106)	(0.058)	(0.039)	(0.045)	(0.055)
OWE	0.018	-0.223	0.243	-0.250	-0.319	0.204	0.155	0.052	0.615**
	(0.284)	(0.308)	(0.255)	(0.226)	(0.471)	(0.313)	(0.274)	(0.271)	(0.256)
N	78,600	18,720	79,920	70,740	16,848	71,928	44,540	10,608	45,288
Fixed effects									
MSCZ pair-year		Y			Y			Y	
BC pair-year			Y			Y			Y
Sample									
All CZ-state	Y			Y			Y		
MSCZ pairs		Y			Y			Y	
Border county pairs			Y			Y			Y

Notes: Estimated using specification described in equation 1 and quarterly data from Quarterly Census of Employment and Wages data. Columns (1)-(3) use the years 1990-2019, (4)-(6) use 1990-2016, (7)-(9) use 2000-2016. Columns (1), (4), and (7) use the sample of all commuting-zones-by-state, and include CZ-state and year fixed effects. Columns (2), (5), and (8) use multi-state commuting-zone-by-state pairs sample and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. Columns (3), (6), and (9) similarly use the sample of border county pairs and include fixed effects for border counties (each time a county is in a pair) and pair-year fixed effects. Outcomes include log weekly earnings, log employment, and own wage elasticity (OWE). OWE is the IV estimate from regressing log employment on log earnings instrumented by log minimum wage. OWE estimates are only reported when there is a strong first-stage effect. Estimates in Panel A are unweighted, while estimates in Panel B are weighted using the working age population. All specifications include log earnings/log employment outside of the restaurant sector, and working-age population as control variables. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: sym* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: TWFE estimates for different samples and specifications, quarterly QCEW data collapsed to annual level

	1990-2019			1990-2016			2000-2016		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Unweighted									
Log wages	0.260***	0.246***	0.250***	0.233***	0.227***	0.227***	0.214***	0.196***	0.208***
	(0.029)	(0.039)	(0.032)	(0.028)	(0.042)	(0.040)	(0.028)	(0.035)	(0.022)
Log employment	-0.256***	-0.071	0.049	-0.306***	-0.101	0.043	-0.079	-0.006	0.136***
	(0.089)	(0.073)	(0.057)	(0.086)	(0.082)	(0.052)	(0.051)	(0.045)	(0.050)
OWE	-0.986***	-0.272	0.192	-1.294***	-0.419	0.179	-0.354	-0.030	0.644**
	(0.300)	(0.271)	(0.229)	(0.320)	(0.323)	(0.226)	(0.219)	(0.208)	(0.241)
B. Weighted by population									
Log wages	0.214***	0.237***	0.205***	0.195***	0.221***	0.172***	0.142***	0.152***	0.175***
	(0.024)	(0.036)	(0.028)	(0.025)	(0.043)	(0.022)	(0.022)	(0.048)	(0.030)
Log employment	0.033	-0.041	0.099	-0.024	-0.051	0.109	0.072	0.025	0.204***
	(0.077)	(0.069)	(0.075)	(0.068)	(0.100)	(0.077)	(0.062)	(0.042)	(0.068)
OWE	0.120	-0.199	0.446	-0.165	-0.237	0.579	0.507	0.122	1.169**
	(0.363)	(0.269)	(0.329)	(0.313)	(0.435)	(0.397)	(0.469)	(0.226)	(0.465)
N	19,680	4,680	20,160	17,712	4,212	18,144	11,152	2,652	11,424
Fixed effects									
MSCZ pair-year		Y			Y			Y	
BC pair-year			Y			Y			Y
Sample									
All CZ-state	Y			Y			Y		
MSCZ pairs		Y			Y			Y	
Border county pairs			Y			Y			Y

Notes: Estimated using specification described in equation 1 and quarterly data from Quarterly Census of Employment and Wages data. Columns (1)-(3) use the years 1990-2019, (4)-(6) use 1990-2016, (7)-(9) use 2000-2016. Columns (1), (4), and (7) use the sample of all commuting-zones-by-state, and include CZ-state and year fixed effects. Columns (2), (5), and (8) use multi-state commuting-zone-by-state pairs sample and include fixed effects for CZ-state (repeated if a CZ-state is part of multiple pairs) and pair-year fixed effects. Columns (3), (6), and (9) similarly use the sample of border county pairs and include fixed effects for border counties (each time a county is in a pair) and pair-year fixed effects. Outcomes include log weekly earnings, log employment, and own wage elasticity (OWE). OWE is the IV estimate from regressing log employment on log earnings instrumented by log minimum wage. OWE estimates are only reported when there is a strong first-stage effect. Estimates in Panel A are unweighted, while estimates in Panel B are weighted using the working age population. All specifications include log earnings/log employment outside of the restaurant sector, and working-age population as control variables. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: sym* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B Construction of QCEW data

In this appendix, we describe how we cleaned and used QCEW data for the analysis in this paper. We downloaded county-level data from QCEW for 1990-2019, merged in working age population from the National Cancer Institute's SEER program, state level minimum wage data from Vaghul and Zipperer (2022) and commuting zones from JNR's replication files. For most analyses, we use annual QCEW data, except for robustness analysis we do in Tables A2 and A3. We keep only fully balanced counties - those with restaurant employment and earnings data for al years from 1990-2019. We define the restaurant sector as the three-digit code 722 ("Food service and drinking places"). Our CZ-state data collapses the balanced counties to the CZ-state level. For our pairs samples, we start with JNR's replication files (i.e. their pair definitions) and merge in our data to only keep the complete pairs with balanced panels. Our analyses all exclude Alaska, Hawaii and Washington DC.[24]

Table B1 shows the effects of retaining only fully balanced counties. Since the QCEW suppresses some data to maintain anonymity, we cannot use all of the counties or MSCZ pairs that are reported in the CBP dataset. However, the counties with suppressed information are usually small and more sparsely populated. Thus the 55 percent of counties account for over 91 percent of the sample population. Similarly, our MSCZ pairs and border county pairs samples cover 83 to 88 percent of the population.

Table B1: Coverage and population shares of QCEW samples

	Total units (counties/pairs)	Units in sample (counties/pairs)	Percentage of units in sample	Percentage of population covered
Counties	3109	1704	54.81	91.49
MSCZ pairs	151	95	62.91	87.85
Border county pairs	1165	430	36.91	83.36

This table shows the coverage of QCEW data by number of units (pairs/counties) and populations. Row 1 shows the balanced QCEW sample sizes. Rows 2 and 3 show coverage for the complete pairs in the border county and MSCZ pairs regressions. The first column reports the total number of units/pairs in JNR's sample. Samples exclude Alaska, Hawaii, and DC.

