IRLE WORKING PAPER
#110-15
April 2015

The Effect of Minimum Wages on Employment: A Factor Model Approach

Evan Totty

Cite as: Evan Totty (2015). "The Effect of Minimum Wages on Employment: A Factor Model Approach". IRLE Working Paper No. 110-15. http://irle.berkeley.edu/workingpapers/110-15.pdf

Berkeley
UNIVERSITY OF CALIFORNIA

The Effect of Minimum Wages on Employment: A Factor Model Approach*

Evan Totty†

April 29, 2015

Abstract

This paper resolves issues in the minimum wage-employment debate by using factor model econometric methods to address concerns related to unobserved heterogeneity. Recent work has shown that the negative effects of minimum wages on employment found using traditional methods are sensitive to the inclusion of controls for regional heterogeneity and selection of states that experience minimum wage hikes, leaving the two sides of the debate in disagreement about the appropriate approach. Factor model methods are an ideal solution for this disagreement, as they allow for the presence of multiple unobserved common factors, which can be correlated with the regressors. These methods provide a more flexible way of addressing concerns related to unobserved heterogeneity and are robust to critiques from either side of the debate. The factor model estimators produce minimum wage-employment elasticities that are much smaller than the traditional OLS results and are not statistically different from zero. These results hold for many specifications and two datasets that have been used in the minimum wage-employment literature. A simulation shows that unobserved common factors can explain the different estimates seen across methodologies in the literature.

*I would like to thank my advisers, Mohitosh Kejriwal and Kevin Mumford, for their advice and discussion. I would also like to thank Jack Barron, Tim Bond, and Justin Tobias for their helpful comments.

†Purdue Univeresity, Department of Economics, 403 W. State Street, West Lafayette, IN 47907. Email: etotty@purdue.edu. Phone: 765-494-7128. Fax: 765-494-9658.

1 Introduction

Understanding the effect of minimum wages on employment has long been of interest to economists, with empirical work on the subject dating back approximately 100 years (Obenauer and von der Nienburg, 1915). Despite the long history of attention, economists are still very much divided on this issue. The last two decades, in particular, have produced an abundance of work on the subject, without providing a consensus. The empirical evidence in these studies differs depending on both the datasets used and the methodology[1]. The goal of this paper is to resolve the issues in the minimum wage-employment literature by using panel data econometric methods that are robust to critiques from either side of the debate. Specifically, this study uses the common correlated effects estimators developed by Pesaran (2006) and the interactive fixed effects estimator developed by Bai (2009). These estimators are applied to two datasets and many specifications that have recently been used in the literature. The factor model methods used in this paper are well suited for a wide variety of empirical studies, although they have not yet received much use.

The modern debate in the minimum wage-employment literature is concerned with how to best address issues related to unobserved heterogeneity in large panel data studies[2]. Specifically, the issues are related to regional heterogeneity in employment patterns and selectivity of states that experience minimum wage hikes. The recent literature is summarized in Table 1. The traditional approach to estimating the minimum wage-employment elasticity with

[1] The theory is also ambiguous. While the competitive model predicts a decrease in employment in response to a minimum wage increase, the monopsony model can predict no effect or even a small positive effect. If there is no significant impact of minimum wages on employment, then the effect of minimum wages must be occurring through other channels, such as decreased costly labor turnover, improved organizational efficiency, increased worker effort, or small price increases (Schmitt, 2013). Evidence on these other channels is sparse and mixed, although recent work by Dube et al. (2014) finds significant reductions in employee turnover following minimum wage increases.

[2] A detailed history of the minimum wage-employment debate can be found in Brown (1999), Neumark and Wascher (2008), and Baleman and Wolfson (2014). Past studies on minimum wages and employment are usually either local case studies focusing on employment in a particular low-skill industry or national studies using panel data on teenage or restaurant employment. Local case studies commonly find no effect. The national panel data studies using traditional methodology typically find statistically significant negative employment effects with minimum wage-employment elasticities in the range of -0.1 to -0.3, while new methods meant to address issues related to unobserved heterogeneity commonly find no effect.

panel data has been to use ordinary least squares with unit (either state or county) and period fixed effects included to address unobserved heterogeneity. This approach produces large and statistically significant elasticities in the range of -0.1 to -0.3 (Neumark and Wascher, 1992, 2007, 2008; Sabia, 2009). Dube, Lester, and Reich (2010, hereafter DLR) and Allegretto, Dube, and Reich (2011, hereafter ADR) raise concerns about this two-way fixed effects approach. They show that there is significant regional heterogeneity in employment patterns and that states that raise their minimum wage tend to have negative pre-existing trends in employment, relative to states that do not raise their minimum wage. DLR and ADR argue that the two-way fixed effects approach does not properly address the regional heterogeneity and pre-existing trends. They use census division-by-period fixed effects, state-specific linear time trends, and a border discontinuity approach in addition to two-way fixed effects and find no effect of minimum wages hikes on employment.

This approach has not been met without criticism. Neumark, Salas and Wascher (2014a, hereafter NSW) argue that the implicit assumption of these methods that geographically proximate places are better controls is not supported by the data. The core of this argument is that local areas - be they state or county - are not frequently picked as donors when they implement a synthetic control-type approach based on Abadie et al. (2010). They also show that large negative effects for teenagers return when you extend the linear state-specific time trends to higher order polynomial trends. NSW therefore conclude that neither the methods nor the results in DLR and ADR are supported by the data, and they should be disregarded in favor of the traditional two-way fixed effects results. These authors have argued in a subsequent paper that census division-by-period fixed effects and the border discontinuity approach produce positive endogeneity bias towards zero by changing the identifying variation to within census division or across state borders (Neumark et al., 2014b).

The use of synthetic controls has also become an area of debate in this literature, starting with NSW. NSW admit that the desire for a more flexible way to capture unobserved heterogeneity is a valid concern, and use a synthetic control-type estimator[3]. This approach produces negative and statistically significant elasticities that are especially large for teenagers. However, the application of synthetic controls to the minimum wage-employment debate is not straightforward. Synthetic controls were intended for the scenario in which there is a single state receiving a one time policy treatment, making the use of continuous and recurring minimum wage treatment problematic. Synthetic controls also leave some details up to the discretion of the researcher, such as choice of matching variables and length of pre- and post-treatment windows. Allegretto, Dube, Reich, and Zipperer (2013) attempt to address these issues by performing the synthetic control approach from Abadie et al. (2010) separately for many instances of a minimum wage increase, along with a better matching variable and longer pre- and post-treatment windows, and find no significant effect of minimum wages on employment[4]. They also provide additional support for the use of census division-by-period fixed effects and the border discontinuity approach[5]. Neumark, Salas, and Wascher (2014b) then respond by continuing to argue that the evidence does not support the use of local controls and that there are large, negative elasticities, particularly for teens. The use of synthetic controls has not resolved the debate.

Clearly, the literature is divided in three areas: (1) The appropriateness of census division-by-period fixed effects and the border discontinuity approach, which assume that local areas provide better counterfactuals, as additional controls for unobserved heterogeneity, (2) the robustness of results to the presence and order of state-specific time trends, and (3) how

[3]In the synthetic control approach from Abadie et al. (2010), the effect of the policy is simply the difference between the synthetic control in the post-treatment period and the treated state in the post-treatment period. NSW construct a synthetic control for every instance of a minimum wage increase and then pool all of the real and synthetic data together and perform OLS with a fixed effect for each set of real and synthetic observations.

[4]Because the synthetic control approach requires that treated states receive no additional treatment during the pre- and post-treatment windows and that donor states be untreated, Allegretto, Dube, Reich, and Zipperer (2013) are able to use only 19 out of 89 instances of minimum wage increases between 1997 and 2007.

[5]Allegretto, Dube, Reich, and Zipperer (2013) refute the interpretation in NSW that the selection of donor units in the synthetic control approach suggests that local areas are not good controls. They also provide a placebo law experiment which allows for longer windows and more local donors and find a clear negative relationship between distance and weights.

to best implement the synthetic control approach. The factor model approach proposed in Pesaran (2006) and Bai (2009) is perfectly suited to reconcile these issues. The methods developed in these papers allow for consistent estimation under the presence of multiple unobserved common factors, which can be correlated with the regressors. The presence of unobserved common factors can cause cross-section dependence, which is problematic for inference, but can also cause bias when the unobserved common factors are correlated with the regressors. In the context of the minimum wage-employment literature, such unobservable common factors could be macroeconomic shocks that influence both low-skill employment and minimum wage policy, and could do so heterogeneously over cross sections. Additionally, state-specific trends and census division-by-period fixed effects can be rewritten as a special case of the factor model structure. This means that the factor model setup can capture these factors if they exist, but does not impose this, or any other, fixed form for the unobserved heterogeneity a priori. The factor model approach can therefore be seen as a more flexible way of addressing the concerns related to unobserved heterogeneity. The factor model approach also has several advantages over the synthetic control approach. It is more amenable to panel regression with continuous and recurring treatment, allowing all of the data to be used. Additionally, the factor model appraoch from Bai (2009) has been shown to perform better than synthetic controls (Gobillon and Magnac, 2013).

The factor model estimators are robust to critiques from either side of the debate. The critique from DLR and ADR is that two-way fixed effects cannot capture the regional heterogeneity and pre-existing trends. As mentioned above, state-specific time trends and census division-by-period fixed effects can be captured by the factor model setup if they exist. The difference is that the factor model approach lets the data determine the nature of the unobserved heterogeneity. The critique from NSW is that local areas are not always better controls and that changing the identifying variation to within census division or across state borders produces positive bias towards zero. The factor model approach tackles the issue of local areas being similar in a natural way, by letting factor loadings, which represent the effect of the unobserved common factors on each cross-section unit, pick out cross-section correlation in the data. This makes the factor model estimators unique because they allow areas to be close in economic dimensions which depart from geographic proximity. This is the case, for instance, when two areas are affected by the same industry-specific shocks because of industry specialization, even if these two areas are not neighbors. Additionally, the factor model approach does not change the identifying variation. The factor model estimators will be applied to the same two-way fixed effects approach that NSW support. Therefore, the factor model approach provides a middle ground between the two sides of the debate.

This study uses data on both restaurant and teenage employment. Tests show strong cross-section dependence in the raw employment data for both datasets, which suggests the presence of unobserved common factors. Initial results using OLS with unit and period fixed effects find statistically significant minimum wage-employment elasticities of -0.138 and -0.178 for restaurant and teenage employment, respectively, in line with other estimates in the literature using these methods. The factor model estimators from Pesaran (2006) and Bai (2009) are then applied to the same data and specification. For the restaurant employment, the factor model estimates of the minimum wage-employment elasticity are in the range of -0.013 to -0.042. For teenage employment, they produce estimates in the range of -0.036 to -0.065. None of these estimates are statistically different from zero. Residual diagnostics confirm that the factor model methods do a better job of removing the cross-section dependence than OLS. As a robustness check, these methods are applied to several time trend specifications from the literature. The results remain small and, in most cases, not statistically different from zero. The factor model estimators do find positive and statistically significant effects of minimum wage hikes on the earnings of both restaurant and teenage workers.

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. This supports the argument from NSW that local areas do not always provide ideal control groups and shows off the flexibility of the factor model approach. A simulation experiment is provided to assess the relative ability of OLS and the factor model estimators to estimate the minimum wage-employment elasticity under different assumptions about the unobserved heterogeneity in the data. Results show that common factors in the true underlying data-generating process can cause the different estimates seen across approaches in the literature.

The remainder of the paper is organized as follows: Section 2 describes the factor model setup that the Pesaran (2006) and Bai (2009) estimators make use of and then describes the estimators themselves. Section 3 describes how the variables are constructed and provides summary statistics for the data. Section 4 reports the results. Section 5 reports the simulation experiments. Section 6 concludes.

2 Empirical Approach

2.1 Unobserved Heterogeneity and Factor Models

The factor model approach provides an alternative way to address issues related to unobserved heterogeneity in the minimum wage-employment literature. The factor model setup is based on a model in which the error term is characterized by a multi-factor error structure. Specifically, the traditional error term in a regression equation is decomposed into time-specific "common factors" that can affect all cross-section units, heterogeneous "factor loadings" that represent how a common factor affects each cross-section unit, and an idiosyncratic error term. This paper adopts the traditional unit and period fixed effects specification and then adds a factor model structure to the error term and uses the factor model estimators from Pesaran (2006) and Bai (2009).

The traditional specification for estimating the effect of minimum wages on employment, originating from Neumark and Wascher (1992), is given by

$$ln(E_{it}) = \beta ln(MW_{it}) + \Gamma X_{it} + \alpha_i + \delta_t + \varepsilon_{it}. \quad (1)$$

The dependent variable, E_{it}, is either the number of restaurant employees in county i and period t or the fraction of teenagers employed in state i and period t, depending on the dataset. MW_{it} is the higher of the federal and state minimum wage in state or county i and period t. Employment and the minimum wage are measured in logs so that β represents the minimum wage-employment elasticity. The term X_{it} is a vector of control variables defined in section 3 that are intended to proxy for supply and demand forces on employment. Unit and period fixed effects are represented by α_i and δ_t, respectively. Several studies have estimated this unit and period fixed effects specification with OLS and no additional controls for regional heterogeneity and found large negative effects of minimum wages on employment (Neumark and Wascher, 1992, 2007; Sabia, 2009).

