DISCUSSION PAPER SERIES

IZA DP No. 9867

Labor Market Effects of US Sick Pay Mandates

Stefan Pichler
Nicolas R. Ziebarth

April 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Labor Market Effects of US Sick Pay Mandates

Stefan Pichler
ETH Zurich, KOF Swiss Economic Institute

Nicolas R. Ziebarth
Cornell University and IZA

Discussion Paper No. 9867
April 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9867
April 2016

ABSTRACT

Labor Market Effects of US Sick Pay Mandates

This paper exploits temporal and spatial variation in the implementation of US sick pay mandates to assess their labor market consequences. We use the Synthetic Control Group Method (SCGM) and the Quarterly Census of Employment and Wages (QCEW) to estimate the causal effect of mandated sick leave on employment and wages. Our findings do not provide much evidence that employment or wages were significantly affected by the mandates which typically allow employees to earn one hour of paid sick leave per work week, up to seven days per year. Joint tests for all treatment regions let us exclude, with 90% statistical probability, that wages decreased by more than 1% as a result of the mandates. With 92% probability, we can exclude that employment decreased by more than 1%.

JEL Classification: I12, I13, I18, J22, J28, J32

Keywords: sick pay mandates, sick leave, medical leave, employer mandates, employment, wages, synthetic control group, United States, Quarterly Census of Employment and Wages (QCEW)

Corresponding author:

Nicolas R. Ziebarth
Department of Policy Analysis and Management (PAM)
Cornell University
106 Martha Van Rensselaer Hall
Ithaca, NY 14850
USA
E-mail: nrz2@cornell.edu

1 Introduction

As a component of the first federal health insurance legislation, paid sick leave was one of the first social insurance pillars worldwide. The *Sickness Insurance Law of 1883* implemented federally mandated employer-provided health insurance in Germany, which covered up to 13 weeks of paid sick leave as well as medical care. Insurance against wage losses due to health shocks was a crucial element of health insurance at that time, and valued by employees and unions alike. Given the limited availability of expensive medical treatments in the 19^{th} century, expenditures for paid sick leave initially accounted for more than half of all health insurance expenditures (Busse and Riesberg, 2004). Increasingly more European countries implemented paid sick leave and today, virtually every European country has some form of universal access to paid sick leave—with varying degrees of generosity.

To date, the US is the only industrialized country that does not provide universal access to paid sick leave, which is largely provided as a fringe benefit by employers on a voluntary basis (Heymann et al., 2009). Coverage rates among full-time workers are around 65%; low-income, part-time and service sector workers have coverage rates of less than 20% (Lovell, 2003; Boots et al., 2009; Susser and Ziebarth, 2016). Susser and Ziebarth (2016) estimate that, in a given week of the year, the total demand for paid sick leave sums to ten percent of all US employees. In addition to concerns related to inequality, worker well-being, and productivity concerns, a lack of sick leave coverage may induce contagious employees to work sick and spread diseases. Pichler and Ziebarth (2015) model the trade-off between shirking and contagious presenteeism when coverage is expanded. They show that infection rates decrease and shirking increases when employees gain access to paid sick leave.

In the last decade, support for sick leave mandates has grown substantially in the US. On the city level, sick leave mandates were passed and implemented in San Francisco (2007), Washington D.C. (2008), Seattle (2012), New York City (2014), Portland (2014), Newark (2014), Philadelphia (2015), and Oakland (2015).

On the state level, Connecticut was first to mandate paid sick leave in 2012. However, the bill excludes businesses with less than 50 full time employees and only applies to the service sector. Subsequently, it only covers about 20% of the workforce (Miller and Williams, 2015; Connecticut Department of Labor, 2015). In contrast, California passed a much more comprehensive bill—covering all employees—effective July 1, 2015. Massachusetts and Oregon also passed relatively comprehensive sick leave mandates, effective July 2015 and January 2015, but exempted small businesses (see Appendix B1).

Lastly, at the federal level, reintroduced in Congress in March 2013, the *Healthy Families Act* proposes a federal sick leave mandate that would cover employees in businesses with more than 15 employees (US Congress, 2015). Similar to the mandates already in place at the state or city level, the *Healthy Families Act* proposes that employees 'earn' one hour of paid sick leave per 30 hours worked, up to 56 hours (or 7 days) per year. Paid sick leave—at the standard wage rate of 100%—can then be taken in case of own sickness or sickness of a relative, in most cases sickness of children.