We also clean state level QCEW data - which is useful as a baseline for constructing minimum wage increase events (although we do not do any TWFE or event study analysis with this data). Here, we use population from the CPS.

[24]The only place where Alaska and Hawaii are included is Figure 1, which is adapted from Dube and Lindner (2024), who include both these states in their analysis.

More details about cleaning of both state and county level data can be found in the readme document of the replication package Dube et al. (2025b) in the Harvard Dataverse, https://doi.org/doi:10.7910/DVN/GTOOOC.

Appendix C Monte Carlo simulation: TWFE event studies may not detect pre-trends, while producing biased estimates

JNR report event study results using a variant of the TWFE model that codes events as binary policy changes rather than as continuous changes (which we refer to as "TWFE-binary"). Implemented via the `xtevent` package in Stata, the distributed lags TWFE-binary produces coefficient paths that do not indicate pre-existing trends and suggest employment reductions post-treatment. On this basis, JNR assert that these TWFE estimates can be interpreted like a conventional difference-in-differences (DiD) event study. Their assertion is incorrect.

A switch from continuous to discrete treatment does not fix the well-known problems of TWFE under staggered adoption. Indeed, almost all of the modern DiD literature showing biases in the TWFE model have done so with binary treatments. The authors of the `xtevent` package explicitly acknowledge this point (Freyaldenhoven et al., 2021).[25] Since these problems were established for binary treatments, they remain in the TWFE-binary specification. Importantly, with a TWFE model, violations of parallel trends that arise even outside the event window can leak into the coefficients that lie inside the plotted window.[26]

This appendix demonstrates these points with a simulation adapted from Dube and Lindner (2024). The simulation provides a "proof by example" that the TWFE-binary estimation using `xtevent` can produce highly misleading inference for detecting pre-trends and estimating treatment effects. In contrast, modern DiD event study methods (such as the LP-DiD approach we use) perform well.

[25] Freyaldenhoven et al. (2021) write: "Under staggered adoption, a two-way fixed effects estimator of β_m in (1) can be represented as a weighted average of policy effects for different event times and cohorts (Sun and Abraham (2021)). However, this estimator can put nonzero weight on effects at event times other than m, and can put negative weight on the effect at event time m for some cohorts."

[26] One clarification: In Cengiz et al. (2019), the baseline panel regression did not rely on "clean controls." However, their baseline results were similar to their results from a stacked DiD approach that used only clean controls. Why were the CDLZ baseline estimates unaffected by parallel trends violations from the 1990s while JNR's estimates using TWFE-binary are not? AS CDLZ explain, they did not bin up leads prior to -3 and lags after 4; they exclude them. This approach creates a closed event window, which mitigates cross-term contamination from the distant past. When the bias in the TWFE model is generated by possible parallel trends violations from one part of the sample, the solution is a design (like BCP but not MSCZ) that does not suffer from that parallel trends violation, or to reduce the bias by limiting the effects of such violations by using clearly delineated event windows that prevent cross-contamination.

C.1 Data generating process

We simulate a balanced panel of $S = 50$ states observed over $T = 15$ periods ($t = 1, \dots, 15$). States belong to one of three cohorts: $\mathcal{G}_0$ (never treated, $|\mathcal{G}_0| = 20$), early adopters $\mathcal{G}_1$ (12 states), and late adopters $\mathcal{G}_2$ (18 states). Early adopters implement the policy at $t = 9$, whereas late adopters do so at $t = 12$. Let

$$D_{st} = \mathbf{1}\{\text{state } s \text{ is treated at time } t\}.$$

Potential outcomes in the absence of treatment follow

$$Y_{st}^0 = \gamma_s + \tau_t + \text{Trend}_{st} + \nu_{st},$$

where $\gamma_s \sim \mathcal{N}(0, 0.05^2)$, $\tau_t \sim \mathcal{N}(0, 0.05^2)$, and $\nu_{st} \sim \mathcal{N}(0, 0.05^2)$. The realized outcome is $Y_{st} = Y_{st}^0 + \beta D_{st}$ with $\beta = 0$ (no true effect of treatment).

Importantly, the parallel trends assumption holds starting with calendar date $t = 6$. The assumption holds locally around the times of adoption and subsequent to adoption. However, there are deviations from parallel trends that occur before $t = 6$, which are all outside the three-lead/three-lag event window:

$$\text{Latent trend}_{st} = \begin{cases} -0.2 \times t & \text{if } s \in \text{early-adopter } (G_1) \text{ and } t < 6 \\ -1 & \text{if } s \in \text{early-adopter } (G_1) \text{ and } t \geq 6 \\ 0.9 \times t & \text{if } s \in \text{late-adopter } (G_2) \text{ and } t < 6 \\ 4.5 & \text{if } s \in \text{late-adopter } (G_2) \text{ and } t \geq 6 \\ 0 & \text{if } s \in \text{never-treated } (G_0) \end{cases}$$

The mean actual and potential outcomes for the three groups are plotted by calendar time in Figure C1, panel A. The key takeaways: averaged across all treatment events, there is a positive pre-trend prior to event date -3; however, for all events, the parallel trends assumption is valid after event date -3. Importantly, the causal effect of treatment is zero for all events.

C.2 Estimators compared

The simulation estimates JNR's preferred TWFE-binary distributed lag specification via the `xtevent` command with a $[-3, 3]$ window where the end points are "binned up." In addition, we employ the Local Projections DiD estimator, also with a $[-3, +3]$ window around each adoption date. For comparability, here we report the "binned up" estimates for the first and last bins when reporting LP-DiD. For example, the LP-DiD estimate associated with event date ≤ -3 is a long difference between the mean outcome for dates -3 or earlier and the outcome at date -1. This captures any early period violations of pre-trends before date -2. For both LP-DiD and XTEVENT, the coefficients are normalised relative to event date -1.

C.3 Simulation results

Across 500 replications, the XTEVENT estimator produces sizable negative post-treatment coefficients, even though the true effect β is zero. The mean and the 95 percent bounds for the coefficients are plotted in panel B of Figure C1. The fall in outcome between event dates 1 and 2 has no correspondence with the true dynamic effect of the treatment, which is flat. As Table C1 shows, at least 99% of the simulations mistakenly find a negative, statistically significant effect (at the 5 percent level) for both event dates 2 and ≥ 3.

At the same time, all lead coefficients appear flat, including the "binned up" ≤ -3 term, thereby failing to detect the early period parallel trends violations that actually drive the bias in the estimated treatment effect. As Table C1 shows, less than 1 percent of the simulations find a statistically significant term for the "binned up" third lead term (event dates ≤ 3), even though in truth that coefficient should be highly negative, capturing a positive pre-trend. In other words, a TWFE model would not indicate a problem with the research design. However, weights in the TWFE model transmute the early adopters' pre-period shocks (calendar dates $t = 1$–5) into a negative, longer-run estimate for lags 2 and 3+, which as we saw are negative and statistically significant at least 99% of the time.