The difference between the OLS estimate of equation (1) and the factor model estimate is that OLS assumes that ε_{it} is an idiosyncratic error term; there are no missing variables that are correlated with the minimum wage. The factor model approach allows for the possibility that there may be unobserved common factors in the error term, which can be correlated with the regressors. In this case, the error term in (1) takes on a multi-factor error structure:

$$\varepsilon_{it} = \lambda_i' f_t + u_{it} \quad (2)$$

where f_t is an (r x 1) vector of unobserved common factors and λ_i is an (r x 1) vector of factor loadings that capture unit-specific responses to the common shocks. The presence of unobserved common factors can cause employment across counties and states to be interdependent because they are being affected by common unobserved shocks. In the minimum wage-employment data, these unobserved common shocks could be macroeconomic shocks. This interdependence across areas is commonly referred to as cross-section dependence. Cross-section dependence is problematic for inference (Andrews, 2005), but will also cause bias if the common factors are correlated with the regressors.

It is easy to think of a common factor that would cause bias for each direction. Technological change, for example, could produce negative bias. Smith (2011) studied teenage employment rates from 1980 to 2009 and showed that job polarization, or the removal of routine, middle-skill tasks due to technological change, pushes middle-skill adults into traditionally teenage jobs, lowering teenage employment. Allegretto, Dube, Reich, and Zipperer (2013) show that between 1990 and 2007, high minimum wage states experienced greater job polarization, on average, than low minimum wage states. Combining these results suggests that high minimum wage states have experienced greater job polarization, which puts downward pressure on teenage employment. This technological change example would cause negative bias in the OLS estimate of the unit and period fixed effects approach, but could be captured in the factor model approach as an unobserved common factor.

The factor model estimators also provide a solution to possible sources of positive endogeneity bias. NSW argue that state minimum wages are more likely to be increased when local labor markets are strong, which would cause positive endogeneity bias towards zero, citing Baskaya and Rubinstein (2011) who find that when the interaction between the federal minimum wage and the propensity for the federal minimum wage to be binding in each state is used as an instrumental variable for a state's minimum wage, larger negative effects are found than when using the traditional two-way fixed effects approach. Regional labor market shocks that influence state minimum wages and employment can also be captured by the factor model estimators as an unobserved common factor with factor loadings equal to zero for all areas unaffected by the regional shocks.

The advantage of the factor model approach over the approach in DLR and ADR is that it imposes no specific form for the unobserved heterogeneity a priori. If no common factors exist in the data and the traditional two-way fixed effects approach is the correct specification, the factor model approach still performs well. This will be confirmed for the minimum wage-employment datasets in the simulations at the end of the paper. However, if there are unobserved common factors in the data, the factor model approach can capture them. In addition to the technological change example given above, the state-specific time trends and census division-by-period fixed effects applied in DLR and ADR can be regarded as special cases of the factor model setup, because they can be rewritten as the inner product of a vector of time-specific common shocks, $f_t = (t, \delta_t)'$, and unit-specific factor loadings, $\lambda_i = (\alpha_i, \zeta_c)'$, where ζ_c is a census division dummy variable. The difference is that the form of the common shock, f_t, and the nature of the spatial correlation in λ_i will be determined by the data, whereas state-specific trends and census division-by-period fixed effects apply a fixed form. Additionally, if the more general point that local areas are more similar in terms of their unobservable characteristics and thus better controls, which is most strongly assumed in DLR's border discontinuity approach, is true, then this can be captured in the factor model approach by nearby areas having similar factor loadings. Therefore, the factor model approach provides a middle ground between the two sides of the debate by being able to capture state-specific time trends and regional heterogeneity without assuming that local areas are better controls or changing the identifying variation.

The factor model approach also has several advantages over synthetic controls. The synthetic control approach from Abadie et al. (2010) that has been used by NSW and Allegretto, Dube, Reich, and Zipperer (2013) also allows for multiple unobserved common factors with heterogeneous factor loadings in the underlying data-generating process. While the synthetic control approach has the advantage of being semi-parametric, the linear factor model approach adopted in this paper is much more amenable to panel regression with continuous and recurring treatment. The synthetic control approach was designed for use in individual case studies, in which there is a single instance of policy treatment. The use of synthetic controls with continuous and recurring minimum wage treatment is problematic for two reasons. First, the researcher must decide how to handle the difference in intensity of treatment across states when estimating the effect of treatment. Second, because the synthetic control approach requires that the donor states receive no treatment and that the treated states receive no additional treatment during the pre- and post-treatment windows, much of the minimum wage variation gets discarded. Federal minimum wage variation is nearly impossible to use because very few states are not affected. State minimum wage variation is also often discarded because state minimum wages are increased so frequently that states often receive additional treatment during the pre- and post-treatment windows. Indeed, Allegretto, Dube, Reich, and Zipperer (2013) are able to use only 19 of 89 minimum wage increases from 1997 to 2007 in their synthetic control approach[6]. Gobillon and Magnac (2013) also shed light on the relative merits of the synthetic control approach and the factor model estimator from Bai (2009). They show that, under the assumption that the true model is a linear factor model, synthetic controls are unbiased only when the exogenous covariates and factor loadings for treated areas belong to the convexified support of exogenous covariates and factor loadings for control areas. Monte Carlo simulations and an empirical application favor the factor model approach from Bai (2009) over synthetic controls.

It is worth noting that the factor structure shown above can be rewritten to incorporate lagged common factors. This is appealing in the context of the minimum wage-employment data, given that it is reasonable to assume that both employment and minimum wages may be slow in responding to economic influences due to social norms against laying off workers and the delay between when minimum wage increases are approved and actually implemented. In this sense, it is intuitive to think that employment and minimum wages may respond to lagged common factors. The factor model can be rewritten to incorporate this by rewriting a dynamic factor model as a static factor model, with the error term in equation (1) now taking the form $\varepsilon_{it} = \Lambda_i' F_t + u_{it}$ where $F_t = (f_t', f_{t-1}', ..., f_{t-s}')'$ is an $(r(s+1)$ x 1) vector of common factors, $\Lambda_i = (\lambda_{i0}', \lambda_{i1}', ..., \lambda_{is}')'$ is an $(r(s+1)$ x 1) vector of factor loadings, and s represents the number of lagged factors.

Clearly, the factor model approach is well-suited to resolve issues in the minimum wage-employment debate, as it can capture regional heterogeneity and time trends without chang-

[6]See Dube and Zipperer (2014) for a detailed discussion on the application of synthetic controls with continuous and recurring treatment.

ing the identifying variation. It also has an important advantage over the synthetic control approach, which is that it can use all of the data. The next section will describe the factor model estimators.

2.2 Factor Model Estimators

There are two commonly used approaches for linear factor model estimation, each of which will be used in this paper. The first method is Pesaran's (2006) common correlated effects estimators. This method does not attempt to estimate the common factors and factor loadings. Rather, Pesaran shows that you can proxy for the factors with cross-sectional averages of the dependent and independent variables. This estimator has the added benefit that it can be computed by ordinary least squares applied to regressions where the observed explanatory variables are augmented with cross-sectional averages of the dependent and independent variables. Pesaran proposes two versions of this method: the common correlated effects mean group (CCEMG) estimator and the common correlated effects pooled (CCEP) estimator. CCEMG allows minimum wages to have heterogeneous effects across units and reports the mean of the individual slope coefficients. CCEP still allows the common factors to have heterogeneous effects, but pools all of the cross-section units together for estimating the regression coefficients. Standard errors are calculated using equations (58) and (69) in Pesaran (2006) for the CCEMG and CCEP estimators, respectively. Confidence intervals and significance reported in the results section are based on bootstrapped t-statistics using the *wild cluster bootstrap-t* procedure from Cameron et al. (2008), clustered at the state level. More details are provided in the appendix.

The second method is Bai's (2009) interactive fixed effects (IFE) estimator. This approach does involve estimating the common factors and factor loadings. This involves the use of principal component analysis[7]. The IFE approach is based on the fact that, given the

[7]Because the IFE estimator actually estimates the factor structure, the number of common factors, r, must be pre-specified. One approach is to use the information criteria from Bai and Ng (2002), which estimates the number of strong factors in the data. Alternatively, IFE results could be provided for different

common factors and factor loadings, the regression coefficients can be estimated using OLS after subtracting the factor structure from the data, and given the regression coefficients, the factors and factor loadings can be estimated by using principal component analysis on the regression residuals. However, the regression coefficients and factor structure are both unknown in practice. Therefore, Bai proposes an iterative procedure in which, given an initial guess of either the regression coefficients or the common factors and factor loadings, one iterates between estimating the regression coefficients using OLS on the de-factored data and estimating the factor structure using principal components on the OLS residuals[8]. This iteration continues until the percent change in the sum of squared residuals is below a specified threshold. A threshold of 10^{-9} is used in this paper. Bias-correction for serial correlation, cross-sectional correlation, and heteroskedasticity is performed using equations (23) and (24) in Bai (2009). Standard errors are calculated using Theorem 4 in Bai (2009). Confidence intervals and significance reported in the results section are based on bootstrapped t-statistics using the *wild cluster bootstrap-t* procedure from Cameron et al. (2008), clustered at the state level. More details are provided in the appendix.

The relative merits of the IFE and CCE estimators should be considered. An important difference between the two is the assumption about the number of common factors. The IFE estimator requires that the true number of common factors be less than or equal to the pre-specified number of common factors. The CCE estimators require that the number of common factors be less than or equal to the number of dependent and independent variables, whose cross-sectional averages proxy for the common factors. Pesaran (2006) and other papers have shown that the CCE estimator continues to perform well even when the number of common factors exceeds the number of cross-sectional averages, although this flexibility comes at a cost: the factor loadings must take on a random coefficients form. In

numbers of common factors.

[8]Bai (2009) suggests initiating the iteration with the OLS estimates of the regression coefficients, ignoring the factor structure, and the principal components estimates of the factor structure from the raw data, ignoring the independent variables, and keeping whichever results have the lowest final sum of squared residuals. Analysis was also performed using the CCEMG estimates as an initial guess of the regression coefficients. The results were almost identical, although the iteration did converge faster.

applications, it is not obvious which of these two assumptions for the CCE estimator will be more restrictive. In this sense, the IFE estimator is more flexible because any number of common factors can be specified.

Westerlund and Urbain (2015) also address the relative merits of the two approaches. The authors are actually comparing the relative merits of principal components versus cross-sectional averages rather than IFE and CCE per se. The estimators they consider are slightly different from those considered in Pesaran (2006) and Bai (2009), and they therefore caution against extrapolating too widely the conclusions to the use IFE and CCE. Nonetheless, they find that while the relative bias of the two estimators depends on the parametrization of the true underlying model, both the principal components approach and the cross-sectional averages approach perform best when $N > T$.

3 Data

3.1 Data Sources

Two different low-skill groups are used in this study: restaurant workers and teenagers. Restaurant workers and teenagers are the two most commonly studied populations in the minimum wage-employment literature (Baleman and Wolfson, 2014) and they have been the focus of the recent debate in the literature. Construction of the variables and summaries of the datasets are provided below. Each dataset is merged with a quarterly minimum wage variable which is always the higher of the federal and state minimum wage.

Quarterly data on restaurant employment is constructed for the years 1990-2010 from the Quarterly Census of Employment and Wages. The QCEW provides quarterly county-level payroll data by industry based on ES-202 filings that establishments submit for the purpose of calculating payroll taxes related to unemployment insurance. The county-quarter restaurant employment dependent variable is constructed from both Full Service Restaurants (NAICS 7221) and Limited Service Restaurants (NAICS 7222) and measures the total number of full service and limited service restaurant employees. The control variables are the county-quarter total private sector employment and the county population. The employment variables are constructed from the QCEW and the county population comes from the county-level Census Bureau population data which is produced annually. Data is available for the entire time frame of analysis for 1,371 counties[9]. Summary statistics for the dataset of analysis on restaurant workers are shown in Table 2.

Quarterly data on teenage employment is constructed for the years 1990-2013 from the Current Population Survey Outgoing Rotation Groups. State-quarter observations are constructed by aggregating the CPS-ORG individual level data up to the state-quarter level. The state-quarter teenage employment dependent variable is the fraction of teenagers (ages 16-19) that are employed. The control variables are the state-quarter relative size of the teenage population and state-quarter unemployment rate, also constructed from the CPS-ORG. Summary statistics for the dataset of analysis on teenagers are shown in Table 2.