A main source of controversy related to government mandated sick leave is the possibility that such policy would hurt employment or wage growth. At first sight, it seems obvious that an expansion in sick leave coverage would lead to moral hazard and higher labor costs for employers (Ziebarth and Karlsson, 2010). When employees earn one hour of paid sick leave per 30 hours worked—ignoring administrative costs—this would equal a wage increase of 1/30 or 3.3% per week for full-time employees. This static calculation assumes that all employees would fully exhaust their annual sick leave credit and would have worked sick with full productivity (or taken unpaid leave) in the counterfactual scenario. However, empirically assessing and directly measuring labor productivity under the two scenarios is extremely challenging (if not impossible). The effects of the policy also mostly likely differ by type of job and industry. To our knowledge, there exists no credible empirical causal evidence on how work productivity changes when employees gain access to paid sick leave. It seems likely that sick employees cannot maintain full work productivity when working sick and that employees on sick leave will compensate for lost productivity after recovery. Hence, the calculated static wage increase of 3.3% appears to be an upper bound for marginal firms.

Consequently—abstaining from administrative costs, changes in work productivity, psychological costs or benefits, and reductions in presenteeism and increases in shirking—one would hypothesize that wages would grow at lower rates after the implementation of sick pay mandates. One could also hypothesize that marginal employees might not get hired or even be laid off, and that employers may convert full-time jobs to part-time when mandates become implemented in non-small businesses above a certain threshold of full-time employees.

This paper empirically assesses how wages and employment have changed after the implementation of sick pay mandates in most of the US regions listed above (also see Table B1). We generate two datasets from the Quarterly Census of Employment and Wages (QCEW) which is provided by the Bureau of Labor Statistics (BLS). The first dataset records total monthly employment and quarterly wages at the county–industry level from 2001 to 2015. The second dataset records total monthly employment and quarterly wages at the state–industry–firm-size level from 2001 to 2015. Econometrically, we exploit the quasi-random nature of the implementation of the sick pay mandates across US regions and over time. To mimic pre-treatment trends as closely as possible, we follow Abadie and Gardeazabal (2003) and Abadie et al. (2010) and build synthetic control groups using the many untreated regional units available.[1] Finally, we use the approach suggested in Dube and Zipperer (2015) for hypothesis testing with single and multiple events.

The setting exploited in this paper carries additional advantages over the standard example of one treated region and a limited number of potential 'donor' control regions to choose from. First, because we evaluate reforms at the county level, we can build synthetic controls out of a pool of more than 3,000 US counties. This allows us to replicate the labor market dynamics of the treated counties very closely. Second, because the treated units are rather small, the assumption of no general equilibrium or spillover effects to neighboring regions seems to be justified. Third, we are able to exploit a long pre-reform time horizon allowing us to built synthetic controls that match the labor market dynamics of the treated county for a long time period, thereby increasing their validity. Fourth, we evaluate the labor market dynamics of six different counties and one state. All these US regions were treated with similar reforms. The treatments were implemented subsequently over a decade and the counties are very heterogeneous in term of size and labor markets. As such, this setting allows us to base our findings on a broad region of common support with a high degree of external validity for other US counties. Finally, the necessary identification assumption of no post-reform unobserved labor market shocks has a high degree of credibility when assessing seven different treatment regions.

Our findings do not provide much evidence that either wages or employment significantly and systematically increased or decreased post-reform. The point estimates for private sector employment as a share of the total county population have ambiguous signs and sizes that vary between -1.7 and +2.8%. With the exception of NYC where the effect is positive and marginally significant, these point estimates are never statistically significant. Joint tests for all six treated counties let us exclude with 92% statistical probability that a potential decrease in employment was larger than 1%. The wage estimates are very similar, but the RMSPE fit is slightly worse as compared to employment. Still, joint tests let us exclude with 90% statistical probability that wage decreases were larger than 1%. When assessing counties separately, for most treatment counties, we can exclude with more than

[1]Other papers that apply this relatively new method are Billmeier and Nannicini (2013); Bohn et al. (2014); Bauhoff (2014); Bassok et al. (2014); Karlsson and Pichler (2015); Restrepo and Rieger (2016).

90% probability that potential wage decreases were larger than 3%. However, the signs of the point estimates are not consistent and even positive in some cases.