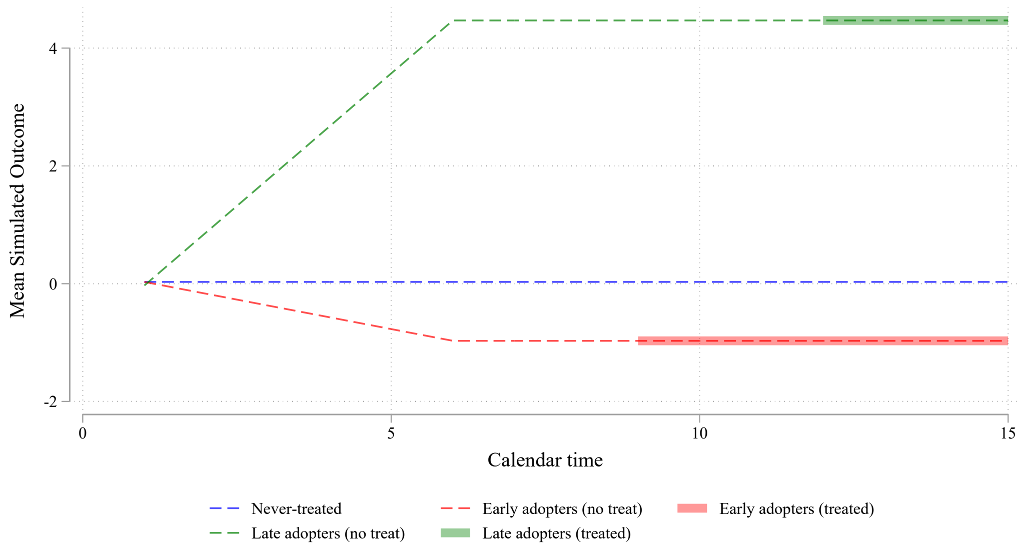
In contrast, the LP-DiD estimator correctly recovers coefficients that are close to zero for event dates -2 or later (Figure 4); the share of statistically significant coefficients ranges between 5% and 7% (Table C1). Moreover, in 100% of the simulations, the LP-DiD recovers a highly negative estimate for the ("binned up") event date ≤ 3, thereby correctly detecting the early period parallel trends violation.

The simulation clarifies that a flat profile of lead coefficients in a TWFE regression does not guarantee that parallel trends hold in the pre-treatment period. The TWFE weighting scheme can conceal violations, even as such violations produce biased estimates of treatment effects. These results arise in the context of a discrete treatment measure and the `xtevent` package used by JNR. In short, JNR's preferred event-study design cannot credibly test the parallel-trends assumption or estimate causal effects.

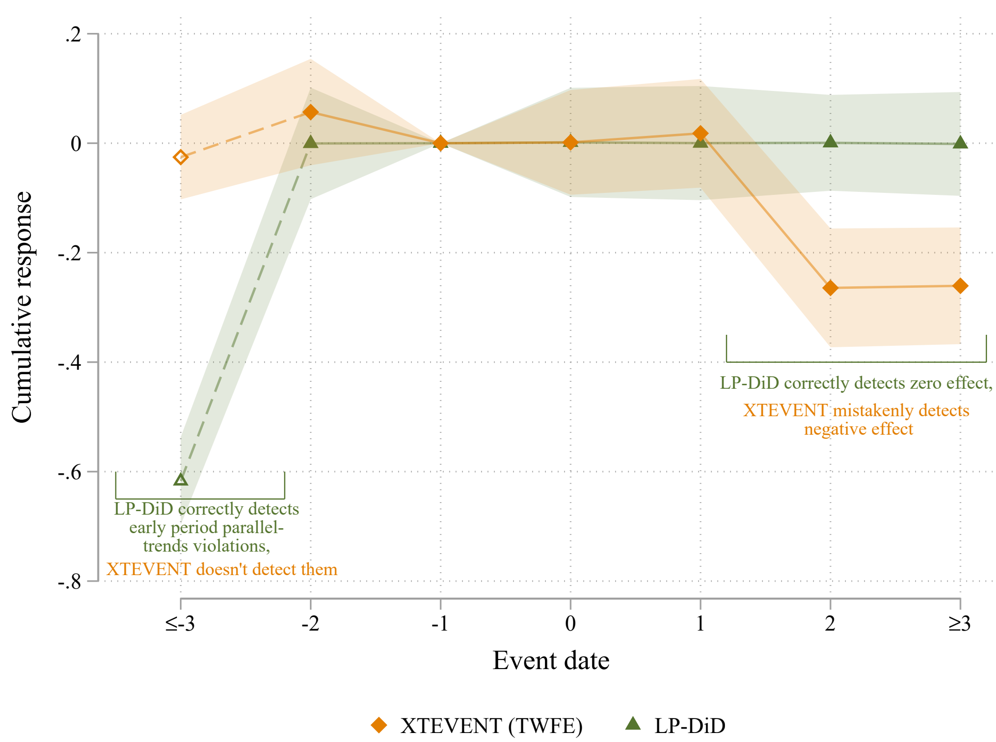
This simulation is based on a highly stylized DGP to demonstrate the qualitative problems with a TWFE-based event study. In Section 5 we use simulations based on actual minimum wage data in MSCZ and BCP samples. We show that even when the true effect of minimum wages is zero, the pre-trends violations within the MSCZ sample (that we empirically document) lead the TWFE-logMW-MSCZ and TWFE-binary-MSCZ estimators to yield substantial negative estimates.

	Coefficients for event times						
	≤ -3	-2	-1	0	1	2	≥ 3
Panel A: LP-DiD							
Positive and sig.	0.000	0.024	0.000	0.018	0.034	0.020	0.028
Negative and sig.	1.000	0.022	0.000	0.028	0.030	0.028	0.038
Panel B: XTEVENT							
Positive and sig.	0.000	0.214	0.000	0.026	0.066	0.000	0.000
Negative and sig.	0.000	0.000	0.000	0.034	0.014	0.992	0.998

Table C1: Share of 500 simulations with a significant coefficients (at the 5% level), by event times: XTEVENT (TWFE) versus LP-DiD.



Panel A. Data generating process



Panel B. Simulation results

Figure C1: Monte-Carlo study with heterogeneous pre-trends.