3.2 Cross-Section Dependence and Data Size

As discussed in the introduction and section 2.1, the presence of unobserved common factors can cause outcomes across areas to be interdependent, known as cross-section dependence, which can be problematic for inference and estimation. The factor model approach is commonly used to model the presence of strong, as opposed to weak, cross-section dependence. Weak cross-section dependence can be thought of as arising from the fact that geographically proximate places will have similar characteristics (geography, culture, demographics, education-level, political institutions, etc.), which will cause correlation in outcomes across areas. Strong cross-section dependence, on the other hand, is typically thought of as arising from unobserved forces ("factors") that influence outcomes across areas in heterogeneous ways. Spatial econometric methods, which deal with cross-section dependence by assuming some specific (often geographic) nature on the interdependence of areas a priori, are an al-

[9]For consistency with DLR and NSW, results are based on a balanced panel.

ternative way to address issues related to cross-section dependence. However, factor models have two important advantages over spatial econometric methods. The first is that the factor model approach assumes no geographic relationship (or any other type of spatial relationship) on the interdependence of areas before estimation. The second is that factor models are intended to capture the presence of strong cross-section dependence, whereas spatial econometric methods typically require that the cross-section dependence only be weak.

Table 2 shows the results of a test for strong cross-section dependence in the data using the cross-section dependence (CD) test from Pesaran (2015), which is based on the average of pair-wise correlations in the data. The null hypothesis of the test, which is distributed standard normal, is that there is only weak cross-section dependence, while the alternative is that the cross-section dependence is strong. The null hypothesis of weak cross-section dependence is rejected at the one-percent level for both the restaurant and teenage employment datasets, with test statistics of 60.06 and -5.64, respectively. This suggests the presence of common factors in each dataset and validates the factor model approach. The fact that there is less strong cross-section dependence in the teenage employment dataset is not surprising, given that the unit of analysis occurs at a more aggregated level.

Because the CPS data does not allow for reliable estimates at the county level, the teenage dataset has a relatively small cross-section dimension of 51 (50 states plus Washington D.C.). As discussed in section 2.2, one of the results from Westerlund and Urbain (2015) is that both the IFE and CCE estimators perform better when $N > T$, which is not the case for the teenage dataset. It is therefore reasonable to assume that the factor model estimators may not be able to capture the common factors as reliably for the teenage dataset as they can for the restaurant dataset and therefore may not be able to fully remove the bias and cross-section dependence that is being caused by common factors. Nonetheless, IFE and CCE estimators still provide significant improvements over traditional OLS methods when common factors exist in the data and still perform well without the presence of common factors in the data even when $N < T$, as shown in Pesaran (2006) and Bai (2009) and in the simulations in section 5.

4 Results

4.1 Minimum Wage-Employment Elasticity

Results are based on the traditional two-way fixed effects specification in equation (1). Results are reported separately for each of the datasets. The results tables first show the OLS estimate of the two-way fixed effects specification, and then show results using each of the factor model estimators described in section 2. Confidence intervals and significance reported in the tables are is based on bootstrapped t-statistics using the *wild cluster bootstrap-t* procedure from Cameron et al. (2008), clustered at the state level. Robustness to the number of common factors, different specifications from the literature, and the sample period is also considered, as well as earnings effects and a falsification test.

Table 3 shows the effect of raising the minimum wage on restaurant employment. Column 1 shows that the OLS estimate of the traditional two-way fixed effects specification is in line with other estimates from the literature, with a large elasticity of -0.138 that is statistically different from zero. The factor model estimators report very different results. The CCEMG, CCEP, and IFE estimators report elasticities of -0.013, -0.013, and -0.042, respectively, each of which are more precisely measured than OLS. None of the factor model estimates are statistically different from zero. The table also shows residual diagnostics which test for the presence of strong cross-section dependence in the residuals. As discussed in section 3, the null hypothesis that there is only weak cross-section dependence in the restaurant data was rejected at the one-percent level, with a test statistic of 60.06. There is still strong cross-section dependence in the OLS residuals, with a test statistic of 26.98, suggesting that OLS has not controlled for the common factors in the restaurant employment data. The factor model estimators do a much better job of capturing the common factors, with test statistics of 4.43, 17.19, and -0.30 for the CCEMG, CCEP, and IFE estimators, respectively.

Table 4 shows the effect of raising the minimum wage on teenage employment. Column 1 shows that the OLS estimate of the minimum wage-employment elasticity is once again in line with other estimates from the literature, with a large elasticity of -0.178 that is statistically different from zero. Just as with the restaurant employment dataset, the factor model estimators report very different results than OLS. The CCEMG, CCEP, and IFE estimators report elasticities of -0.040, -0.065, and -0.036, respectively, none of which are statistically different from zero. Residual diagnostics for the teenage employment data show that the factor model estimators are not able to further remove the strong cross-section dependence from the residuals, as the test statistics for the factor model residuals remain similar to test statistics for the OLS residuals.

The failure of the factor model estimators to further remove the strong cross-section dependence from the teenage employment residuals is likely be due to the relatively small size of the cross-section dimension in the teenage employment dataset, as discussed in section 3.2. While the IFE estimator is $\sqrt{NT}$-consistent, the estimates of the factors themselves via principal components are only $\sqrt{N}$-consistent. Given that the teenage employment dataset has a relatively small cross-section dimension but a large time dimension, the IFE regression coefficients converge more quickly than the estimates of the factors themselves. Thus, imprecise estimates of the factors could lead to the reduction of bias while still leaving the residuals contaminated with cross-section dependence. Nonetheless, the main concern in this paper is removing bias from the minimum wage-employment elasticity, rather than cross-section dependence in the residuals, and the factor model estimators are clearly still capturing the presence of common factors, because the factor model estimates are very different from the OLS estimates. Section 5 will show that the factor model estimators would produce elasticities similar to OLS if there were no common factors present, even for the teenage dataset.

Results for the IFE estimator shown in Table 3 and Table 4 were based on 4 factors and 8 factors, respectively. The information criteria from Bai and Ng (2002) were uninformative in selecting the number of common factors[10]. Therefore, the results in Table 3 were based on 4 factors because that was the minimum number of factors that removed the presence of strong cross-section dependence from the residuals. For the teenage employment dataset, results were based on 8 common factors because that was the number of factors required to explain a large fraction of variance in the residuals. This can be seen in Table 5, which shows the relative importance of each common factor. The relative importance of each common factor is calculated as the fraction of the total variance in the residuals explained by factors 1 to p, given as the sum of the first p largest eigenvalues of the second moment matrix of the OLS residuals divided by the sum of all eigenvalues. The first 4 factors for restaurant employment explain about 86% of the variance in the OLS residuals. For the teenage employment dataset, 8 common factors are required to explain a similar amount of the variance.

Although 4 common factors for restaurant employment and 8 common factors for teenage employment appear to sufficiently capture the unobserved heterogeneity, robustness to the number of factors is also considered. Table 6 shows the IFE estimates for 2-8 common factors. The restaurant minimum wage-employment elasticity is invariant to the number of factors, with elasticities that remain small and not statistically significant for different numbers of factors. The teenage minimum wage-employment elasticity is not entirely invariant to the number of common factors. While the effect is not statistically significant for any number of common factors, the IFE estimate is somewhat large when only 2 or 3 common factors are included, with estimates of -0.104 and -0.093, respectively. Once 4 common factors are included, the IFE estimates are much more similar to the CCE estimates reported in Table 4 and they remain relatively small for up to 8 common factors[11]. The fact that the

[10]The information criteria proposed in Bai and Ng (2002) always picked the maximum number of factors allowable for the restaurant dataset. For the teenage employment dataset, some criteria picked the maximum number allowable, while others picked the minimum number. The same issue occurred in Kim and Oka (2014), who used the IFE estimator to resolve issues in the literature on the effect of divorce law reform on divorce rates, and in Bailey et al. (2014), who study real house price changes across MSAs in the USA. Additionally, the information criteria in Bai and Ng (2002) do not always perform well when $\min\{T, N\} < 60$, which is the case for the teenage employment dataset.

[11]The IFE results for teenage employment remain in the -0.01 to -0.04 range for higher numbers of factors

IFE estimates are more stable for higher numbers of factors supports the evidence in Table 5 that many factors are needed in order to capture the unobserved heterogeneity in the teenage employment data.

It is clear that the factor model estimators produce results that are very different than the traditional two-way fixed effects approach. Each of the factor model estimators produces elasticities that are smaller than OLS and not statistically different from zero. The next section will consider robustness to other specifications from the literature and sub-samples of the data.

4.2 Robustness Checks - State-Specific Trends, Sub-Samples, Earnings, and a Falsification Test

4.2.1 State-Specific Time Trends

Because one of the debates in the minimum wage-employment literature has been that the results are not robust to the presence and order of state-specific time trends, Table 7 and Table 8 show estimates for specifications that include state-specific time trends. In this setup, the multi-factor error structure is applied to the error term of an equation which already models state-specific trends in addition to unit and period fixed effects. As discussed previously, the factor model setup can capture deterministic trends, but will not suffer from any potential bias resulting from specifying state-specific trends when they are not appropriate. Therefore, from a theoretical standpoint, the inclusion of time trends is not necessary, although it would me more efficient to include them if they are part of the true underling data-generating process.

Table 7 shows that the OLS estimate of the restaurant minimum wage-employment elasticity does vary across specifications. The large negative effect disappears when a linear time trend is included, but a smaller negative effect that is statistically different than zero returns

as well.

for flexible state-specific trends. The factor model estimators, on the other hand, are much more robust to the presence and order of state-specific time trends. Only the CCEP estimate with a 3rd order polynomial state-specific time trend shows any considerable change in magnitude or significance and all of the factor model estimates remain smaller than the OLS estimate of the traditional two-way fixed effects specification in Table 3.

Table 8 shows the results for teenage employment. The OLS estimate of the teenage minimum wage-employment elasticity is highly sensitive to the presence and order of state-specific time trends, with large negative effects for no state-specific trends or 3rd order and higher polynomial trends, but no effect for linear or quadratic trends. The factor model estimates also show some variance across specifications, but not nearly as much as the OLS estimates. The factor model estimators remain smaller than OLS for each specification and are always smaller than the traditional OLS estimate in Table 4. The variance across specifications for the factor model estimates of the teenage minimum wage-employment elasticity is likely another result of the relatively small cross-section dimension of the data. The results from Table 4 remain the primary results, given that they will capture time trends only if they show up in the data. Results in section 4.4 will suggest that time trends are not among the common factors for teenage employment.

In summary, the factor model results are more robust to the presence and order of state-specific time trends than OLS, and the pattern of the factor model estimates producing smaller minimum wage-employment elasticities than OLS continues to hold.

4.2.2 Sub-Samples of the Data

Results based on sub-samples of the data are also considered. Sub-sample analysis provides a robustness check to one critique of the factor model approach, which is the assumption of time-invariant factor loadings. While the factor model approach provides a more flexible way to model unobserved heterogeneity than applying any fixed form a priori, it cannot control for a situation in which cross-section units have frequent time variation in their factor loadings. However, if factor loading variation were infrequent and thus time-invariant within sub-samples of the time dimension of the data, sub-sample analysis could help address this issue.

Sub-sample results are shown in Tables 9 and 10 for restaurant and teenage employment, respectively. Three sub-samples are considered for each dataset: (1) the first half of the time dimension of the data, (2) the second and third quarters of the time dimension of the data, and (3) the last half of the time dimension of the data[12]. For restaurant employment, OLS shows a large negative effect in the 1995q2-2005q3 sub-sample. The factor model estimates do not show any large negative effects. The CCEP estimate is negative and statistically significant in the 1990q1-2000q2 sub-sample, but the magnitude is -0.060, which is much smaller than the traditional negative effects found in the literature. For teenage employment, OLS shows large negative effects in earlier sub-samples. The factor model estimators find similar results, although the magnitude of the negative effect in the first half of the sample is smaller.

Once again, the pattern of smaller elasticities for the factor model approach than OLS continues to hold.

4.2.3 Earnings Effects and a Falsification Test

Table 11 shows the effect of minimum wages on the earnings of restaurant workers and teenagers. The QCEW dataset provides the average weekly rate of pay for restaurant workers, calculated as the total restaurant payroll in each county in a given quarter divided by the total restaurant employment level in the county for that quarter. For restaurant earnings results, the total private sector employment control variable is replaced by a measure of the average private sector earnings. For the teenage employment dataset, the dependent variable is the log average hourly wage for teenagers in each state for a given quarter, based only on those who were working and paid between \$1 and \$100 per hour in 2009 dollars.

[12]The dimension of the sub-samples are $N = 1371$, $T = 42$ for restaurant employment and $N = 51$, $T = 48$ for teenage employment.

The factor model estimators all find positive and statistically significant effects of minimum wage increases on the earnings of restaurant and teenage workers that are similar to the OLS estimate.

