IRLE
Institute for Research on Labor and Employment

IRLE WORKING PAPER
#116-15
September 2015

Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas and Wascher

Sylvia Allegretto, Arindrajit Dube, Michael Reich and Ben Zipperer

Cite as: Sylvia Allegretto, Arindrajit Dube, Michael Reich and Ben Zipperer (2015). "Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas and Wascher". IRLE Working Paper No. 116-15.
http://irle.berkeley.edu/workingpapers/116-15.pdf

Berkeley
UNIVERSITY OF CALIFORNIA

Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas and Wascher

Sylvia Allegretto, Arindrajit Dube, Michael Reich and Ben Zipperer*

September 29, 2015

Abstract

We assess the Neumark, Salas and Wascher (NSW) critique of our minimum wage findings. Recent studies, including one by NSW, obtain small employment elasticities for restaurants, -0.06 or less in magnitude. The substantive critique in NSW thus centers primarily upon teens. Using a longer (1979-2014) sample than used by NSW and in our own previous work, we find clear evidence that teen minimum wage employment elasticities from a two-way fixed-effects panel model are contaminated by negative pre-existing trends. Simply including state-specific linear trends produces small and statistically insignificant estimates (around -0.07); including division-period effects further reduces the estimated magnitudes toward zero. A LASSO-based selection procedure indicates these controls for time-varying heterogeneity are warranted. Including higher order state trends does not alter these findings, contrary to NSW. Consistent with bias in the fixed-effects estimates from time-varying heterogeneity, first difference estimates are small or positive. Small, statistically insignificant, teen employment elasticities (around -0.06) obtain from border discontinuity design with contiguous counties. Contrary to NSW, such counties are more similar to each other than to other counties. Synthetic control studies also indicate small minimum wage elasticities (around -0.04). Nearby states receive significantly more weight in creating synthetic controls, providing further support for using regional controls. Finally, NSW's preferred new matching estimates are plagued by a problematic sample that mixes treatment and control units, obtains poor matches, and shows the largest employment drops in areas with relative minimum wage declines.

1 Introduction

Recent controversies in minimum wage research have centered on how to credibly estimate employment effects. Since the inception of the "new minimum wage" literature in the early 1990s, the source of identifying variation in the United States has largely come from state-level differences in minimum wage policy—either directly, or in interaction with federal policy. As shown in panel A of Figure 1, state minimum wages proliferated substantially over the past three decades. Between the years 1979 and 1985, only one state

*Allegretto: Institute for Research on Labor and Employment, University of California, Berkeley; Dube: Department of Economics, University of Massachusetts Amherst and IZA; Reich: Department of Economics and Institute for Research on Labor and Employment, University of California, Berkeley; Zipperer: Washington Center for Equitable Growth. We are grateful to Doruk Cengiz, Zachary Goldman, Carl Nadler, Thomas Peake and Luke Reidenbach for excellent research assistance. Financial support for this paper came entirely from the University of California, Berkeley and the University of Massachusetts Amherst.

(Alaska) had a minimum wage exceeding the federal standard. Subsequently, as the federal wage was unchanged for extended periods of time, states stepped in. The number of states exceeding the federal reached local peaks of 17 in 1990, 13 in 1997 and 33 in 2008, and stood at 22 in 2014.

On the one hand, the extensive state-level variation makes the U.S. an attractive laboratory for studying the effects of the minimum wage. On the other other hand, this distribution of minimum wage policies has been far from random. If we divide the states into two equally sized groups—"high" versus "low" groups based on their average real minimum wages over the 1979-2014 period—we find that minimum wage policies are highly spatially clustered. (See panel B of Figure 1).[1] High minimum wage states are concentrated on the west coast, the northeast and parts of the midwest. As a casual glance at panel B of Figure 1 suggests, high minimum wage states also tend to be Democratic-leaning. For example, 92 percent of high minimum wage states voted for Barack Obama in the 2012 presidential election, as compared to 15 percent of low minimum wage states. High minimum wage states also have unionization rates that are nearly twice as high, and they experienced proportionately smaller reductions in these rates over the past three decades. These differences raise the possibility that trends in other policies and economic fundamentals may also differ between these groups of states.

The non-random distribution of state minimum wage policies in the U.S. poses a serious challenge to the canonical two-way fixed effects approach, which relies on the assumption of parallel trends across all states. Specifically, that model assumes all such heterogeneity can be explicitly controlled by using common time effects and time-invariant state effects, plus a small number of controls which typically include the overall unemployment rate. However, the political economic differences and regional clustering of high and low minimum wage states suggest that the two-way fixed effects model may mis-estimate the counterfactual employment levels absent a minimum wage increase.

To account for such heterogeneity, our past minimum wage research—Dube, Lester and Reich (2010), hereafter DLR (2010), and Allegretto, Dube and Reich (2011), hereafter ADR—has used either border discontinuities or coarser regional and parametric trend controls, as nearby areas tend to experience similar shocks. When using such strategies, the estimated employment impact for highly affected groups such as restaurant workers or teens tends to be small, and often statistically indistinguishable from zero, even though there are sizable earnings effects for these groups. Moreover, these employment results stood in sharp contrast to those from the two-way fixed effects model, which typically suggested more substantial disemployment effects. Importantly, DLR (2010) and ADR also used distributed lags and leads in minimum

[1] If we consider the *change* in the value of the real minimum wage since since 1979q1 and average that value across all subsequent periods, the categorization of "high" and "low" groups is similar: 43 out of 51 states have the same high/low categorization using the two approaches. The patterns documented in panel B of Figure 1 are therefore closely related to the variation used to estimate the effect of minimum wage policies in panel models.

wages to show that the disemployment effects estimated in the two-way fixed-effects model often reflected pre-existing trends rather than changes in employment that occurred after the policy was implemented. This evidence directly contradicted the "parallel trends" assumption made by the two-way fixed-effects model.[2]

In two papers, Neumark, Salas and Wascher (NSW 2014a, 2014b) critique the use of local area controls. From our perspective, they make **three** important claims.

First, they defend the results from the two-way fixed effects estimator, arguing against the evidence that pre-existing trends contaminate those estimates. Using quarterly leads and lags over a 25 quarter window around minimum wage changes, they claim that there is "no evidence of a large accumulated negative effect in the period up to the minimum wage increase" for teen employment when using the two-way fixed effects model. They argue that at the quarterly frequency, the coefficients for the leading terms are sometimes positive, even though they acknowledge that they are "more negative than positive." They also argue that the inclusion of controls for spatial heterogeneity does not produce smaller pre-existing trends.

Second, they argue that the use of local area controls throws away too much useful information. Regarding our specific controls, NSW argue that limiting the identifying variation to be within census divisions, or within bordering areas, is unwarranted. They reach this conclusion primarily by assessing whether their synthetic control method puts more weight on nearby areas. While they do not argue against the use of state-specific trends *per se*, they claim that the small magnitudes of the employment estimates in ADR from specifications with state-specific linear trends are driven by an "endpoint bias" owing to the presence of recessions in the beginning and end of the sample. They also argue that estimates for models that include third, fourth, or fifth order polynomial time trends by state suggest sizable disemployment effects, with elasticities exceeding -0.180 in magnitude.

Third, NSW propose a new "matching" estimator loosely based on the synthetic control approach. They argue that this matching estimator suggests substantial employment effects, at least for teens. They claim that this "data driven" approach provides a superior alternative to methods we have proposed to account for time-varying confounders of minimum wage policies.

We respond to each of these claims. We begin by noting that of the two groups discussed in this exchange (restaurant workers and teens), a substantive disagreement remains mainly for teens. NSW's preferred matching estimators suggest a small impact on restaurant employment, with elasticities no more than -0.063 in magnitude, and smaller than the employment estimates from the two-way fixed effects model. Indeed, the preferred estimates of employment effects for restaurants are similarly small in most of the recent literature, spanning a wide array of methods and samples, including NSW (2014a), Addison, Blackburn and

[2] Other minimum wage researchers—e.g., Aaronson, French and Sorkin (2015), Magruder (2013) and Huang, Loungani and Wang (2014)—have subsequently used the border discontinuity design to estimate causal effects of minimum wage policies in both U.S. and international contexts.

Cotti (2014), DLR (2010, 2015), and Totty (2015). In contrast, the earnings effects for restaurants are sizable across most specifications and samples. To move the discussion forward, we focus most of our attention on teens.[3]

In this paper, we use 36 years of Current Population Survey (CPS) basic monthly data from 1979 through 2014 to estimate the impact of minimum wages on teen employment. We find unmistakable evidence of pre-existing trends that contaminate the two-way fixed effects model. A sizable part of the -0.219 employment elasticity estimate from the two-way fixed effects model accrues prior to actual increases in minimum wages. We show that this bias is visible even in the estimates in NSW (2014b), which are based on the 1990-2011 period. In the expanded sample, just the inclusion of state-specific trends dramatically reduces the magnitude of the effect in the full sample (to -0.065), and renders the estimates statistically insignificant. Accounting for regional variation—by allowing for time effects to differ by the nine census divisions—also reduces the magnitude of the effect (-0.130). Together, these two sets of controls produce an estimate of 0.006. None of these three estimates are statistically significantly different from zero, in contrast to the two-way fixed effects model. Moreover, the estimates from specifications with more controls are not generally less precise than the estimates from the two-way fixed effects model.

By using a substantially longer sample, we are also able to assess several claims made by NSW. We show that it is not the "end points" in the original 1990-2009 sample of ADR that produce more positive estimates in models with state-specific trends, contradicting a central claim by NSW (2014b). We progressively expand the sample and exclude downturn periods. Neither produces clear evidence of disemployment in those models. We also show that in the full sample, adding higher order state-specific polynomial trends makes little difference: estimates including higher order trends continue to suggest small effects on employment. Therefore, NSW's finding of a substantial negative effect using third, fourth, or fifth order polynomials appears to be driven by greater imprecision of the higher order trends estimated with a shorter panels. We demonstrate as much by progressively expanding the sample.[4] Importantly, the employment estimates are also very small in magnitude when we use first-differences instead of deviations-from-means to estimate the two-way fixed effects model (allowing up to three years of lags in the minimum wage). This result provides yet another demonstration of a bias resulting from time-varying state effect that is correlated with minimum wages. The negative employment estimates from the deviations-from-means variant of the two-way fixed effects model appears to derive from a comparison of post-treatment employment to a baseline substantially

[3]While we focus on teens, we also do respond in this paper to the key criticisms made by NSW (2014a, 2014b) of DLR (2010), which focused on restaurants. In section 4.1, we discuss the validity of the border discontinuity design by considering if contiguous counties are indeed more similar. In section 6, we provide new results using updated 1990-2014 QCEW data. Finally, in online Appendix C, we respond to the claim in NSW (2014a, 2014b) that the falsification test in DLR (2010) that used a spatially correlated "placebo" law was invalid.

[4]ADR noted that estimation of parametric trends may be difficult with short panels. Estimation of higher order trends further amplifies the problem, as we discuss in section 2.2 and Appendix B.

far back in the past.

We also present results using a new data-driven approach to choosing controls: the double-selection post-LASSO estimator advocated by Belloni, Chernozhukov and Hansen (2014). This approach optimally chooses the set of controls (beyond the basic two way fixed effects) using sparsity as a criterion—without pre-selecting any such controls for inclusion or exclusion. We find that the model chosen by the data produces employment estimates that are close to zero (-0.013). Importantly, a handful of states drive the negative effect: accounting for state-specific trends for as few as five (mostly coastal) states is sufficient to reduce the magnitude of the estimated effect to zero. Moreover, the LASSO-based criteria never picks any higher-order polynomial time trends. All of these conclusions hold whether we consider the full 1979-2014 sample or a more recent sample (1990-2014). Again, the finding of a small employment effect in more saturated models is not driven by "throwing away too much information." A very sparse set of controls chosen by the data delivers the same verdict.

These state-level CPS findings for teens are corroborated with a more fine-grained research design that uses contiguous counties straddling state borders. We review evidence from DLR (2015), who use the county-level Quarterly Workforce Indicators (QWI) dataset and find a small minimum wage elasticity for teen employment of -0.059, statistically indistinguishable from zero. And in contrast to NSW's claim, DLR (2015) finds that—on average—neighboring counties are indeed more similar in levels and trends of covariates than are counties that are farther away. Moreover, they find clear evidence of reduced teen turnover (i.e., hires plus separations) in response to the minimum wage changes. Along with the earnings effects, the turnover findings contradict the notion that the border discontinuity design is unable to detect effects because it discards too much information. Finally, state-level evidence using the QWI also shows that including state-specific trends renders the minimum wage elasticity for teen employment indistinguishable from zero (Gittings and Schmutte 2015).

As an alternative to using spatial controls, we also review evidence using the synthetic control approach presented in Dube and Zipperer (2015). This approach chooses a weighted average of potential control ("donor") states to match pre-intervention outcomes in the treated state and its synthetic control. They pool all state-level minimum wage changes between 1979 and 2014 with at least 3 years of pre- and 1 year of post-intervention data. The pooled estimate using the resulting 29 events shows a sizable teen wage effect, but an pooled employment elasticity of -0.036.

We also assess the evidence from the the NSW (2014a) matching estimator, which is loosely based on the synthetic control method. We find that NSW's estimator contains a number of serious problems. Most importantly, their sample of events is flawed because both treatment and control units experience sizable minimum wage changes, making the treatment/control distinction nearly meaningless. Since the synthetic control approach requires a clean pre-intervention period and untreated donors to estimate the donor weights, the violation of these assumptions makes their estimated donor weights unreliable. Finally, they use a very short pre-intervention window (4 quarters) to calculate synthetic control matches, which raises questions about the quality of their matches. Re-analyzing their data, we find that most of the events they study contain very small net minimum wage increases in the "treatment" states, as compared to their synthetic controls. While employment does fall in the "treated" states as compared to the control ones, it does so mostly in states in which the minimum wage differential between treatment and control units did not actually increase. For the subset of the events with a proper minimum wage treatment, there was no indication of sizable employment loss. To further assess the quality of matches obtained in the NSW sample, we consider the sensitivity of the estimates to using a slightly earlier pre-treatment period. With this earlier pre-treatment period, we find that the employment estimates switch signs and become positive, suggesting that their synthetic controls did not track the treated states very well. While we consider the use of synthetic controls to study minimum wages to be a useful strategy, the results from the NSW matching estimator are not reliable.

Although most of our paper focuses on teens, section 6 presents new evidence on restaurant employment, for which our substantive disagreements with NSW's estimates are far more muted. Using updated 1990-2014 QCEW data, we confirm the existence of pre-existing trends in the two-way fixed effects estimates for restaurant employment, and the lack of such pre-existing trends with the border-discontinuity design. We show that the medium-run (3 years) and long-run (4 or more years) estimates for restaurants using the border-discontinuity design suggest employment estimates that are small—less than -0.1 in magnitude; and while the long-run estimates are not very precise, medium run estimates are reasonably so.

Overall, NSW's critiques of our work and their proposed estimators do not withstand scrutiny. In the U.S. data over the past three decades, minimum wage effects estimated using the two-way fixed effects model favored by NSW appear to be biased toward finding a negative impact on low-wage employment. A wide variety of approaches used by us as well as by other researchers find that adjusting for such time-varying heterogeneity leads to employment elasticity estimates that are much smaller than those from the two-way fixed effects model; indeed, our estimates are often not very different from zero.

The rest of the paper is structured as follows. Section 2 presents our core results on teen employment using CPS data, including controls for time-varying heterogeneity, model selection using LASSO, and testing for pre-existing trends. Section 3 presents evidence on teen employment using a border discontinuity design, drawing from DLR (2015). Section 4 presents evidence from pooling synthetic control estimates from 29 state-level minimum wage changes, drawing from Dube and Zipperer (2015). Section 5 assesses the NSW matching estimator. Section 6 discusses the evidence on restaurant employment. Section 7 concludes.

2 Effects on teen employment: CPS data using state-level variation

Teens have been extensively studied in the minimum wage literature because they are heavily affected by minimum wage policies. Based on the Current Population Survey Outgoing Rotations Group (CPS ORG) data, during the 1979-2014 period, 40.2 percent of working teens earned within 10 percent of the statutory minimum wage (higher of state or federal), as compared to 7.7 percent of workers overall. The relatively large proportion of minimum wage workers among teens makes it relatively easy to detect an effect of the policy on outcomes for this group, thus making them an attractive group to study.

At the same time, the lessons from teens may be limited, for several reasons. First, for an understanding of the impact of the policy more generally, teens are not representative of all minimum wage workers. Second, teens comprise a shrinking share of low-wage workers. Among workers earning within 10 percent of the statutory minimum wage, the teen share has fallen over time, from 32.2 percent in 1979 to 22.7 percent in 2014.[5] Finally, labor-labor substitution may imply that some of the teen disemployment effects represent employment gains by other groups.

Nonetheless, the high incidence of minimum wage work among teens suggests that if one is to find disemployment effects of the policy, it will likely be for teens. Therefore, the debate on teen employment still has relevance today.

In this section, we estimate teen employment and wage elasticities of the minimum wage using individual-level CPS data from 1979 through 2014. We begin with a description of the sample and variable definition, and then develop our main empirical specifications. For teen employment, we use individual-level records from the Unicon extracts of the full basic monthly sample (`https://www.unicon.com/cps.html`), and for wage outcomes we use the NBER Merged Outgoing Rotation Groups (ORG) (`http://www.nber.org/morg/`).[6] All individual-level regressions are weighted by the basic monthly sample weights or earnings sample weights. Our primary sample ends in 2014, the most recent complete year of data, and begins in 1979 because that is the first year of the ORG data containing earnings outcomes.

Teens are observations aged 16 through 19. In regression specifications with individual-level data, covariates include the overall state quarterly unemployment rate, the quarterly teen share of the population and dummies for sex, age, marital status, race, and Hispanic ethnicity. We define race as white,

[5] The teen share is calculated for all workers (hourly or otherwise) with positive hourly earnings that are not imputed in the CPS ORG data.

[6] A previous (2013) version of this paper used only state-level aggregations of the ORG sub-sample of the CPS, for the period beginning in 1990. While the ORG sub-sample is necessary to estimate wage outcomes, we use the much larger basic monthly sample for employment outcomes. Our preferred sample in this current paper uses individual-level data to avoid any aggregation issues. However, for computational feasibility we use the aggregated data in the LASSO specifications in Section 3.2.

black, or other, and interact these dummies and an indicator for Hispanic ethnicity with an indicator for period 2003 and later, as there was a large race and ethnicity classification change in the CPS after 2002. We calculate quarterly teen shares of the 16 and over population using the full basic monthly sample. We use as the quarterly state unemployment rate the quarterly mean of the non-seasonally-adjusted monthly unemployment rate from the Bureau of Labor Statistics Local Area Unemployment series (`http://download.bls.gov/pub/time.series/la/`). We define wages as the reported hourly wage for workers paid hourly wages or, for other workers, usual weekly earnings divided by usual weekly hours worked. When estimating wage effects, we exclude from our sample all observations with wages imputed by the BLS, following the guidance of Hirsch and Schumacher (2004) to define allocated earnings.[7] State-level minimum wages are quarterly means of monthly state-level minimum wage levels, or federal minima when they exceed the state law, for all fifty states and the District of Columbia. Monthly state minimum wage data are from Allegretto and Nadler (2015) for 1984 through 2014, and from Autor, Manning, and Smith (2015) for 1979 through 1983.

We begin with estimating a canonical model with time (t) and place (j) fixed effects. Here i denotes an individual, while j denotes the state of residence of individual i:

$$Y_{it} = \alpha + \beta MW_{jt} + \mathbf{X}_{it}\Lambda + \gamma_j + \delta_t + \nu_{it} \tag{1}$$

The key independent variable is the log of minimum wage (MW_{jt}), which takes on the higher of the federal minimum wage or the minimum wage in state j, while $\mathbf{X_{it}}$ is a vector of controls. The dependent variable Y_{it} is either the log of hourly earnings, or a dummy for whether person i is currently working. Since the hourly earnings variable is available only for those in the outgoing rotation groups, those regressions are estimated using the CPS ORG data. Moreover, we discard all observations with imputed wage data. The employment regressions are estimated using the full basic monthly CPS samples. The vector of covariates X_{it} includes dummies for gender, race/Hispanic origin, age, and marital status; the teen share of the population in the state; and the non-seasonally adjusted quarterly state unemployment rate.[8] We report all the results as elasticities: for earnings equations, the elasticity is simply the coefficient estimate of β, and for employment equations, we divide this coefficient by the weighted sample mean of employment.

In our most saturated specification, we additionally include (up to fifth order) state-specific time trends,

[7] We define wage imputations as records with positive allocation values for hourly wages (for hourly workers) and weekly earnings or hours (for other workers) during 1979-1988 and September 1995-2014. For 1989-1993, we define imputations as observations with missing or zero "unedited" earnings but positive "edited" earnings (which we also do for hours worked and hourly wages). We do not label any observations as having imputed wages during 1994-August 1995, when there are no BLS allocation values for earnings or wages.

[8] In the prior 2013 version of this paper, we used educational attainment as a covariate. We omit education dummies in this paper because the minimum wage may influence schooling decisions and because the CPS changed education classifications during the 1979-2014 period.

and also allow the time effects to vary by each of the nine census divisions, denoted by d:

$$Y_{it} = \alpha + \beta MW_{jt} + \mathbf{X}_{jt}\Lambda + \gamma_j + \delta_{dt} + \sum_k \left(\phi_{jk} \times t^k\right) + \nu_{it} \tag{2}$$

We report the intermediate specifications with just the state-specific trends and the division-period effects as well as the most saturated specification. Altogether, these twelve specifications—with common or division-period fixed effects, and with polynomial trends of degree $k = 0, ..., 5$— include the four key specifications used in ADR, which only used linear and not higher order trends. Three of these specifications—those with linear trends and/or division-period effects—are the ones criticized by NSW (2014a, 2014b).

2.1 Main results for teens

In the panel A of Table 1 we report the wage results from the sample of teens with earnings in the individual-level CPS ORG data from 1979-2014. The outcome variable here is the natural log of the hourly wage. All regressions include state fixed effects. The first row includes common time effects, while the second row includes time effects that vary by the nine census divisions. Column 1 contains no allowance for state-specific trends, while columns 2 through 6 add state-specific polynomial trends of successively higher orders. We find that the estimated wage effects are always economically substantial and statistically highly significant. This result holds across the twelve specifications. The wage elasticities are remarkably uniform, ranging between 0.226 and 0.271 for the common time specification and between 0.211 and 0.253 when including division-period effects. The addition of division-period effects or higher-order trends does not substantially diminish these estimates.

In panel B of Table 1, we report analogous results for teen employment using the full basic monthly CPS. Importantly, the employment elasticity is substantial and negative only in the specifications without any state-specific trend controls. Simply including state-specific linear trends reduces the common-time specification estimate in magnitude from -0.219 to -0.065 and renders it statistically insignificant. The finding in ADR (2011) that including state-specific trends diminishes the magnitude of the estimated employment effect is replicated in this expanded sample, whose end points (1979, 2014) are notably not recessionary years. The replication of the results in the expanded sample refutes NSW's key argument that the findings in ADR were driven by "endpoint bias" in the estimation of state trends owing to the the presence of recessionary years.

In Appendix B, we provide additional evidence that the "endpoint bias" explanation is incorrect. To summarize those findings, Appendix Figure B1 shows estimates from 72 different samples with alternative starting and ending dates varying between 1979 and 1990, and 2009 and 2014, respectively, for specifications with and without state-specific linear trends. Extending the sample by considering end points away from recessionary periods does not produce more negative estimates when state trends are included. Moreover, we also show in Appendix B that excluding downturns—either using the official NBER definition or a much more expansive one—does not produce evidence of substantial disemployment effects in models with state trends.

Continuing with the common time effect models in the first row of Table 1, panel B, when we include state-specific trends of higher order, the coefficients are always smaller than -0.1 in magnitude and are not statistically significant. Four out of five of these estimates are less than -0.07 in magnitude. These results refute the claim in NSW that inclusion of higher order (third or greater) state-specific trends restores the finding of a sizable negative effect. Estimation of cubic, quartic or quintic trends by state places greater demand upon the data, especially when the panel is short. By using a substantially longer panel, we are able to estimate these trends more reliably. We find that the estimates from including 3rd and 5th order polynomials, -0.066 and -0.068, respectively are virtually identical to the estimate with just a linear trend (-0.065). The estimate from the 2nd order trend is slightly smaller in magnitude (-0.044) while the estimate from the 4th order trend is slightly larger in magnitude (-0.091). However, in all cases, the estimates are under -0.1 in magnitude and never statistically significant. Overall, these results suggest that including higher order trends are unlikely to change the conclusions reached in ADR. We provide additional evidence on the suitability and reliability of higher order trends below in section 2.2.