Appendix D Difference-in-differences event-study methodology

This appendix explains how we construct combined minimum wage events and how we implement the difference-in-differences (DiD) event study that underpins the paper's estimates. Section D.1 details the rules we used to identify events, define post-periods, and select clean controls. Section D.2 lays out the econometric specification, Section 4.2 provides additional evidence on pre-trends by extending our event study specification, while Section D.3 summarizes supplementary robustness checks.

D.1 Construction of minimum wage events

Here we describe the steps to identify each of the 45 combined state-level events in our analysis.

1. We first identify all instances of a prominent *state* minimum wage increase of at least $0.25 and 5 percent.[27]

2. We then exclude events that occurred in years with a federal minimum wage increase. We do so because an event that begins in a federal increase year cannot have a "clean" counterfactual—there are no states that did not experience minimum wage increases in that year or in the "event period." Thus, we remove events in 1990, 1991, 1996, 1997, 2007, 2008 and 2009. We label the remaining events as our set of "admissible events."

3. We then classify admissible events as "provisional combined events" if they do not include any admissible events in the preceding three years. This step allows us to pool multi-year minimum wage increases into a single combined event.

4. In some cases, this procedure does not properly capture an event's start date. Specifically, if a multi-phased increase begins with too small an increment, we would incorrectly code the start year later than when the event actually began. To remedy this possibility, we use the following algorithm:

   - For each of the provisional combined events that we identified in the previous step, we check for the presence of any minimum wage increases (large or small) in the year prior to the

[27] This definition differs slightly from Cengiz et al. (2019) for data-driven reasons. However, in Appendix Table D3 we show robustness to the use of the Cengiz et al. (2019) events as well.

provisional start year, while also checking that a federal increase did not occur in the prior year.

- If we find such an increase and the state did not have indexation at that time, we consider the event a legislated minimum wage increase. For such events, we assign the start year as one year earlier. (Minimum wage indexation years are from Brummund and Strain (2020).)
- If the state did have indexation, we still check if the small increase was legislated. (Some small increases were legislated, rather than the result of indexation, even in states that had indexation at the time). If the small increase is indeed legislated, we designate the event start date as the year prior to the provisional start year. If it is just a small increase due to indexation, we keep our provisional start year.
- We then repeat this process until none of our start years have a legislated, non-federal-year minimum wage increase in the preceding year.

5. The steps above yield 49 events and a maximum post-period of 6 years. In four states, however, an event fell within the post-period of an earlier event in the state: Delaware 2014 and 2019; New Jersey 2014 and 2019; Oregon 1998 and 2003; and Rhode Island 1999 and 2004. In these cases, we removed the later event and subsumed it with the earlier one, provided that the later event's post-period minimum wage increase was smaller than the increase for the earlier event. In practice, this condition held in all four cases. As a result, our final list includes 45 events.

We display these 45 combined state events as filled, dark blue circles in Figure D1. These 45 events represent 98 prominent increases and 37 smaller increases—a total of 135 underlying events.[28] On average, events span a 4.2-year post-treatment period.

While our baseline event study analysis examines these 45 combined state minimum wage increases, the exact set of usable events varies by dataset, as described in Table D1. Using balanced panels of QCEW county data, we can study 34 events within MSCZs and 42 events within BCPs. The average minimum wage increase over the post-treatment period is nearly identical for both: 24.3 percent and 24.6 percent, respectively.

[28]Overall, the 45 events capture 72 percent of all prominent state increases. Appendix D3 shows robustness to various event definitions.

D.2 DiD event study design

Our pooled event study analysis aggregates individual case studies. We use a local projections DiD (LP-DiD) event study model, which produces the same estimates as using the stacked event study approach in Cengiz et al. (2019). However, the LP-DiD approach does not require us to construct a stacked dataset, so we can use the original panel data.[29]

D.2.1 Defining "clean controls"

Our event study contrasts outcome changes in treated units with those in "clean controls," defined as areas that experience *no* state-level minimum-wage increases during the three years before an event and throughout its post-treatment window, which can extend up to six years. Because a federal minimum wage increase would invalidate the counterfactual, we exclude any calendar year containing a federal increase from the post-treatment period.[30]

We implement a *maximum* six-year post-period, corresponding to event times 0 through 5, with the actual length depending on the treatment cohort. A 2013 event contributes a full six-year window (2013–2018); a 1994 event contributes only two post years because the 1996 federal increase truncates the period; and a 2017 event contributes three post years because the sample ends in 2019. Table D2 reports the distribution of post-period lengths across cohorts. Under these rules, a state qualifies as a clean control for a 2013 event only if it records no state-level increases between 2010 and 2018, whereas a clean control for a 1994 event must have no increases between 1991 and 1995. These restrictions ensure that treatment and control units share the same minimum-wage path over the relevant pre- and post-event horizons, bolstering the credibility of the parallel-trends assumption.

D.3 Supplementary findings for event studies

In this subsection, we describe modifications to the baseline event study specification to evaluate the robustness of our estimates presented in Table 3 of the main paper.

[29] See Dube et al. (2025a) for the general case and Dube and Lindner (2024) for the minimum wage context.

[30] No area remains entirely "untreated" during the entire 1990-2019 period given the recurring nature of minimum-wage adjustments. Appendix D.3 presents extensive robustness checks that vary the selection of events, controls, and window lengths.

D.3.1 Robustness to clean control definitions

In the baseline specification, we define clean controls as states with no state minimum wage changes in the three years prior to the event. Since the three-year cutoff is arbitrary, we assess robustness using alternative clean control definitions: states with no minimum wage increases in the two or four years before the treatment event. We present the results in Columns 1, 2, 4 and 5 of Table D4. The wage and employment estimates remain consistent, with OWEs ranging from -0.388 to 0.006, depending on geographic unit (MSCZ or BCP) and weighting. For comparison, the OWE range using the three-year pre-period is -0.381 to 0.009. All OWEs are statistically insignificant across these specifications.

D.3.2 Adding federal minimum wage increases

Our primary event study estimates focus on state-level minimum wage increases, where treated states experience increases while clean control states do not. In most cases (37 out of 45), all control states have lower minimum wages than the treated state at the end of the post-period. Across all events, 99 percent of all control units have a lower minimum wage than their corresponding treated state at the end of the evaluated period.

Federal increases are more challenging to analyze, as they require comparing states that experienced larger increases to those that experienced smaller ones. Using federal events requires assuming the absence of significant heterogeneity in treatment effects from federal increases. Despite this limitation, we conduct an analysis of federal events under these assumptions.

To classify states as bound or unbound by a federal increase, we adopt the methodology of Clemens and Wither (2019). A state is bound if its minimum wage was below $6.55 at the start of 2008 (the second year of the three-year phased federal increase). Similarly, states are bound for the 1996-1997 federal increases if their minimum wage was below $4.75 at the beginning of 1996. We exclude the 1990-1991 federal increase because of insufficient pre-period data.