Table 11 also shows the results of a falsification test based on the manufacturing industry. Only 2.8% of manufacturing workers earn within 10% of the minimum wage (Dube et al., 2010). The manufacturing industry therefore should not experience significant employment effects from minimum wages hikes. Reassuringly, the factor model estimators find no statistically significant effect on employment, although the CCEP estimator does find a positive and statistically significant effect on earnings.

Robustness checks for state-specific time trends and sub-samples continue to show the pattern of smaller elasticities for factor model estimators than OLS, as first seen in Tables 3 and 4. While there is some variance across time trend specifications, the results remain smaller than the traditional OLS estimate in all cases and not statistically different from zero in most cases. The sub-sample results for teenage employment do show elasticities that are close to the traditional OLS range in a couple of instances, but this is only in the first half of the time dimension of the data and thus less relevant for current policy considerations. Earnings results show that, while the factor model estimators do not find any effect of minimum wage hikes on employment of low-skill workers, they do find the expected positive earnings effects.

4.3 Pre-Existing Trends

Support for the census division-by-period fixed effects, state-specific trends, and border discontinuity approach in DLR and ADR comes from evidence that these controls remove negative pre-existing trends in employment for states that raise their minimum wage relative to states that do not. The way that the authors show this is to include leads of the minimum wage variable, in which case the traditional two-way fixed effects approach produces negative lead effects that are removed when the additional controls are included. The authors argue that finding negative effects prior to the policy change reflects spurious pre-trends due to minimum wage changes tending to occur at times and places with unusually low employment growth. The leading effect results have been another topic of debate in the follow up work by NSW, Allegretto et al. (2013), and Neumark et al. (2014b), so similar tests are provided here for comparison.

Results are based on the same model specified in equations (1) and (2), except with 1-, 2-, and 3-year leads in the minimum wage variable included[13]. The elasticity reported at each point in time is the cumulative elasticity, which is the sum of the contemporaneous elasticity for each year until that point in time.

Figure 1 shows the results for restaurant employment. Similar to the results reported in DLR and ADR, OLS produces a negative trend in the leading effect of the minimum wage. The elasticity one year before the minimum wage increase occurs is -0.15, which is similar to the contemporaneous effect shown in Table 3. This trend is removed by the factor model estimators, which produce flat leading effects that are closer to zero. Figure 2 shows the results for teenage employment. There is again a negative trend in the leading effect for OLS, with elasticities of nearly -0.2 three years before the minimum wage increase and -0.3 one year before the increase. The factor model estimators do not fully remove the negative leading effects for teenage employment, but they do decrease the magnitude of the leading effects.

The failure of the factor model estimators to fully remove the negative leading effects for teenage employment suggests that there could still be some leftover negative bias in the factor model estimates from Table 4 and 6 due to minimum wages changes tending to occur at times and places with unusually low employment growth. The results in Table 4 and 6 showed elasticities for teenage employment that were smaller than OLS and not statistically different from zero, but were not as close to zero as the restaurant results and the results in ADR. Additionally, some of the robustness checks did show somewhat larger effects for

[13]At the time of this analysis, state-level minimum wage data is only available for every state through 2015. Therefore, the data on teenage employment is shortened from 1990-2013 to 1990-2012 for these results.

teenagers. One explanation for this is that there may be very small negative effects of minimum wage hikes on teenage employment, although not as large as the traditional -0.1 to -0.3 range. An alternative explanation is that the true effect is closer to zero, but the factor model estimators cannot fully remove the negative bias in the case of the teenage employment dataset because of the relatively small cross-section dimension issues discussed previously. The fact that the factor model estimators fail to fully remove the negative lead effects for teenage employment provides some support for the latter explanation[14].

4.4 Understanding the Unobserved Heterogeneity

The previous results suggested that common factors exist in both the restaurant and teenage employment data and that they produce negative bias in the traditional OLS two-way fixed effects approach. This section will analyze the factor structure that the IFE approach estimates in an attempt to shed some light on what the factor structure is capturing. The factor structure analyzed here is from the IFE estimates of the two-way fixed effects specification in Table 3 and 4. DLR and ADR argue that controls for regional heterogeneity and time trends need to be added to the traditional two-way fixed effects approach. NSW argue against the use of these controls and prefer the traditional approach. The purpose of this section is to see if there is any support for the DLR and ADR approach by looking at what is leftover that the factor model approach captures when only two-way fixed effects are included as controls for unobserved heterogeneity.

No direct economic interpretation can be given to the factors that the IFE approach estimates. Rather, they are defined purely in a statistical sense. They are the eigenvectors that correspond to the largest eigenvalues of the second moment matrix of the regression residuals. Additionally, while $\lambda_i' f_t$ is identifiable, the factors themselves are identifiable only up to a sign change. However, it is possible to plot the time series behavior of each factor and

[14] It could be the case that businesses do have an employment response before the minimum wage increase actually occurs, since minimum wage increases are usually announced at least one year in advance. However, an elasticity of nearly -0.2 three years before the policy change seems extreme, and it is not necessarily obvious why there would be a leading effect in teenage employment but not restaurant employment.

map the effect of the common factors on employment for each unit in a given time period to check for regional heterogeneity.

Figure 3 shows the first common factor for restaurant employment. This factor resembles a time trend in employment, which produces county-specific time trends when multiplied by the first factor loading for each county in the restaurant dataset. The purpose of this figure is to confirm that the factor model estimators can capture time trends, which they clearly do for the restaurant dataset. The second factor for restaurant employment, not shown here, resembles a quadratic time trend. None of the common factors for teenage employment exhibit obvious time trend patterns. Plots of the teenage common factors and additional restaurant common factors are not shown here because they have no clear interpretation. Table 5 shows the AR(1) coefficients for each factor to illustrate how persistent each of the factors is. The first two restaurant factors appear to be very persistent, although this is due to the fact that these factors appear to be time trends. The AR(1) coefficients after the first two factors have been de-trended are 0.69 and 0.79, respectively.

While the factors for the restaurant employment dataset do capture time trends, including time trends in the traditional OLS approach does not fully capture the factor structure. As shown in Table 6, 4 common factors are required to remove the strong cross-section dependence from the restaurant employment data. While the first two factors resemble a linear and a quadratic time trend, the next two factors do not appear to be higher order time trends. For teenage employment, none of the common factors appear to be time trends, and including flexible time trends produces estimates that are very different from the factor model estimates.

As described above, the full factor structure, $\lambda_i' f_t$, is identifiable. This represents the combined effect of each of the common factors on the log of employment in county i and period t. Plotting the combined effect of the factors on employment for each county/state on a map will show if there is any regional heterogeneity in the effect of the common factors on employment. The factor structure makes no geographic assumptions on the form of the unobserved heterogeneity a priori. Rather, any geographic relationship that shows up in the factor structure is the result of spatial correlation in the residuals that is captured by the factor loadings.

Figures 4-6 plot the effect of the common factors on restaurant employment for 1990q1, 2000q1, and 2010q1, respectively. Looking first at 1990q1, there is clear evidence of regional heterogeneity. The common factors are producing a positive effect on employment throughout much of the West Coast, Florida, and the Northeast. Ohio and much of the South are experiencing negative effects on employment from the common factors. The maps for 2000q1 and 2010q1 show that the nature of the regional heterogeneity changes over time. In 2010q1, much of the West Coast is now experiencing negative employment effects from the common factors, while much of the South is now experiencing positive effects. Results are similar for the teenage employment dataset, shown in figures 7-9.

The fact that the factor structure does capture time-varying regional heterogeneity in the effect of the common factors on employment provides support for the argument in DLR and ADR that regional heterogeneity should be addressed when estimating the effect of minimum wages on employment. However, the maps also lend some support to the criticisms from NSW, which is that neighboring counties or same-division states may not always be ideal control groups. For example, the reason why Ohio stands out in Figure 4 is because the counties on the opposite side of the state border are in fact very different in terms of the effect that they experience from the common factors. This is true for the teenage employment dataset as well. Figure 7 shows that states in the Pacific census division (Washington, Oregon, California) and the West North Central census division (Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas) experience similar effects from the common factors, while states in the South Atlantic division (Delaware, Maryland, Washing D.C., Virginia, West Virginia, North Carolina, South Carolina, Georgia, and Florida) are not always similar in their effect of the common factors on employment.

Despite the fact that no direct economic interpretation can be given to the factors, the appearance of both regional heterogeneity and, in the case of restaurant employment, county-specific time trends is reassuring in the sense that the common factors do appear to make economic sense. The exact geographic nature of the regional heterogeneity is surprisingly precise in some cases. For example, in Figure 6, the counties containing and surrounding the cities of San Francisco, Los Angeles, Dallas, San Antonio, Houston, Kansas City, St. Louis, Minneapolis, Chicago, Atlanta, New York, and Boston all stand out from the counties around them because of spatial correlation in the factor loadings. This is quite remarkable, given that the principal components procedure for estimating the factor structure does not impose any geographic relationship on the factor loadings.

5 Simulation

This section assesses the relative ability of OLS and the factor model estimators to estimate the minimum wage-employment elasticity under different assumptions about the unobserved heterogeneity in the data. Specifically, the performance of OLS, CCEMG, CCEP, and IFE are compared with and without the presence of common factors in the data-generating process. The goals of these simulations are (1) to confirm the precision of the factor model estimators when no common factors exist in the data and (2) to confirm the direction of the bias in the OLS estimate of the minimum wage-employment elasticity caused by the common factors. The simulation uses the same data from the results section in the DGP, but imposes different assumptions on the unobserved heterogeneity to simulate new employment observations.

The first simulation analyzes the performance of the OLS, CCEP, CCEMG, and IFE estimators with only state/county and period fixed effects representing the unobserved heterogeneity in the DGP. This DGP uses the OLS results as the true value of the coefficients, with independent and identically distributed (IID) normal errors. For the restaurant employment DGP, these coefficients come from the OLS estimates in Table 3 and the error variance is computed using the residuals from this specification. For the teenage employment DGP, the true value of the coefficients and the error variance come from the OLS estimates in Table 4. The simulation is performed for 1,000 repetitions for each dataset[15].

The results of this simulation are shown in Table 12. Columns (1) and (4) report the median of each of the estimators for restaurant and teenage employment, respectively, and columns (2)-(3) and (5)-(6) report the 95% confidence range of the estimates. The true value of the coefficient for the minimum wage-employment elasticity in the DGP is shown in the first row of the table. All four estimators perform well without the presence of factors in the DGP, with median estimates near the true value. The factor model estimators also perform well in terms of the 95% range of the estimates, with only slightly wider ranges than OLS. There are two important results from this simulation: First, the factor model estimators perform well without the presence of factors in the DGP, even with the small cross-section dimension of the teenage dataset. Second, the pattern of results in this simulation does not match the pattern of results in section 4.1. These results show that the factor model estimators would produce minimum wage-employment elasticities similar to OLS if there were no common factors in the data.

The second simulation analyzes the performance of the OLS, CCEP, CCEMG, and IFE estimators with state/county and period fixed effects and common factors representing the unobserved heterogeneity in the DGP. This DGP uses the coefficients, common factors, and factor loadings from the IFE estimation, with independent and identically distributed (IID) normal errors. For the restaurant employment DGP, the true value of the coefficients comes from Table 3 and the error variance is computed using the residuals from this specification. For the teenage employment DGP, the true value of the coefficients and the error variance come from the IFE estimates in Table 4. The simulation is performed for 1,000 repetitions

[15] The DGP is $y_{it} = \hat{\beta} ln(MW_{it}) + X_{it}\hat{\Gamma} + \hat{\alpha}_i + \hat{\delta}_t + v_{it}$, where the independent variables are the same variables from the main results section, the parameters are from the OLS results reported in Tables 3 and 4, and v_{it} is an idiosyncratic error term whose variance is determined by the variance of the OLS residuals. State/county and period fixed effects are included in the OLS, CCEP, CCEMG, and IFE estimation for this simulation.

for each dataset[16].

The results of this simulation are shown in Table 13. For restaurant employment, the CCEP, CCEMG, and IFE estimators all perform well. The OLS estimator, however, shows consistent and severe negative bias across repetitions. In fact, the true value of the coefficient for the minimum wage-employment elasticity is not even in the 95% range of the OLS estimates. For teenage employment, the OLS estimator once again shows significant negative bias, with the 95% confidence range not containing the true value of the minimum wage-employment elasticity coefficient. The CCEP and CCEMG estimators also show some negative bias for the teenage dataset, although not as much as OLS. This is consistent with the discussion in Section 3.2 that the factor model estimators may not be able to remove all of the bias caused by common factors in the teenage employment data due to the relatively small cross-section dimension of the data. This simulation produces three important results: First, OLS shows significant negative bias when the common factors for restaurant and teenage employment are included in the DGP. Second, the pattern of results from this simulation matches the pattern of results seen in Table 3 and Table 4: the OLS estimates of the minimum wage-employment elasticity are much larger in magnitude than the factor model estimates both in Table 3 and Table 4 and in simulations with common factors included in the DGP. Third, the factor model estimators perform better than OLS when common factors are present even with the small cross-section dimension of the teenage employment data.