In the case of Connecticut—the only state that can be evaluated to date—we do find some evidence that employment dynamics for affected industries (private service sector firms with more than 49 employees) have developed slightly weaker after the sick pay mandate became effective. However, the negative 2.5% point estimate is only statistically significant at the 12% level. Moreover, employment in smaller firms seems to have slightly increased as a response to the reform. We interpret this as suggestive evidence that some large companies reduce employment or split up to avoid the regulation (or that employment growth is less dynamic) when sick pay mandates are not comprehensive and allow for exceptions in the same industry and state. Note that we are able to compare labor market dynamics for non-small private sector firms in Connecticut with a synthetic control group of non-small private sector employment using all other US states as potential donors.

The next section summarizes the existing literature. Section 3 discusses the US sick pay mandates in more detail, and Section 4 provides details on the data. The empirical approach and identification assumptions are discussed in Section 5. Section 6 discusses the empirical findings, and Section 7 concludes.

2 Existing Research on Sick Leave

Existing economic research on sick leave almost exclusively focuses on countries other than the US. The reason is simply a lack of local and state government sick leave mandates in the US as well as a lack of appropriate data. Whereas high-quality administrative sick leave data exist in most Scandinavian countries (Andrèn, 2007; Markussen et al., 2011; Dale-Olsen, 2014), actual sick leave behavior in the US is largely unobservable. There are a few exceptions. One exception is Gilleskie (1998) who exploits 1987 MEPS data both on work absence behavior and demand for medical care to structurally model work absence behavior and simulate the effects of alternative policies. According to Gilleskie (1998), about a quarter of all males employees would not take sick leave when ill. Susser and Ziebarth (2016) use the representative 2011 ATUS Leave Supplement to estimate that, in a given week of the year, two percent of US employees—mostly low-income female employees—would go to work sick. In almost half of all cases, the reasons indicated for such presenteeism behavior were directly related to a lack of paid sick leave coverage. Ahn and Yelowitz (2016) confirm that US employees take more sick leave when they have paid sick leave coverage. Colla et al. (2014) find that

73% of all firms in San Francisco offered paid sick leave before the sick pay mandate in 2006, and that this share increased to 91% in 2009. Pichler and Ziebarth (2015) exploit the same variation as this paper to show that influenza-like infection rates decreased by about 5% in US cities that mandated sick pay. Some reports suggest that the early mandates in San Francisco and DC did not have negative employment effects (Boots et al., 2009; Appelbaum and Milkman, 2011; van Kammen, 2013). Using 2009 to 2012 data from the American Community Survey (ACS) Ahn and Yelowitz (2015) come to a similar conclusion for Connecticut.

Outside the US, several empirical papers estimate the causal effects of variation in sick pay, and find that employees adjust their intensive labor supply in response to such cuts (Johansson and Palme, 1996, 2005; Ziebarth and Karlsson, 2010, 2014; De Paola et al., 2014; Dale-Olsen, 2014; Fevang et al., 2014). The focus of these papers naturally differs from others that study extensive labor supply effects of disability insurance (Autor and Duggan, 2006; Burkhauser and Daly, 2012; Kostol and Mogstad, 2014; Borghans et al., 2014; Burkhauser et al., 2015). Yet, it is closer in nature to US studies on work-related accidents and diseases covered by Workers' Compensation (Meyer et al., 1995; Campolieti and Hyatt, 2006; McInerney and Bronchetti, 2012; Hansen, 2016).

Other papers on sickness absence investigate general determinants (Barmby et al., 1994; Markussen et al., 2011; Dale-Olsen, 2014), probation periods, known to reduce absenteeism (Riphahn, 2004; Ichino and Riphahn, 2005), culture (Ichino and Maggi, 2000), gender (Ichino and Moretti, 2009; Gilleskie, 2010; Herrmann and Rockoff, 2012), income taxes (Dale-Olsen, 2013), union membership (Goerke and Pannenberg, 2015), and unemployment (Askildsen et al., 2005; Nordberg and Røed, 2009; Pichler, 2015). There is also research on the impact of sickness absence on earnings (Sandy and Elliott, 2005; Markussen, 2012). Finally, a few papers study the phenomenon of presenteeism explicitly (Aronsson et al., 2000; Chatterji and Tilley, 2002; Brown and Sessions, 2004; Pauly et al., 2008; Barmby and Larguem, 2009; Markussen et al., 2012; Pichler, 2015; Pichler and Ziebarth, 2015).[2] For example, Pauly et al. (2008) ask 800 US managers about their views on employee presenteeism with *(a)* chronic and *(b)* acute diseases.