Our approach requires adjustments to defining clean controls and post-periods, since unbound states often implement state-level minimum wage increases. For example, during the 2007-2009 federal increases, requiring clean controls to have no state-level increases over a six-year post-period and a three-year pre-period results in no qualifying states. Another issue arises with defining the post-period. Previously, the post-period extended until the next federal increase or the end of the sample period. For federal increases, extending the post-period reduces the number of control states as unbound states often implement state increases, while retaining more controls requires shortening the post-period.

To address these challenges, we identify the maximum post-period for all unbound states during the federal increases, defined as the longest period (up to six years) without a prominent (at least \$0.25 and 5%) state-level increase. We then apply the same post-period to the corresponding treated (bound) state for each pairwise comparison. For example, consider State A, which is bound by the federal increase in 2008 and shares MSCZs with two unbound states, B and C. If State B implemented a significant state increase in 2011 and State C did so in 2015, the comparison between A and B would use a post-period of 2008-2010, while the comparison between A and C would use a post-period of 2008-2013.

Our control group is ultimately a subset of Clemens and Withers' unbound states, as we exclude unbound states from the control group once they implement state-level increases following federal events.

There are 44 bound states in 1996 and 27 in 2008, creating 71 potential "federal events" in addition to the 45 state-increase events we defined earlier. However, as we have pair-year fixed effects in our event study specification, events can only be evaluated if there is treatment variation within a pair-year. In other words, for federal events, we can evaluate the increase in a bound state only if one of its neighbors is an unbound state with at least one clean post-period year.

Thus, a total of 18 additional federal events are evaluated when using MSCZs and 27 are evaluated for BCPs.[31] As Columns 1 and 4 of Appendix Table D3 show, across both MSCZ and BCP designs and regardless of weighting, OWEs range between -0.108 and 0.499. None of these estimates are statistically significant, ruling out sizable effects. Including federal minimum wage increases in the event studies thus does not alter our key findings.

D.3.3 Excluding federal increases in the pre-treatment period

This robustness check refines our analysis by including only events where control areas did not experience state *or* federal minimum wage increases in the three years before the event. In the baseline specification, states qualify as clean controls if they meet two conditions: (1) no state minimum wage increases in the three years preceding the event and (2) no state or federal increases in the post-period. To satisfy the

[31] These include 6 and 12 events from 1996 and 2008 respectively for MSCZs; and 10 and 17 respectively for BCPs

second condition, the post-period is truncated to avoid overlapping with federal increases, resulting in varying post-periods across event cohorts. Once the post-period is defined, clean controls must also have no state increases within that period. The asymmetry in how our baseline specification treats federal and state minimums in the pre-treatment period arises because federal increases tend to affect most states, including the treated and control states. This robustness check addresses the bias that could still arise from the differential bites of the federal increase and lagged effects.

In the baseline, a small number of events are evaluated using control states that experienced federal increases in the three-year pre-period. For instance, events in 2010 are included, with clean controls defined as states with no state increases from 2007 to 2015. However, these controls may have experienced federal increases during 2007-2009. In this robustness check, we exclude such events, dropping 12 events in total from 1992-1994, 1998-2000, and 2010-2012; only two of these events occur in 2000 or later.[32]

Columns 2 and 5 of Table D3 present the results for the MSCZ and BCP samples. After excluding events with federal increases in the pre-period, we continue to observe significant wage effects, null employment effects and OWEs between -0.076 to 0.462, depending on the geographic unit and weighting.

D.3.4 Cengiz et al. (2019) events, uncombined

Our baseline estimates combine multiple phases of state minimum wage increases into single events. This approach evaluates the total effect of phased-in policy changes without requiring controls for additional increases during the post-treatment period. Furthermore, the baseline analysis excludes post-periods with federal increases, ensuring clean control groups without minimum wage changes. This methodology refines the event construction in Cengiz et al. (2019), who treated each prominent increase as a separate event and controlled for small and federal increases parametrically, requiring stronger assumptions.

To evaluate whether our conclusions are robust to this refinement, we analyze 103 CDLZ events from 1991 to 2016 that are available in our data.[33] For this specification, we retain the baseline definition of clean controls but include additional covariates: 1) An indicator for whether another large state increase

[32] This stricter criterion excludes the widely studied 1992 New Jersey minimum wage increase analyzed by Card and Krueger (1994), as the federal increase in 1990-1991 affected both New Jersey and Pennsylvania.

[33] Cengiz et al. (2019) originally identified 138 events from 1979-2016, with 121 events occurring in 1991 or later. Collapsing their quarterly dataset to annual data reduces this to 120 events, as Pennsylvania had two increases in 2007 during different quarters. Excluding Alaska, Hawaii, and Washington, DC yields 103 potentially usable events. The exact number of events analyzed depends on the availability of balanced data and clean cross-border controls. Details on the number of events used in MSCZ and border county analyses are provided in Table D3.

occurred in the treated state during the pre-treatment period (instead of combining events). 2) Indicators for small state increases in the treated state during the pre- or post-treatment periods. 3) Indicators for federal increases affecting control states in the pre- or post-treatment periods (instead of excluding events with interfering federal increases).

Columns 3 and 6 of Table D3 present the results for this specification. The findings remain consistent, showing strong wage effects, near-zero employment effects and OWEs ranging from -0.247 to 0.102, none of which are statistically significant.

D.3.5 Log employment as outcome

Our preferred event study specification uses log employment rate instead of log employment, with controls for log population. This approach eliminates the need for log population as a covariate. Recent advances in the difference-in-differences literature have highlighted potential biases when treatment effects vary by covariate levels (e.g., policy impacts differing between high- and low-population areas).

For robustness, we modify our baseline event study specification to follow the methodology used in JNR and DLR (and in TWFE estimates in this paper). Specifically, we use log employment as the outcome variable instead of log employment ***rate***, while incorporating controls for log employment and log earnings in non-restaurant sectors, as well as log population. All outcome and control variables are defined as *post - pre* differences, as described in our event-study equation (Equation 2).