In summary, the simulations show that the CCEP, CCEMG, and IFE estimators would produce minimum wage-employment elasticities similar to OLS if state/county and period fixed effects fully represented the unobserved heterogeneity in the underlying data generating process. When common factors are included in the DGP, the OLS estimate of the minimum

[16] The DGP is $y_{it} = \hat{\beta} ln(MW_{it}) + X_{it}\hat{\Gamma} + \hat{\alpha}_i + \hat{\delta}_t + \hat{\lambda}_i' \hat{f}_t + v_{it}$, where the independent variables are the same variables from the main results section, the parameters are from the IFE estimation reported in Tables 3 and 4, and v_{it} is an idiosyncratic error term whose variance is determined by the variance of the IFE residuals. State/county and period fixed effects are included in the OLS, CCEP, CCEMG, and IFE estimation for this simulation.

wage-employment elasticity is negatively biased, while the factor model estimators perform much better. These results suggest that the presence of common factors in the true underlying DGP can cause the different estimates of the minimum wage-employment elasticity seen across approaches in Table 3 and Table 4.

6 Discussion and Conclusion

The minimum wage-employment literature is clearly in disagreement about the effect of raising the minimum wage on employment. The two sides of the debate find different results, despite analyzing the same datasets. The conflicting results arise from issues related to how to properly address regional heterogeneity in employment trends and selection of states that experience minimum wage hikes. Specifically, the two sides are divided on whether census division-by-period fixed effects and contiguous county analysis are appropriate, on robustness to the presence and order of state-specific trends, and on how to best implement a synthetic control approach.

The factor model approach is perfectly suited to resolve these issues. The factor model estimators used in this paper have been shown to perform well without the presence of common factors in the data, but they can also capture census division-by-period fixed effects and state-specific time trends as a special case of the factor model structure. Indeed, there is evidence that the factor model estimators capture both time trends and regional heterogeneity when only two-way fixed effects are included. The difference is that the factor model approach lets the data determine the nature of the unobserved heterogeneity, rather than imposing any specific form a priori. This flexibility is evident from the fact that several of the common factors appear to be something other than time trends and from the fact that, while there is regional heterogeneity in the effect of the common factors, same-census division states and contiguous counties are not always similar in their effect of the common factors. The factor model approach also has several advantages over the use of synthetic controls, such as a more straightforward implementation and the ability to use all of the data. Thus, the factor model approach is robust to critiques from either side of the debate because it can capture the regional heterogeneity and time trends that DLR and ADR want to control for without assuming that local areas are better controls or changing the identifying variation to within census division or across state borders, which NSW argue produces positive bias towards zero. The factor model approach does find positive and statistically significant effects on the earnings of both restaurant workers and teenagers.

The presence of common factors in the data is supported by cross-section dependence tests which reject the null hypothesis that there is only weak cross-section dependence in the data. The factor model estimators find little to no effect of minimum wage increases on employment. For restaurant employment, the factor model estimates of the minimum wage-employment elasticity are in the range of -0.013 to -0.042. For teenage employment, the factor model estimates produce an elasticity in the range of -0.036 to -0.065. None of the factor model estimates are statistically different from zero. These results are generally robust to the presence and order of state-specific time trends, sub-sample analysis, and the number of common factors specified for the IFE estimator. The simulation confirms that the presence of common factors in the underlying DGP can cause the different estimates seen across approaches.

While much attention has been given to the effect of minimum wages on employment levels, less attention has been given to other potential adjustment mechanisms. The lack of evidence supporting negative employment effects suggests that the effect of minimum wages is occurring through other channels, such as employment flows, worker effort, non-wage benefits, hours, or prices. Evidence on these channels is sparse and mixed, and should be the focus of future research.

References

Abadie, A., A. Diamond, and J. 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.

Allegretto, S. A., A. Dube, and M. Reich, 2011: Do minimum wages really reduce teen employment? accounting for heterogeneity and selectivity in state panel data. *Industrial and Labor Relations Review*, **50 (2)**, 204–240.

Allegretto, S. A., A. Dube, M. Reich, and B. Zipperer, 2013: Credible research designs for minimum wage studies. Working Paper 148-13, Institute for Research on Labor and Employment.

Andrews, D. W., 2005: Cross-section regression with common shocks. *Econometrica*, **73**, 1551–1585.

Bai, J., 2009: Panel data models with interactive fixed effects. *Econometrica*, **77**, 1229–1279.

Bai, J. and S. Ng, 2002: Determining the number of factors in approximate factor models. *Econometrica*, **70**, 191–221.

Bailey, N., S. Holly, and M. H. Pesaran, 2014: A two stage approach to spatiotemporal analysis with strong and weak cross-section dependence. Working Paper 4592, Center for Economic Studies and Ifo Institute.

Baleman, D. and P. Wolfson, 2014: *What do Minimum Wages Do?* Kalamazoo, Michigan: W.E. Upjohn Institute.

Baskaya, Y. S. and Y. Rubinstein, 2011: Using federal minimum wage effects to identify the impact of minimum wages on employment and earnings across u.s. states, unpublished paper, Central Bank of Turkey.

Brown, C., 1999: Minimum wages, employment, and the distribution of income. *Hankbook of Labor Economics*, O. Ashenfelter and D. Card, Eds., New York: Elsview.

Cameron, A. C., J. B. Gelbach, and D. L. Miller, 2008: Bootstrap-based improvements for inference with clustered standard errors. *The Review of Economics and Statistics*, **90 (3)**, 414–427.

Dube, A., T. W. Lester, and M. Reich, 2010: Minimum wage effects across state boarders: Estimates using contiguous counties. *The Review of Economics and Statistics*, **92 (4)**, 945–964.

Dube, A., T. W. Lester, and M. Reich, 2014: Minimum wage shocks, employment flows, and labor market frictions. *Journal of Labor Economics (forthcoming).*

Dube, A. and B. Zipperer, 2014: Pooled synthetic control estimates for continuous treatment: An application to minimum wage case studies, unpublished paper.

Gobillon, L. and T. Magnac, 2013: Regional policy evaluation: Interactive fixed effects and synthetic controls. Working Paper 7493, IZA.

Kim, D. and T. Oka, 2014: Divorce law reform and divorce rates in the usa: An interactive fixed-effects approach. *Journal of Applied Econometrics*, **29**, 231–245.

Neumark, D., J. I. Salas, and W. Wascher, 2014a: Revisiting the minimum wage-employment debate: Throwing out the baby with the bathwater. *Industrial and Labor Relations Review*, **67 (3)**, 608–648.

Neumark, D., J. I. Salas, and W. Wascher, 2014b: More on recent evidence on the effects of minimum wages in the united states. Working Paper 20619, NBER.

Neumark, D. and W. Wascher, 1992: Employment effects of minimum and subminimum wages: Panel data on state minimum wage laws. *Industrial and Labor Relations Review*, **46 (1)**, 55–81.

Neumark, D. and W. Wascher, 2007: Minimum wages, the earned income tax credit and employment: Evidence from the post-welfare reform era. Discussion Paper 2610, IZA.

Neumark, D. and W. Wascher, 2008: *Minimum Wages*. Cambridge, MA: MIT Press.

Obenauer, M. L. and B. von der Nienburg, 1915: Effect of minimum-wage determination in oregon. Bureau of labor statistics bulletin no. 176, United States Department of Labor.

Pesaran, M. H., 2006: Estimation and inference in large heterogeneous panels with a multi-factor error structure. *Econometrica*, **74**, 967–1012.

Pesaran, M. H., 2015: Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, **34**, 1089–1187.

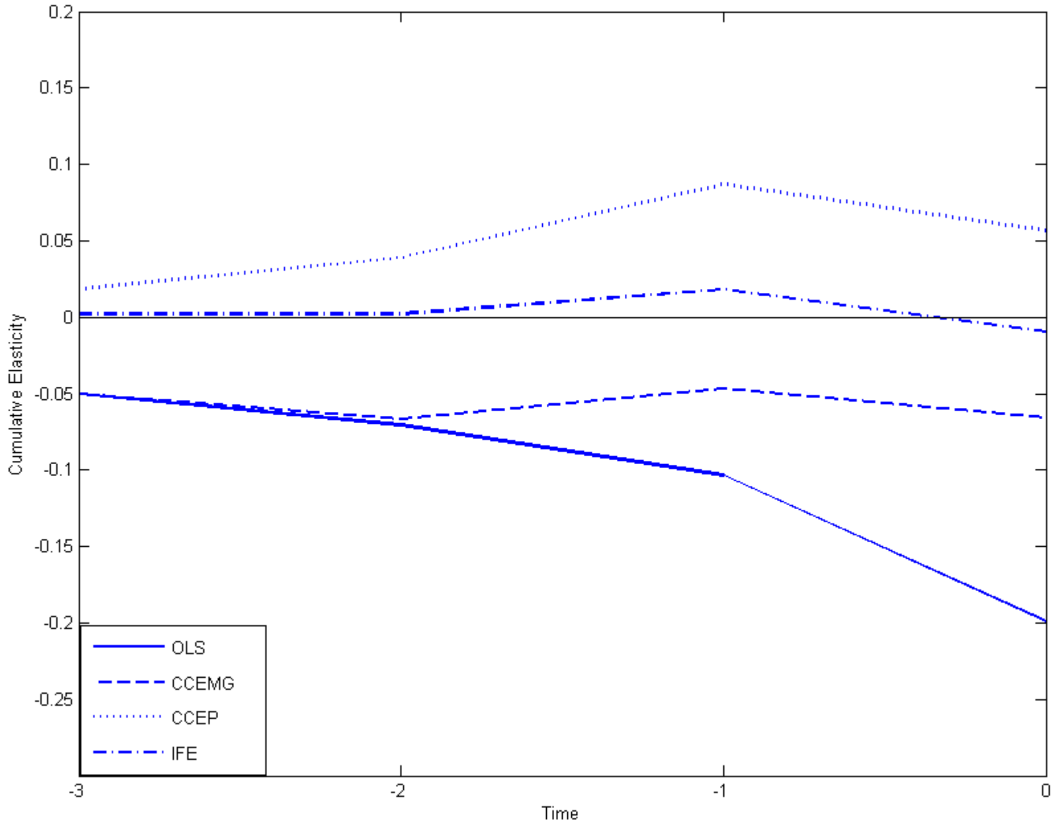
Sabia, J., 2009: The effects of minimum wage increases on retail employment and hours: New evidence from monthly cps data. *Journal of Labor Research*, **30 (1)**, 75–97.

Schmitt, J., 2013: Why does the minimum wage have no discernible effect on employment? Tech. rep., Center for Economic and Policy Research.

Smith, C. L., 2011: Polarization, immigration, education: What's behind the dramatic decline in youth employment? Finance and Economics Discussion Series 2011-41, Board of Governors of the Federal Reserve System.

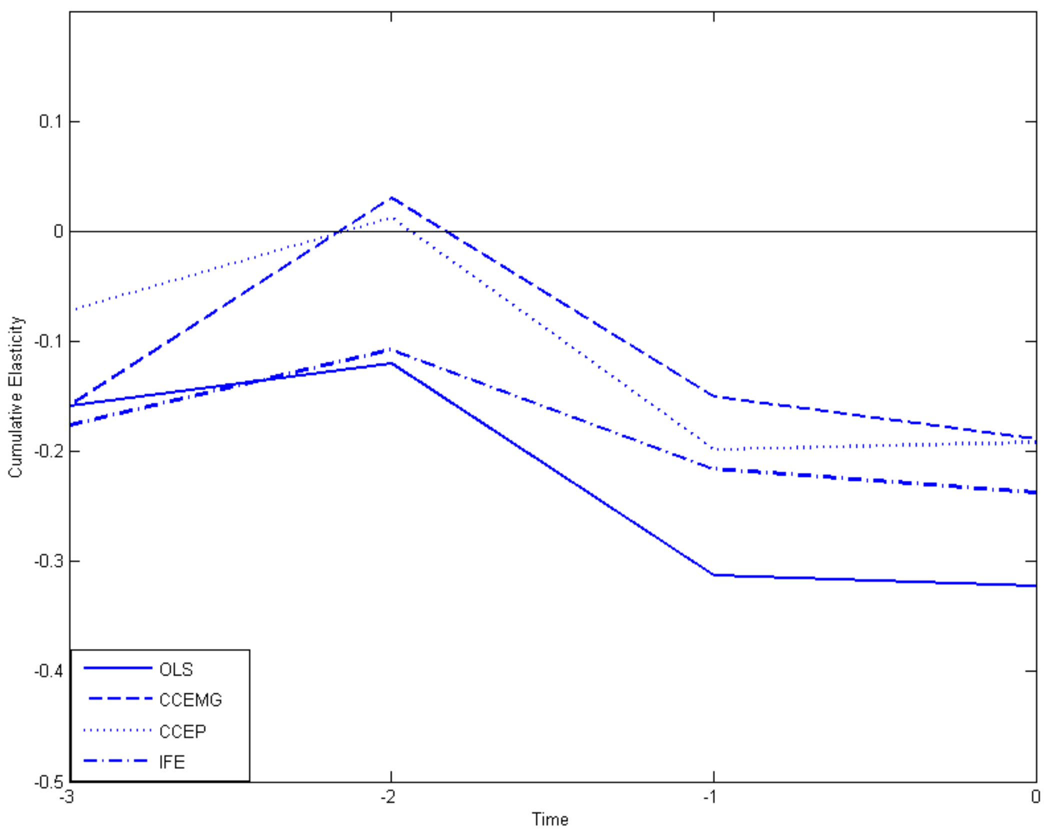
Westerlund, J. and J.-P. Urbain, 2015: Cross-sectional averages versus principal components. *Journal of Econometrics*, **185**, 372–377.