Finally, note that paid sick leave differs from paid vacation or paid maternity leave in both scope and aim (Gruber, 1994; Ruhm, 1998; Waldfogel, 1998; Rossin-Slater et al., 2013; Lalive et al., 2014; Thomas, 2015; Baum and Ruhm, 2016; Dahl et al., 2016). Whereas paid sick leave is an insurance against wage losses due to health shocks, paid vacation and maternity leave mostly aim to help balance family obligations with work and address gender inequality in the workplace. Sick leave

[2]Outside of economics, the literature on 'presenteeism' is richer (Dew et al., 2005; Schultz and Edington, 2007; Hansen and Andersen, 2008; Johns, 2010; Böckerman and Laukkanen, 2010; Peipins et al., 2012)

mandates, on the other hand, can also be justified from a public health perspective—because access to paid sick leave reduces contagious presenteeism and the negative externalities associated with the spread of contagious diseases (Pichler and Ziebarth, 2015).

3 US Sick Pay Mandates

The US is the only industrialized country without universal access to paid sick leave. About half of the workforce lacks access to paid sick leave, particularly low-income employees in the service sector (Heymann et al., 2009; Susser and Ziebarth, 2016).

Table B1 in the Appendix provides a summary of recent US sick pay reforms at the city and state level. The details of the bills differ from city to city and state to state, but basically all sick pay schemes represent employer mandates. Several mandates exclude small firms or offer exemptions. Employees "earn" paid sick leave credit (typically one hour per 30-40 hours worked) up to one week per year and, if unused, the credit rolls over to the next calendar year. Because employees need to accrue sick pay credit, most sick pay schemes explicitly state a 90 day accrual period. However, the right to take *unpaid* sick leave is part of most sick pay schemes.

[Insert Figure 1 about here]

As Table B1 shows, San Francisco was the first city to mandate paid sick leave effective February 5, 2007 (Figure1c). Washington DC followed on November 13, 2008 and expanded the mandate on Feb 22, 2014 to include temporary workers and tipped employees. Seattle (September 1, 2012), Portland (Jan 1, 2014), New York City (April 1, 2014), and Philadelphia (May 13, 2015) followed (see Figure 1).

Connecticut was the first state to mandate paid sick leave on January 1, 2012 (Figure 1b). However, the law only applies to service sector employees in non-small businesses and covers only about 20% of the workforce. The recently introduced schemes in California (July 1, 2015), Massachusetts (July 1, 2015), and Oregon (Jan 1, 2016) are much more comprehensive (see Table B1).

4 Quarterly Census of Employment and Wages (QCEW): 2001-2015

The paper makes use of publicly available data from the Quarterly Census of Employment and Wages (QCEW) provided by the Bureau of Labor Statistics (BLS) (2016). The QCEW comes from an establishment census. All establishments covered by US Unemployment Insurance (UI)—97% of all US civilian employment—are included.[3] Using the quarterly UI contribution reports filed by the establishments, the BLS calculates the number of actually filled jobs per month as well as the average weekly wage per quarter.

The BLS reports the data in different levels of spatial and timely disaggregation. To evaluate reforms at the *(a)* county, and *(b)* state level (see Table B1) we generate two different datasets, one at the *(a)* county level and one at the *(b)* state level. The raw data are reported by industry. Because the sick pay mandates mostly apply to private sector jobs, we generate variables that measure private sector employment and private sector wages. The county level data are provided for the time period from January 2001 to July 2015, and the state level data are provided for the time period from January 2001 to April 2015.

County Level Data. Table 1 provides the summary statistic for the *(a)* county level data and all variables generated and employed in the analysis. The table shows summary statistics for the universe of 3,069 counties between 2001 and 2015. In total, as of July 2013, the United States counted 3,142 counties or county-equivalents. The 73 missing counties in our data are counties without any official establishment location, e.g., in very rural counties in Alaska (United States Census Bureau, 2016a). As for employment, we have monthly data points for each county over a total of 174 months, yielding 534,006 county-month observations. As for the quarterly wage data, we have 178,002 county-quarter observations. Population counts vary at the annual level and yield 42,966 total county-year observations.