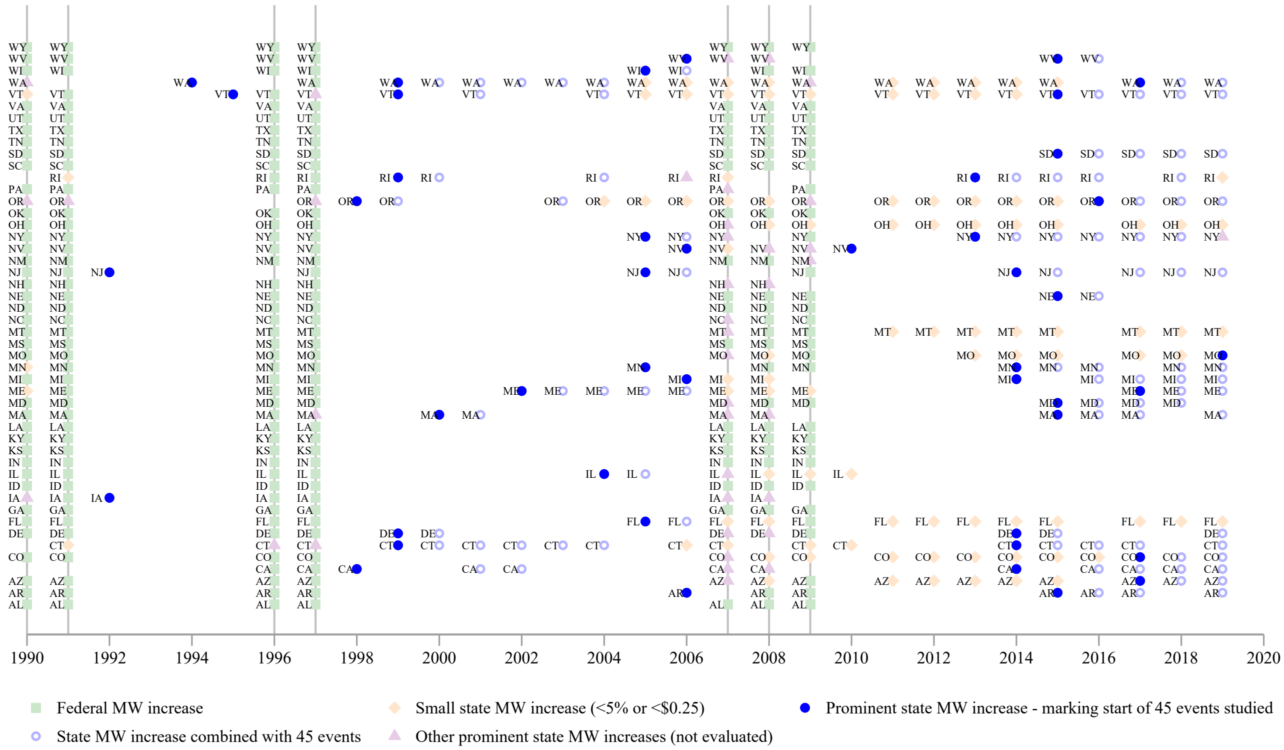
As columns 3 and 6 of Table D4 show, this modification has no noticeable impact on our findings. The earnings and employment estimates remain nearly the same, with OWEs between -0.346 and 0.033.

D.3.6 Accounting for composition effects in event study estimates

The lengths of post-treatment periods vary across events. Among the 42 events used for BCP analysis, 25 have a post-treatment window of at least 4 years, while 17 have 3 years or less. Although the average effect (calculated across all post-treatment periods, regardless of length) is well estimated, the dynamic response may be influenced by composition issues. These occur when different events contribute unequally to estimates for different event times, potentially causing concern if the effects vary across different event cohorts.[34] To evaluate this issue, Tables D5 and D6 present estimates for each event time from 0 to 5, holding event composition constant. For example, for the BCP sample in Table D6, column 6 presents effects for event times 0 to 4, considering only the 22 events with at least 5 post-treatment periods (i.e., extending through event time 4). Other columns follow a similar approach. For both the MSCZ and BCP samples, we find no evidence that composition effects influence outcome dynamics. Wage effects increase over time, while employment effects remain close to zero, even when composition is held constant.

[34] The distributed lag version of TWFE also assigns different weights to events for estimating various leads and lags, but in a much more opaque fashion.

Figure D1: All minimum wage increases from 1990-2019



Notes: This figure plots all state minimum wage increases. The blue solid circles represent the start of 45 events that we study in our main analysis. The blue hollow circles represent state minimum wage increases that we evaluate as part of the post-periods of these 45 events. Some state increases are neither the start of an event, nor are evaluated as part of another event. These increases are either very small (some reflect indexation) - represented by orange diamonds; or they are prominent increases, mostly in years when the federal minimum wage also increased - represented by purple triangles.[35]. The green squares show federal minimum wage increases. Any state without a green square in a federal increase year either increased its own minimum wage that year (small or prominent), making its binding minimum wage higher than the federal; or its minimum wage was higher than the federal in the previous year as well, making the federal increase redundant.

Table D1: Number of events evaluated in different event study specifications

	State	MSCZ pairs	BC pairs
Combined MW increases	45	34	42
All underlying MW increases	135	97	122
Pairs in regression	.	91	420
Pairs with treatment variation	.	62	261

This table shows the number of events evaluated and the number of border county and MSCZ pairs in QCEW baseline event study regressions. These are the regressions in Table 3. We report the state sample too to show total number of potential events. The numbers reported are for the full period regressions for each sample. The first row reports the total number of combined events, while the second row gives the number of all MW increases evaluated as part of the combined events included. The third and fourth rows report the number of commuting zone/county pairs used. Our estimation sample for all years is all units which are either newly treated or a clean control for that particular year. Thus, unit-years which are neither are dropped. Moreover, the Stata command reghdfe, which we use with pair-year fixed effects drops singletons - pair-years where one side is either treated or control, but the other side is neither. Therefore, the pair-years that remain in the regression are those where either 1) one side is treated and the other is a clean control; or 2) both are treated; or 3) both are controls. The total unique number of such pairs (belonging to all three categories) is reported in the third row. However, considering the fixed effects are at the pair-year level, the coefficients are calculated using the pair-years with variation in treatment (only category 1 above). Thus, the final row reports all such pairs. Row 1 in this table, and any table in the paper which shows ”events evaluated” reports the number of events using these pairs only.

Table D2: Details for the 45 combined minimum wage events, by year of the initial increase

	Number of events	Post period length(years)	No. of clean control states	Mean increase in log MW
1992	2	4	39	.131
1994	1	2	41	.142
1995	1	1	44	.057
1998	2	6	39	.249
1999	5	6	38	.269
2000	1	6	33	.251
2002	1	5	29	.271
2004	1	3	30	.233
2005	5	2	31	.245
2006	4	1	32	.2
2010	1	6	17	.089
2013	2	6	22	.368
2014	6	6	24	.302
2015	7	5	24	.26
2016	1	4	24	.196
2017	4	3	24	.305
2019	1	1	24	.091
Overall	45	4.2	29.4	.246

This table provides information by cohort on the 45 state-level events used in the event-study analysis. The 45 events are defined as per state minimum wage changes. Availability of data and variation across county/commuting zone pairs dictate which events are actually used in different specifications. The first column reports the number of events in that cohort/year. The second column gives the length of the post-period for that cohort in years. Post-periods are a maximum of 6 years unless interrupted by a federal MW increase year or end-of-sample. The third column reports number of states that serve as clean controls for the cohort - a clean control needs to have no state MW increase in the three years before the cohort year, and also no state MW increase in the relevant post-period. Finally, the last column reports the average log MW increase (average of the difference between log MW in the last year of the post-period and the log MW in year *t-1*, where *t* is the event year) across all events in a cohort. The last row gives the total number of events and means *by event* (not by cohort) of post-period length, clean control states, and log MW increase.