The bottom section of panel B, Table 1 additionally includes division-period effects, isolating the identifying variation to within the nine census divisions. Including division-period effects typically produces estimates that are even less negative. For example, without any state trends (column 1) the estimate falls from -0.219 to -0.130 in magnitude, and is marginally significant. However, inclusion of state trends renders the estimates close to zero and not statistically significant, with point estimates ranging between -0.040 and 0.006.

To sum up to this point, of these twelve specifications, only two produce significantly negative employment estimates, and both of these lack any state-specific trends. In contrast to claims in NSW, the attenuation of the estimate from inclusion of linear trends is not driven by presence of recessions in the end points of the sample, nor by exclusion of higher order polynomial trends by state.

2.2 Model selection using LASSO

The results from Table 1 provide unambiguous evidence that the inclusion of controls for time-varying heterogeneity substantially reduces the magnitude of minimum wage effects on employment. Of the 12 specifications, only the two specifications without any state-specific trends produce estimates that are either substantial (exceed -0.1 in magnitude) or statistically significant. The other specifications tend to generate estimates that are smaller in magnitude. Moreover, including the division-period effects tends to produce estimates that are less negative. This variation raises a fundamental question: what is the best set of controls to include in these regressions?

In this section, we address this question by applying the double-selection post-LASSO approach advocated by Belloni, Chernozhukov and Hansen (2014). This method uses sparsity as a criterion for covariate selection. The LASSO regression is able to identify a small set of key confounders from a large set of potential covariates, assuming such a sparse representation is feasible. The central innovation of the LASSO method is to augment the mean squared error objective function with an additional penalty term that is a weighted sum of the absolute value of the regression coefficients. The resulting minimization typically zeroes out many of the coefficients, leading to a small set of the most important predictors. The double-selection criteria applies the LASSO to a program evaluation context, in which the LASSO is used to select covariates that either predict the outcome (in our case teen employment), or the treatment (log minimum wage). After having selected the covariates using these two LASSO regressions, Belloni et al. suggest running a simple OLS regression of the outcome on the treatment and the double-selected set of controls (hence the term “post-LASSO”).[9]

Computational challenges in estimating LASSO with a large number of observations and variables require us to use data aggregated at the state-quarter level. As a first step, we first estimate all the specifications in Table 1 using aggregated data. These regressions are similar to those estimated in NSW (2014a, 2014b). We regress the log of the teen employment-to-population ratio on the log of the minimum wage, the state unemployment rate and the teen share of population, while additionally controlling for state fixed effects, either common (or division-specific) period effects, and possible state-specific time trends. We also include demographic group shares analogous to covariates in the individual-level regressions: shares by gender, age groups, race categories, and marital status. We additionally weight all regressions by the size of the teen population. These results are reported in Table 2, panel A, which shows that in most cases aggregation does not make much of a difference.[10] Of the 12 estimates from columns 1-6 in Table 2 (with and without division-period effects, and up to fifth order state polynomial trends), only the two-way fixed effects model produces an elasticity that is substantial and statistically significant. All the other 11 coefficients are under -0.09 in magnitude and are not statistically significant.

[9]This post-LASSO approach leverages the advantages of LASSO-based selection of the most important controls, while guarding against the “shrinkage bias” in LASSO coefficients due to the penalization term.

[10]One notable difference between Tables 1 and 2 occur in the specification with division-period dummies and no state trends: -0.130 and marginally significant with the micro-data with individual-level covariates, and -0.043 and not significant with the aggregated data. Since the regressions with micro-data control for individual-level covariates, the estimates in Table 1 constitute our preferred set.

For model selection, we estimate two LASSO regressions of the log of teen EPOP and the log minimum wage over a set of covariates: the unemployment rate, teen share of population, demographic group shares as specified above, division-period dummies, and state-specific time trends of orders 1 through 5. The LASSO regressions always automatically account for state and time fixed effects by partialing them out prior to the LASSO estimation. With the superset of controls chosen by these two LASSO regressions, we estimate an OLS regression—which also always includes state and time fixed effects.[11]

In column 8 of Table 2, we report the estimates from our double-selection post-LASSO regression allowing the full set of controls. First, we note (although not shown in the table) that with the default, recommended penalization parameter ($\lambda = 940$)[12], the double-selection criteria for teen employment picks division-period effects from one census division (the Pacific division), 29 state-specific linear trends, and no higher order trends. The resulting point estimate (-0.012) is numerically close to, and statistically indistinguishable from zero. The results from this exercise confirm that the controls for time-varying heterogeneity used in ADR—especially state trends—should be included, and that the data-driven set of controls suggests a minimum wage elasticity for teen employment that is close to zero.[13]

The estimates in the top panel of Table 2 are based on a penalization parameter λ that is chosen optimally, using the default plug-in method. To assess how inclusion of the most important controls (as deemed by the double-selection criteria) affects the minimum wage estimate, we also vary λ to go from the most saturated specification to the simple two-way fixed effects model. Since no higher order tends were picked by the LASSO-based criteria using the default λ, for this exercise we limit ourselves to linear trends only, which also eases the computational burden. The double-selection post-LASSO estimate that just allows for linear trends is essentially the same as when allowing up to fifth order trends: it is 0.013 instead of -0.012, as shown in column 7 of Table 2, with the small difference stemming from the a slightly smaller value of the optimal λ when the maximum number of controls is larger ($\lambda = 934$). Appendix Table A1 shows the point estimates and the confidence intervals associated with varying λ between 0 (the most saturated model) and 3500 (which only picks the state unemployment rate as a control beyond the manually-specified two-way fixed effects). The point estimate quickly falls under -0.045 in magnitude as λ is lowered to 2,000 or below. As also shown

[11] We adapted the STATA code for the post-LASSO regressions from Christian Hansen's web page: `http://faculty.chicagobooth.edu/christian.hansen/research/JEPStata.zip`, including the `lassoshooting.ado` file which estimates the LASSO regressions. To account for the fact that our OLS regressions using aggregated data weight the regressions by teen population, we pre-multiplied the data by the square root of teen population prior to estimating the LASSO regressions. Results using unweighted version of the double-selection post-LASSO were quantitatively similar. The unweighted post-LASSO elasticity estimate for the 1979-2014 sample is -0.016, as compared to the weighted elasticity estimate of -0.012 reported in Table 2. In `lassoshooting.ado`, we include state and time fixed effects in the "controls( )" option, which partials out these variables prior to estimating the LASSO regressions.

[12] The default level for the penalization parameter λ in the Belloni et al. program `lassoshooting.ado` is set $\lambda = 2.2 \times \sqrt{N} \times \sqrt{2 \ln\left(\frac{2p}{0.1/\ln(N)}\right)}$, where p is the number of covariates and N is the sample size.

[13] Although we do not report the results in tables, the wage estimates from the double-selection post-LASSO regression (0.229) are quantitatively close to the estimates in Table 1.

in Appendix Table A1, for $\lambda = 2000$, the LASSO double-selection procedure includes just 5 state-specific linear trends lowers the elasticity in magnitude to -0.040. In other words, merely adding state-specific linear trends for these 5 states (which happen to be CA, SD, OR, WA and VT) to the fixed effects model produces an estimate that is close to zero, and not statistically significant.[14] We stress that this highly sparse model, which adds only five controls for unobserved heterogeneity beyond the canonical two-way fixed effects model, nonetheless delivers the same qualitative finding as in ADR. This result contradicts the suggestion of NSW that the findings in ADR were driven by "throwing away too much information."[15]

We have seen that in the longer sample, the magnitude of the point estimates is not very sensitive to the inclusion of higher order state-specific trends, unlike the findings in NSW (2014a). For comparability to the results in NSW (2014a), we also report in the bottom panel of Table 2 the double-selection post-LASSO estimates for the sample restricted to 1990 and later. The estimates across specifications in this shorter sample exhibit greater variation. For example, similar to NSW, the specifications with third or higher order state-specific trends (but without division-period effects) exhibit sizable and statistically significant negative elasticities. Here, too, however, the double-selection post-LASSO estimate is small in magnitude (-0.031) and not statistically distinguishable from zero. The estimate for this shorter sample is based on 20 state-specific linear trends; importantly, as before, no non-linear trends are picked. Therefore, while the shorter sample produces more varied estimates using OLS and alternative trend specifications—likely due to imprecision of estimating many higher order trends—a data-driven choice of predictors that considers higher order trends produces an estimate that is close to zero in this sample as well.

We provide additional evidence and discussion of the unreliability of estimates with higher order trends in short panels in Appendix B. There we discuss the evidence from 72 sample periods, varying the start dates between 1979 and 1990, and end dates between 2009 and 2011. To summarize, Figure B2 shows that, in general, employment estimates are much more sensitive to the order of the polynomial for state-specific trends in samples with fewer years. When we expand the sample progressively, the variability diminishes, and the employment estimates from specifications with any type of state-specific trend become much smaller in magnitude than the two-way fixed-effects estimate.

Overall, model selection techniques that make no prior assumptions about which controls should be included in a regression confirm our approach of including controls for time-varying heterogeneity and support our original conclusion about the size of the minimum wage elasticity for teen employment.

[14]Four of the five states are coastal, showing the importance of obtaining a valid counterfactual for the high minimum wage Pacific division. When estimating state-specific trends, the omitted state is Alabama.

[15]Christian Hansen's 2013 NBER Econometric Lecture reports 5 possible asymptotically equivalent calculations for λ, which, in our case of $p = 1207, N = 7344$, range between 12.562 and 1161.99. As shown in Appendix Table A1, this implies a range of double-selection post-LASSO estimates for the minimum wage elasticity between -0.023 and 0.055.

2.3 Timing of the employment effects

To assess the validity of a research design, it is customary to consider the timing of the putative effects from treatment. Estimates from a given research design are less credible if the effects appear to occur substantially prior to treatment—such a pattern indicates the likelihood of contamination from pre-existing trends. In the minimum wage context, Krueger (1994) tested for pre-existing employment trends and revealed the shortcomings of time series evidence for Puerto Rico, which suggested that the island's minimum wage implausibly reduced employment two years prior to an actual increase in the minimum. Showing the absence of leading effects can also help validate a research design, as in the Autor (2003) difference-in-differences study of the effects of state employment protections on temporary employment.

The leading effects falsification test is particularly relevant for studies of U.S. minimum wages, where the policy variation is not uniformly distributed across states. In prior work (DLR 2010, ADR) we used a distributed lag model to demonstrate that pre-existing trends contaminate the estimates of the conventional two-way fixed effects model, which often exhibits sizable and statistically significant leading effects. Nonetheless, NSW (2014b) raise questions about our findings on pre-existing trends for teen employment. First, they argue that pre-existing trends are not clearly indicated in the two-way fixed effects model. Second, they argue that even after differencing out the leading effects, the subsequent cumulative effects remain negative, sizable and comparable to the static estimates. Third, they argue that the inclusion of controls for spatial heterogeneity does not produce better results, in the sense of passing the leading effects falsification test. Here we assess all of these claims.

To shed light on this disagreement, we use exactly the same distributed lag structure as in NSW (2014b). That is, we add to our prior static specifications in equations (1) and (2) twelve quarters of leading and twelve quarters of lagged minimum wages. We estimate these regressions using the individual-level CPS data and control sets we used before for teens in the 1979-2014 period using four specifications. Beginning with the two-way fixed effects model

$$Y_{it} = \alpha + \sum_{k=-12}^{12} \beta_k MW_{j,t-k} + \mathbf{X}_{it}\Lambda + \gamma_j + \delta_t + \nu_{it} \tag{3}$$

we increasingly saturate the model to include state-specific linear time trends and division-period fixed effects

$$Y_{it} = \alpha + \sum_{k=-12}^{12} \beta_k MW_{j,t-k} + \mathbf{X}_{it}\Lambda + \gamma_j + \delta_{dt} + \phi_j \times t + \nu_{it}. \tag{4}$$

We also report estimates from the two intermediate specifications—with just divison-time fixed effects and state-specific linear trends. We calculate the cumulative employment response from these four models by summing the coefficients for individual leads and lags, and convert them to elasticities by dividing by the sample mean of teen employment rate: therefore, the cumulative response elasticity at event time τ is calculated as $\rho_\tau = \sum_{k=-12}^{\tau} \eta_k = \frac{1}{\bar{Y}} \sum_{k=-12}^{\tau} \beta_k$. Note that these cumulative responses are from a default baseline of $\tau < -12$; we will consider alternative baselines below by subtracting out leading coefficients from the cumulative responses. We estimate these models with the full 1979-2014 sample and individual level data—these are our preferred estimates. But we also compare these results with the aggregated 1990-2011q1 data from NSW (2014b) to shed light on the disagreement. Below, we begin by assessing the performance of the two-way fixed effects model in terms of the presence of leading effects. Subsequently we turn to the performance of models with greater controls for time-varying heterogeneity.

Performance of the two-way fixed effects model

The top left graph in Figure 2 plots these cumulative responses from the two-way fixed effects model, along with 95 percent confidence bands for the full sample. First, although somewhat noisy, there is a clear visual pattern: every pre-treatment point estimate for the two-way fixed effects model is negative, and 5 of the 12 coefficients are statistically significant at the 5 percent level. To reduce noise and more easily extract a signal from the data, columns 2 and 4 of Figure 2 show four-quarter averages of these quarterly cumulative response elasticities: $\overline{\rho}_{[\tau,\tau+3]} = \frac{1}{4} \sum_{m=0}^{3} \rho_{\tau+m}$, along with the 95 percent confidence bands. These averaged cumulative response elasticities and standard errors are also reported in the first column of Table 3. For the two-way fixed effects model, the four-quarter averages of the leading cumulative response elasticity $\overline{\rho}_{[-12,-9]}$ is -0.145, and is statistically significant at the five percent level (row A of Table 3). In other words, during the third year prior to the minimum wage increase, the magnitude of the average cumulative response elasticity is implausibly large, and roughly two-thirds the size of the static employment elasticity of -0.219 (see Table 1). The average cumulative response elasticities during the second and the first year preceding the minimum wage increase ($\overline{\rho}_{[-8,-5]}$ and $\overline{\rho}_{[-4,-1]}$) are even more negative, -0.187 and -0.202, respectively; both are statistically significant at the 5 percent level. We find that in the full sample of 1979-2014, there is unmistakable evidence that the two-way fixed-effects model fails the falsification test that leading coefficients during 1, 2 or 3 years prior to treatment are zero.

Second, and relatedly, we find that a sizable portion of the two-way fixed effects estimate accrues prior to treatment. In Table 3, we calculate estimates for 3 and 4+ year effects from the policy. For the "medium term" or 3 year estimates, we begin by calculating the average cumulative response elasticity in the third year following the minimum wage increase $\overline{\rho}_{[8,11]}$, and subtracting from this the baseline value. We use three different baselines: the average cumulative response in the first, second, or third year preceding the increase, i.e., $\overline{\rho}_{[-4,-1]}$, $\overline{\rho}_{[-8,-5]}$, or $\overline{\rho}_{[-12,-9]}$, respectively. For example, using the first year before treatment as the baseline, the 3-year estimate is: $\overline{\rho}_{[8,11]} - \overline{\rho}_{[-4,-1]}$. We also construct 4+ year or "long term" estimates as $\rho_{12} - \overline{\rho}_{baseline}$, where the baseline can again be $\overline{\rho}_{[-4,-1]}$, $\overline{\rho}_{[-8,-5]}$, or $\overline{\rho}_{[-12,-9]}$.[16]

The 3 and 4+ year estimates for the fixed effects model are reported in Column 1 of Table 3, in rows labeled to clarify how the estimates are being calculated. For example, F-C subtracts the value in row C ($\overline{\rho}_{[-4,-1]}$) from the value in row F ($\overline{\rho}_{[8,11]}$). Overall, these results show that for the two-way fixed-effects model, either 3 or 4+ year estimates are substantially smaller than the estimate from the static specification. While the static estimate from Table 1 is -0.219, the 3 year and 4+ year estimates range between -0.11 and -0.13 when using $\tau \in [-4, -1]$ or $\tau \in [-8, -5]$ averages as baselines. Although these estimates are statistically significant, there is a 40-50 percent reduction in the effect size, as compared to the static estimate, which implicitly uses a mixture of earlier and later baselines over $\tau < 0$. Using an earlier baseline ($\tau \in [-12, -9]$) produces 3 and 4 year estimates around -0.17 (rows F-A and G-A), while using an even earlier baseline of $\tau < -12$ (i.e., the average cumulative response elasticities in rows F and G themselves) produces estimates exceeding -0.3 in magnitude.This pattern of more negative estimates when using earlier baselines is consistent with a bias due to pre-existing trends that are unaccounted for by the two-way fixed-effects model.[17] We provide additional evidence on the role of earlier baselines when we present first-difference estimates in section 2.4.

These results appear to differ from those in NSW 2014(b), where the authors deny that there is evidence of pre-existing trends in the two-way fixed-effects model. They also argue that netting out the leading coefficients does not alter the estimates very much. To assess their conclusions, we estimate analogous regressions using their data and specification (i.e., state-by-quarter level data from 1990q1-2011q1).[18] In the left panel of Figure 3, we show the cumulative responses from the two-way fixed effects model using their data, which visually match Figure 6 in NSW (2014b).[19] While individual leading coefficients from the two-way fixed-effects model do vary, they are mostly negative in sign—especially during the eight quarters preceding the minimum wage increase. In the right panel of Figure 3, as well as in column 1 of Table 4, we report the four-quarter averages. The four-quarter averaged cumulative response elasticities $\overline{\rho}_{[-4,-1]}$ and $\overline{\rho}_{[-8,-5]}$ are sizable, and are -0.118 and -0.126, respectively, although they are not statistically significant at conventional levels. Importantly, however, as shown in Table 4 (column 1), the estimated 3 year and 4+

[16] We say "4+ year" because ρ_{12} reflects the cumulative response at or after the 12th quarter following a minimum wage increase.

[17] While netting out the leading effects should presumably reduce bias due to pre-existing trends, there is no guarantee that it removes it sufficiently. If employment were falling prior to the increase in minimum wage, it may continue to do so in the post-treatment period even absent treatment. Netting out the leading effects does not guard against this possibility. For this reason, we think it is useful to compare the estimates with and without netting out the leading effects as a diagnostic tool. But if we find that a particular model (like the two-way fixed effects model) produces very different estimates after netting out the leading effects, we should search for models that perform better in such a diagnostic test.

[18] We use the replication data on Ian Salas' website: `https://sites.google.com/site/jmisalas/data-and-code`.

[19] This model is estimated using exactly the same data, sample, and specification that produce NSW 2014 (b) Figure 6: they include controls for unemployment rate, state and period fixed effects.

year effects range between -0.040 and -0.074 when using baselines that are 1 or 2 year prior to treatment,[20] and are much smaller than their estimate from the static specification (-0.165, reported in Table 1, column 1 of NSW (2014a)). In other words, the estimates from two-way fixed effects model in the sample used in NSW (2014b) show that roughly between 50 and 75 percent of the reduction in employment implicit in the static minimum wage employment estimate appears to occur prior to the minimum wage increase—whether the baseline is defined as one or two years prior to the minimum wage change.

Yet, when NSW (2014b) analyze estimates from the same regression and compare cumulative responses beginning at date $\tau = -2$, they conclude that "the contemporaneous elasticities are close to -0.2, building to a maximum of about -0.4 five quarters after the increase—a period around which the estimates are significantly different from zero." Their claim appears to contradict the finding in our Table 4 reported above, even though they are based on the same exact regression. Why do they reach such a different conclusion? As shown in Figure 3, the difference arises entirely from due to their choice of $\tau = -2$ as the baseline (highlighted in red in the figure). If one instead constructs the baseline as the average of the cumulative response during four quarters just prior to treatment ($\overline{\rho}_{[-4,-1]}$), or the four preceding quarters ($\overline{\rho}_{[-8,-4]}$)—as we do in Table 4—one confirms the conclusion reached in our analysis using the full sample (and the conclusion reached in ADR) that a sizable portion of the two-way fixed-effects estimate reflects leading effects. Only by using a peculiar baseline quarter could NSW reach their conclusion that the minimum wage estimates from the two-way fixed-effects model do not reflect leading (or pre-existing) effects. For example, if they had chosen $\tau = -4$ as the baseline, they would have reached the conclusion that the 3 year or 4+ year effects are positive in sign (compare $\overline{\rho}_{[8,11]}$ or ρ_{12} to ρ_{-4} in Figure 3). Since the leading coefficients appear to exhibit a seasonal pattern—or at least considerable variability—one can reach very different conclusions by picking particular quarters. When we average across quarters to construct a more robust baseline—as we do in Table 3—we find a clear signal that netting out the leading coefficients substantially reduces the medium and long term estimates from the two-way fixed-effects model.[21]

Performance of models with controls for spatial heterogeneity

How do models with greater controls for spatial heterogeneity perform on the leading effects falsification test? The 4-quarter averaged coefficients $\overline{\rho}_{[\tau,\tau+k]}$ for these additional models are shown in Table 3, columns 2, 3 and 4. In almost all cases the magnitudes of the leading averaged cumulative responses are smaller: of

[20] These estimates are from rows labeled "F-C" "F-B" "G-C" and "G-B"

[21] Because quarterly leads and lags can be noisy, ADR and DLR (2010) included leads and lags at 4-quarter or 2-quarter frequency for the purpose of smoothing. However, we do acknowledge that there is some arbitrariness in any choice of smoothing. In this paper, for the purpose of comparability with NSW (2014b), we have used their quarterly lead/lag structure in the regression, and simply take four quarter averages of these coefficients to construct $\overline{\rho}_{[\tau,\tau+3]}$. This procedure has the added advantage of reducing arbitrariness in choosing the lead/lag frequency.

the nine leading coefficients from the three models, only 1 is statistically significant at the 5 percent level ($\overline{\rho}_{[-8,-5]}$ in column 2 with just division-period controls), in contrast to the two-way fixed effects model where all three of the averaged leads are significant. In general, in our sample the specification with just division-period controls (column 2) does not perform much better than the two-way fixed effects model. However, the specification with just state-specific linear trends (column 3) performs substantially better: here the leading average cumulative responses are much smaller in magnitude and not statistically significant. Finally, the more saturated specification with both divison-period effects and state linear trends exhibits the smallest leading averaged cumulative responses..

For comparison, we report analogous results using the NSW (2014b) data and specifications in Table 4, columns (2), (3) and (4). NSW (2014b) argue that models with additional controls for spatial heterogeneity do not outperform the two-way fixed effects model on the leading effects falsification test. First, their conclusion is based only on their evaluating the model with just division-period effects (i.e, column 6). As in the full 1979-2014 sample, this specification exhibits a substantial negative bias in the leading coefficients. However, the inclusion of state-specific trends produces much smaller leading coefficients—with or without the inclusion of division-period effects (columns 4 and 3, respectively).

Of course, since the exact nature of the bias may differ from sample to sample, researchers should consider a variety of specifications and see which perform the best when matching the trends in treated and control units. We should put more weight on specifications that do not exhibit pre-existing trends. A robust conclusion from Table 4 is that models with greater controls for time-varying heterogeneity tend to perform better in accounting for pre-existing trends.

What do these models with controls for state-specific trends and division-period effects imply about medium (3 year) and longer run (4+ year) effects from the policy? In our full sample, when using either four quarters just prior to treatment ($\overline{\rho}_{[-4,-1]}$), or the four preceding quarters ($\overline{\rho}_{[-8,-4]}$) as the baseline, the medium or long run estimates range between -0.065 and 0.255 (rows F-B, F-C, G-B, G-C from Table 3, columns 2-4).[22] In other words, there is scant indication of medium or long term disemployment effects in any of these models. By virtue of including leads and lags of minimum wage as regressors, the state-specific trends in these models are estimated using variation outside of the 25 quarter window around the minimum wage increase. Consequently, these trends are unlikely to reflect either lagged or anticipation effects that occur within 12 quarters of treatment. Since dynamic versions of the models with state-trends do not indicate more negative effects than the static version, it is unlikely that the results in the static specification conflate trends with delayed effects (see Wolfers 2006 for more on this issue). Using distributed lag models, Meer

[22]This conclusion is qualitatively similar in the NSW sample (columns 2, 3 and 4) in our Table 4, where the equivalent range is (-0.033, 0.395).

and West (2015) nonetheless argue that aggregate employment falls two or three years after minimum wage increases. Yet, for teens—a highly affected group—our estimates from models with state-specific trends do not suggest such a fall in employment. Our results hold even when the trends are estimated only using data outside of 25 quarters surrounding minimum wage increases, thereby ensuring that delayed effects up to three years are not affecting the trend estimation and hence the minimum wage effects. We provide additional evidence on this question in the next section 2.4, where we estimate distributed lag models using first differences, just as in Meer and West (2015).