Figure 1: Pre-Existing Trends - Restaurant Employment



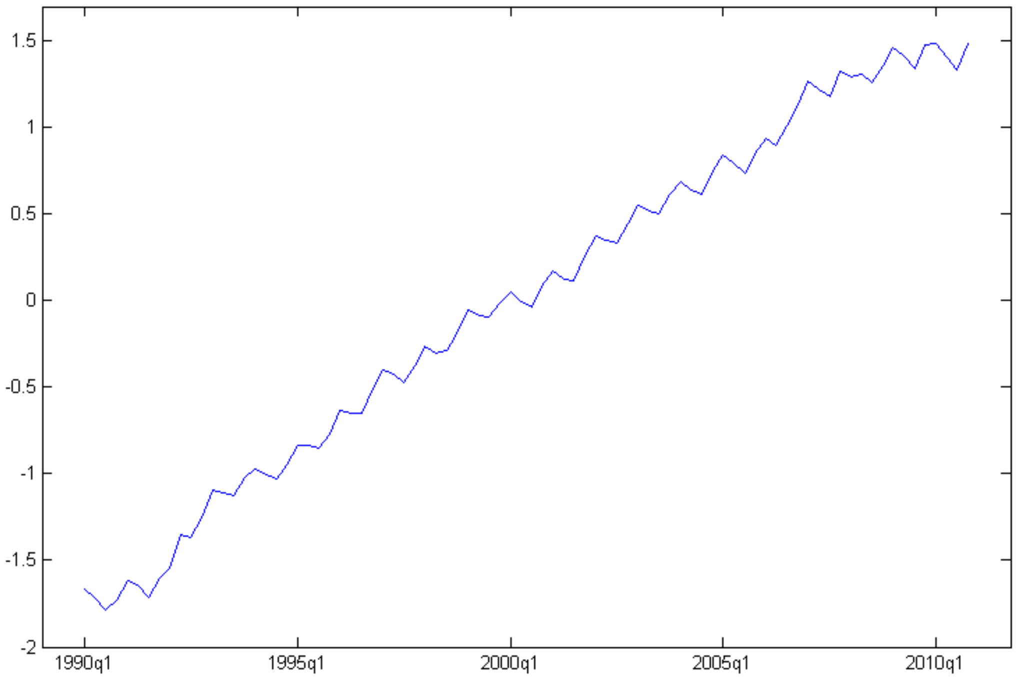
This figure plots lead effects of minimum wage hikes. The results are based on equation (1) with 1-, 2-, and 3-year leads in the minimum wage variable included. The elasticity reported at each point in time is the cumulative elasticity, which is the sum of the contemporaneous elasticity for each year until that point in time.

Figure 2: Pre-Existing Trends - Teenage Employment



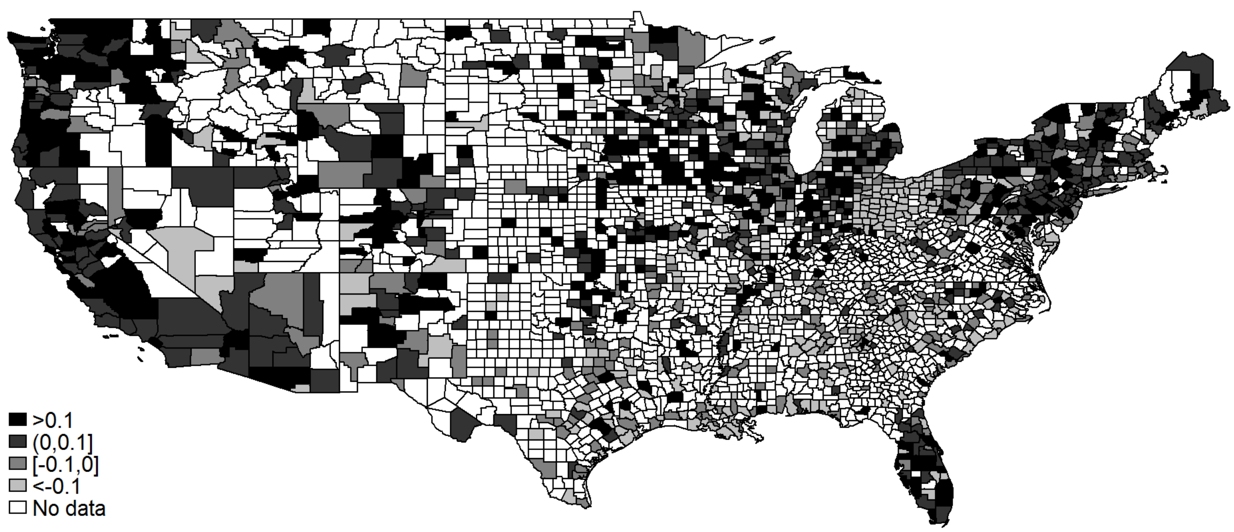
See Figure 1 for details.

Figure 3: First Common Factor for Restaurant Employment



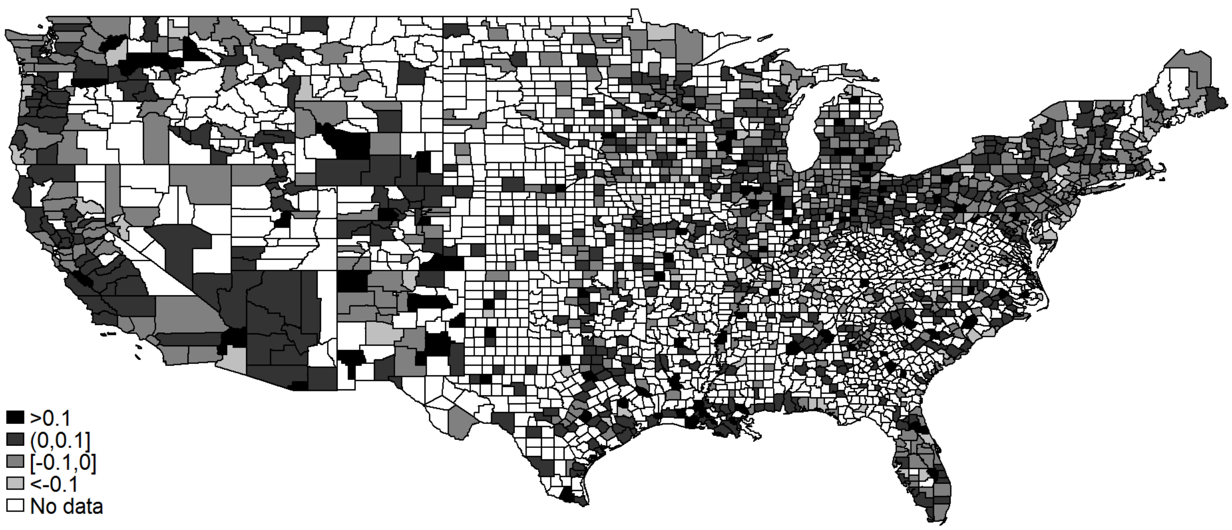
This figure plots the first common factor for restaurant employment, estimated from the IFE approach.

Figure 4: Effect of Common Factors - Restaurant Employment, 1990q1



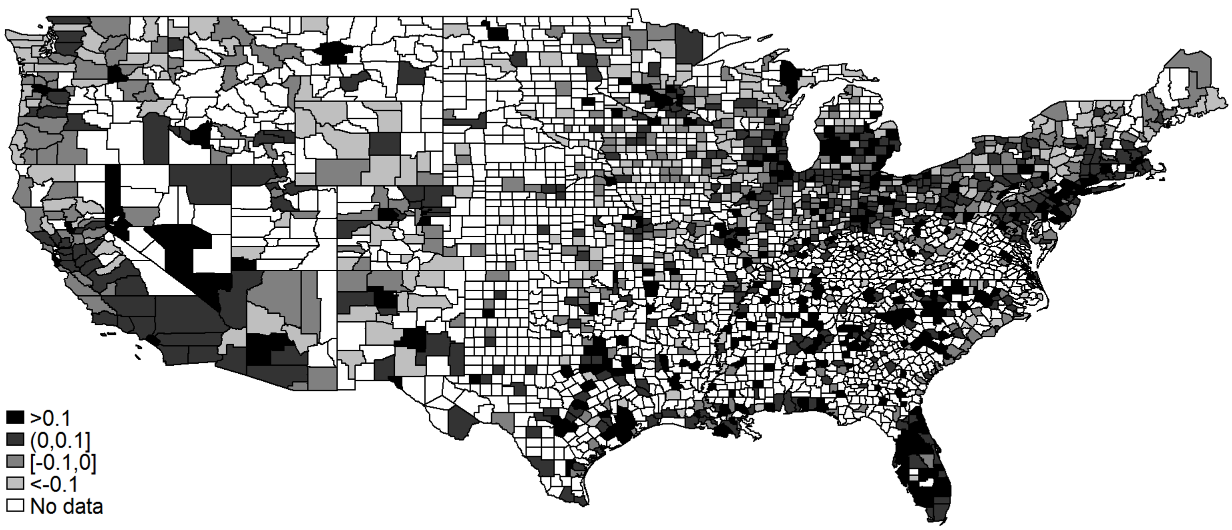
This figure plots $\lambda_i' f_t$ for the specified period, estimated from the IFE approach, which represents the combined effect of the estimated unobserved common factors on log employment in the given county or state.

Figure 5: Effect of Common Factors - Restaurant Employment, 2000q1



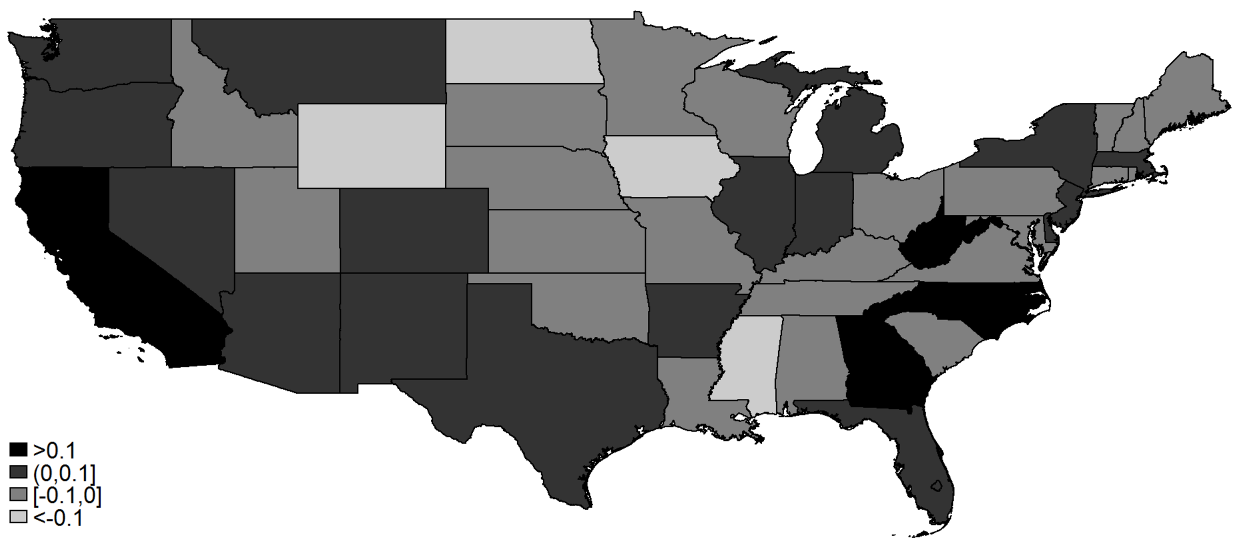
See Figure 4 for details.