Table D3: Event study estimates from QCEW data: Robustness to event definitions

	MSCZ pairs			BC pairs		
	(1) Federal events included	(2) No fed increase in pre-period	(3) CDLZ events + CDLZ controls	(4) Federal events included	(5) No fed increase in pre-period	(6) CDLZ events + CDLZ controls
A. Unweighted						
Log wages	0.026***	0.031***	0.043***	0.025***	0.029***	0.045***
	(0.006)	(0.008)	(0.010)	(0.005)	(0.006)	(0.006)
Log employment rate	0.006	0.015	0.000	-0.003	-0.001	-0.011
	(0.009)	(0.009)	(0.015)	(0.008)	(0.011)	(0.017)
MWE	0.032	0.064	0.001	-0.014	-0.006	-0.043
	(0.049)	(0.041)	(0.061)	(0.041)	(0.043)	(0.066)
OWE	0.238	0.462*	0.008	-0.108	-0.052	-0.247
	(0.365)	(0.259)	(0.336)	(0.308)	(0.370)	(0.374)
B. Weighted by population						
Log wages	0.022***	0.024**	0.040***	0.019***	0.027***	0.030***
	(0.005)	(0.010)	(0.011)	(0.006)	(0.010)	(0.011)
Log employment rate	0.002	-0.002	-0.008	0.010	0.003	-0.005
	(0.009)	(0.006)	(0.011)	(0.009)	(0.007)	(0.006)
MWE	0.011	-0.007	-0.032	0.046	0.010	-0.021
	(0.044)	(0.024)	(0.044)	(0.041)	(0.024)	(0.025)
OWE	0.108	-0.076	-0.205	0.499	0.102	-0.176
	(0.447)	(0.262)	(0.270)	(0.554)	(0.247)	(0.213)
N	2,468	2,086	2,264	11,972	10,178	10,550
Events evaluated	52	27	79	69	31	93

Notes: This table checks robustness of our baseline event study estimates in Table 3 by changing event definitions. Columns 1-3 report estimates for MSCZ pairs, while Columns 4-6 have results for BC pairs. Columns 1 and 4 include federal increase events, as discussed in sub-section D.3.2. Columns 2 and 5 remove events that have a federal increase up to three years prior. Finally, Columns 3 and 6 use the uncombined event definitions from Cengiz et al. (2019), also using their specification to control for other events in the event window (see sub-section D.3.4). Outcomes are defined similarly to Table 3. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D4: Event study estimates from QCEW data: Robustness to control sets and covariates

	MSCZ pairs			BC pairs		
	(1) Clean controls: No increase 2 years prior	(2) Clean controls: No increase 4 years prior	(3) Outcome is log employment	(4) Clean controls: No increase 2 years prior	(5) Clean controls: No increase 4 years prior	(6) Outcome is log employment
A. Unweighted						
Log wages	0.032***	0.031***	0.031***	0.028***	0.027***	0.027***
	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)
Log employment (rate)	0.001	-0.000	0.001	-0.004	-0.009	-0.008
	(0.010)	(0.011)	(0.010)	(0.009)	(0.010)	(0.009)
MWE	0.004	-0.001	0.004	-0.016	-0.039	-0.033
	(0.046)	(0.049)	(0.045)	(0.039)	(0.045)	(0.039)
OWE	0.029	-0.009	0.033	-0.134	-0.343	-0.293
	(0.315)	(0.341)	(0.319)	(0.324)	(0.382)	(0.327)
B. Weighted by population						
Log wages	0.020*	0.020*	0.020*	0.020*	0.015*	0.020*
	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)	(0.010)
Log employment (rate)	-0.008	-0.008	-0.007	0.000	-0.003	-0.001
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.004)
MWE	-0.030	-0.030	-0.028	0.000	-0.011	-0.006
	(0.024)	(0.025)	(0.023)	(0.018)	(0.023)	(0.017)
OWE	-0.381	-0.388	-0.346	0.006	-0.179	-0.083
	(0.288)	(0.303)	(0.287)	(0.239)	(0.317)	(0.211)
N	2,220	2,074	2,132	10,788	10,114	10,414
Events evaluated	34	33	34	42	41	42

Notes: This table checks robustness of our baseline event study estimates in Table 3 by changing covariates or how control areas are defined. Columns 1-3 report estimates for MSCZ pairs, while Columns 4-6 have results for BC pairs. Columns 1 and 4 (2 and 5) require controls areas to have no state minimum wage increases 2 (4) years prior, instead of the baseline 3. Columns 3 and 6 change the dependent variable from log employment *rate* to log employment. They also add controls for log population, and log employment and log earnings in non-restaurant sectors. This is similar to JNR's TWFE specification. However, keeping in line with the specification in the baseline event-study equation 2, our controls are also in differenced form. Outcomes, such as MWE and OWE are defined similarly as Table 3. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D5: Dynamic event study estimates for cohorts with different post-periods - MSCZ pairs