When using the 4 quarters prior to treatment as baseline (Table 3, panel C, row G-C), the long-run estimates for models with some controls for time-varying heterogeneity range between -0.064 (column 2) to the somewhat large 0.152 (column 4). These estimates compare to an estimate of -0.123 from the two-way fixed effects model (column 1). It is useful to keep several issues in mind when interpreting these longer term effects. Most importantly, the variation to estimate these effects is more limited. This reduced variation generates standard errors for the 4+ year effects that are sizably larger than the standard errors for the 3 year effects. Moreover, different from short and medium term effects, the 4+ year effects are estimated jointly with the state-specific trends, which reflects another possible limitation. For these reasons, we can probably learn less from these effects than we can from medium term effects. Nonetheless, with those caveats in mind, we find little indication of more negative impacts in the longer run—as is sometimes suggested (or assumed) in the literature. Indeed, the only indication of lower employment occurs within the first two years of treatment. Such a temporary fall in employment is likely not a causal effect of the policy, but may instead reflect the selectivity of times and places of minimum wage increases. But whatever the interpretation of the short run estimate, we do not find evidence of more serious disemployment effect for teens in the medium or longer run.

2.4 First-difference versus deviations-from-means estimators

A fixed effects model can be estimated by including dummy variables for each state in the regression, or equivalently by estimating the regression without fixed effects using data that is in deviations from group means. This approach—typically called the "deviations-from-means" of "within-group" estimator—is what we have utilized thus far. When using state-aggregated data, an alternative method of purging the state fixed effects is by first differencing. If the fixed-effects model represents the true data generating process, both approaches produce consistent estimators. The deviations-from-means estimator is more efficient with homoscedastic and non-serially correlated errors (see Angrist and Pischke 2009). However, the first-difference estimator is less prone to bias if the state effects are not "fixed" and are time-varying instead. This aspect of the estimator is attractive, given the likely bias in the fixed-efects model that we documented above. On the other hand, the first-difference estimator is more likely to miss delayed effects of the policy; this shortcoming can addressed by explicitly including lags in the minimum wage.

Therefore, as an alternative, we estimate the model in first differences using state (j) by year (t) aggregated data, while including up to 3 annual lags in the average minimum wage. The baseline specification corresponding to the two-way fixed-effects model is:

$$\Delta Y_{jt} = \alpha + \sum_{k=0}^{3} \eta_k \Delta MW_{j,t-k} + \mathbf{\Delta X}_{jt}\Lambda + \delta_t + \nu_{jt} \tag{5}$$

where we include common period effects and state-year-aggregated controls $\boldsymbol{X}_{jt}$ (unemployment rate, teens share of population and shares of race, gender, age and marital categories). With aggregate data, the outcome Y_{jt} is log of the annual teen employment-to-population ratio, and as before MW_{jt} is log of the annual mean minimum wage; so the regression coefficients η_k are elasticities of contemporaneous or lagged minimum wages. As before, we saturate this baseline model to account for division-period effects, as well as state-specific trends. Note that in the first-differenced specification, adding state fixed-effects is analogous to including state-specific linear trends in the deviations-from-means (since the first differencing purges the state fixed effects). The most saturated specification is as follows:

$$\Delta Y_{jt} = \alpha + \sum_{k=0}^{3} \eta_k \Delta MW_{j,t-k} + \Delta \mathbf{X}_{jt}\Lambda + \phi_j + \delta_{dt} + \nu_{jt}. \tag{6}$$

We also report two intermediate specifications with just state fixed-effects or just division-period effects. The four specifications are very close to the specifications estimated by Meer and West (2015), who argue that the delayed effects of minimum wages on total employment mostly occur within 2-3 years of the implementation of the policy. We report estimates both with and without teen population weights. We also estimate variants of the models including three annual leads in log minimum wage. Finally, as a point of comparison, we also estimate analogous models using deviations-from-means: these are very similar to models underlying the estimates in Table 3, except that they are estimated using annualized data (and sometimes without leading minimum wages).

In Table 5, we report the cumulative 3 year minimum wage elasticities for teen employment $\rho_3 = \sum_{k=0}^{3} \eta_k$, as well as the contemporaneous elasticity η_0. Panel A reports estimates from the models using the deviations-from-means estimator—as in previous sections, and broadly reproduces the results in Table 3 using annual data. In column 1, the contemporaneous and the three-year cumulative elasticity are sizable and negative, ranging between -0.231 and -0.154 depending on weights, and three out of the four estimates are statistically significant at the 5 percent level. In contrast, the estimates for columns 2, 3 and 4 with controls for state trends and division-period effects range between -0.135 and 0.243; none of the negative point estimates are statistically significant. Similar to what we found in Table 3, including leading minimum wages reduces the magnitude of the estimates for contemporaneous and 3 year cumulative effects: three of the four estimates in column 5 are less than -0.1 in magnitude and not statistically significant.

Panel B of Table 5 reports the first-difference estimates. Now, the two-way fixed effects model in column 1 produces estimates ranging between -0.014 and 0.147, and none of these estimates are statistically significant. To emphasize, the sizable negative estimates of the two-way fixed effects model obtain only when the model is estimated using deviations-from-means, and not first-differences. This is true even when we account for up to three years of lags in minimum wages. This result is consistent with the idea that the first-difference estimates are less likely to be picking up time-varying heterogeneity correlated with the minimum wage (at least those that occur prior to the increase in minimum wage)—a point that was also made in Neumark and Wascher (1992).

The first-difference estimates in columns 2, 3 and 4 of Table 5 further control for state fixed-effects (i.e., allow for trend differences between states) and division-period effects. These estimates range between -0.036 and 0.369; again, none of the negative point estimates are statistically significant. A similar conclusion emerges from the estimates in columns 5-8 that additionally control for up to 3 years of leading minimum wages. Interestingly, the first-difference estimates are not very sensitive to the inclusion of state fixed-effects (which capture trend differences between states). For example, the three-year cumulative elasticity when using weights is 0.147 without fixed effects (Panel B, column 1) and 0.167 with fixed effects (Panel B, column 3). The estimates with division-period effects tend to be more positive, especially so when using population weights, and especially for longer run effects.[23] Overall, none of the first-difference estimates in Table 5 that include additional controls for trends by using divison-period interactions or state fixed effects suggest substantial employment loss, even three years after the increase in minimum wage.

We make one additional observation about the results in Table 5: the negative employment effects for total employment found in Meer and West (2015) do not show up in analogous specifications for teen

[23] In general, we have chosen to weight the state-aggregated regressions by teen population weights in most parts of the paper. With aggregated data, a rationale for using weighted estimates is that they correspond more closely to estimates using individual level data (see Angrist and Pischke 2009 for a discussion). Since we prefer individual-level regressions with controls whenever possible, we have reported estimates with population weights when using a deviations-from-means estimates, which are estimatable using individual-level data. The first-difference specification, however, does not have a corresponding individual level representation, and here there is less clear rationale for using weights. For this reason, we have reported weighted and unweighted variants of regressions in Table 5. We note that both panels A and B of Table 5 suggest that unweighted estimates tend to generally be somewhat more negative (or less positive), and this is especially so for specifications with division-period effects. Moreover, this pattern is more pronounced for first-difference models. Since the division-period effects do reduce variation, they may be more sensitive to use of population weights, especially when using state-aggregated data with a small number of states within a division. For this reason, it is probably best to consider both weighted and unweighted estimates if they tend to differ.

employment, at least with state-level CPS data from 1979 to 2014. For their baseline specification, they find 3-year cumulative elasticities for total private-sector employment of -0.074 (column 1 of their Table 4). In contrast, our closest first-difference specification (unweighted, with state fixed effects, without leads) in column 2 of panel B suggests an elasticity for teen employment of around 0.073. Table 5 thus raises some questions about the causal interpretation of the findings in Meer and West (2015).[24]

3 Evidence for teens using a border discontinuity design with Quarterly Workforce Indicators data

The previous section drew upon CPS data, which limited our spatial controls to the nine census divisions and state-specific trends. In this section we examine evidence on the effects of minimum wages upon teens using the Quarterly Workforce Indicators (QWI) data. The QWI allows for a finer grained spatial approach, namely comparing contiguous counties across state borders. To motivate the advantages of leveraging proximity to construct a counterfactual, Section 3.1 reports results in DLR (2015) that show contiguous areas are indeed more similar. Subsequently, continuing to draw on DLR (2015), section 3.2 describes the border discontinuity design, the QWI data, and the results for teen employment.

3.1 Similarity of local areas: Are contiguous county pairs more alike?

The border discontinuity approach assumes that neighboring areas make good controls because they are more similar in economic structure and in the shocks they face. NSW (2014a) argue that this assumption is unwarranted. Based on their synthetic control donor weights—problems with which we discuss at greater length in section 5—they state (p. 632): "the cross-border county is a poor match—no better than a county chosen at random from the list of all potential comparison counties."

DLR (2015) use the county-level Quarterly Workforce Indicators (QWI) dataset to assess whether adjacent county pairs are indeed more alike in terms of covariates than are non-adjacent county pairs. They consider six key covariates: log of overall private sector employment, log population, private-sector employment-to-population ratio (EPOP), log of average private sector earnings, overall turnover rate and teen share of population. None of these covariates is likely to be substantially affected by the treatment status. Therefore, a finding that contiguous counties are more alike in these dimensions cannot be attributed to having more similar minimum wages.

For each of these six covariates, DLR (2015) calculate the mean absolute differences between (1) a county

[24] The lack of evidence for teen disemployment using the first-difference specification holds whether or not we include the state-level unemployment rate as a control, and whether we restrict the sample to 1990 and later (results not shown).

in their border sample and its contiguous cross-state-border pair, and (2) a county in their border sample and every non-contiguous pair outside of the state. For the latter, each of the 972 counties in 966 cross-border pairs is paired with every possible out-of-state county, for a total of 1,737,884 pairings. For each time period, they calculate the absolute differences in levels and changes of these variables between the county and (1) its cross-border pair and (2) its non-contiguous pair, respectively. Subsequently, they collapse the dataset back to the county-pair-period level and calculate the means of the absolute differences in covariates between counties within pairs.

Table 2 in DLR (2015) shows the results for these variables in levels, as well as 4 and 12 quarter changes. In all cases, the mean absolute differences are larger for non-contiguous pairs; and in all cases but one, the gaps are statistically significant at the 1 percent level. The average percentage gap in absolute differences for the six variables in levels, 4 quarter, and 12 quarter changes is 22.7 percent. Many of the gaps are substantial (above 25 percent): notably, for levels of employment, population and earnings; for 4 quarter change in the EPOP; and for 12 quarter changes in the EPOP and the turnover rate. This evidence shows that that cross-border counties offer an attractive control group that better balances observed covariates,and therefore reducing the scope for bias due to omitted confounders. These results contradict the NSW claim that contiguous counties are not more similar to each other than two counties chosen at random.

3.2 Border discontinuity results using QWI data

DLR (2015) also estimate minimum wage elasticities for teen employment using a border discontinuity approach, and QWI data from 2000 through 2011. The QWI data, which are produced though a partnership between the U.S. Census Bureau and the state Labor Market Information (LMI) offices, provide a public-use aggregation of the matched employer-employee Longitudinal Employer Household Dynamics (LEHD) database. These in turn are compiled from administrative records collected by 49 states and the District of Columbia for both jobs and firms. The operational unit in the QWI is a worker-employer pair, i.e., a job. The primary source of information in the micro-data is the near-universe of employer-reported Unemployment Insurance (UI) records, covering around 98 percent of all private-sector jobs. The UI records provide details on employment, earnings as well as place of work and industry. The Census Bureau uses other data—primarily from Social Security records—to either match or impute demographic information of workers.

The public-use QWI series offers monthly employment counts and average earnings by detailed industry at the county level for specified age and gender groups, and as well quarterly figures for hires, separations and turnover rates.[25] The QWI sample used in DLR (2015) consists of the 1,130 counties that border

[25] Teens in the QWI are ages 14 to 18. For more documentation on the QWI, see DLR (2015) and Abowd et al. (2009).

another state. Collectively, these border counties comprise 1,181 unique county pairs. They merge data on the county's overall and teen population, and the value of each state's minimum wage in each quarter, with the QWI county-pair panel dataset. While most counties in the border pair sample are geographically proximate, some counties in the western United States have much larger land area, and the geographic centroids of the counties in these pairs lie far apart. DLR (2015) uses a data-driven procedure to choose a cutoff (75 kilometers) for the distance between centroids of the adjacent counties—which retains about 81 percent of the sample.

The outcome variables are log earnings and log total teen employment, denoted as Y_{ct}, in county at time period t. Since a single county can be a part of multiple cross-border pairs, the data is stacked by pairs; the standard errors are clustered by state and by border pair to account for multiple instances of counties in the dataset.

Using a dataset of border county pairs level data, they regress Y_{ct} (e.g., log of teen employment) on log of minimum wage in state j containing county c (MW_{jt}), and allow the period effects δ_{pt} to vary by the pair associated with that county in that observation.

$$Y_{ct} = \alpha + \eta MW_{jt} + \mathbf{X}_{ct}\Lambda + \gamma_c + \delta_{pt} + \nu_{cpt} \tag{7}$$

Controls $\mathbf{X}_{ct}$ include log of overall private sector employment and log population. The inclusion of the county-pair specific period effect sweeps out all the variation between pairs, and only uses variation within local areas surrounding a policy border. This border discontinuity specification controls for time-varying heterogeneity in the outcomes across local areas. Unbiased estimates using the two-way fixed effects model require the strong assumption that minimum wage differences between *any* locations j are uncorrelated with residual outcomes. In contrast, the pair-period effects δ_{pt} in model (7) significantly weaken this assumption, only requiring it to hold for locations within a given local area around a state border. The error term ν_{cpt} is indexed by c and p because a single county can be in the dataset multiple times if it is part of multiple county pairs.

Table 6 reproduces the estimated minimum wage elasticities, η, for teen employment originally reported in Table 3 in DLR (2015) . The estimates from column (1) are based on the two-way fixed-effects model that uses the same sample and covariates as equation (7), except that it uses common set of period effects δ_t. The estimates from specification (2) are from equation (7) using pair-period effects, δ_{pt}.

The estimates on earnings from the QWI for both columns (1) and (2) in Table 6 are positive, sizable, and statistically significant at the 1 percent level. The effects are somewhat larger in column (2) with the pair-period effects, increasing from 0.177 to 0.222. These results are similar to the wage elasticities using CPS data reported in Table 1: for example, the teen wage elasticity from the CPS data is 0.253 in the specification with both state-specific linear trend and division-time fixed effects. Turning to teen employment, the estimated minimum wage elasticity from the two-way fixed effects model in column (1) of Table 6 is -0.173 and statistically significant at the 1 percent level. In contrast, the estimated employment elasticity from column (2) with the county-pair period effects falls in magnitude to -0.059 and is statistically indistinguishable from zero. The results reported in Table 6 are consistent with those found in Section 2.1 using CPS data: Controlling for spatial heterogeneity using either coarser and more parametric trend controls, or finer controls in a border discontinuity design, both suggest employment effects for teens that are substantially smaller than the two-way fixed-effects model.

Low-wage labor markets have long been characterized by high turnover, with very short employment spells and frequent shifts between labor market participation and non-participation. Among teens, DLR (2015) reports that jobs with less than one quarter of tenure account for 30 percent of all jobs, and 81 percent of all separations. Table 6 reports minimum wage elasticities on turnover using the QWI. The elasticity of turnover equals -0.204 in the border discontinuity specification (column 2) and is statistically significant at the 1 percent level. A 10 percent minimum wage increase reduces teen turnover rates by around 2.0 percent. DLR (2015) show that the relatively small change in employment levels, but substantially larger reductions in employment flows, is similar to what one would expect from a calibrated model with search frictions and on-the-job search. Importantly, in conjunction with the strong earnings effects, the turnover findings also undermine the notion that this research design throws away too much information to detect any effects of the policy on outcomes.

In NSW (2014b), the authors additionally argue that limiting variation to local areas exacerbates the endogeneity problem because more of the variation in this design comes from state-level changes, which they conjecture are more likely affected by endogeneity problems than federal changes. They do not actually provide any direct evidence that limiting variation to local areas increases reliance on state-level changes. As a matter of accounting, the entirety of the identifying variation in all models that use log of minimum wage as the independent variable come from state-level variation in policy—either on their own or in interaction with federal minimum wage. So if endogeneity of state policies is a concern, it is likely to affect all the estimates in a serious fashion. Moreover, the border discontinuity method is particularly useful when policy endogeneity is a worry—a point elaborated on in DLR (2015). Minimum wage policies may react to shocks affecting the whole state, not just those affecting counties right at the border. Therefore, policy differences within contiguous border county pairs are unlikely to reflect endogeneity concerns that may bias studies using state-level variation.

4 Controlling for endogeneity using synthetic controls

The synthetic control approach of Abadie et al. (2010) offers another way to account for the time-varying heterogeneity that may contaminate the estimation of the minimum wage effect. Unlike a two-way fixed effects regression, the data-generating process underlying the synthetic control approach allows multiple time factors with different factor loadings across states. This allows states to follow trends that are not parallel—generalizing the assumptions of the two-way fixed effects model. Dube and Zipperer (2015) use the synthetic control approach to estimate minimum wage effect for 29 state minimum wage increase events, and pool the results from these individual case studies. In section 4.1 we review the basic approach and the results found in that paper. In section 4.2 we assess how the synthetic control donor weights vary by distance between the treated and donor units—a key part of the the argument in NSW (2014a). In section 4.3 we discuss related findings in Totty (2015), another recent paper that uses an approach similar to synthetic controls to evaluate the effects of the minimum wage on teens.

4.1 Pooled synthetic control estimates for teens

For a single treated state (denoted by $j = 1$), the synthetic control procedure constructs a vector of weights $\boldsymbol{W}$ over J untreated donor states, such that the weighted combination of donor states closely matches the treated state in pre-intervention outcomes, thereby balancing the average factor loadings in the treated and synthetic control groups. With such a vector of donor weights, $\boldsymbol{W}$, the synthetic control estimate for the effect of the treatment can be written as:

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} W_j Y_{jt}$$

The algorithm chooses $\mathbf{W}$ to minimize the mean squared error in pre-treatment characteristics between the treated state and the weighted average of these characteristics among the donor states, summed over K predictor variables. Abadie et al. (2010) do not specify exactly which pre-intervention outcomes should be used. Dube and Zipperer (2015) consider four candidate models and achieve best out-of-sample fit for donors by using annualized averages of pre-treatment outcomes in addition to pre-treatment covariates: annualized pre-treatment averages of teen EPOPs and average teen wages; pre-treatment average values of the shares of teens who are white, black, and female, and the pre-treatment average age of teens; pre-treatment averages of the overall state unemployment rate and unionization rate; and pre-treatment averages of ten industry shares.[26]

[26]Dube and Zipperer implement the synthetic control approach in STATA using the `synth` package with `nested` optimization and `allopt` starting point checks for robustness: `http://www.mit.edu/~jhainm/synthpage.html`. While the consistency of the

For the minimum wage context, Dube and Zipperer construct elasticities to account for the variable treatment intensity across states. They first calculate the average percent difference between the treated and synthetic control outcomes in the post-intervention period $t = t', \dots, T$, and divide this by the the percent minimum wage increase over the full post-treatment period to obtain the elasticity:

$$\hat{\eta}_1 = \frac{\frac{\frac{1}{T}\sum_{t=t'}^{T}\left(Y_{1t} - \sum_j W_j Y_{jt}\right)}{\frac{1}{T}\sum_{t=t'}^{T}\sum_j W_j Y_{jt}}}{\frac{MW_T - MW_{t'-1}}{MW_{t'-1}}}$$

The synthetic control method requires treatment events with well-defined pre- and post-treatment periods, which limits the usable amount of minimum wage variation as compared to conventional regression techniques. To date, the vast majority of states have been affected by the federal minimum wage increases, making federal increases unsuitable for use with the synthetic control method. Few untreated donors are available to construct a synthetic control for states affected by the federal minimum wage. For example, 45 states changed their minimum wage at some point during the year of the 2007 federal minimum wage increase, leaving only 5 states as potential donors to form synthetic controls. As we discuss in section 5 below, inclusion of federal changes is one of the problematic aspect of NSW (2014a) sample of 493 events because most of the donors are also receiving treatments.

Dube and Zipperer consider the 1979-2013 period and select those events with no minimum wage changes two years prior to treatment and with at least one year of post-treatment data—which they consider to be the minimal requirement for reliably measuring the policy's impact. They also exclude events with minimum wage increases of less than 5 percent, as these have little bite. Finally,they discard events with less than five untreated potential donor states, since it is difficult to obtain a reliable control with so few donors. Of the 215 state-level minimum wage changes during their 1979-2013 study period, 29 meet these minimal criteria. The eligible 29 events have pre- and post-treatment periods of varying length. To simplify choices, for each event they select its "maximal" pre-treatment period available from 8-32 quarters; having done so, they then select each event's maximal post-treatment window from 4-12 quarters. On average, the 29 events have 19 quarters of pre-intervention data, and 10 quarters of post-intervention data.

Table 7 reports the resulting teen employment and wage elasticities, reproduced from Dube and Zipperer (2015), Table 6. To summarize the results, the estimates from the 29 case studies are pooled together. The first two columns, which report the mean and median elasticities across the cases, show that the minimum wage changes were binding in the sample: the median of the wage elasticities is 0.237 and the mean is

synthetic control estimator holds for a variety of predictor weight vectors, $\mathbf{V}$, such weights can also be chosen optimally using a data-driven procedure (see Abadie et al. 2010 for details). Therefore, the preferred algorithm dynamically optimizes over both the vectors of predictor weights, $\mathbf{V}$, as well as the vector of donor weights $\mathbf{W}$.

0.368. Indeed, Dube and Zipperer report that 25 of the 29 wage elasticities are positive. In contrast, the median and mean employment elasticities for the 29 treatment events are relatively small: -0.051 and -0.058, respectively.

Dube and Zipperer also pool the estimates from the case studies using the ranked percentile of the true effect against the placebo effects from donors. Pooling the estimates using their ranks is particularly useful since the exact distribution of the sum (or mean) of the ranks under the null is known, so that inference is exact and is valid for small samples. Additionally, they invert the mean rank statistic to construct the Hodges-Lehmann confidence interval for the pooled estimate. As discussed in Dube and Zipperer (2015), collapsing this confidence interval yields their preferred pooled elasticity—the Hodges-Lehmann point estimate—which has desirable small sample properties.

Across treatment events, the mean employment rank is 0.470, close to the 0.500 that would be expected under the null of a zero treatment effect. The pooled Hodges-Lehman elasticity is small in magnitude at -0.036 and is statistically insignificant, as the mean rank falls between the cutoffs (0.395, 0.605) derived from the 2.5th and 97.5th percentiles of the mean of 29 uniformly distributed random variables. The associated 95 percent confidence interval is (-0.170, 0.087). Pooling across the 29 events allows one to draw economically meaningful inferences and rules out a substantial portion of the old "consensus" estimate range of -0.1 to -0.3. (Brown 1999). These small aggregated employment effects also contrast sharply with those for wages. The pooled wage elasticity of 0.266 is statistically significant at the 1 percent level, as is the mean rank of 0.758. The associated 95 percent confidence interval rules out wage effects smaller than 0.169 and larger than 0.414.

Figure 4 illustrates these aggregate effects by showing the time path of the annualized Hodges-Lehmann employment and wage elasticities and the 95 percent confidence intervals, both before and after the minimum wage increase.[27] The top panel shows 4-quarter averages of the wage elasticity estimates ranging from 7 years prior to the minimum wage increase (i.e., quarters -28 through -25 in event time) to 3 years afterward (i.e., quarters 8 through 11). The bottom panel shows analogous estimates for employment.