Figure 6: Effect of Common Factors - Restaurant Employment, 2010q1



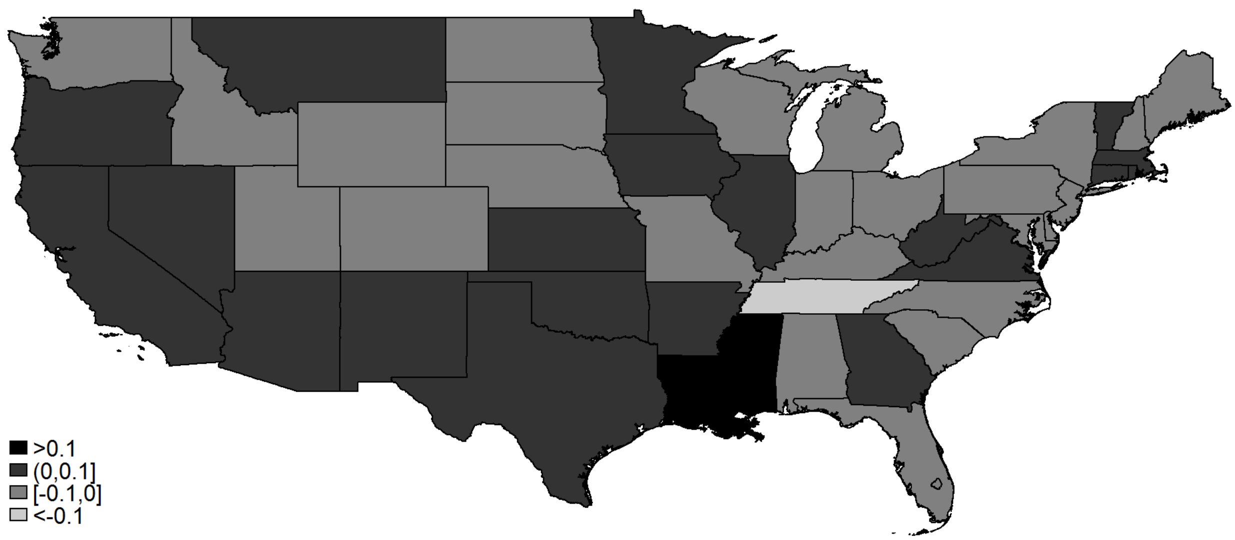
See Figure 4 for details.

Figure 7: Effect of Common Factors - Teenage Employment, 1990q1



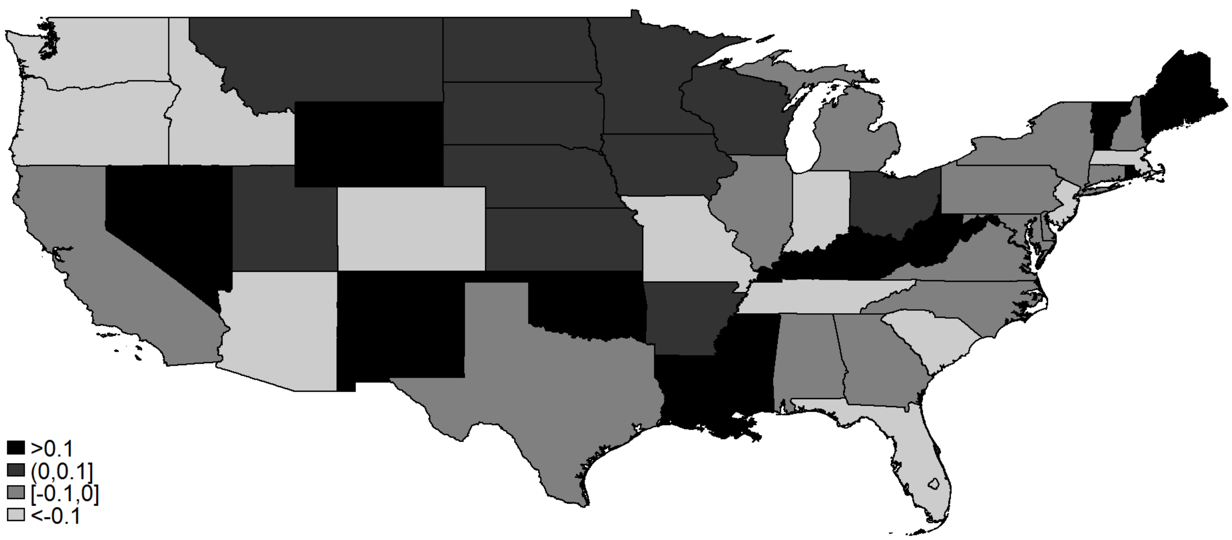
See Figure 4 for details.

Figure 8: Effect of Common Factors - Teenage Employment, 2000q1



See Figure 4 for details.

Figure 9: Effect of Common Factors - Teenage Employment, 2010q1



See Figure 4 for details.

Table 1: Review of Recent Literature

Study	Population	Approach	Elasticity	Effect
Sabia (2009)	Retail	Traditional	-0.106***	Negative
Neumark and Wascher (2007)	Teenagers	Traditional	-0.136*	Negative
DLR (2010)	Restaurants	Census division-by-period fixed effects	-0.023	None
		CDxP FE & state-specific linear trends	0.054	None
		Contiguous county pairs	0.016	None
ADR (2011)	Teenagers	Census division-by-period fixed effects	-0.036	None
		State-specific linear trends	-0.034	None
		CDxP FE & state-specific linear trends	0.047	None
NSW (2014a)	Restaurants	Synthetic Controls	-0.063***	Negative
	Teenagers	State-specific 5th-order polynomial trend	-0.185**	Negative
		Synthetic Controls	-0.145**	Negative
ADRZ (2013)	Teenagers	Synthetic Controls	-0.028	None

The elasticity result is taken directly from the results reported in each study. For Neumark and Wascher (2007), this elasticity is constructed using the employment-population ratio in Table 1 and the employment coefficient in Table 2, specification 1. Sabia (2009) is not able to analyze sub-sectors of retail employment such as restaurant employment, but restaurant employment is included in the retail sector. The "traditional" approach refers to using two-way fixed effects, with no additional controls for regional heterogeneity or selection of states experiencing minimum wage hikes. DLR=Dube et al. (2010), ADR=Allegretto et al. (2011), NSW=Neumark et al. (2014a), ADRZ=Allegretto et al. (2013). The "synthetic control" approach in NSW pools all synthetic and real data together, and then estimates the two-way fixed effects approach with a fixed effect for each set of synthetic and real observations. Significance levels are as follows: *10 percent, **5 percent, ***1 percent.

Table 2: Summary Statistics

	Mean (1)	Standard Deviation (2)
*Restaurant employment (CD test statistic: 60.06***)*		
Restaurant Employment	4,766	11,144
Total Private Sector Employment	70,959	181,564
Population	180,971	422,618
Minimum Wage	$5.49	$1.22
Periods	84	
Number of Counties	1,371	
*Teenage employment (CD test statistic: -5.64***)*		
Fraction of Teenagers Employed	0.41	0.12
Unemployment Rate	5.68	2.14
Relative Size of Teenage Population	0.09	0.01
Minimum Wage	$5.58	$1.26
Periods	96	
N	51	

Data on restaurant employment and total private sector employment come from the Quarterly Census of Employment and Wages. County population data comes from the Census Bureau. Teenage employment data and controls come from the CPS Outgoing Rotation Groups and are aggregated to the state-quarter level. The minimum wage is always the higher of the federal and state minimum wage, reported in nominal dollars. Tests for strong cross-section dependence are based on the test in Pesaran (2015).

Table 3: Minimum Wage-Employment Elasticity - Restaurant Employment

	OLS (1)	CCEMG (2)	CCEP (3)	IFE (4)
	Dependent variable: ln(employment)			
ln(MW)	**-0.138***	**-0.013**	**-0.013**	**-0.042**
	[-0.297,0.019]	[-0.042,0.026]	[-0.046,0.028]	[-0.085,0.015]
ln(Total private sector emp.)	0.512***	0.704***	0.585***	0.519***
	[0.430,0.595]	[0.667,0.742]	[0.515,0.653]	[0.424,0.601]
ln(Population)	0.587***	0.373***	0.412***	0.296***
	[0.432,0.742]	[0.184,0.566]	[0.285,0.547]	[0.138,0.436]
Period fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
CD test statistic	26.98***	4.43***	17.19***	-0.30
TxN	115,164	115,164	115,164	115,164

IFE results are based on 4 common factors. OLS standard errors are clustered at the state level. Standard errors for CCEMG, CCEP, and IFE are calculated following the description in the appendix. The confidence intervals and significance reported for CCEMG, CCEP, and IFE are based on bootstrapped t-statistics following the *wild cluster bootstrap-t* procedure in Cameron et al. (2008), clustered at the state level. Significance levels are as follows: *10 percent, **5 percent, ***1 percent. Residual diagnostics for strong cross-section dependence are based on the test in Pesaran (2015).

Table 4: Minimum Wage-Employment Elasticity - Teenage Employment

	OLS (1)	CCEMG (2)	CCEP (3)	IFE (4)
	Dependent variable: ln(employment/population)			
ln(MW)	**-0.178****	**-0.040**	**-0.065**	**-0.036**
	[-0.323,-0.033]	[-0.214,0.135]	[-0.191,0.061]	[-0.157,0.097]
Unemployment rate	-3.608***	-2.660***	-2.805***	-1.787***
	[-4.243,-2.973]	[-3.162,-2.158]	[-3.415,-2.195]	[-2.457,-1.347]
Teen population share	-0.154	0.482	0.274	0.249
	[-0.709,0.401]	[-0.104,1.068]	[-0.223,0.771]	[-0.233,0.773]
Period fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
CD test statistic	-5.96***	-6.01***	-6.03***	-6.69***
TxN	4,896	4,896	4,896	4,896

IFE results are based on 8 common factors. See Table 3 for additional details.

Table 5: Summary Statistics for Common Factors

	Factor #p									
	1	2	3	4	5	6	7	8	9	10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Restaurant employment									
AR1($\hat{f}_{pt}$)	0.990	0.975	0.527	0.447	0.988	0.969	0.954	0.884	0.882	0.867
R_p^2	0.538	0.707	0.785	0.857	0.899	0.928	0.951	0.971	0.987	0.999
	Teenage employment									
AR1($\hat{f}_{pt}$)	0.496	0.295	-0.014	-0.009	0.021	0.031	-0.015	0.008	0.122	0.028
R_p^2	0.195	0.374	0.470	0.561	0.648	0.730	0.806	0.879	0.944	0.975

The first 10 common factors of restaurant and teenage employment come from the IFE estimates of the factor structure with 10 pre-specified common factors. R_p^2 is the relative importance of each factor, calculated as the fraction of the total variance of the residuals explained by factors 1 to p. This is given as the sum of the first p largest eigenvalues of the sample second moment matrix of the OLS residuals divided by the sum of all eigenvalues. AR1($\hat{f}_{pt}$) is the first order autocorrelation coefficient for the given factor. AR(1) coefficients for the detrended version of the first two restaurant factors are 0.69 and 0.79, respectively.

Table 6: IFE Estimates for Different Numbers of Common Factors

	Number of common factors						
	2	3	4	5	6	7	8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Restaurant employment						
ln(MW)	-0.016	-0.035	-0.042	-0.023	-0.035	-0.008	-0.007
	[-0.056,0.029]	[-0.072,0.019]	[-0.085,0.015]	[-0.068,0.028]	[-0.075,0.016]	[-0.044,0.030]	[-0.052,0.023]
CD test statistic	29.02***	29.39***	-0.30	-0.22	0.40	0.61	0.74
	Teenage employment						
ln(MW)	-0.104	-0.093	-0.069	-0.064	-0.061	-0.018	-0.036
	[-0.233,0.038]	[-0.224,0.045]	[-0.196,0.076]	[-0.179,0.072]	[-0.175,0.079]	[-0.134,0.125]	[-0.157,0.097]
CD test statistic	-6.45***	-6.45***	-6.48***	-6.52***	-6.64***	-6.66***	-6.69***

This table reports IFE results for different numbers of pre-specified common factors. See Table 3 for additional details about standard errors, inference, and CD test statistics.

Table 7: State-Specific Trend Robustness Checks - Restaurant Employment

	Polynomial order for state-specific trend					
	None	1st	2nd	3rd	4th	5th
	(1)	(2)	(3)	(4)	(5)	(6)
OLS	-0.138*	-0.041**	-0.024	-0.052***	-0.040**	-0.031*
	[-0.297,0.019]	[-0.075,-0.008]	[-0.058,0.009]	[-0.085,-0.020]	[-0.071,-0.009]	[-0.067,0.003]
CCEMG	-0.013	-0.003	-0.012	-0.013	0.001	-0.007
	[-0.042,0.026]	[-0.039,0.022]	[-0.044,0.019]	[-0.045,0.018]	[-0.032,0.029]	[-0.035,0.022]
CCEP	-0.013	-0.003	-0.018	-0.041***	-0.013	-0.007
	[-0.046,0.028]	[-0.036,0.025]	[-0.049,0.014]	[-0.073,-0.021]	[-0.038,0.022]	[-0.034,0.024]
IFE	-0.042	-0.048	-0.041	-0.044	-0.033	-0.025
	[-0.085,0.015]	[-0.087,0.014]	[-0.084,0.015]	[-0.085,0.013]	[-0.080,0.019]	[-0.078,0.023]

This table reports results with controls for state-specific time trends. Column (1) is the traditional two-way fixed effects result without state-specific time trends shown in Tables 3 and 4. Columns (2)-(6) add flexible state-specific time trends, beginning with linear state-specific time trends and extending to 5th-order polynomial state-specific time trends. IFE results are based on 4 common factors. See Table 3 for additional details about standard errors and inference.