	Full sample	Balanced sample - different post-period lengths					
	(1)	(2) ≥ 1	(3) ≥ 2	(4) ≥ 3	(5) ≥ 4	(6) ≥ 5	(7) $= 6$
A. Log earnings							
k = 0	0.019***	0.019***	0.024***	0.023**	0.016	0.017	0.020
	(0.007)	(0.007)	(0.009)	(0.009)	(0.011)	(0.012)	(0.019)
k = 1	0.037***		0.037***	0.036***	0.028**	0.033***	0.040**
	(0.009)		(0.009)	(0.010)	(0.011)	(0.012)	(0.017)
k = 2	0.045***			0.045***	0.037**	0.046***	0.055**
	(0.012)			(0.012)	(0.016)	(0.016)	(0.022)
k = 3	0.045***				0.045***	0.053***	0.068***
	(0.014)				(0.014)	(0.014)	(0.018)
k = 4	0.051***					0.051***	0.070***
	(0.015)					(0.015)	(0.019)
k = 5	0.065***						0.065***
	(0.021)						(0.021)
Post-averaged effect	0.031***	0.019***	0.031***	0.035***	0.032**	0.040***	0.053***
	(0.008)	(0.007)	(0.008)	(0.010)	(0.012)	(0.013)	(0.018)
B. Log employment rate							
k = 0	0.001	0.001	0.000	0.002	0.004	0.018	0.024*
	(0.007)	(0.007)	(0.010)	(0.011)	(0.014)	(0.011)	(0.014)
k = 1	0.000		0.000	-0.000	-0.000	0.009	0.007
	(0.010)		(0.010)	(0.010)	(0.012)	(0.012)	(0.015)
k = 2	-0.005			-0.005	-0.006	0.006	0.017
	(0.015)			(0.015)	(0.018)	(0.018)	(0.025)
k = 3	-0.007				-0.007	0.010	0.009
	(0.025)				(0.025)	(0.024)	(0.035)
k = 4	0.016					0.016	0.021
	(0.028)					(0.028)	(0.040)
k = 5	0.034						0.034
	(0.045)						(0.045)
Post-averaged effect	-0.000	0.001	0.000	-0.001	-0.002	0.012	0.019
	(0.010)	(0.007)	(0.009)	(0.012)	(0.016)	(0.018)	(0.028)
OWE	-0.004	0.032	0.009	-0.030	-0.074	0.297	0.357
	(0.335)	(0.380)	(0.304)	(0.338)	(0.538)	(0.384)	(0.460)
N (for average effect)	2,132	2,132	1,786	1,474	1,182	906	740
Events evaluated	34	34	28	23	18	16	9

Notes: Coefficients are from specification in Equation 3 estimated using MSCZ pairs data from the QCEW. k in the row headers refers to years from the event. Column 1 uses the full sample of events, while the other columns restrict the regression to balanced samples based on length of post-period. For example, Column 4 only calculates an effect for 3 post-periods using events that have at least 3 post-periods. Thus, all estimates for any particular column come from the same events. The OWE reported in the last row is the own-wage elasticity for the post-averaged effect. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D6: Dynamic event study estimates for cohorts with different post-periods - BC pairs

	Full sample	Balanced sample - different post-period lengths					
	(1)	(2) ≥ 1	(3) ≥ 2	(4) ≥ 3	(5) ≥ 4	(6) ≥ 5	(7) $= 6$
A. Log earnings							
k = 0	0.016***	0.016***	0.020***	0.020***	0.018***	0.022***	0.024**
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)
k = 1	0.030***		0.030***	0.031***	0.026***	0.034***	0.033***
	(0.008)		(0.008)	(0.008)	(0.009)	(0.009)	(0.012)
k = 2	0.045***			0.045***	0.038***	0.047***	0.048***
	(0.012)			(0.012)	(0.013)	(0.012)	(0.016)
k = 3	0.039***				0.039***	0.053***	0.057***
	(0.012)				(0.012)	(0.011)	(0.015)
k = 4	0.054***					0.054***	0.056***
	(0.011)					(0.011)	(0.015)
k = 5	0.059***						0.059***
	(0.013)						(0.013)
Post-averaged effect	0.027***	0.016***	0.025***	0.032***	0.030***	0.042***	0.046***
	(0.006)	(0.005)	(0.007)	(0.008)	(0.009)	(0.009)	(0.012)
B. Log employment rate							
k = 0	-0.004	-0.004	-0.003	0.002	0.001	0.007	0.009
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.011)
k = 1	-0.012		-0.012	-0.006	-0.005	-0.000	0.007
	(0.010)		(0.010)	(0.008)	(0.009)	(0.010)	(0.012)
k = 2	0.000			0.000	0.002	0.008	0.021
	(0.011)			(0.011)	(0.013)	(0.015)	(0.019)
k = 3	0.002				0.002	0.010	0.017
	(0.016)				(0.016)	(0.018)	(0.025)
k = 4	0.001					0.001	0.012
	(0.022)					(0.022)	(0.030)
k = 5	0.016						0.016
	(0.033)						(0.033)
Post-averaged effect	-0.006	-0.004	-0.007	-0.001	0.000	0.005	0.014
	(0.009)	(0.005)	(0.008)	(0.008)	(0.011)	(0.014)	(0.021)
OWE	-0.235	-0.226	-0.285	-0.042	0.003	0.125	0.293
	(0.331)	(0.330)	(0.294)	(0.238)	(0.357)	(0.329)	(0.445)
N (for average effect)	10,414	10,414	8,784	7,258	5,818	4,402	3,576
Events evaluated	42	42	36	30	25	22	14

Notes: Coefficients are from specification in Equation 3 estimated using BC pairs data from the QCEW. k in the row headers refers to years from the event. Column 1 uses the full sample of events, while the other columns restrict the regression to balanced samples based on length of post-period. For example, Column 4 only calculates an effect for 3 post-periods using events that have at least 3 post-periods. Thus, all estimates for any particular column come from the same events. The OWE reported in the last row is the own-wage elasticity for the post-averaged effect. Standard errors, reported in parentheses, are clustered at the state level. Stars indicate statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Brummund, Peter and Michael R Strain (2020) “Does Employment Respond Differently to Minimum Wage Increases in the Presence of Inflation Indexing?” *Journal of Human Resources*, 55 (3).

Card, David and Alan B. Krueger (1994) “Minimum Wages and Employment: A Case Study of the New Jersey and Pennsylvania Fast Food Industries,” *American Economic Review*, 84 (4), 772–793.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) “The effect of minimum wages on low-wage jobs,” *The Quarterly Journal of Economics*, 134 (3), 1405–1454.

Clemens, Jeffrey and Michael Wither (2019) “The minimum wage and the Great Recession: Evidence of effects on the employment and income trajectories of low-skilled workers,” *Journal of Public Economics*, 170, 53–67.

Dube, Arindrajit, Daniele Girardi, Oscar Jorda, and Alan M Taylor (2025a) “A Local Projections Approach to Difference-in-Differences,” *Journal of Applied Econometrics*.

Dube, Arindrajit and Attila Lindner (2024) “Minimum Wages in the 21st Century,” in *Handbook of Labor Economics, volume 5-6*.

Dube, Arindrajit, Michael Reich, Akash Bhatt, and Denis Sosinskiy (2025b) “Replication Data for: Restaurant Employment, Minimum Wages, and Border Discontinuities,” Harvard Dataverse, https://doi.org/10.7910/DVN/GTOOOC.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M Shapiro (2021) “Visualization, identification, and estimation in the linear panel event-study design,”Technical report, National Bureau of Economic Research.

Sun, Liyang and Sarah Abraham (2021) “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects,” *Journal of Econometrics*, 225 (2), 175–199.

Vaghul, Kavya and Ben Zipperer (2022) "Historical State and Sub-state Minimum Wages," Version 1.4.0, https://github.com/benzipperer/historicalminwage/releases/tag/v1.4.0.