Average teen wages rise sharply at the time of, and after the minimum wage increase. The Hodges-Lehmann point estimate for the teen wage elasticity lies between 0.2 and 0.4 in the post-treatment period. Approaching 0.4, the pooled wage elasticity is high after three years of treatment. The size of the wage estimate may, in part, reflect a positive and statistically significant (but small) pre-treatment elasticity for

[27] Specifically, we annualize actual treated state and synthetic control outcomes by taking the event-specific mean of these values at every pre- and post-treatment four-quarter interval. The percent difference between these values, divided by the actual minimum wage increase, forms the event-specific elasticity at each time interval. Performing the analogous calculation for the donors, we then construct event-time-specific percentile ranks, which we invert to calculate Hodges-Lehmann point estimates and 95 percent confidence intervals, where the latter use cutoffs from the mean of N independent uniforms, where N is the appropriate number of treatment events.

wages. Positive estimates of the pre-treatment elasticity indicate difficulty in finding a convex combination of donors to very closely match the treated units, likely because the states receiving treatment were particularly high wage ones. Importantly, however, all pre-treatment point estimates for employment are small in magnitude and hover around zero, adding validity to our research design. After the minimum wage increase, employment nominally falls, but the elasticity remains less than 0.1 in magnitude. There is no indication of a more negative effect further out in time: the impact during the third year after treatment (quarters 8 through 11) is close to zero, albeit less precise. Thus, even accounting for a small positive wage bias, there is clearly a sizable wage increase following treatment; and importantly, the employment estimates do not exhibit any similar difficulty in matching the pre-treatment outcomes.

Overall, the synthetic control results from Dube and Zipperer (2015) suggest that for the 29 minimum wage events with a clear and substantial pre-intervention period and a clean set of donors, we see a clear impact of the minimum wage on average teen wage, but little effect on teen employment. The pooled (Hodges-Lehmann) minimum wage employment elasticity for teens of -0.036 is similar to the results using a border discontinuity design (Dube Lester Reich (2015)), state panel results with controls for state-specific trends, and state panel results using the first-difference estimator without weights (with weights the estimates are somewhat more positive).

4.2 Link between local area and synthetic controls

We have argued that local area controls tend to be more reliable, but not that all reliable controls are local. It is possible that synthetic controls and local area controls both provide unbiased estimates, even though the synthetic controls are not always proximate. Nonetheless, it is still informative to examine how the donor weights for constructing synthetic controls vary with distance to the treated state. This variation provides information on the extent of spatial correlation among the loadings from the time-varying factors. With strong spatial correlations in loadings, nearby areas are likely to receive higher donor weights.

To assess this issue, we examine here the relationship between the relative weights assigned to states in the synthetic control donor pool and the geographic distance between donor and treated states for the set of minimum wage increases analyzed by Dube and Zipperer and discussed in section 4.1.[28] The analysis in this section is original to this paper and is not done in Dube and Zipperer (2015).

[28] In an earlier (2013) version of this paper we performed an analogous exercise but used synthetic control weights from a set of randomized placebo laws and found the average weights declined sharply by distance between the donor and the (placebo) treated state. We used placebo laws to assess whether nearby states are indeed more similar, which is the key contention of NSW. NSW (2014b, p.26) then criticized our use of placebo and not actual minimum wage increases, stating that "this approach strikes us as uninformative about the question at hand—whether a particular subset of states provides a more valid set of controls for states where the minimum wage actually does increase." By looking at the actual set of minimum wage treatments, the analysis in this section addresses their concerns and shows that, on average, donor weights do indeed decline with distance.

For each donor j from a given treatment event, we define the relative donor weight equal to the synthetic control weight $0 \leq w_j \leq 1$ divided by the average donor weight for that event (equal to one divided by the number of donors). Figure 5 non-parametrically plots (using `lowess`) the mean relative donor weight as a function of the distance between the geographic centroids of the donor and treated states. The confidence bands are based on standard errors clustered by event. Since the measure of distance is less meaningful when dealing with Hawaii and Alaska, we drop these two non-contiguous states from treatment and donor samples for this exercise. (The key findings are similar when they are included.) For the resulting 25 events, the relationship between distance and donor weights is clearly negative, especially for the first 500 miles. For example, a donor state whose centroid was 100 miles away from the treated state receives, on average, about 2.5 times the weight of a donor state that was 500 miles away from the treated state, and nearly 3.8 times the weight of a donor that was 1000 miles away. In a related exercise, Dube and Zipperer (2015, Table 4) show that donors that are in the same Census division as the treated state receive on average 3.2 times the weight compared to states that are outside of the division.[29] Overall, the greater weight for nearby donors constitutes evidence for the similarity of factor loadings between states that are nearer geographically, providing added support for leveraging proximity when constructing controls.

NSW (2014a, 2014b) argue that donor weights for states within the same division are not greater than weights for donors that are outside of the division. This claim contrasts with the evidence analyzed above from Dube and Zipperer (2015). Below, in section 5, we discuss the relatively short pre-intervention period NSW use to estimate the synthetic control donor weights, and the likely poor match quality they obtain. In contrast, Dube and Zipperer (2015) use a longer pre-intervention period and a data-driven set of predictors to estimate the synthetic controls, and formally assessed the quality of matches—making their findings on this question more informative than those in NSW.

4.3 Evidence from factor model estimators

The data generating process underlying the synthetic control approach is a factor model. States may vary in their loadings for (possibly many) time factors, allowing for a rich pattern of time-varying unobserved heterogeneity. Such a factor model can also be estimated using other approaches, such as the Bai (2009) interactive fixed effects estimator. Totty (2015) estimates minimum wage effects on teen employment using three types of factor model estimators, including the Bai estimator, and two variants of the common

[29] NSW (2014b) criticize the calculations we performed in an earlier (2013) version of this paper of the within-division donor weights implied by the findings reported in NSW (2014a). Their argument appears to be that the mean of ratios differs from the ratio of means. We used the latter because that is the information provided in NSW (2014a). However, which of these averages is more informative is beside the point. As we discuss in this paper, we have serious questions about the quality of matches obtained by NSW. In contrast, by using longer estimation samples for synthetic control weights, along with a reasonably high number of donors and a data-driven choice of predictors, we confirm in this section that nearby donor states receive greater weight when constructing synthetic controls for minimum wage events.

correlated effects estimator of Pesaran (2006).

Totty finds that accounting for time-varying heterogeneity in this fashion substantially reduces the size of the minimum wage employment estimates, consistent with the evidence presented in this paper. In his 1990-2010 sample, the two-way fixed effects estimate for the minimum wage elasticity of teen employment is -0.178 (statistically significant at the 5 percent level). In contrast, the estimates from the three factor models range between -0.040 and -0.065 and are not statistically significantly different from zero. Totty (p. 32) concludes that "[t]he factor model estimators find little to no effect of minimum wage increases on employment."

While his estimated factor structures do not always neatly line up with geography, there are clear regional patterns to the factor loadings. Totty writes (pp. 6-7): "Analysis of the factor structure suggests that the factor model estimators are capturing time trends and regional heterogeneity in the error term of the traditional two-way fixed effects specification, which supports the approach in DLR and ADR. But, in some cases, same-division states and cross-border counties appear very different in their unobservables." We interpret Totty's results as being consistent with the idea that—on average—more proximate areas are better matches, as is also indicated by our evidence that covariates in contiguous border county pairs are more similar than among other pairs (section 3.1) and by the fall in synthetic control donor weights by distance (section 4.2). As we noted at the beginning of this section, the factor based structure underlying the synthetic control and the factor model estimators need not always correspond to the local area control approach. It is possible that local area controls, on average, remove the bias from the two-way fixed effects model even if the factor structure does not map onto proximity in every case. Estimators may pick different control groups while still being unbiased; the similarity of the minimum wage estimates from using local area controls, synthetic controls and factor models suggest that this may indeed be the case here. Overall, these approaches are complementary, and together they offer a robust set of tools to account for the time-varying confounders that clearly plague the two-way fixed effects model.

5 The problems with the NSW matching estimator

NSW (2014a) also propose a matching estimator based on synthetic control weights that produces fairly different estimates fromn Dube and Zipperer (2015) and Totty (2015). In this section, we discuss their event selection and empirical specifications and find that they are both deeply flawed. We show that most of their "treatment states" do not experience sizable minimum wage increases and many receive none at all, calling into question the value of their entire exercise. Subsequently, we re-analyze NSW's results and show that for the states in which there was an actual treatment, there is no evidence of employment loss. We also find other evidence that the match quality is poor. Using a slightly earlier pre-intervention period than they use produces positively signed employment estimates, indicating that the treatment and control units did not track each other very well prior to the intervention.

NSW begin with all 544 federal and state minimum wage increases between 1990 and 2011. They require the events to have a four quarter pre-treatment period (t=-4, -3, -2, and -1 in event time), along with a four quarter treatment period (t=0, 1, 2, 3), reducing their list to 493 usable minimum wage events. With this four quarter pre-treatment period, they estimate synthetic control donor weights for each of the 493 treatment events using a sample of donors that includes every other state—including states that saw increases in minimum wage during dates (t=-4,...,-1,1,...,3).[30] For each event, then, they have a matched synthetic control unit for their period. They stack this matched data and subsequently estimate standard two-way fixed effects panel regression with this stacked dataset.

The most fundamental shortcoming of the NSW matching estimator concerns their sample. Of the 493 events studied by NSW, 129 comprise what they call a "clean sample," in which there are no minimum wage changes in the control units during 4 quarters prior or subsequent to treatment. They do not, however, just use this "clean sample," but add to it an additional 364 events in which both treatment and potential control units are experiencing minimum wage increases during treatment periods.[31] As a result, in their full 493 event unclean sample—which they use for their main estimation—there are: 1) minimum wage changes in the treated units in the pre-intervention period (t=-4,...,-1), and 2) minimum wage changes in the donor (or potential control) states in the pre- and post-intervention periods (t=-4,...,0,...,3). This sample construction thus renders the distinction between "treatment" and "control" units nearly meaningless. As shown in column 1 of Table 8, for their full sample of 493 events, the treated units experienced an average 0.098 log point increase in the minimum wage.[32] But during the same time period, the control units experienced a 0.071 log point increase in the minimum wage, yielding only a 0.027 log point (approximately 2.7 percent) net increase in the "treated" versus "control" units. This increase is very small: for comparison, in the 29 events analyzed by Dube and Zipperer (2015), the minimum wage rose 19.3 percent more in the treated areas as compared to the control areas.

[30] To estimate the donor weights for each event, NSW match on residual employment, after partialing out state and time fixed effects, as well as the minimum wage. This method is not standard, as the authors note. For one, the synthetic control estimator's consistency is based on being able to match the actual outcome, and not residuals. More importantly, and as NSW discuss, including the minimum wage in constructing residual employment is problematic because the minimum wage effect is what one is trying to estimate. However, they find the exclusion of the minimum wage in constructing the residual does not substantially affect the findings. To keep our results comparable, in our re-analysis of their data we follow their practice and construct donor weights to match residual employment net of time and state fixed effects as well as the minimum wage.

[31] Interestingly, NSW (2014a) find a small, statistically insignificant minimum wage elasticity for teen employment of -0.06 when they apply their method only to the "clean sample." They nonetheless dismiss these results, arguing that in this sample, even the two-way fixed effects estimate is not sizably negative. This argument strikes us as indefensible. The two-way fixed-effects estimate in their "clean sample" may simply be less biased than in the expanded (unclean) sample. In general, we see little justification in expanding the sample to include events inappropriate for the synthetic control approach, just because the two-way fixed-effects estimate in that sample matches that from the full state panel sample.

[32] We used the programs and dataset posted here:`http://j.mp/datacodeILRR`.

To assess their sample further, we divide the 493 events into quartiles by the extent of treatment: $\Delta \ln(MW_{treated,j}) - \Delta \ln(MW_{SC,j})$, the differential increase of the log minimum wage in the treated versus in the synthetic control units. As shown in the first column of Table 8, the bottom quartile (quartile 1) actually received a net negative treatment: the "treated" units experienced a 0.024 net *decrease* in log minimum wage as compared to their synthetic controls. The second quartile received essentially no net treatment (a very small increase of 0.005), while the third quartile received a 0.028 increase in log minimum wage. Only the fourth quartile received a substantial treatment—a net minimum wage increase of around 0.099 log points (approximately 10.4 percent). Most of NSW's events thus are ill-suited for studying the effect of minimum wage increases using the synthetic control approach. There is little point in defining events, treatment groups and synthetic controls if most of these events entail so little net variation in minimum wages. In addition, it is not clear why we should expect a reliable counterfactual for the "treated state" by matching past outcomes in the treated and synthetic control units in the pre-intervention period, when the treatment status was in reality changing in both groups during that period in arbitrary ways.[33] Finally, NSW use only four quarters of pre-treatment data, a very short window to estimate synthetic control donor weights. Other existing work using synthetic control methods use several years of pre-treatment data (Abadie et al. 2010; Bohn et al. 2014). Overall, the nature of NSW's sample raises serious questions about their findings.

What does their sample imply about the employment effects of the minimum wage? A difference-in-differences approach provides a straightforward way to estimate an employment elasticity using these 493 events. In Table 8, we show the pattern construct difference-in-differences estimates for log of teen employment, $\frac{1}{J}\sum_j^J [\Delta \ln(E_{treated,j}) - \Delta \ln(E_{SC,j})]$, where J is the total number of events. For the full sample (top panel, Treatment - Control row), log employment changes by -0.007 in the treatment units differentially following the minimum wage increase; this decline is statistically significant at the 10 percent level. By scaling this employment effect by the differential increase in log minimum wage (0.027), we obtain an elasticity of $\frac{\sum_j [\Delta \ln(E_{treated,j}) - \Delta \ln(E_{SC,j})]}{\sum_j [\Delta \ln(MW_{treated,j}) - \Delta \ln(MW_{SC,j})]} = -0.247$.[34] This difference-in-differences elasticity estimate is somewhat larger in magnitude than the -0.145 elasticity estimate obtained by NSW using a panel regression. Nonetheless, both results suggest at least moderate sized employment losses.[35] However, several pieces of

[33] Matching on residual employment, after partialing out minimum wage effects, may guard against the bias from aligning employment in the treatment and synthetic control groups in a pre-treatment period where treatment status was in reality changing. But this approach relies on having the right estimate for the minimum wage, which is unknown, and is estimated using the very two-way fixed effects model that is in contention. NSW acknowledge that their logic has a "potential circularity," but argue that their results are similar whether the synthetic control weights are constructed from residual employment after partialing out minimum wages, or not. However, this does not guarantee that the weights are constructed correctly. For example, if the mean minimum wage effect is small, but with some heterogeneity, the weights constructed from matching the treatment and control units will be wrong, and the resulting estimates may be biased. More generally, it is problematic to use an "unclean sample" in which treatment status is changing in the pre- and post-treatment periods in both the treatment and control groups.

[34] Standard errors for the elasticity were computed using the `suest` command in STATA, clustering on state.

[35] NSW do not conduct this type of difference in differences (or perhaps a simple post-treatment difference), even though such an approach corresponds most closely to the standard application of the synthetic control method (e.g., in Abadie et al.

evidence suggest that these disemployment estimates are unreliable. First, the largest fall in employment (-0.012) occurs in quartile 1, which experiences a minimum wage *decline*, implying a positive minimum wage employment elasticity of 0.490. Second, for quartile 4—the only quartile with a substantial increase in the minimum wage—the employment fall is more muted (-0.007) and it is not statistically significant. The implied minimum wage elasticity based on the fourth quartile is -0.074, substantially smaller in magnitude than what NSW find. Third, for many events (e.g., quartile 2) the change in the minimum wage is virtually the same in treated and control groups: these observations provide little usable information to identify the effect of the treatment. To summarize to this point, when using NSW's sample of events and their synthetic controls, some events suggest more sizable job loss, and some suggest sizable minimum wage increases, but these typically happen to be different events; and many events are simply uninformative.

Given the shortcomings of the NSW sample we discussed above, we should worry about the quality of matches obtained by their procedure. To assess the impact of match quality on the estimates, we performed the following exercise. The synthetic control weights in NSW (2014a) are estimated using quarters $t = -4, ..., -1$ in event time, and the minimum wage estimates were constructed by taking the difference between the post- and pre-treatment periods. As a check, we use a slightly earlier pre-intervention period $t = -8, ..., -5$ to form the difference-in-differences estimates. Since this earlier period was not used to estimate the synthetic controls, it provides a validation test: if control groups are constructed well and provide a valid counterfactual, then using this earlier pre-intervention data should provide broadly similar results.[36] As shown in column 4 of Table 8, when using this earlier pre-treatment period, the relative increase in the minimum wage between treatment and synthetic control groups is somewhat larger, 0.038 as opposed to 0.027; this occurs in each of the four quartiles as well. In other words, using this earlier pre-treatment period does not cause any attenuation in the extent of treatment. Using this earlier pre-treatment period, however, the employment estimate is now positive in sign (0.008) and not statistically significant; and the implied overall minimum wage elasticity is 0.205. If the NSW synthetic control weights were reliable, this earlier pre-

2009). Instead, they create a sample where they stack the synthetic controls and treated units, and regress log employment on log minimum wage, controlling for time period dummies, event-by-state dummies, state unemployment rate, and teen share of population. Their estimate of -0.15 is somewhat smaller than the -0.25 difference-in-differences elasticity we report in Table 8. The difference-in-differences estimate presented here is based on the actual variation in minimum wages induced by the treatments in 493 events. In contrast, the NSW panel regression additionally uses variation in minimum wages (1) between synthetic control units, and (2) between treated units, which seems contrary to the purpose of defining treatment events. Moreover, the difference-in-difference formulation allows us to transparently diagnose what drives the mean estimate by considering different subsets of events, as we do in Table 8. We cluster the standard errors at the level of events; this likely understates the standard errors by not accounting for estimation of synthetic controls, and also the possible correlation in the control units across events. However, clustering by events strikes us as more accurate than clustering by the narrower event-by-treatment status, as is done in NSW (2014a). For this reason, we should not be cautious about the statistical significance of findings from the NSW matched estimator.

[36]The sample of events shrinks somewhat from 493 to 442 when using the earlier period since the events in 1991 in the NSW dataset do not have a balanced earlier period (t= -8,...,-5). However, this sample restriction has little impact on the baseline difference-in-differences estimates (results not shown). For example, while the overall minimum wage elasticity for teen employment using the sample of 493 events is -0.247 (Table 8, column 3), the analogous elasticity for the restricted sample is -0.271 (results not shown in tables).

intervention period should not suggest such different estimates of the employment effect of the policy. The sensitivity of the results to an earlier pre-intervention period suggests serious problems with match quality and violation of the assumption that the synthetic control and treatment units exhibit parallel trends.

In quartile 4, the only quartile with a sizable treatment magnitude, the earlier pre-treatment estimates are close to zero, with a minimum wage elasticity of 0.029, compared to an estimate of -0.074 using the later pre-treatment period. In other words, for arguably the most informative events in the NSW sample, neither pre-intervention period suggests a substantial disemployment effect.

Overall, our re-analysis of the NSW (2014a) data suggests serious flaws with their sample construction and their estimation of synthetic control groups. Most tellingly, while they use minimum wage increases as events for which to construct synthetic controls, most of the events do not exhibit a clear treatment. In combination with the short pre-intervention period used to estimate the synthetic control weights, the unclean nature of the sample appears to produce poor matches. Moreover, in the cases where the treatment group actually experiences a sizable increase in the relative minimum wage as compared to their synthetic control, there is not an indication of a sizable reduction in employment.[37] The conceptual problems with NSW matching estimator, the problems with their sample construction, and the discouraging findings from simple diagnostic tests all strongly suggest that the estimates they present are unreliable.

6 Effects on restaurant employment

Up to this point, we have focused on the effect of minimum wages on teen employment. We pivot here to the other group in question—the restaurant workforce. The minimum wage literature (e.g., Card and Krueger 1994, 2000; DLR 2010, 2015) has studied restaurant employment because the restaurant industry is by far both the most intensive user of minimum wage workers. Based on CPS ORG data, during the 2000-2014 period (which provides a NAICS-based industry classification), 28.3 percent of workers in "food services and drinking places" (NAICS 722) earned within 10 percent of the statutory minimum wage. During the same period, this sector also accounted for 28.6 percent of workers earning within 10 percent of the minimum wage. NSW (2014a) devotes substantial attention to critiquing the methods and details of our work in DLR (2010) on the effects of minimum wages on restaurant employment. Yet the substance of their disagreement turns out to be surprisingly modest, which is why we have focused our attention first on teens. In this section, we review the estimates from our previous work on restaurants using border discontinuity design and show updated results using an expanded sample between 1990 and 2014. We then compare these results to other

[37]We note in passing that the results from this sub-sample with a clearer treatment may still suffer from the other shortcomings of the NSW matching estimator, such as the short estimation window, that we have discussed in this section.

existing work that try to account for time-varying heterogeneity, including NSW (2014a, 2014b), and find a remarkably narrow range in the estimated employment elasticities.

Table 9 provides our summary of previous restaurant studies. Dube, Lester and Reich (2010) used QCEW data between 1990 and 2006q2 and a border discontinuity design that we described in detail in section 3.2. The dependent variable in the estimating equation (7) is log of restaurant employment; controls include log of overall population and (in some specifications) log of overall private sector employment. Using 504 counties with balanced panels, DLR (2010) found a minimum wage elasticity of 0.016 for employment in full and limited service restaurants using the border discontinuity design. We reproduce these estimates in this paper's Table 9. In contrast, the two-way fixed effects estimate for restaurants in the full sample of counties with balanced panels was -0.176, and the two-way fixed effects estimate in the sample of border county pairs was -0.112. More recently, Dube, Lester and Reich (2015) use QWI data between 2000 and 2011 for a slightly broader sectoral category "food services and drinking places" (NAICS 722) and find estimate an elasticity of -0.022 using the border discontinuity design; these results are also reported in Table 9.[38] In both of these studies, the minimum wage elasticities for average restaurant earnings using the border discontinuity design were large—ranging between 0.188 and 0.207—and statistically significant (results not shown in tables in this paper). [39]

Table 10 shows updated results using the QCEW data for the "food services and drinking places" (NAICS 722) from 1990 to 2014.[40] The first column of Table 10 reports estimates from the sample of all 1,535 counties with balanced panels in the data, using a model with period and county fixed effects. The second column reports estimates from the same specification, but for the contiguous border county pair sample, using 568 border counties with balanced panels. The third column reports estimates using the same sample as column 2, but with pair-period effects; this is the preferred border-discontinuity specification.

The earnings estimates in panel A are consistently large, ranging between 0.208 and 0.246 across the three columns, roughly what one would expect given that CPS ORG data show that 28.3 percent of workers in this sector earned within 10 percent of the statutory minimum wage in the 2000-2014 period. In panel A, we also report minimum wage elasticities for restaurant employment using the specification in equation (7), where the treatment variable is just the contemporaneous log minimum wage. For the two-way fixed

[38]As a refinement, DLR (2015) use a data-driven procedure to further limit the set of county pairs to include only those whose centroids are no more than 75 kilometers apart.

[39]Additionally, as DLR (2015) shows, with the border discontinuity design we find substantial reductions in both separations and hires—and hence employee turnover—following a minimum wage increase. This result is similar to the findings for teens discussed in section 3.2.