Table 8: State-Specific Trend Robustness Checks - Teenage Employment

	Polynomial order for state-specific trend					
	None	1st	2nd	3rd	4th	5th
	(1)	(2)	(3)	(4)	(5)	(6)
OLS	-0.178**	-0.074	-0.045	-0.142**	-0.108	-0.125*
	[-0.323,-0.033]	[-0.194,0.045]	[-0.167,0.077]	[-0.303,-0.019]	[-0.275,0.059]	[0.270,-0.020]
CCEMG	-0.040	-0.025	0.029	-0.129	-0.088	-0.106
	[-0.214,0.135]	[-0.198,0.147]	[-0.119,0.193]	[-0.321,0.075]	[-0.293,0.129]	[0.310,0.089]
CCEP	-0.065	-0.043	0.009	-0.090	-0.018	-0.065
	[-0.191,0.061]	[-0.173,0.088]	[-0.106,0.125]	[-0.260,0.080]	[-0.192,0.155]	[-0.221,0.091]
IFE	-0.036	0.041	0.028	-0.079	-0.064	-0.094
	[-0.157,0.097]	[-0.053,0.137]	[-0.074,0.128]	[-0.220,0.054]	[-0.174,0.062]	[-0.208,0.098]

IFE results are based on 8 common factors. See Table 7 for additional details.

Table 9: Sub-Sample Robustness Checks - Restaurant Employment

	1990q1-2000q2 (1)	1995q2-2005q3 (2)	2000q3-2010q4 (3)
OLS	-0.018	-0.137**	-0.021
	[-0.177,0.141]	[-0.264,-0.010]	[-0.095,0.019]
CCEMG	0.120	0.088	-0.014
	[-0.335,0.095]	[-0.219,0.042]	[0.013,-0.049]
CCEP	-0.060**	-0.003	-0.020
	[-0.004,-0.118]	[0.039,-0.043]	[0.015,-0.059]
IFE	-0.052	-0.015	0.002
	[0.015,-0.013]	[0.056,-0.055]	[0.063,-0.053]

This table reports results based on three sub-samples of the data. The dimensions of the data in the sub-samples are $N = 1371$ and $T = 42$. IFE results are based on 4 common factors. See Table 3 for additional details about standard errors and inference.

Table 10: Sub-Sample Robustness Checks - Teenage Employment

	1990q1-2001q4 (1)	1996q1-2007q4 (2)	2002q1-2013q4 (3)
OLS	-0.205**	-0.102	-0.069
	[-0.409,-0.001]	[-0.249,0.045]	[-0.261,0.131]
CCEMG	-0.020	-0.113	-0.045
	[-0.519,0.485]	[-0.364,0.145]	[-0.349,0.258]
CCEP	-0.156	-0.107	-0.066
	[-0.376,0.073]	[-0.310,0.112]	[-0.282,0.159]
IFE	-0.142	-0.102	-0.069
	[-0.416,0.186]	[-0.303,0.109]	[-0.268,0.143]

This table reports results based on three sub-samples of the data. The dimensions of the data in the sub-samples are $N = 51$ and $T = 48$. IFE results are based on 8 common factors. See Table 3 for additional details about standard errors and inference.

Table 11: Minimum Wage Effects on Earnings and a Falsification Test

	OLS (1)	CCEMG (2)	CCEP (3)	IFE (4)
		Restaurants		
Earnings	0.209***	0.231***	0.222***	0.145***
	[0.160,0.257]	[0.175,0.264]	[0.199,0.251]	[0.091,0.210]
		Teenagers		
Earnings	0.104***	0.097**	0.110***	0.158***
	[0.041,0.167]	[0.010,0.188]	[0.034,0.184]	[0.083,0.232]
		Manufacturing		
Employment	-0.013	0.015	0.026	0.007
	[-0.176,0.149]	[-0.028,0.051]	[-0.021,0.063]	[-0.038,0.041]
Earnings	-0.096	0.086	0.085**	-0.045
	[-0.314,0.121]	[-0.078,0.235]	[0.013,0.161]	[-0.139,0.063]
Period fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes

Results for restaurant and teenage earnings are based on the specification in equation (1) and estimated in Table 3 and 4, except with worker earnings as the dependent variable. Results for the manufacturing industry are also based on this specification, except with manufacturing employment/earnings as the dependent variable. IFE results are based on 8 factors for teenagers and 4 factors for restaurant and manufacturing. See Table 3 for additional details about standard errors and inference.

Table 12: Minimum Wage-Employment Elasticity Simulation Results - No Factors in DGP

	Restaurant Employment			*Teenage Employment*		
	Median	2.5%	97.5%	Median	2.5%	97.5%
	(1)	(2)	(3)	(4)	(5)	(6)
True value	***-0.138***			***-0.178***		
OLS	**-0.139**	-0.152	-0.128	**-0.177**	-0.223	-0.134
CCEMG	**-0.139**	-0.167	-0.111	**-0.179**	-0.248	-0.107
CCEP	**-0.138**	-0.159	-0.120	**-0.177**	-0.234	-0.124
IFE	**-0.138**	-0.153	-0.124	**-0.177**	-0.225	-0.131

This table reports simulation results for the case without common factors in the data generating process. The DGP is $y_{it} = \hat{\beta} ln(MW_{it}) + X_{it}\hat{\Gamma} + \hat{\alpha}_i + \hat{\delta}_t + v_{it}$, where the independent variables are the same variables used in the results section, the parameters are from the OLS results for the traditional state and period fixed effects specification reported in Table 3 and Table 4, and v_{it} is an idiosyncratic error term whose variance is determined by the variance of the OLS residuals. The number of repetitions is 1,000.

Table 13: Minimum Wage-Employment Elasticity Simulation Results - Factors in DGP

	Restaurant Employment			*Teenage Employment*		
	Median	2.5%	97.5%	Median	2.5%	97.5%
	(1)	(2)	(3)	(4)	(5)	(6)
True value	***-0.042***			***-0.036***		
OLS	**-0.175**	-0.203	-0.135	**-0.175**	-0.222	-0.129
CCEMG	**-0.045**	-0.092	0.007	**-0.134**	-0.198	-0.058
CCEP	**-0.041**	-0.087	-0.002	**-0.101**	-0.158	-0.047
IFE	**-0.043**	-0.072	-0.011	**-0.038**	-0.088	0.026

This table reports simulation results for the case with common factors in the data generating process. The DGP is $y_{it} = \hat{\beta} ln(MW_{it}) + X_{it}\hat{\Gamma} + \hat{\alpha}_i + \hat{\delta}_t + \hat{\lambda}_i' \hat{f}_t + v_{it}$, where the independent variables are the same variables used in the results section, the parameters are from the IFE results for the traditional state and period fixed effects specification reported in Table 3 and Table 4, and v_{it} is an idiosyncratic error term whose variance is determined by the variance of the IFE residuals. The number of repetitions is 1,000.

A Appendix

A.1 Common Correlated Effects

The individual slope coefficients for the CCE estimator are given by

$$\hat{b}_i = (X_i'\bar{M}_\omega X_i)^{-1}X_i'\bar{M}_\omega y_i,$$

where X_i and y_i are the independent and dependent variables, $\bar{M}_\omega$ is given by

$$\bar{M}_\omega = I_T - \bar{H}_\omega(\bar{H}_\omega'\bar{H}_\omega)^{-1}\bar{H}_\omega',$$

and $\bar{H}_\omega = (D, \bar{Z}_\omega)$, with D being a (T x 1) vector of ones, which produces an individual fixed effect, and $\bar{Z}_\omega$ being the (T x (k+1)) matrix of cross-sectional averages of the dependent and independent variables.

Pesaran (2006) then suggests two versions of the CCE estimator. The Commom Correlated Effects - Mean Group (CCEMG) is the average of the individual CCE coefficients,

$$\hat{b}_{CCEMG} = N^{-1}\sum_{i=1}^{N}\hat{b}_i,$$

while the Common Correlated Effects - Pooled (CCEP) is given by

$$\hat{b}_{CCEP} = \left(\sum_{i=1}^{N}\theta_i X_i'\bar{M}_\omega X_i\right)^{-1}\sum_{i=1}^{N}\theta_i X_i'\bar{M}_\omega y_i.$$

The CCEP assumes that $\beta_i = \beta$ for all i, although it does allow the slope coefficients of the common effects to differ across i.

The variance of the CCEMG estimator is given by

$$\hat{\Sigma}_{CCEMG} = (N-1)^{-1}\sum_{i=1}^{N}(\hat{b}_i - \hat{b}_{CCEMG})(\hat{b}_i - \hat{b}_{CCEMG})'.$$

The variance of the CCEP estimator is given by

$$\hat{\Sigma}_{CCEP} = N^{-1}\hat{\Psi}^{-1}\hat{R}\hat{\Psi}^{-1},$$

where $\hat{\Psi}$ and $\hat{R}$ are given by

$$\hat{\Psi} = \sum_{i=1}^{N} N^{-1}\left(\frac{X_i'\bar{M}_\omega X_i}{T}\right)$$

and

$$\hat{R} = (N-1)^{-1}\sum_{i=1}^{N} \frac{1/N}{\sqrt{N^{-2}}}\left(\frac{X_i'\bar{M}_\omega X_i}{T}\right)(\hat{b}_i - \hat{b}_{CCEMG})(\hat{b}_i - \hat{b}_{CCEMG})'\left(\frac{X_i'\bar{M}_\omega X_i}{T}\right).$$

A.2 Interactive Fixed Effects

Bai (2009) notes that, given F and Λ, the regression coefficients could be estimated in the usual way after subtracting the factor structure out of the data,

$$\hat{\beta}(F) = \left(\sum_{i=1}^{N} X_i'X_i\right)^{-1} \sum_{i=1}^{N} X_i'(Y_i - F\lambda_i),$$

and, given β, principal component analysis could be used to compute F and Λ from the pure factor model,

$$W_i = Y_i - X_i\beta.$$

In practice, both the regression coefficients and the factor structure are unknown. Therefore, Bai (2009) proposes an iterative procedure:

Step 1: Ignore the factor structure and estimate the regression coefficients. Given these regression coefficients, estimate the factors and factor loadings. Given these factors and factor loadings, re-estimate the regression coefficients. Iterate until the percent change in the sum of squared residuals in the regression coefficient estimation is less than a specified threshold[17].

Step 2: Ignore the regression coefficients and estimate the factors and factor loadings. Given these factors and factor loadings, estimate the regression coefficients. Given these regression coefficients, re-estimate the factors and factor loadings. Iterate until the percent change in the sum of squared residuals in the factor and factor loading estimation is less than a specified threshold.

Step 3: Keep the estimates from the previous step that had the lowest final sum of squared residuals.

The data is demeaned in both directions before estimation to account for unit and period fixed effects. Bias correction is performed using equations (23) and (24) in Bai (2009).

The variance of the IFE estimator is given by

$$\hat{\Sigma}_{IFE} = \frac{1}{NT} D_0^{-1} D_Z D_0^{-1},$$

where $D_0 = (NT)^{-1} \sum_{i=1}^{N} Z_i' Z_i$, $D_Z = N^{-1} \sum_{i=1}^{N} \hat{\sigma}_i^2 (T^{-1} \sum_{t=1}^{T} z_{it}' z_{it})$, with $\hat{\sigma}_i^2 = T^{-1} \sum_{t=1}^{T} \hat{\epsilon}_{it}^2$, $Z_i = M_{\hat{F}} X_i - N^{-1} \sum_{k=1}^{N} [\hat{\gamma}_i' (\hat{L}'\hat{L}/N)^{-1} \hat{\gamma}_k] M_{\hat{F}} X_k$, and $\hat{L} = (\hat{\gamma}_1, ..., \hat{\gamma}_N)'$.

[17]The CCEMG estimates were also used as an alternative way to initiate step 1. In this case, the initial estimate of the regression coefficients is based on the CCEMG results. Given these estimates, the factor structure is estimated using principal components. The regression coefficients are then re-estimated, given these factors and factor loadings, using the IFE estimate of the regression coefficients described above. Iteration then continues as described in step 1. This approach convergences much faster than when the initial estimate of the regression coefficients is based on traditional OLS and finishes with approximately the same final sum of squared residuals. The results are nearly identical.