[40]We start in 1990 here because the NAICS industry-based QCEW data is not available at the county level for previous years. We use the NAICS 722 category rather than just the four-digit limited and full-service restaurants (NAICS 7221, 7222), which was used in DLR (2010), both because the former is a broader category allowing for possible substitution between sub-sectors, and because there were classification changes to these four-digit industries beginning in 2011. For convenience we use the term "restaurant" instead of "food services and drinking places" below.

effects model in the all county sample, the estimate is -0.242 and statistically significant at the 1 percent level (Table 10, Panel B, column 1). In the contiguous border county pair sample, the two way fixed effects estimate on employment is -0.186, and is statistically significant at the 5 percent level (Table 10, Panel A, column 2). Finally, for the border discontinuity design with pair-period effects, the estimate is 0.022 and is not statistically significantly different from zero.[41]

Panel B of Table 10 displays estimates from a dynamic model, in which we modify equation (7) to include as regressors 12 quarters of leads and 12 quarters of lags of log minimum wage.

$$Y_{ct} = \alpha + \sum_{k=-12}^{12} \eta_k MW_{j,t-k} + \mathbf{X}_{ct}\Lambda + \gamma_c + \delta_{pt} + \nu_{cpt} \tag{8}$$

Controls $\mathbf{X}_{ct}$ include log of overall private sector employment and log population. We also estimate a two-way fixed-effects variant of equation (8) with common time effects, δ_t, instead of δ_{pt}. We calculate the cumulative response elasticity by successively summing the η_k coefficients: $\rho_\tau = \sum_{k=-12}^{\tau} \eta_k$. As before, to reduce the noise, we report 4-quarter averages of the cumulative response elasticity, $\overline{\rho}_{[\tau,\tau+3]} = \frac{1}{4}\sum_{m=0}^{3} \rho_{\tau+m}$. In the all-counties sample, for the two-way fixed-effects model, the average cumulative response elasticity in the four quarters preceding the minimum wage increase is -0.201 and is statistically significant at the 1 percent level (Table 10, panel B, row C, column 1). Between the 9th and 12th quarter preceding the minimum wage, the cumulative response averages -0.119 and is statistically significant at the 1 percent level. In other words, we find unambiguous evidence of pre-existing trends that contaminate the two-way fixed-effects estimate for employment in the food services and drinking places sector over the 1990-2014 period. In the border county pair sample, we still find evidence of pre-existing trends, though they are somewhat more muted. For example, in the four quarters preceding the minimum wage increase, the cumulative response elasticity is -0.124 and is statistically significant at the 10 percent level. The estimates for earlier quarters are negative in sign, but neither as large nor statistically significant. In contrast, the specification with pair-specific period effects in column 3 show no indication of pre-existing trends: the point estimates are all less than 0.02 in magnitude, and none are statistically significant. This result provides strong evidence that the border-discontinuity design is able to provide more reliable estimates by using more similar comparisons.

Table 10 also reports medium run (3 year) and long run (4+ year) effects for restaurant employment using alternative baselines: 1-4 quarters, 5-8 quarters, or 9-12 quarters prior to the minimum wage increase. The estimates for the two-way fixed-effects model are larger when using earlier baseline periods, consistent with pre-existing trends. Moreover, the long run, (4+ year) estimate is substantially larger in magnitude

[41] The point estimates from the border discontinuity design are quite similar in specifications in which we do not control for overall private sector employment (0.023), or refine the sample (along the lines of DLR 2015) by limiting county pairs to those whose centroids are within cutoffs ranging between 50 and 150 kilometers of each other; these estimates range between -0.047 and -0.010.

than the 3 year estimates. While such a long-run estimate could be causal, it is also possible that this long run estimate reflects confounding long run trends. The border discontinuity design estimates suggest the latter is the more likely explanation: while imprecise, the point estimates range between -0.073 and -0.081, depending on the baseline—much smaller in magnitude than the two-way fixed-effects estimates. Medium run estimates from the border discontinuity specification are more precise and the point estimates range between -0.024 and -0.032, depending on the baseline.

How do our border discontinuity design estimates for restaurant employment compare with other existing work? In addition to our own previous work, Table 9 summarizes the key minimum wage elasticities for restaurant employment from a number of other recent papers. Where available, we have provided estimates using the two-way fixed-effects model, as well as models controlling for time-varying heterogeneity. These papers include Addison, Blackburn and Cotti (2014) using parametric trend controls, Totty (2015) using factor models, and NSW (2014a) using their "matching estimator."

Addison, Blackburn and Cotti (2014), also responding to the concerns raised in NSW (2014a) about using linear trends, use first to fifth order polynomial trends and the 1990-2005 QCEW sample for employment in "food services and drinking places" (NAICS 722). They find that when controlling for county-specific trends, the minimum wage elasticities range between -0.062 and -0.006, while the two-way fixed-effects model suggests an estimate of -0.101 (see Table 9 in this paper). For the 1990-2012 period, their estimates with county-specific trends range between -0.040 and -0.010.[42] Addison et al. further estimate elasticities using alternative controls for trends (specifically, trends estimated peak-to-peak or in the post-1993 sample, or with a Hodrick-Prescott filter) for both sample periods; these additional 6 estimates range between -0.042 and 0.058, depending on the sample and specification.[43] In other words, even accounting for the concerns raised by NSW with respect to inclusion of linear trends, the restaurant employment estimates are found to be small in magnitude and similar to those from a border discontinuity design.

Another recent paper, Totty (2015), uses factor model estimators to estimate the impact of minimum wages in the restaurant sector. He uses 1990-2010 QCEW data on total employment in NAICS 7221 and NAICS 7222. As we noted in the context of teen employment in section 4.2, these estimators (interactive fixed effects, and common correlated effects) allow states to differ in their loadings for multiple time factors, similar to synthetic controls. Totty's results on restaurants, also included in Table 9, clearly confirm both the importance of unobserved time-varying heterogeneity and the negative bias in two-way fixed effects estimate; Totty's estimates allowing for time-varying heterogeneity range between -0.013 and -0.042 and

[42] In our sample of all counties with balanced panels between 1990-2014, using log of overall private sector employment and log population as controls, and using up to fifth order county-specific polynomial trends, we obtain minimum wage elasticities for restaurant employment (NAICS 722) between -0.070 and -0.020 (results not shown in tables).

[43] Addison et al. also show that their conclusion is not affected by estimating the regressions in differences (short or long)—echoing our findings for teens in this paper.

are not statistically significantly different from zero. These results confirm the conclusions reached in DLR (2010, 2015) and Addison, Blackburn and Cotti (2014).

Finally, while we have discussed the methodological problems with NSW's matching estimator in section 5, here we note that it, too, suggests a small elasticity for restaurant employment of -0.063 when using population weights, and 0.008 without. They concede as much when discussing their results.[44] The main thrust of NSW's concerns about our restaurant employment estimate focus on whether our use of the border discontinuity design is appropriate. But from the perspective of most economists and policymakers interested in how restaurant employment is affected by minimum wage policies, the key point is as follows.Tables 9 and 10 report a total of 18 estimates that include additional controls for heterogeneity beyond the two-way fixed-effects model from this paper, Addison, Blackburn and Cotti (2014), DLR (2010, 2015), NSW (2014a, using their matched estimator), and Totty (2015). Taking these together, we obtain a minimum wage elasticity range for restaurant employeemnt between 0.039 and -0.063.[45] Since the earnings elasticity is typically around 0.2, these results imply labor demand elasticities generally smaller than -0.3 in magnitude. Moreover, all of these estimators—including NSW's preferred matching estimator—suggest employment effects that are usually substantially smaller than the two-way fixed-effects model (an exception is Addison et al.'s 1990-2012 sample—as reported in this paper's Table 9, they find a zero effect even for the two-way fixed effects model). While there may be disagreement about the merits of specific estimators, these results comprise a highly robust set of findings. They confirm: (1) at most a modest impact of minimum wages to date on restaurant employment[46], and (2) the violation of the parallel trends assumption in the two-way fixed-effects model, and likely bias toward finding evidence of job loss. Moreover, while NSW (2014a) acknowledge that their preferred matching estimates are small in magnitude, they do not offer a clear explanation for the sizable

[44] NSW (2014a), p. 643: "What do we conclude? First, the evidence of disemployment effects we obtained is clearly not as strong for restaurant employment in the QCEW, when using the synthetic control weights, as for teen employment in the CPS." NSW (2014b, pp. 13-14)) also discusses pre-existing trends in the restaurant sample using the 1990-2010 NAICS 7221 and NAICS 7222 QCEW data. Here they concede that when using the two-way fixedeffects model, there is an indication that employment had fallen prior to the the minimum wage, though they argue that there is also an indication of a "downward employment shift" after the minimum wage increase, especially for the sample of all counties. Their Figure 5, panel A, however, shows that for the all county sample, the fall in employment by $\tau = -1$ was larger than the reduction in the next 11 quarters. Only with the inclusion of the 12th quarter lag—which is the long term effect—does the estimate become substantially negative. And that estimate is particularly susceptible to long-term trend differences between treated and control areas. Moreover, their Figure A corroborates that when using county-pair-quarter fixed effects, the leads are close to zero, again validating the research design in DLR (2010).

[45] Another recent paper studies restaurant employment using a border discontinuity design for a small number of states. Aaronson, French and Sorkin (2015, Table 2, row C) obtain short-run elasticities for continuing workers of -0.05 for for limited service restaurants and -0.12 for full service restaurants. Neither elasticity is statistically significant. They also find increases in exit and entry in limited service restaurants, with an overall elasticity of -0.1 that they describe as being "very imprecise" (they do not report a standard error). Overall, they argue that the short run employment estimates are in line with findings in DLR (2010). They interpret the increased exit and entry using a putty-clay model and project that there may be larger disemployment effects in the longer run. Such effects are possible, but as an empirical matter, as we show in this paper and in DLR (2010), a border discontinuity design similar to the one used in Aaaronson, French and Sorkin does not suggest sizable employment losses in restaurant sector in the "medium run," i.e., after 12 or 16 quarters following the minimum wage.

[46] It is possible that the composition of a restaurant workforce may change in response to a minimum wage increase, even when overall restaurant employment does not. However, at least for age and gender, DLR (2015) does not find evidence for such compositional shifts.

difference between their matching estimates and the estimates from the two-way fixed-effects model.[47] NSW's own estimates for restaurant employment effects thus suggest that the two-way fixed-effects model is biased toward finding a substantially larger disemployment effect than is accurate—a key argument in DLR (2010).

To be sure, there are remaining disagreements on the details of the restaurant findings. For instance, NSW (2014a, 2014b) criticize a falsification test we performed in DLR (2010) to demonstrate the unreliability of the two-way fixed effects estimates; we respond to these criticisms in Appendix C. However, the key takeaway remains: the research literature seems to be reaching an agreement on the medium-run effects of minimum wages on restaurant employment.

7 Conclusion

Much of the minimum wage research on employment effects has focused on teens and on restaurant workers because these two groups are especially affected by minimum wage policies. We review here a wide variety of recent restaurant studies—including NSW (2014a, 2014b), DLR (2010, 2015), Addison, Blackburn and Cotti (2014) and Totty (2015)—that use different datasets, time periods and estimators, and yet arrive at similar findings. In these studies, the preferred elasticities of employment with respect to minimum wages lie within a fairly narrow range of -0.063 and 0.039,suggesting at most a small effect of the policy. The main substantive disagreement—and most of this paper—thus centers upon the effects on teens.

Teens, it should be repeated, today represent a declining and smaller share of minimum wage workers than when Brown (1982, 1999) wrote his review articles on minimum wages. Moreover, if a rising minimum wage leads employers to substitute adults for teens, evidence of negative effects on teens does not necessarily imply negative effects on overall low-wage employment. The magnitudes of the labor demand elasticity implied by minimum wage effects on teen employment are, therefore, likely to be an upper bound for the low wage workforce overall. At the same time, these limitations also imply that if there are substantial disemplyoment effects of minimum wages, we should probably be able to detect them for teens.

Our key contention is that when using cross-state variation to estimate minimum wage effects, it is critically important to account for time-varying heterogeneity. A casual inspection of a map of minimum

[47]NSW (2014a, p. 643) includes the following discussion. "[T]the evidence of disemployment effects we obtained is clearly not as strong for restaurant employment in the QCEW, when using the synthetic control weights, as for teen employment in the CPS. As noted earlier, this is not surprising, nor is it in any way contradictory with the existing literature. Most of the existing evidence focuses on teenagers or other low-skill groups, rather than on any particular industry; and at the industry level, labor-labor substitution seems more likely to mask the full extent of the disemployment effects for the least-skilled. This difference in findings may also stem, in part, from difficulties in applying our method to the QCEW data. Second, and perhaps more significant, when we weight the estimates we find some evidence of disemployment effects when the synthetic control weights are used. These are, arguably, the most defensible estimates." In footnote 48 they write: "We do not find these negative estimates if we exclude the private-sector employment control." In short, NSW are able to obtain statistically significant estimates only when using population weights and when including overall private sector employment as a control, and even here the estimates are no greater in magnitude than -0.063.

wages across the U.S. shows a high degree of spatial clustering in minimum wage policies since the 1980s. This clustering coexists with an array of potential confounds that vary between high and low minimum wage states. The chances seem small that all of these factors happen to balance each other.

When we use state-level variation, the sizable negative estimates of minimum wage elasticities for teen employment from the two-way fixed-effects model largely disappear once we control for time-varying heterogeneity in any numberof ways. Importantly, the estimates fall substantially in magnitude once we include controls for state-specific trends. This reduction is even greater in longer samples that allow more precise estimation. These results hold when we include higher-order state-specific trends in our specifications, contradicting findings in NSW for their shorter sample. Adding division-period effects further reduces the magnitude of the estimates. Across specifications with division-period controls, and linear or higher order state-specific trends, the minimum wage elasticity estimates for teen employment range between -0.040 and 0.006. Using the double-selection post-LASSO approach that chooses controls using a data-driven procedure without assuming which ones should be included or excluded, we find that controls for time-varying heterogeneity (especially state-specific linear trends) should be included and that the resulting employment elasticity is close to zero (-0.012). Additionally, this approach establishes that the inclusion of just a handful of state trends reduces the employment estimate to close to zero—contradicting the assertion in NSW that our findings in ADR were driven by throwing away too much data. By using a longer panel than in NSW (2014a, 2014b), and 72 different subsamples with alternative starting and ending dates, we refute their conjecture that the presence of downturns in the "endpoints" imparts a positive bias in estimates from models with state-specific trends.

We also confirm that the two-way fixed-effects research design is contaminated by pre-existing trends. Using the same 25 quarter window around minimum wage changes as used in NSW (2014b), we show that teen employment was unusually low (and falling) prior to minimum wage increases, especially during the 8 quarters prior to the minimum wage increase. Indeed, this pattern is visible even in the results in NSW (2014b); by failing to average the leading coefficients across quarters, NSW focused on the noise instead of the signal. Pre-existing trends in the form of sizable leading coefficients for the minimum wage should make researchers nervous about that particular research design. Consistent with the idea that the state effects are not truly "fixed," we find that when we estimate the two-way fixed-effects model in first-differences instead of deviations-from-means (while allowing for three years of lags in minimum wages), the employment estimates become close to zero, or positive. Finally, we note that by including up to 12 quarters of lagged minimum wages and using CPS data, we show that state linear trends capture pre-existing trends and not delayed effects as discussed in Wolfers (2006) and Meer and West (2015), a point confirmed by the small or positive estimates for teen employment from the first-difference distributed lag specification recommended by Meer and West (2015).

Moreover, these results for teen employment are corroborated by Gittings and Schmutte (2015), who use state-level variation and QWI data, by DLR (2015), who use a border discontinuity design with contiguous counties and QWI data, and by Dube and Zipperer (2015), who use a synthetic controls approach. Totty (2015) also obtains similar results using factor model estimators. The matching estimator for teens in NSW (2014a) thus stands out as somewhat of an outlier in the recent minimum wage literature on teen employment. As we show here, NSW's results flow from the problematic nature of their sample and the particularities of their estimation strategy. Their estimator is riddled with internal inconsistencies, most importantly mixing treatment and control groups. Most of their "treatment events" do not actually receive a sizable increase in the minimum wage relative to controls. When we examine the small number of events in NSW's sample that do exhibit a sizable increase in the minimum wage relative to controls, we find that the reductions in employment are substantially less than in NSW's overall estimate. Finally, as a demonstration of the poor quality of matches in NSW's sample (likely driven by mixing treatment and control events, as well as a very short period used to estimate synthetic control weights), we show that the employment patterns are highly sensitive to the choice of their pre-intervention period. This sensitivity indicates that NSW's treatment and control groups were likely not following parallel trends, contrary to what they had assumed.

To conclude, the key claims made by NSW (2014a, 2014b) do not withstand scrutiny. There is, indeed, a substantial negative bias in the minimum wage employment elasticities estimated using the canonical two-way fixed-effects model. This model incorrectly assumes parallel trends across states in the U.S. A variety of methods of addressing that bias has led us, as well as other researchers, to the conclusion that the employment effect of U.S. minimum wage policies on low-wage employment to date have been fairly small; and these effects are fairly precisely estimated for the "medium run," including 3-4 years after minimum wage increases.

To be sure, researchers' recent engagement with these issues has had a positive impact on the minimum wage literature. In particular, the controversies have shed light on the nature of identifying variation, on the more reliable methods, and on narrowing the range of credible employment estimates.

Of course, these studies are based upon state and federal minimum wage changes between 1979 and 2014, a period when the federal minimum wage was relatively low by both historical and international standards (Dube 2014). For instance, only about 6 percent of workers were directly affected by the minimum wage increases over this period (Autor, Manning and Smith 2015). In 2014 and 2015, a number of cities around the U.S. enacted local minimum wage policies that will ultimately have much greater bite, in some cases directly affecting up to 35 percent of covered employment (Reich et al. 2015). Future research will tell us whether the impacts of such high minimum wages will differ from the effects of the policies studied in this paper.

References

[1] Aaronson, Daniel, Eric French and Isaac Sorkin 2015. “Firm Dynamics and the Minimum Wage: a Putty-Clay Approach.” Working Paper 2013-26. Research Department, Federal Reserve Board of Chicago. https://www.chicagofed.org/digital_assets/publications/working_papers/2013/wp2013_26.pdf

[2] Abadie, Alberto, Alexis Diamond and Jens Hainmueller 2010. “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program.” *Journal of the American Statistical Association* 105, 490: 493-505.

[3] Abowd, John et al. 2009. ”The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators.” Pp. 149-230 in Timothy Dunne, Bradford Jensen and Mark Roberts eds. *Producer Dynamics: New Evidence from Micro Data*. Chicago: University of Chicago and NBER.

[4] Addison, John, McKinley Blackburn and Chad Cotti 2014. “On the Robustness of Minimum Wage Effects: Geographically-Disparate Trends and Job Growth Equations.” IZA Discussion Paper 8420, Institute for the Study of Labor (IZA). http://ftp.iza.org/dp8420.pdf

[5] Allegretto, Sylvia and Carl Nadler 2015. “Tipped Wage Effects on Earnings and Employment in Full-Service Restaurants.” Forthcoming, *Industrial Relations* 54, 4, October 2015.

[6] Allegretto, Sylvia, Arindrajit Dube and Michael Reich 2011. “Do Minimum Wages Really Reduce Teen Employment? Accounting for Heterogeneity and Selectivity in State Panel Data.” *Industrial Relations* 50, 2: 205-40.

[7] Allegretto, Sylvia, Arindrajit Dube, Michael Reich and Ben Zipperer 2013. “Credible Research Designs for Minimum Wage Studies.” Working Paper 148-13. Institute for Research on Labor and Employment, UC Berkeley. http://www.irle.berkeley.edu/workingpapers/148-13.pdf

[8] Angrist, Joshua and Jörn-Steffen Pischke 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton: Princeton University Press.

[9] Autor, David 2003 .”Do Temporary Help Jobs Improve Labor Market Outcomes for Low-Skilled Workers?” *Quarterly Journal of Economics*118, 4: 1279-1334.

[10] Autor, David and David Dorn 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103, 5: 1553–1597.

[11] Autor, David, Alan Manning and Christopher Smith 2015. "The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment." MIT Department of Economics Working Paper `http://economics.mit.edu/files/3279`

[12] Bai, Jushan 2009. "Panel Data Models With Interactive Fixed Effects." *Econometrica* 77: 1229–1279.

[13] Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen. 2012. "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain." *Econometrica* 80,6: 2369–2429.

[14] Belloni, Alexandre, Victor Chernozhukov and Christian Hansen 2014. "High-Dimensional Methods and Inference on Treatment and Structural Effects in Economics." *Journal of Economic Perspectives* 28, 2: 29-50.

[15] Bohn, Sarah, Magnus Lofstrom and Steven Raphael 2014. "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" *Review of Economics and Statistics* 96, 2: 258-69.

[16] Brown, Charles 1999. "Minimum Wages, Employment, and the Distribution of Income." in Orley C. Ashenfelter and David Card, eds., Vol. 3, Part B of *Handbook of Labor Economics*, Elsevier, pp. 2101 – 2163.

[17] Dube, Arindrajit. 2014. "Designing Thoughtful Minimum Wage Policy at the State and Local Levels." The Hamilton Project. `http://www.hamiltonproject.org/papers/designing_minimum_wage_policy_at_state_and_local_levels/`

[18] Dube, Arindrajit, T. William Lester and Michael Reich 2010. "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics* 92, 4: 945-64.

[19] Dube, Arindrajit, T. William Lester and Michael Reich 2015. "Minimum Wage Shocks, Employment Flows and Labor Market Frictions." *Journal of Labor Economics*, forthcoming. `https://arindube.files.wordpress.com/2014/11/dlr2r3_fullpaper.pdf`

[20] Dube, Arindrajit and Ben Zipperer 2015. "Pooling Multiple Case Studies using Synthetic Controls: An Application to Minimum Wage Case Studies." IZA Discussion Paper 8944. `http://ftp.iza.org/dp8944.pdf`

[21] Gittings, Kaj and Ian Schmutte 2015. "Getting Handcuffs on an Octopus: Minimum Wages, Employment, and Turnover." *Industrial and Labor Relations Review*, forthcoming. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2234118

[22] Hansen, Christian 2013. "Econometric Methods for High-Dimensional Data." NBER Summer Institute. http://www.nber.org/econometrics_minicourse_2013/

[23] Hirsch, Barry T. and Schumacher, Edward J. 2004. "Match Bias in Wage Gap Estimates Due to Earnings Imputation." *Journal of Labor Economics*, 22, 3: 689-722.

[24] Huang, Yi, Prakash Loungani and Gewei Wang 2014. "Minimum Wages and Firm Employment: Evidence from China." IMF Working Paper 14/184. http://www.imf.org/external/pubs/ft/wp/2014/wp14184.pdf

[25] Krueger, Alan 1994. "The Effect of the Minimum Wage When It Really Bites: A Reexamination of the Evidence from Puerto Rico." NBER Working Paper 4757. http://www.nber.org/papers/w4757

[26] Magruder, Jeremy 2013. "Can Minimum Wages Cause a Big Push? Evidence from Indonesia." *Journal of Development Economics* 100, 1: 48-62.

[27] Meer, Jonathan and Jeremy West 2015. "Effects of the Minimum Wage on Employment Dynamics." *Journal of Human Resources*, forthcoming. http://econweb.tamu.edu/jmeer/Meer_West_Minimum_Wage.pdf

[28] Neumark, David and William Wascher 1992. "Employment Effects of Minimum and Sub-minimum Wage Laws:Panel Data on State Minimum Wage Laws." *Industrial and Labor Relations Review* 46,1: 55-81.

[29] Neumark, David and William Wascher 2011. "Does a Higher Minimum Wage Enhance the Effectiveness of the Earned Income Tax Credit?," *Industrial and Labor Relations Review* 64, 4: 712-746.

[30] Neumark, David, J. M. Ian Salas and William Wascher 2014a. "Revisiting the Minimum Wage and Employment Debate: Throwing out the Baby with the Bathwater?" *Industrial and Labor Relations Review*, 67 (Supplement): 608-48.

[31] Neumark, David, J. M. Ian Salas and William Wascher 2014b. "More on Recent Evidence on the Effects of Minimum Wages in the United States." NBER Working Paper 20619. http://www.nber.org/papers/w20619

[32] Pesaran, M. Hashem. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74,4: 967–1012.

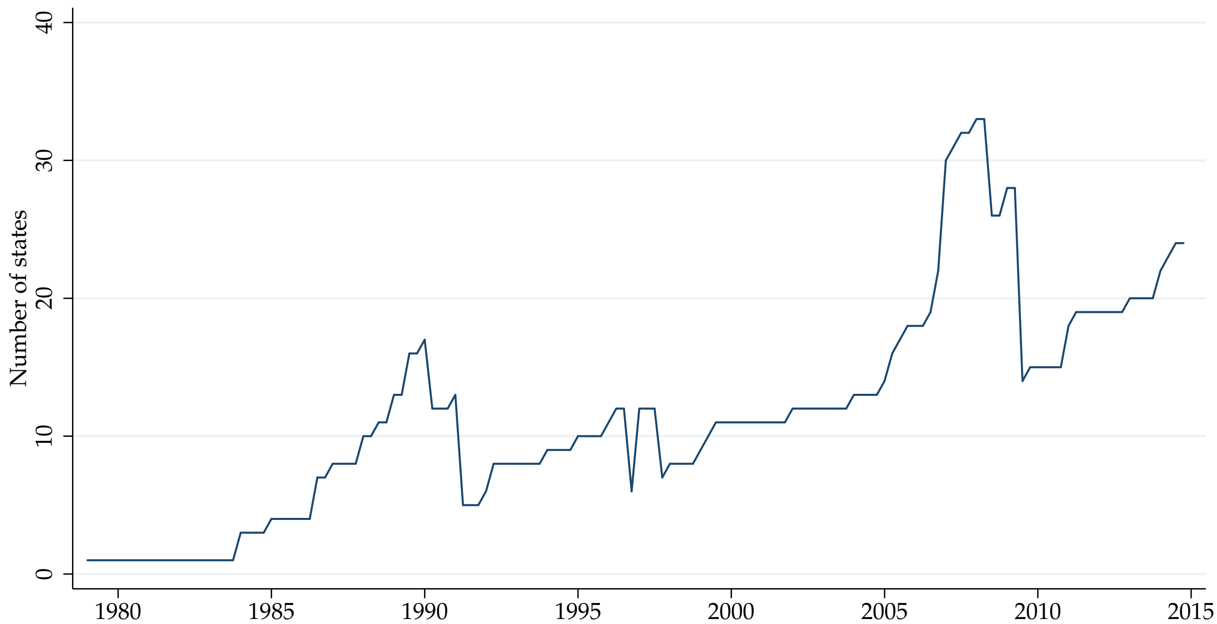
[33] Reich, Michael, Ken Jacobs, Annette Bernhardt and Ian Perry 2015. “The Proposed Minimum Wage Law for Los Angeles: Economic Impacts and Policy Options. Policy Brief, Center on Wage and Employment Dynamics, Institute for Research on Labor and Employment, UC Berkeley. http://irle.berkeley.edu/cwed/briefs/2015-01.pdf

[34] Totty, Evan 2015. “The Effect of Minimum Wages on Employment: A Factor Model Approach.” IRLE Working Paper 110-15. Institute for Research on Labor and Employment. UC Berkeley. http://www.irle.berkeley.edu/workingpapers/110-15.pdf

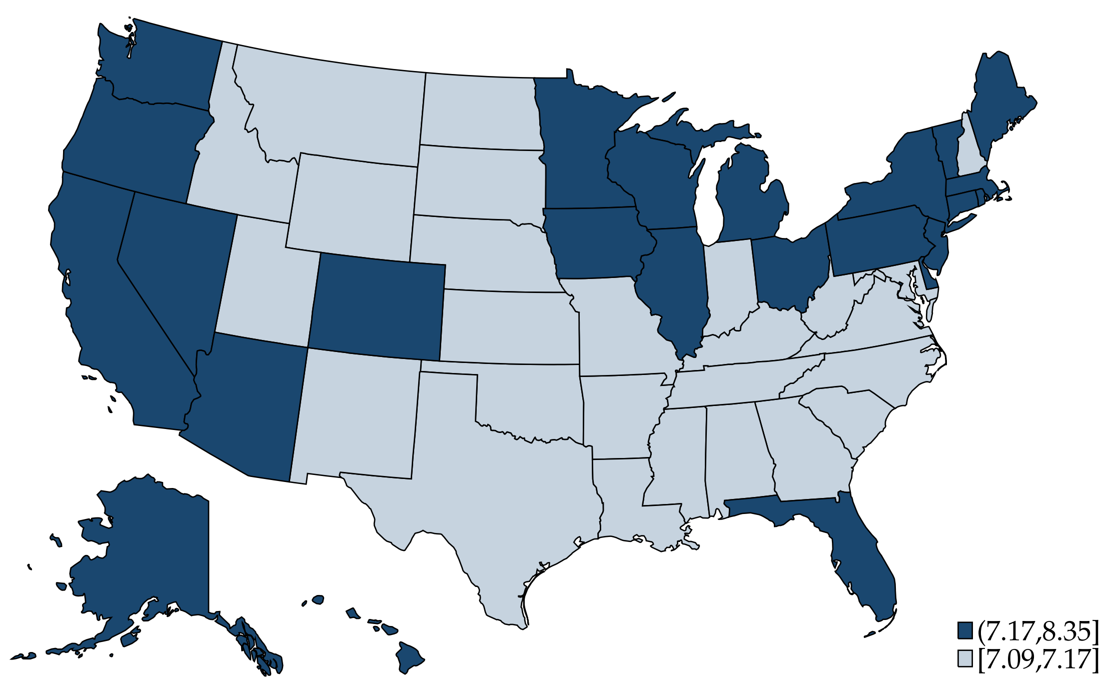
[35] Wolfers, Justin 2006. “Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96, 5: 1802-1820.

Figure 1: State-level minimum wage variation, 1979-2014

Panel A: States with minimum wages exceeding the federal floor



Panel B: High and low minimum wages



Notes: Panel A shows the number of states with quarterly minimum wages exceeding the quarterly federal minimum wage. Panel B divides states into those with high or low real minimum wages, where quarterly minimum wages are inflation-adjusted using the CPI-U-RS and averaged over 1979-2014. High (low) minimum wage states have a minimum above (below) the sample median.

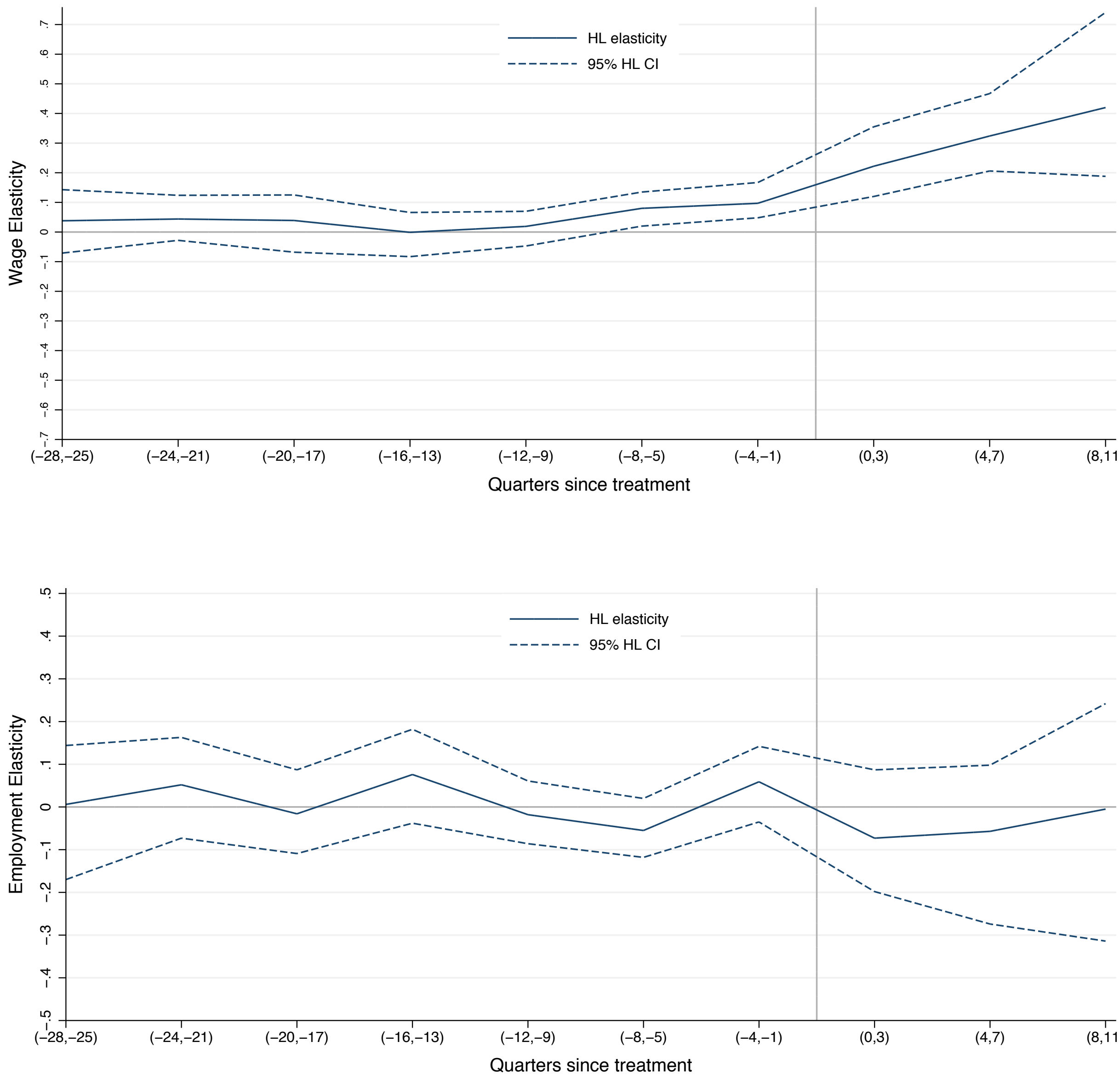
Figure 2: Cumulative response of teen employment to minimum wages, individual-level CPS data, 1979-2014

Notes: The figure shows cumulative response elasticities of teen employment to the minimum wage; 95% confidence intervals are from the individual-level distributed lag regressions for the 1979-2014 period described in Section 2.3. For each of the four regression models, the figure shows the quarterly effects and confidence intervals in blue and the 4-quarter averaged effects and confidence intervals in green. Standard errors are clustered by state.

Figure 3: Cumulative response of teen employment to minimum wages, state-quarter data from NSW (2014b), 1990-2011q1

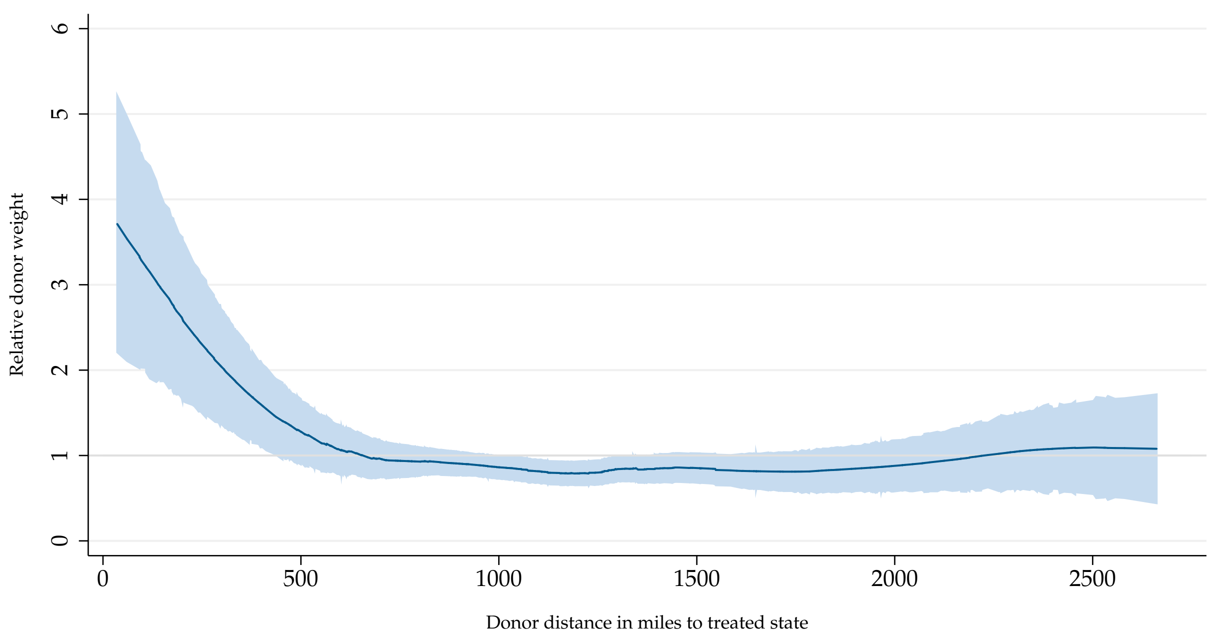
Notes: The figure shows cumulative response elasticities of teen employment to the minimum wage; 95% confidence intervals are from the distributed lag regressions described in Section 2.3, with the state-quarter aggregated data and control set used in NSW (2014b), where the outcome is log of the teen employment-to-population ratio. For each of the four regression models, the figure shows the quarterly effects and confidence intervals in blue and the 4-quarter averaged effects and confidence intervals in green. Standard errors are clustered by state. Red colored marker for $t = -2$ indicates the baseline used by NSW.

Figure 4: Time paths of synthetic control teen wage and employment elasticities



Notes: Reproduced from Dube and Zipperer (2015), Figure 3. The figure shows the time paths of the pooled synthetic control minimum wage elasticities across 29 events. The top panel shows the time path for the average teen wage, and the bottom panel for teen employment. Elasticities are averaged over 4-quarter increments in event time, where 0 is the quarter of the minimum wage increase. For estimation details see section 4.1. The solid blue line shows the Hodges-Lehmann point estimate and the dashed blue line shows the associated 95 percent confidence interval.

Figure 5: Donor distance and relative weights



Notes: The figure shows the locally weighted regression (lowess) of the relative donor weights on donor distance to treated states for the 25 treatment events, excluding Alaska and Hawaii, as described in Section 4.2, where the 95 percent confidence interval is calculated from 1000 cluster bootstrap iterations at the treatment event level.

Table 1
Minimum wage elasticities for average teen wage and employment controlling for time varying heterogeneity, individual-level CPS data 1979-2014

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average teen wage						
Common time FE	0.266***	0.228***	0.226***	0.271***	0.268***	0.267***
	(0.038)	(0.020)	(0.022)	(0.032)	(0.031)	(0.037)
N	295,835	295,835	295,835	295,835	295,835	295,835
Division-period FE	0.245***	0.253***	0.232***	0.227***	0.211***	0.212***
	(0.036)	(0.033)	(0.037)	(0.037)	(0.034)	(0.035)
N	295,835	295,835	295,835	295,835	295,835	295,835
Panel B: Teen employment						
Common time FE	-0.219***	-0.065	-0.044	-0.066	-0.091	-0.068
	(0.043)	(0.041)	(0.061)	(0.066)	(0.065)	(0.072)
N	3,534,924	3,534,924	3,534,924	3,534,924	3,534,924	3,534,924
Division-time FE	-0.130*	0.006	-0.012	-0.023	-0.040	-0.039
	(0.077)	(0.047)	(0.048)	(0.040)	(0.039)	(0.041)
N	3,534,924	3,534,924	3,534,924	3,534,924	3,534,924	3,534,924
State-specific trend type:						
Linear		Y	Y	Y	Y	Y
Quadratic			Y	Y	Y	Y
Cubic				Y	Y	Y
Quartic					Y	Y
Quintic						Y

Notes: The table reports minimum wage elasticities for average teen wage and employment, using individual-level Current Population Survey data from 1979-2014 (basic monthly data for employment, and Outgoing Rotation Groups for wage). The dependent variable is either log wage, or a binary employment indicator. For the wage outcome, the table reports the coefficients on log quarterly minimum wage. For employment, the estimates are converted to elasticities by dividing the coefficients on log minimum wage (and standard error) by the sample mean employment rate. All regressions include controls for the quarterly state unemployment rate, the quarterly teen share of the working age population, dummies for demographic variables as described in Section 2, and state fixed effects. As reported in the table, specifications either include common period fixed effects or Census division-period fixed effects, with up to fifth order state-specific polynomial trends. Regressions are weighted by sample weights, robust standard errors (in parentheses) are clustered at the state level and significance levels are indicated by *** 1%, ** 5%, * 10%.

Table 2
Model selection: Minimum wage elasticities for teen employment, state-quarter aggregated CPS data 1979-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full sample (1979-2014)								
Common time FE	-0.172**	0.022	0.001	-0.055	-0.087	-0.073		
	(0.066)	(0.082)	(0.076)	(0.079)	(0.082)	(0.086)		
N	7,344	7,344	7,344	7,344	7,344	7,344		
Division-period FE	-0.043	0.056	0.056	0.035	0.004	0.004		
	(0.086)	(0.057)	(0.057)	(0.049)	(0.049)	(0.056)		
N	7,344	7,344	7,344	7,344	7,344	7,344		
LASSO-selected division-period FE							0.013	-0.012
							(0.083)	(0.083)
N							7,344	7,344
Panel B: Post-1990 sample (1990-2014)								
Common time FE	-0.106	0.003	-0.060	-0.148**	-0.178**	-0.211***		
	(0.066)	(0.079)	(0.066)	(0.067)	(0.068)	(0.064)		
N	5,100	5,100	5,100	5,100	5,100	5,100		
Division-period FE	-0.029	0.069	0.043	-0.013	-0.022	-0.060		
	(0.092)	(0.063)	(0.061)	(0.057)	(0.070)	(0.062)		
N	5,100	5,100	5,100	5,100	5,100	5,100		
LASSO-selected division-period FE							-0.009	-0.031
							(0.073)	(0.070)
N							5,100	5,100
State-specific trend type:								
Linear		Y	Y	Y	Y	Y		
Quadratic			Y	Y	Y	Y		
Cubic				Y	Y	Y		
Quartic					Y	Y		
Quintic						Y		
LASSO-selected trends (linear only)							Y	
LASSO-selected trends (up to quintic)								Y

Notes: The table reports minimum wage elasticities for teen employment, using state-quarter aggregated Current Population Survey basic monthly data from 1979-2014. The dependent variable is the log of the state-quarter sample-weighted mean of teen employment. The reported estimates are coefficients for log quarterly minimum wage. All regressions include controls for the overall quarterly state unemployment rate, the quarterly teen share of the working age population, and state-quarter means for demographic controls used in Table 1 and described in Section 2, and state fixed effects. Specifications include either common period effects, or Census division-period effects, and up to fifth order polynomial trends by state. Columns 7-8 report double-selection post-LASSO estimates where controls (besides state and period effects) are selected using LASSO regressions predicting teen employment and minimum wage: these include demographic controls, division-period effects, and state-specific trends (linear in specification 7; up to quintic in specification 8). Regressions are weighted by teen population. Robust standard errors in parentheses are clustered by state; significance levels are *** 1%, ** 5%, * 10%.

Table 3
Dynamic minimum wage elasticities for teen employment, individual-level CPS data 1979-2014

		(1)	(2)	(3)	(4)
Panel A: 4-quarter averages of cumulative response elasticities					
A.	$\bar{\rho}_{[-12,-9]}$	-0.145**	-0.092	-0.059	-0.006
		(0.072)	(0.057)	(0.050)	(0.046)
B.	$\bar{\rho}_{[-8,-5]}$	-0.202**	-0.204**	-0.105	-0.096
		(0.090)	(0.079)	(0.072)	(0.065)
C.	$\bar{\rho}_{[-4,-1]}$	-0.187**	-0.153	-0.057	0.006
		(0.085)	(0.113)	(0.062)	(0.094)
D.	$\bar{\rho}_{[0,3]}$	-0.278***	-0.207	-0.113**	-0.000
		(0.068)	(0.130)	(0.051)	(0.098)
E.	$\bar{\rho}_{[4,7]}$	-0.379***	-0.303*	-0.183***	-0.043
		(0.075)	(0.164)	(0.056)	(0.129)
F.	$\bar{\rho}_{[8,11]}$	-0.315***	-0.219	-0.117*	0.067
		(0.094)	(0.159)	(0.062)	(0.118)
G.	ρ_{12+}	-0.310***	-0.217	-0.004	0.158
		(0.114)	(0.191)	(0.064)	(0.127)
Panel B: Medium run (3 year) elasticities					
F-A	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-12,-9]}$	-0.170***	-0.127	-0.058	0.073
		(0.049)	(0.119)	(0.050)	(0.089)
F-B	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-8,-5]}$	-0.113***	-0.015	-0.012	0.164**
		(0.041)	(0.096)	(0.054)	(0.068)
F-C	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-4,-1]}$	-0.128***	-0.065	-0.060	0.061
		(0.038)	(0.070)	(0.043)	(0.050)
Panel C: Long run (4+ years) elasticities					
G-A	$\rho_{12+} - \bar{\rho}_{[-12,-9]}$	-0.165**	-0.126	0.055	0.164
		(0.068)	(0.153)	(0.063)	(0.101)
G-B	$\rho_{12+} - \bar{\rho}_{[-8,-5]}$	-0.108*	-0.013	0.101	0.255***
		(0.058)	(0.133)	(0.072)	(0.083)
G-C	$\rho_{12+} - \bar{\rho}_{[-4,-1]}$	-0.123*	-0.064	0.053	0.152**
		(0.062)	(0.108)	(0.060)	(0.065)
Division-period FE			Y		Y
State-specific linear trends				Y	Y

Notes: The table reports cumulative response elasticities of teen employment with respect to minimum wages using individual level CPS basic monthly data from 1979-2014. Regressions include the contemporaneous, 12 quarterly leads and 12 quarterly lags of log minimum wage. The dependent variable is a binary employment indicator and estimates are converted to elasticities by dividing the log minimum wage coefficients and standard errors by the sample mean employment rate. Panel A reports four quarter averages of the cumulative response elasticities starting at t=-12 in quarterly event time, as described in Section 2.3. Panel B reports the cumulative effect in year 3, after subtracting alternative baseline levels at 1, 2 or 3 years prior to treatment, as indicated. Panel C reports the long run cumulative response elasticity at t=12 or later, after subtracting alternative baseline levels. All regressions include controls for the overall quarterly state unemployment rate, the quarterly teen share of the working age population, dummies for demographic controls used in Table 1, and as described in Section 2, and state and period fixed effects. Specifications may additionally include Census division-period fixed effects and state-specific linear trends. Regressions are weighted by sample weights, robust standard errors are clustered at the state level and significance levels are indicated by *** 1%, ** 5%, * 10%.

Table 4
Dynamic minimum wage elasticities for teen employment, state-quarter aggregated CPS data 1990-2011q1

		(1)	(2)	(3)	(4)
Panel A: 4-quarter averages of cumulative response elasticities					
A.	$\bar{\rho}_{[-12,-9]}$	0.015	-0.027	0.081	0.058
		(0.059)	(0.072)	(0.055)	(0.062)
B.	$\bar{\rho}_{[-8,-5]}$	-0.126	-0.205*	-0.038	-0.073
		(0.080)	(0.113)	(0.073)	(0.104)
C.	$\bar{\rho}_{[-4,-1]}$	-0.118	-0.143	0.005	0.056
		(0.085)	(0.152)	(0.083)	(0.133)
D.	$\bar{\rho}_{[0,3]}$	-0.169**	-0.170	-0.006	0.090
		(0.074)	(0.184)	(0.101)	(0.145)
E.	$\bar{\rho}_{[4,7]}$	-0.338***	-0.350	-0.144	-0.024
		(0.066)	(0.216)	(0.110)	(0.173)
F.	$\bar{\rho}_{[8,11]}$	-0.166**	-0.177	0.019	0.176
		(0.081)	(0.226)	(0.108)	(0.168)
G.	ρ_{12+}	-0.192**	-0.159	0.144	0.323*
		(0.092)	(0.289)	(0.144)	(0.187)
Panel B: Medium run (3 year) elasticities					
F-A	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-12,-9]}$	-0.181***	-0.149	-0.062	0.118
		(0.056)	(0.172)	(0.087)	(0.126)
F-B	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-8,-5]}$	-0.040	0.028	0.057	0.249***
		(0.062)	(0.137)	(0.093)	(0.089)
F-C	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-4,-1]}$	-0.047	-0.033	0.014	0.120
		(0.068)	(0.113)	(0.088)	(0.077)
Panel C: Longer run (4+ years) elasticities					
G-A	$\rho_{12+} - \bar{\rho}_{[-12,-9]}$	-0.207***	-0.132	0.063	0.265*
		(0.073)	(0.239)	(0.125)	(0.151)
G-B	$\rho_{12+} - \bar{\rho}_{[-8,-5]}$	-0.066	0.045	0.181	0.395***
		(0.089)	(0.208)	(0.138)	(0.119)
G-C	$\rho_{12+} - \bar{\rho}_{[-4,-1]}$	-0.074	-0.016	0.139	0.267**
		(0.074)	(0.181)	(0.124)	(0.110)
Division-period FE			Y		Y
State-specific linear trends				Y	Y

Notes: The table reports cumulative response elasticities of teen employment with respect to minimum wages using state-quarter aggregated CPS basic monthly data from1990-2011q1, taken from from NSW (2014b) replication package. Regressions include the contemporaneous, 12 quarterly leads and 12 quarterly lags of log minimum wage. The dependent variable is a binary employment indicator and estimates are converted to an elasticity by dividing the log minimum wage coefficients and standard errors by the sample mean employment rate. Panel A reports four quarter averages of the cumulative response elasticities starting at t=-12 in quarterly event time, as described in Section 2.3. Panel B reports the cumulative effect in year 3, after subtracting alternative baseline levels at 1, 2 or 3 years prior to treatment, as indicated. Panel C reports the long run cumulative response elasticity at t=12 or later, after subtracting alternative baseline levels. All regressions include controls for the overall quarterly state unemployment rate, the quarterly teen share of the working age population, and state and period fixed effects. Specifications may additionally include Census division-period fixed effects and state-specific linear trends. Regressions are weighted by sample weights, robust standard errors are clustered at the state level and significance levels are indicated by *** 1%, ** 5%, * 10%.

Table 5
Minimum wage elasticities for teen employment: deviations-from-means versus first-difference estimates, state-year aggregated CPS data 1979-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Deviations-from-means								
Population Weighted								
Contemporaneous MW elasticity	-0.161**	-0.011	0.052	0.106*	0.004	0.074	0.037	0.107*
	(0.077)	(0.086)	(0.098)	(0.063)	(0.083)	(0.071)	(0.079)	(0.062)
3 year cumulative MW elasticity	-0.154	0.006	0.225*	0.243**	-0.086	0.048	0.134	0.231**
	(0.118)	(0.173)	(0.127)	(0.104)	(0.098)	(0.153)	(0.108)	(0.112)
Unweighted								
Contemporaneous MW elasticity	-0.163**	-0.033	0.008	0.110	-0.045	-0.023	-0.009	0.020
	(0.068)	(0.088)	(0.068)	(0.077)	(0.074)	(0.083)	(0.075)	(0.082)
3 year cumulative MW elasticity	-0.231**	-0.135	0.135*	0.174*	-0.160*	-0.092	0.079	0.126
	(0.088)	(0.129)	(0.071)	(0.092)	(0.080)	(0.131)	(0.071)	(0.093)
Division-Period FE		Y		Y		Y		Y
State-specific linear trends			Y	Y			Y	Y
Controls for leads in Min. Wage					Y	Y	Y	Y
Panel B: First-difference								
Population Weighted								
Contemporaneous MW elasticity	0.029	0.089	0.038	0.096	0.021	0.086	0.030	0.093
	(0.085)	(0.059)	(0.089)	(0.059)	(0.080)	(0.060)	(0.084)	(0.061)
3 year cumulative MW elasticity	0.147	0.354**	0.167	0.369**	0.120	0.394**	0.151	0.421**
	(0.141)	(0.147)	(0.147)	(0.150)	(0.134)	(0.166)	(0.147)	(0.177)
Unweighted								
Contemporaneous MW elasticity	-0.014	-0.005	-0.007	0.002	-0.036	-0.003	-0.031	0.005
	(0.064)	(0.073)	(0.066)	(0.075)	(0.074)	(0.079)	(0.077)	(0.081)
3 year cumulative MW elasticity	0.056	0.076	0.073	0.097	-0.016	0.111	0.002	0.137
	(0.078)	(0.116)	(0.081)	(0.120)	(0.083)	(0.119)	(0.090)	(0.127)
Division-Period FE		Y		Y		Y		Y
State FE			Y	Y			Y	Y
Controls for leads in Min. Wage					Y	Y	Y	Y

Notes: The table reports contemporaneous and 3 year cumulative minimum wage elasticities for teen employment using state-year aggregated CPS basic monthly data from1990-2011q1. All specifications include the contemporaneous log annual minimum wage, and three years of lags of the log annual minimum wage, in levels or differences. The dependent variable is the log of the state-year sample-weighted mean of teen employment (in levels or differences). All regressions include controls for the overall quarterly state unemployment rate, the quarterly teen share of the working age population, and state-year means for demographic controls used in Table 1, and as described in Section 2, in levels or differences. The table reports the coefficient on the contemporaneous log minimum and the sum of the contemporaneous and lagged terms. Estimates in Panel A are from the deviation-from-means estimator, and estimates in panel B are from the first-difference estimator. The deviation-from-means specifications always include state fixed effects, and may additionally include state linear trends as indicated. The first difference specifications may additionally include state fixed effects as indicated. All specifications include period fixed effects, and may additionally include division-period effects as indicated. Columns 5-8 additionally control for three years of leading minimum wages (in levels or differences). Regressions are unweighted or weighted by the state-year teen population size, as indicated, robust standard errors (in parentheses) are clustered at the state level and significance levels are indicated by *** 1%, ** 5%, * 10%.

Table 6
Minimum wage elasticities for teens using border discontinuity design, QWI data 2000-2011

	(1)	(2)
Earnings	0.177***	0.222***
	(0.036)	(0.047)
Employment	-0.173***	-0.059
	(0.071)	(0.084)
Turnover	-0.377***	-0.204***
	(0.061)	(0.072)
Contiguous county pair-period FE		Y

Notes: Reproduced from Dube, Lester and Reich (2015), Table 3, using county-level QWI data on 14-18 year olds from 2000-2011. The dependent variables are log quarterly teen employment, quarterly average teen earnings, or the teen turnover rate, and the table reports the coefficients on the log quarterly minimum wage. Column 1 includes period fixed effects, and county fixed effects. Column 2 additionally includes contiguous county-pair-period fixed effects. All regressions include controls for log of overall population and private-sector employment. Standard errors are clustered multi-dimensionally at the state-level and border segment-level. County pairs restricted to those whose centroids lie within 75km of each other; the cut-off is based on a data-driven procedure described in Dube, Lester and Reich (2015). Significance levels are indicated by *** 1%, ** 5%, * 10%.

Table 7
Pooled synthetic control minimum wage elasticities for average teen wage and employment

	Median elasticity	Mean elasticity	Mean rank	Hodges-Lehmann Elasticity	Hodges-Lehmann 95% CI
Average Teen Wage	0.237	0.368	0.758***	0.266	(0.169, 0.414)
Teen Employment	-0.051	-0.058	0.470	-0.036	(-0.170, 0.087)

Notes: The table reproduces the synthetic control minimum wage elasticities for average teen wage and teen employment pooled across 29 cases, as originally reported in Dube and Zipperer (2015), Table 6. The median, mean and Hodges-Lehmann pooled elasticities are reported, as are the 95 percent Hodges-Lehmann confidence intervals, which are based on inverting the mean rank statistic. Critical values for the mean percentile rank are derived from the mean of 29 uniform distributions. Significance levels on the mean rank statistic are indicated by *** 1%, ** 5%, * 10%.

Table 8
Re-analysis of results from NSW matching estimator: difference-in-differences estimates

		NSW pre-treatment period (t=-4,-3,-2,-1)			Earlier pre-treatment period (t= -8,-7,-6,-5)		
		Δ log MW	Δ log teen emp	MW Elasticity	Δ log MW	Δ log teen emp	MW Elasticity
Overall	Treatment	0.098	-0.048		0.160	-0.080	
		(0.003)	(0.008)		(0.006)	(0.012)	
	Control	0.071	-0.042		0.122	-0.088	
		(0.003)	(0.009)		(0.004)	(0.013)	
	Treatment - Control	**0.027*****	**-0.007***	**-0.247***	**0.038*****	**0.008**	**0.205**
		(0.003)	**(0.004)**	**(0.128)**	**(0.005)**	**(0.006)**	**(0.156)**
Quartile 1	Treatment	0.055	-0.058		0.118	-0.119	
		(0.006)	(0.016)		(0.012)	(0.019)	
	Control	0.080	-0.046		0.146	-0.102	
		(0.004)	(0.011)		(0.007)	(0.015)	
	Treatment - Control	**-0.024*****	**-0.012**	**0.490**	**-0.027*****	**-0.018****	**0.646**
		(0.003)	**(0.009)**	**(0.382)**	**(0.009)**	**(0.009)**	**(0.437)**
Quartile 2	Treatment	0.101	-0.029		0.153	-0.050	
		(0.003)	(0.011)		(0.004)	(0.016)	
	Control	0.096	-0.028		0.146	-0.064	
		(0.003)	(0.008)		(0.004)	(0.014)	
	Treatment - Control	**0.005*****	**0.000**	**-0.106**	**0.007****	**0.014**	**1.938**
		(0.000)	**(0.010)**	**(2.043)**	**(0.003)**	**(0.011)**	**(1.458)**
Quartile 3	Treatment	0.103	-0.055		0.171	-0.092	
		(0.005)	(0.021)		(0.007)	(0.030)	
	Control	0.075	-0.049		0.129	-0.121	
		(0.003)	(0.022)		(0.006)	(0.033)	
	Treatment - Control	**0.028*****	**-0.006**	**-0.205**	**0.041*****	**0.029****	**0.695****
		(0.002)	**(0.006)**	**(0.225)**	**(0.003)**	**(0.012)**	**(0.307)**
Quartile 4	Treatment	0.133	-0.044		0.200	-0.048	
		(0.010)	(0.015)		(0.008)	(0.018)	
	Control	(0.033)	-0.037		0.068	-0.052	
		(0.005)	(0.013)		(0.007)	(0.017)	
	Treatment - Control	**0.099*****	**-0.007**	**-0.074**	**0.132*****	**0.004**	**0.029**
		(0.007)	**(0.007)**	**(0.075)**	**(0.009)**	**(0.013)**	**(0.097)**

Notes: The table reports mean differences of log minimum wage and log teen employment rate for both control and treatment groups between post-treatment period (t=0, … , 3) and pre-treatment period (t=-4 , … , -1) using the NSW (2014a) sample of 493 events, as well as between post-treatment period and earlier pre-treatment period (t=-8, … ,-5), using the available sub-sample of 442 events. "Treatment – Controls" rows are difference-in-difference (DD) estimates, in boldface. The top panel reports the estimates for the overall samples. The subsequent panels report estimates from four quartiles of the extent of treatment (i.e., DD in log minimum wage). Minimum wage elasticities are obtained by dividing DD estimate for log teen employment by the DD estimate for log minimum wage. Robust standard errors (in parentheses) of elasticities are clustered at the state level and calculated using "suest" command in STATA. Significance levels are indicated only for the DD estimates by *** 1%, ** 5%, * 10%.

Table 9
Summary of literature: minimum wage elasticities for restaurant employment

	Sample years	Two-way FE	Additional controls										
Addison, Blackburn, and Cotti (2014)[c]	1990-2005	-0.101**	-0.006	-0.051***	-0.041	-0.062*	-0.046						
		(0.039)	(0.033)	(0.014)	(0.027)	(0.033)	(0.033)						
	1990-2012	(0.000)	-0.040*	-0.024	-0.035*	-0.023*	-0.01						
		(0.035)	(0.021)	(0.018)	(0.019)	(0.014)	(0.014)						
Dube, Lester, and Reich (2010)[a]	1990-2006	-0.176*						0.039	0.016				
		(0.096)						(0.050)	(0.098)				
Dube, Lester, and Reich (2015)[b]	2000-2011	-0.073*							-0.022				
		(0.042)							(0.091)				
Neumark, Salas, and Wascher (2014)[a]	1990-2006	-0.120***								-0.063***			
		(0.042)								(0.022)			
Totty (2015)[a]	1990-2010	-0.138*									-0.013	-0.013	-0.042
		[-0.297, 0.019]									[-0.042, 0.026]	[-0.046, 0.028]	[-0.085, 0.015]
Linear state trends			Y	Y	Y	Y	Y	Y					
Quadratic state trends				Y	Y	Y	Y						
Cubic state trends					Y	Y	Y						
Quartic state trends						Y	Y						
Cubic state trends							Y						
Census division-period FE								Y					
Contiguous county pair-period FE									Y				
NSW matching estimator										Y			
CCE-P											Y		
CCE-MG												Y	
IFE													Y

Sources: Addison, Blackburn, and Cotti (2014) Table 1; Allegretto, Dube, and Reich (2013) Table 3; Dube, Lester, and Reich (2010) Table 2; Dube, Lester, and Reich (2015) Table 3; Neumark, Salas, and Wascher (2014a) Table 8; Totty (2015) Table 3. *Notes:* Significance levels are indicated by *** 1%, ** 5%, * 10%. Standard errors in parentheses, except for Totty (2015) which report 95 percent confidence interval based on wild cluster bootstrap-t clustered at state level. Restaurant employment is defined as total employment in NAICS 7221 and NAICS 7222 using QCEW data[a]; employment in NAICS 722 using QWI data[b]; or employment in NAICS 722 using QCEW data[c]. The column labeled "Sample years" shows the years included in the data. CCE-P is common correlated effect, pooled; CCE-MG is common correlated effect, mean group. IFE is interactive fixed effect.

Table 10
Minimum wage elasticities for restaurant earnings and employment, QCEW data 1990-2014

		(1)	(2)	(3)
Panel A Contemporaneous minimum wage elasticities				
Earnings		0.233***	0.246***	0.208***
		(0.025)	(0.030)	(0.062)
Employment		-0.242***	-0.186**	0.022
		(0.076)	(0.076)	(0.129)
Panel B 4-quarter averages of cumulative response elasticities for employment				
A.	$\bar{\rho}_{[-12,-9]}$	-0.119***	-0.047	0.010
		(0.046)	(0.047)	(0.097)
B.	$\bar{\rho}_{[-8,-5]}$	-0.138***	-0.06	0.018
		(0.046)	(0.054)	(0.140)
C.	$\bar{\rho}_{[-4,-1]}$	-0.201***	-0.124*	0.010
		(0.058)	(0.064)	(0.161)
D.	$\bar{\rho}_{[0,3]}$	-0.280***	-0.169**	0.015
		(0.078)	(0.083)	(0.186)
E.	$\bar{\rho}_{[4,7]}$	-0.333***	-0.208**	-0.026
		(0.088)	(0.094)	(0.215)
F.	$\bar{\rho}_{[8,11]}$	-0.360***	-0.208*	-0.014
		(0.106)	(0.107)	(0.235)
G.	ρ_{12+}	-0.510***	-0.353**	-0.063
		(0.148)	(0.158)	(0.307)
Panel C Medium run (3 year) elasticities for employment				
F-A	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-12,-9]}$	-0.241***	-0.161**	-0.024
		(0.075)	(0.079)	(0.168)
F-B	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-8,-5]}$	-0.221***	-0.148**	-0.032
		(0.071)	(0.072)	(0.125)
F-C	$\bar{\rho}_{[8,11]} - \bar{\rho}_{[-4,-1]}$	-0.159***	-0.085	-0.024
		(0.056)	(0.059)	(0.105)
Panel D Long run (4+ years) elasticities for employment				
G-A	$\rho_{12+} - \bar{\rho}_{[-12,-9]}$	-0.391***	-0.306**	-0.073
		(0.116)	(0.129)	(0.249)
G-B	$\rho_{12+} - \bar{\rho}_{[-8,-5]}$	-0.371***	-0.293**	-0.081
		(0.114)	(0.125)	(0.211)
G-C	$\rho_{12+} - \bar{\rho}_{[-4,-1]}$	-0.309***	-0.229**	-0.073
		(0.098)	(0.110)	(0.185)
Sample:		All counties	Border county pairs	Border county pairs
County pair-period FE				Y

Table 10 (Continued)
Minimum wage elasticities for restaurant earnings and employment, QCEW data 1990-2014

Notes: All specifications use quarterly county-level data for Food Services and Drinking Places (NAICS 722) from the 1990-2014 QCEW. The dependent variable is log of county-quarter restaurant employment or average earnings, as indicated, and right-hand side controls include log of county-quarter population and overall private sector employment. Panel A specifications include only the contemporaneous log quarterly minimum wage, whose coefficients are reported. The specifications in Panels B through D include the contemporaneous, 12 quarterly leads and 12 quarterly lags of log minimum wage. Panel B reports 4-quarter averages of the cumulative response elasticities starting at t=-12 in quarterly event time. Panel C reports the cumulative effect in year 3, after subtracting alternative baseline levels at 1, 2 or 3 years prior to treatment, as indicated. Panel D reports long-run cumulative response elasticity at t=12, after subtracting alternative baseline levels. Column 1 uses the balanced panel of counties with common period fixed effects, column 2 uses the contiguous-border-county-pair sample with common period effects, and column 3 uses the contiguous-border-county-pair sample with pair-period effects. Robust standard errors in parentheses are clustered at the state-level in column 1 and clustered multi-dimensionally at the state-level and border segment-level in columns 2 and 3. Significance levels are indicated by *** 1%, ** 5%, * 10%.

Appendix A: The double-selection post-LASSO estimates of minimum wage elasticity for teen employment using alternative penalization parameters

Appendix Table A1 reports the double-selection post-LASSO minimum wage elasticities for teen employment using the 1979-2014 basic monthly CPS, varying the penalization parameter, λ, between the most saturated specification to the least saturated one. All estimates include state and period fixed effects, which are partialed out prior to the LASSO based covariate selection. The covariates that LASSO chooses from include demographic controls, unemployment rate, state-specific linear trends, and division-period effects. Appendix Table A1 shows the point estimates and the confidence intervals associated with varying λ between 0 (the most saturated model) and 3500 (which only picks the state unemployment rate as a control beyond the manually-specified two-way fixed effects). The point estimate quickly falls under -0.045 in magnitude as λ is lowered to 2,000 or below. For $\lambda = 2000$, the LASSO double-selection procedure includes just 5 state-specific linear trends lowers the elasticity in magnitude to -0.040. In other words, merely adding state-specific linear trends for these 5 states (which happen to be CA, SD, OR, WA and VT) to the fixed effects model produces an estimate that is close to zero, and not statistically significant.

Appendix Table A1
Double-selection post-LASSO estimates for minimum wage elasticity for teen employment, for alternative values of the LASSO penalization parameter, state-quarter aggregated CPS data 1979-2014

λ	# of state specific trends chosen	# of divisions selected	# division–period FE chosen	Estimate	Std. Error
0	.	.	.	0.056	0.057
16	50	8	1,139	0.055	0.057
50	48	8	1,113	0.053	0.057
84	47	8	1,091	0.055	0.057
118	48	8	1,053	0.049	0.058
152	48	8	1,012	0.035	0.059
186	47	8	948	0.018	0.061
220	47	8	860	0.005	0.061
254	46	8	779	-0.005	0.061
288	45	8	644	0.001	0.058
322	43	8	552	0.001	0.055
356	43	8	466	-0.004	0.059
390	44	8	387	-0.007	0.056
424	44	8	323	-0.001	0.053
458	43	8	258	-0.013	0.050
492	44	8	200	-0.003	0.060
526	43	8	155	0.006	0.064
560	42	7	116	-0.022	0.059
594	41	7	98	-0.025	0.063
628	41	7	83	-0.027	0.063
662	40	6	73	-0.025	0.063
696	40	5	64	-0.040	0.064
730	39	3	61	-0.038	0.060
764	39	2	53	-0.038	0.061
798	37	2	48	-0.036	0.060
832	36	2	46	-0.041	0.061
866	36	2	44	-0.040	0.061
900	34	1	27	0.010	0.082
934	32	1	20	0.013	0.083
968	30	1	22	0.010	0.083
1,008	29	1	19	0.005	0.084
1,048	25	1	19	0.002	0.084
1,088	24	1	18	0.005	0.084
1,128	20	1	16	0.008	0.084
1,168	18	1	14	-0.023	0.089
1,208	18	1	14	-0.023	0.089
1,248	17	1	14	-0.026	0.09
1,500	14	0	0	-0.030	0.085
1,750	10	0	0	-0.043	0.08
2,000	5	0	0	-0.045	0.081
2,250	3	0	0	-0.165***	0.039
2,500	2	0	0	-0.164***	0.040
2,750	1	0	0	-0.169***	0.040
3,000	1	0	0	-0.241***	0.063
3,250	0	0	0	-0.262***	0.054
3,500	0	0	0	-0.262***	0.054

Notes: The table reports double-selection post-LASSO estimates regressing log teen employment on log of the quarterly minimum wage, using state-quarter aggregated CPS data. All regressions are on data after partialing out state and period fixed effects. LASSO regressions allow state-specific linear trends and division-period fixed effects, and demographic controls (see notes to Table 2 for details). λ is the penalization parameter for the LASSO regressions. Standard errors (in parentheses) are clustered at the state level, and significance levels are indicated by *** 1%, ** 5%, * 10%.

Appendix B: Sample windows, business cycles and parametric trends

In this Appendix, we systematically evaluate the claim in NSW (2014a, 2014b) that the results in ADR are driven by the choice of sample years (1990-2009) because of "endpoint bias." In particular, NSW claim that the presence of recessions in the beginning and end points affects the estimation of the state-specific trends, and biases the minimum wage elasticities for teen employment towards zero when such trends are included.[48] They also claim that inclusion of higher-order (cubic or greater) state-specific trends in that sample restores the more negative estimates by correcting for this "endpoint bias."

Figure B1 compares the four key specifications across 72 samples by varying the starting years between 1979 and 1990, and the ending years between 2009 and 2014. The two-way fixed-effects model stands out in the figure as having more negative estimates in each of the 72 samples than any of the other models. Indeed, the estimates from the two-way fixed effects model are more negative in the longer samples, consistent with the argument that long-run trends may be contaminating the estimates. The estimates from the intermediate specifications (with either state-specific linear trends, or division-period fixed effects) vary somewhat, depending on the sample. In particular, the model with just linear trends produces estimates that are somewhat negative in samples that begin with 1990, but become smaller in magnitude for estimates in samples beginning in 1989 or earlier. Moreover, extending the sample forward also reduces the magnitude of the estimate. In contrast, the most saturated specification is quite robust with respect to the choice of the sample period. While stability of estimates across samples need not indicate accuracy, these results are consistent with the idea that using both state-specific trends and division-period effects guards against estimation errors when either set of control is included on its own. We made a similar observation in ADR, where we specifically warned against the reliability of estimating parametric trends in short samples and suggested the usefulness of including multiple types of controls.[49]

Figure B1 is also informative about "endpoint bias." It does not provide any indication that the 1990-2009 sample used in ADR produced more positive estimates for columns 3 and 4, which include controls for linear trends. Indeed, if anything, the opposite appears to be the case: the estimates are more negative in the 1990-2009 sample than in the other samples. We note here that the estimates with the 1990-2009 sample

[48] Specifically, NSW (2014a, p. 616) states : "In models that include state-specific trends, the recessions at the beginning and end of ADR's sample period could have a large influence on the estimated state-specific trends—a so-called endpoint bias. If the recessions have a purely aggregate influence that is common across all states, this will not happen, as the year effects will absorb this common influence. But if the recessions led to cross-state deviations between teen employment rates and aggregate labor market conditions, then the estimated longer-term trends in teen employment could be biased. This, in turn, could lead to mis-classification of periods in which teen employment was high or low relative to the predicted values net of the minimum wage and hence influence the estimated minimum wage effects for reasons having nothing to do with the longer-run trends for which the specification is trying to control."

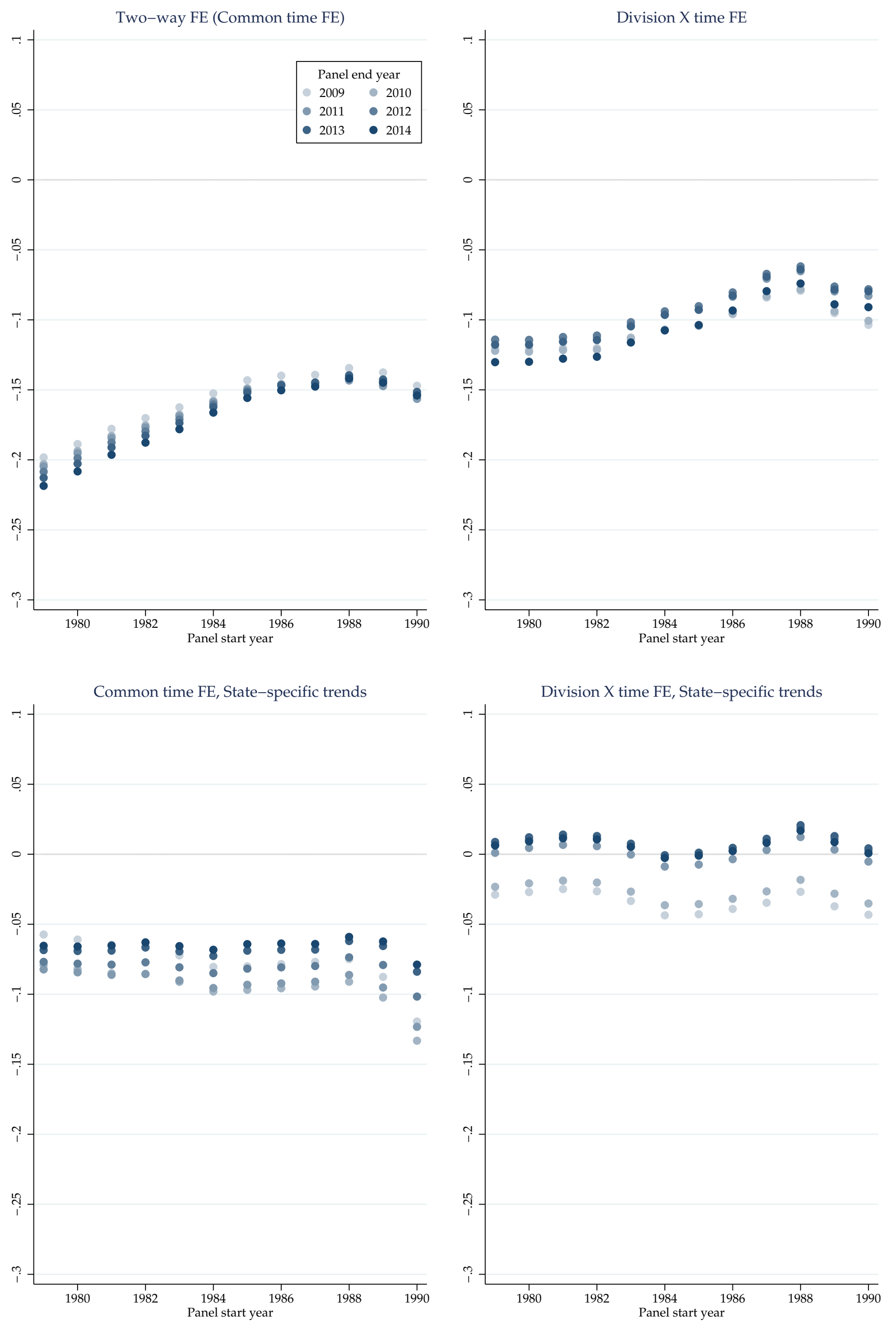
[49] We wrote in ADR, p. 237: "Generally speaking, our preferred specification 4 [with division-period effects] tends to be more stable across time periods than does specification 3 with just state linear trends...While linear trends do a good job of eliminating long-term trend differences across states in longer panels, they are a less valuable means of controlling for spatially correlated shocks, and they are estimated poorly in shorter panels."

using the CPS basic monthly data are somewhat more negative than the estimates reported in ADR, which used the CPS ORG data. However, the conclusion from the most saturated model remains qualitatively the same as in the original sample. Moreover, even small expansions of the sample produce estimates closer to the ones for our full 1979-2014 sample.

As another assessment of the role of business cycles in affecting estimation of trends, Appendix Table B1 shows how the estimates vary when we exclude recessionary periods from the sample. The table includes two definitions of recessions. One consists of the standard NBER-defined recessionary periods. The second expands the NBER concept to include quarters until the national employment reaches the pre-recession peak. This expansive definition excludes the following periods from the sample: 1980q1-1980q4, 1981q3-1983q3, 1990q3-1992q4, 2001q1-2004q4, 2007q4-2014q1. Overall, we find little indication that excluding recessionary quarters produces sizably negative estimates in models with state trends (column 3 and 4). The exclusion of NBER recessions makes little difference to the point estimates in any of the 4 models. Exclusion using the expanded definition produces a point estimate of -0.033 for the specification with state specific trends only (column 3); and -0.078 for for the most saturated specification (column 4); neither are statistically significant at conventional levels.

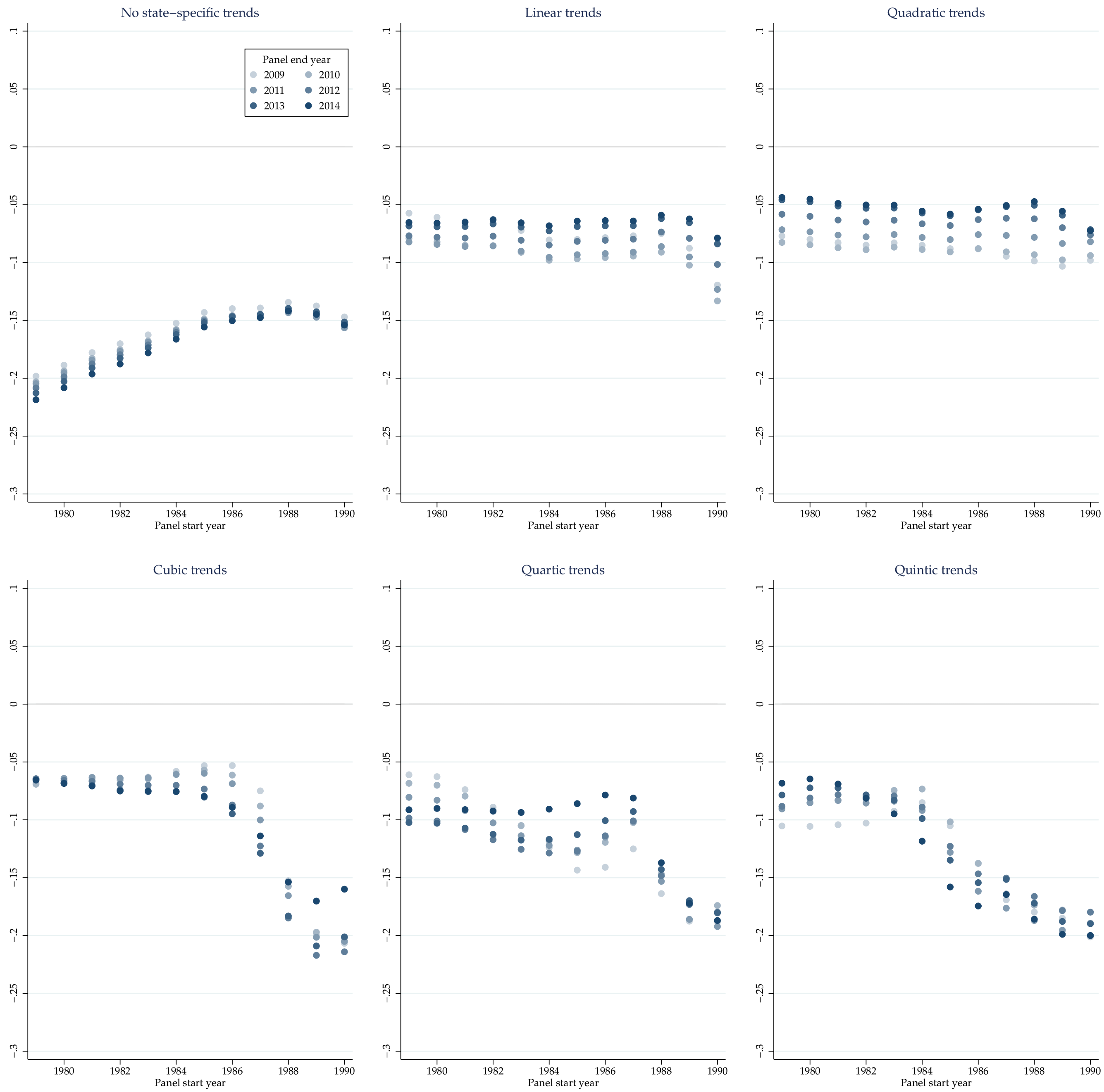
We noted in section 2.2 that the models including cubic or higher order polynomial time trends by state produce estimates that were more negative in the shorter (post 1990) sample, but not in the expanded sample. Figure B2 shows how the results vary when higher order trends are introduced across the 72 samples with start dates varying between 1979 and 1990, and end dates varying between 2009 and 2014. (These specifications use common time fixed effects and do not additionally control for division-period effects.) We find that the estimates using quadratic trends are similar to those with linear trends, are fairly robust to sample choice, and almost never exceed -0.1 in magnitude. However, when we extend the sample by including earlier start dates, we produce estimates that are generally smaller in magnitude. Starting the sample even a few years earlier than 1990 greatly shrinks the estimates from models with trends towards zero, even when higher order trends are included. Recall, however, that the results from LASSO-based double-selection procedure reported in section 2.2suggest that higher order trends are not warranted by the data. This is true in both the full sample as well as the sample beginning in 1990. The combination of these two facts casts serious doubt on the relevance of the finding in NSW (2014a) that inclusion of third or higher order trends in the 1990-2011 sample reproduces more negative employment effect of minimum wages on teen employment.

Figure B1: Minimum wage elasticities for teen employment, by panel start and end year



Notes: The figure reports teen employment elasticities with respect to the minimum wage for the four regression models using individual-level data described in Section 2. From these linear probability models, we calculate elasticities by dividing the log minimum wage coefficient by the sample mean employment.

Figure B2: Minimum wage elasticities for teen employment, by type of state-specific polynomial trend, and panel start/end year



Notes: The figure reports teen employment elasticities with respect to the minimum wage for the two-way fixed-effects model, with state-specific trends of polynomials degree zero through five, using the individual-level data described in Section 2. From these linear probability models, we calculate elasticities by dividing the log minimum wage coefficient by the sample mean employment.

Appendix Table B1
Minimum wage elasticities for teen employment using alternative samples to exclude recessions, individual-level CPS data 1979-2014

	(1)	(2)	(3)	(4)
Full Sample	-0.219***	-0.130*	-0.065	0.006
N=3,534,924	(0.043)	(0.077)	(0.041)	(0.047)
Leave out recessions	-0.207***	-0.130*	-0.064	-0.004
N=2,901,261	(0.039)	(0.076)	(0.043)	(0.052)
Leave out expanded recessions	-0.151***	-0.143*	-0.033	-0.078
N=1,924,468	(0.050)	(0.081)	(0.076)	(0.063)
Division-period FE		Y		Y
State-specific linear trends			Y	Y

Notes: The table reports minimum wage elasticities for average teen wage and employment, using individual-level Current Population Survey data from 1979-2014 (basic monthly data for employment, and Outgoing Rotation Groups for wage). The dependent variable is either log wage, or a binary employment indicator. The reported elasticities are calculated by dividing the coefficients on log minimum wage (and standard errors) by the sample mean employment rate. All regressions include controls for the quarterly state unemployment rate, the quarterly teen share of the working age population, dummies for demographic controls used in Table 1 and described in Section 2, and state and period fixed effects. Specifications additionally include Census division-period effects and state-specific linear trends, as indicated in the table. Samples are either the full sample, the sample excluding recessionary quarters, or the sample excluding expanded recessionary quarters. Recessionary quarters include 1980q1-1980q3, 1981q3-1982q4, 1990q3-1991q1, 2001q1-2001q4, 2007q4-2009q2. Expanded recessions are defined to further include quarters until national employment levels reach pre-recessionary peaks: 1980q1-1980q4, 1981q3-1983q3, 1990q3-1992q4, 2001q1-2004q4, 2007q4-2014q1. Regressions are weighted by sample weights, and robust standard errors (in parentheses) are clustered at the state level and significance levels are indicated by *** 1%, ** 5%, * 10%.

Appendix C: Spatially correlated placebos

Appendix B of DLR (2010) used a placebo-based falsification exercise to provide additional evidence on the bias of the canonical two-way fixed-effects model arising from spatial heterogeneity. The basic idea is to assess whether minimum wages in nearby areas are correlated with own employment—even when variation in the true minimum wage could not possibly affect employment across different counties in the sample. DLR (2010) ensure this by taking a sample of counties for which the minimum wage evolved in exactly the same way—they were always bound by the federal minimum wage. Yet, the results indicated that restaurant employment in these counties was correlated with nearby minimum wages—which we described as a "placebo" since there was no true minimum wage variation across counties in the estimation sample. DLR (2010) interpreted this finding to be consistent with the presence of an omitted variable that is spatial in nature.[50] NSW (2014a, 2014b) argue that this test is invalid, because the effect of the true minimum in these areas is not fully accounted for using time fixed effects. In this Appendix, we describe the original exercise in greater detail and explain why the results do, indeed, demonstrate that spatial heterogeneity contaminates the two-way fixed-effect model's employment estimates. We also discuss why the counter-argument in NSW (2014a, 2014b) is incorrect.

For this exercise, DLR (2010) started with the stacked border county pair sample, and kept only the subset of counties in which the prevailing minimum wage was always equal to the federal minimum wage: $MW_{st}^S = MW_t^F$. Call this the set S. Define also the set S' of cross-border counties adjacent to each of the counties $i \in S$—this will be used to define a placebo minimum wage below.[51] To emphasize, while S-county minimum wage, MW_{it}^S is always equal to the federal minimum, MW_t^F, the same is not true for the minimum wage in their cross-border neighbors, $MW_{jt}^{S'}$.

Now consider the data generating process underlying the two-way fixed effects model:

$$E_{it}^S = \gamma MW_{it}^S + \mu_i + \tau_t + \nu_{it}. \tag{9}$$

where E_{it}^S is log of restaurant employment in county i, MW_{it}^S is the log of the prevailing minimum in that state, μ_i is a county fixed effect and τ_t is a time fixed effect, and ν_{it} is a mean zero disturbance term. Next, consider regressing E_{it}^S, on own county minimum wage, MW_{it}^S, as well cross-border neighbors' minimum wage, $MW_{jt}^{S'}$ along with time dummies (D_t) and county dummies, D_i using the sample S:

[50] Section I.G. and Appendix B of DLR (2010) describe the original exercise.

[51] In this Appendix, for comparability we use a similar notation as in NSW(2014b) instead of the original DLR (2010) notation and the notation elsewhere in this paper. However, we index the variables MW_{it}^S, $MW_{jt}^{S'}$ instead of $MW_{st}^S, MW_{st}^{S'}$ to clarify the county in question (i, or its neighbor, j).

$$E_{it}^{S} = \gamma MW_{it}^{S} + \delta MW_{jt}^{S'} + D_i\theta + D_t\lambda + \epsilon_{it}. \tag{10}$$

Note that in our estimation sample of counties $i \in S$, the prevailing minimum wage is the federal one, and so $MW_{it}^{S} = MW_{t}^{F}$ is only a function of time. As a consequence, the effect of the prevailing minimum wage in S is fully soaked up by the time dummies D_t. To put it differently, by including D_t as covariates, we are "dummying out" the actual prevailing minimum wages in the states S. And so estimating (10) is equivalent to estimating:

$$E_{it}^{S} = \delta MW_{jt}^{S'} + D_i\theta + D_t\lambda + \epsilon_{it} \tag{11}$$

which is what DLR (2010) did estimate. The purpose of choosing the set of border counties S where $MW_{it}^{S} = MW_{t}^{F}$ was to avoid having to obtain a consistent estimate of γ, since in this sample, any effect of own-county minimum wage effect is fully accounted for through the time dummies.

What should we expect $\hat{\delta}$ to be? Barring cross-border spillover, the level of S'-county minimum wages should have no causal effect on S-county employment conditional on the time effects.[52] This is why we considered $MW_{jt}^{S'}$ as a "placebo" minimum wage: a $\hat{\delta} \neq 0$ indicates a likely bias due to an omitted variable, for which $MW^{S'}$ is acting as a proxy. This omitted variables bias is exactly what one expects in the presence of spatial heterogeneity—that minimum wage increases tend to be correlated with certain regional shocks.

As a point of comparison, we further estimated a two-way fixed-effect model regressing actual employment in the S' counties on S' minimum wages in this particular sample:

$$E_{jt}^{S'} = \beta MW_{jt}^{S'} + D_i\theta + D_t\lambda + \epsilon_{jt} \tag{12}$$

If the size of β is similar to δ, then according to the two-way fixed effects model, the "effect" of $MW^{S'}$ on neighboring county employment, E^{S}, is comparable to the "effect" on own employment, $E^{S'}$, even after controlling for the minimum wage in S. DLR (2010) estimated that $\hat{\delta} = -0.123$, which is sizable, compared to $\hat{\beta} = -0.208$. They commented that "we find a negative effect in both samples (though imprecise), with elasticities exceeding -0.1 in magnitude, suggesting bias in the canonical [two-way fixed effects] specification."

NSW (2014a, 2014b) criticize this falsification test. Most importantly, they argue that that the S-county sample is "contaminated." Their argument has two parts. First they (correctly) point out that that even though the minimum wages in S-counties are equal to the federal minimum $MW_{it}^{S} = MW_{t}^{F}$, the minimum wages are changing over time, which can have a real effect. Second, they (incorrectly) argue that because

[52] Dube, Lester and Reich (2010, section VA, failed to find evidence of such spillovers in a test comparing border and interior counties. Moreover, NSW's criticism of our falsification exercise is not based on the possibility of such spillovers. For this reason, we do not discuss the spillover issue further in this Appendix.

the cross-border minimum wage $MW_{jt}^{S'}$ can be correlated with true minimum wage MW_{it}^{S}, δ can reflect some of the effect of MW_{it}^{S} on E_{it}^{S}, thereby "contaminating" the falsification exercise. They write: "But $MW_{st}^{S'}$ in equation (5) [similar to our equation 11 above] varies with the federal minimum wage in a way that is not perfectly correlated with the period fixed effects, because whether the federal minimum wage variation changes the cross-border minimum wage depends on whether the state or federal minimum wage is binding. Thus, federal minimum variation is not swept out by the period fixed effects, and therefore the cross-border minimum wage variation will be correlated with the actual state minimum wage variation." Formally, their argument is that since $Cov(MW_{jt}^{S'}, MW_{it}^{S}) = Cov(MW_{it}^{S'}, MW_{t}^{F}) > 0$, if $\gamma \neq 0$ and hence MW^{F} has a causal effect on E_{it}^{S}, this can be reflected in $Cov(E_{it}^{S}, MW_{jt}^{S'}) < 0$ under the data generating process represented by equation (9).

This argument is fundamentally flawed. By construction, the prevailing minimum wage in the sample of S-counties is the federal one: $MW_{it}^{S} = MW_{t}^{F}$. So, the true minimum wage effect—whatever it may be—is completely "dummied out" by the time dummies D_t, fully accounting for any causal effect of the federal minimum wage on E_{it}^{S}. Therefore, conditional on D_t, if MW_{it}^{S} is still correlated with E_{it}^{S}, it is *only* for a spurious reason, and *not* due to a causal effect of MW_{t}^{FED}. Formally, while *unconditionally* it may be that $Cov(MW_{jt}^{S'}, MW_{it}^{S}) = Cov(MW_{it}^{S'}, MW_{t}^{F}) > 0$, conditional on the fixed effects, we have $Cov(MW_{jt}^{S'}, MW_{it}^{S}|D_t, D_j) = Cov(MW_{it}^{S'}, MW_{t}^{F}|D_t, D_j) = 0$. Therefore, conditioning on the time dummies also conditions on the actual minimum wage in the S-county. So if $Cov(E_{it}^{S}, MW_{jt}^{S'}|D_t, D_j) = Cov(E_{it}^{S}, MW_{jt}^{S'}|D_t, D_j, MW_{it}^{S}) < 0$, by definition it is via something *other than* the effect of the prevailing minimum wage, because the inclusion of the time dummies fully accounts for the impact of the prevailing (federal) minimum wage in the S-counties.

Surprisingly, NSW (2014a, 2014b) insist upon rejecting this argument and claim that the time dummies do not remove the causal effect of MW_{t}^{F} upon E_{jt}^{S}. As supposed evidence for this claim, they provide a decomposition of $MW_{it}^{S'}$ that aims to demonstrate that δ can still reflect true causal effects. In actuality, their decomposition provides an illustration of exactly the opposite argument. In particular, they rewrite equation 11 as:

$$E_{it}^{S} = \delta \left(MW_{t}^{F} \cdot I\left\{ MW_{jt}^{S'} = MW_{t}^{F} \right\} + MW_{jt}^{S'} \cdot I\left\{ MW_{jt}^{S'} > MW_{t}^{F} \right\} \right) + D_i\theta + D_t\lambda + \epsilon_{it} \qquad (13)$$

They then argue that the first term, $MW_{t}^{F} \cdot I\left\{ MW_{jt}^{S'} = MW_{t}^{F} \right\}$, will not be not swept out by the time dummies because the federal minimum wage is multiplied by a dummy of whether the binding minimum wage in neighboring county j is the federal one. They write "[c]learly the federal variation can play a role here because the federal minimum wage is multiplied by a dummy that is sometimes one and sometimes zero, breaking the perfect collinearity with the time fixed effects." They interpret this to mean that δ can reflect the true effect of MW_t^Fon E_{it}^S.

However, NSW miss the implication of this decomposition. Of course, the interaction term $MW_t^F \cdot I\left\{MW_{jt}^{S'} = MW_t^F\right\}$ can be correlated with E_{it}^S even after controlling for the time effects; after all, it is part of $MW_{jt}^{S'}$, which we show is empirically correlated with E_{it}^S. The point is that there in no *plausible causal* interpretation of that correlation under the data generating process represented by equation (9). We have already laid this out above: by definition, conditioning on D_t conditions on $MW_{st}^S = MW_t^F$, so the conditional covariance between E_{it}^S and $M_{jt}^{S'}$ is precisely purged of the impact of the federal policy. So what would be the meaning of a negative correlation between E_{it}^S and the interaction term $MW_t^F \cdot I\left\{MW_{jt}^{S'} = MW_t^F\right\}$? Since MW_t^F is just a function of the time effects, the interaction term simply measures heterogeneity in the time effects by the nature of the minimum wage in the "neighborhood." So a negative correlation between E_{it}^S and the interaction term indicates that when the federal wage rises uniformly across all counties in the sample at date t, and if we take two counties i_1, i_2 that are both bound by the federal wage, employment falls more in i_1 than in i_2 when i_1's neighbor (j_1) is also bound by the federal minimum wage, while $i_2's$ neighbor (j_2) happens not to be. Again, this is for two counties i_1 and i_2 that have the identical (i.e., federal) minimum wage, so the true causal effect of the federal increase should be the same under the data generating process in equation 9. A non-zero coefficient on $MW_t^F \cdot I\left\{MW_{jt}^{S'} = MW_t^F\right\}$ indicates that employment changes in i_1 and i_2 turn out to depend on the characteristics of the neighbors j_1 and j_2—precisely reflecting the evidence of a omitted variable that is spatial in nature. Therefore, NSW's argument of "contamination" of the placebo by federal minimum wage is erroneous and the decomposition they marshal as evidence for "contamination" actually demonstrates the validity of our original exercise.[53]

NSW (2014b) also make a second point regarding this placebo test. They argue that correcting a small data error in DLR (2010)[54] changes $\hat{\beta}$ from -0.208 to -0.114 , while $\hat{\delta}$ is largely unaffected (changing from -0.123 to -0.125). We acknowledge the data error, but note that this correction actually appears to strengthen

[53]In Table 5 of NSW (2014b), the authors artificially change the federal minimum wage, and show that this change has a mechanical effect on the point estimate of $\hat{\delta}$ through changing in some cases the value of the placebo minimum wage ($M_{jt}^{S'}$) . This exercise sheds no light on the validity of our placebo test: transforming the placebo minimum wage and finding that the coefficient is altered is hardly a surprise. We are arguing that $\hat{\delta}$ should be zero under the data generating process of the two-way fixed-effects model, but it is not empirically—reflecting the correlation between employment with the neighboring minimum wage, $M_{jt}^{S'}$. Their exercise of changing the federal minimum wage artificially shows that the measured $\hat{\delta}$ can be changed by artificially changing MW_t^F and hence $MW_{jt}^{S'}$. This is both true and irrelevant. Subsequently, as an effort to "solve" the "contamination" problem, NSW proceed to eliminate over half of the sample by cutting out many of the years and by imposing an arbitrary restriction on cross-border minimum wage variation. One of these restrictions excludes all federal minimum wage increases from the sample; this restriction does not attain their objective since the real minimum wage is changing due to inflation. At any rate, once they eliminate over half of data using these arbitrary criteria, they find that the placebo estimate becomes close to zero. This "solution" does not shed any light on the validity of the placebo exercise, since their assertions about the "invalidity" of the original placebo test are erroneous.

[54]DLR (2010) incorrectly coded Maryland as having raised its wage to $6.15 in q1 and q2 of 2006, when in reality it was $5.15 during those two quarters. We thank NSW for catching this mistake. NSW also note that this coding error does not influence any of the analysis in the paper besides the placebo exercise.

the conclusion we drew in DLR (2010): the point estimate of the placebo minimum wage, $\hat{\delta}$, is essentially of the same size (or slightly larger) than the (corrected) own minimum wage estimate, $\hat{\beta}$, and both exceed -0.1 in magnitude. The corrected sample suggests that nearly the entirety of the negative employment estimate can be explained by the unaccounted spatial heterogeneity in the two-way fixed-effects model.