Employment and Wage Measures. We generate two main outcome variables of interest for the county level analysis. The first is *Private Sector Employment*. We use the total number of filled jobs at the monthly county level—and as reported by the QCEW—and divide by the county level population as reported by the United States Census Bureau (2016b). Hence, we obtain county-specific *Private Sector Employment* for each US county on a monthly basis from 2001 to 2015. Table 1 shows that the average private sector employment share is 27.1%, and the average public sector employment share is 7.7%. This means that, on average, for every 100 residents in a county in the US, 27 private sector jobs paying UI contributions are officially reported.

Note that individuals who hold multiple jobs are counted for every job they hold. In addition, filled jobs are assigned to counties by the physical address of the establishment, not by the county of residence of the jobholder. These are the two reasons (in addition to economic prosperity), why some counties have significantly higher employment ratios than others, and even employment ratios

[3] Not included are self-employed, army members, railroad employees, most elected officials, and most farm workers.

above 100%. Whereas the minimum value for the private sector employment share is a mere 1.1%, the county with the highest employment share reaches a value of 402% (Table 1).

The second variable of interest is *Weekly Wages*. Specifically, employers paying UI contributions report total quarterly gross compensation, including bonus payments and stock options. Wages are then calculated by dividing the total quarterly compensation by the total quarterly employment. Dividing by the number of weeks in a quarter yields the weekly wages displayed in Table 1. Because wages are only reported on a quarterly basis, the number of unique observations decreases to 3,069(counties)$\times$58(quarters)=178,002. The average weekly wage is \$603 (or \$31.4K per year), but the variation ranges from \$158 to \$5,124. Because quarterly Consumer Price Indices are not available at a lower regional level, we use nominal wages as reported by the QCEW.

[Insert Table 1 about here]

Finally, Table 1 shows that the average county population is 98K. However, the standard deviation is large and 338K. Los Angeles County is the largest county with 10.1M population.

State Level Data. Table 2 provides the summary statistic for the *(b)* state level data. When considering all 51 states, we obtain 51(states)$\times$171(months)=8,721 state-month observations for employment and 51(states)$\times$57(quarters)=2,907 state-quarter observations for wages. To date, using the state level data, this paper only evaluates the sick pay mandate in Connecticut. The Connecticut mandate only applies to firms with more than 49 employees in the service sector (Table B1). Because the QCEW data are broken down by industry, we can carve out employment and wage dynamics for the service sector in Connecticut. In addition, the QCEW state level data are also provided by establishment size which helps us to define the Connecticut treatment group in a very precise manner.

To be specific, we aggregate employment and wages for the four establishment categories *<5 employees, 5-9 employees, 10-19 employees, 20-49 employees* and five establishment categories larger than 49 employees. Accordingly we generate variables for the two firm sizes *<50 employees* and *>49 employees*. However, the data by industry and establishment size are only reported for the first quarter of each year. We use the monthly employment and quarterly wage data at the state level to impute values for the rest of the year. To do this, we need to make one reasonable assumption: That the ratios of *<50 employees* vs. *>49 employees* as given in the first quarter of each calendar year remain stable for the rest of the year.[4]

[4]For two firm size categories in Delaware, namely "fewer than 5 employees per establishment" and "500 to 999 employees per establishment" we have missing data in 2014. We impute the missing values by taking the shares of large and small firms in 2013 along with the monthly employment and quarterly wage data.

Employment and Wage Measures. Analogous to the county level data and treatments, we intend to assess the employment and wage dynamics in Connecticut after the sick pay mandate was implemented. Because the mandate only applies to private service sector firms with more than 49 employees, the first measure in Table 2 is *Private Service Sector Employment, >49 employees*. Across all US states and between 2001 and 2015, for every 100 residents of a state, 15 people worked in the service sector and in establishments with more than 49 employees. The same share worked in smaller establishments in the service sector. In contrast, non-service sector (i.e. production) employment amounted only to 4.5 jobs per 100 residents in firms with more than 49 employees. Total public sector employment was 8%. Finally, average nominal weekly private sector wages were $821 (Table 2).

[Insert Table 2 about here]

Treatment Regions. Table B1 in the Appendix provides the list of cities and states that implemented sick pay schemes between 2006 and 2016. Regions in black are treatment regions that we evaluate in this paper, whereas regions in gray cannot be evaluated for the following reason: The last five regions (Philadelphia and Oakland as well as California, Massachusetts and Oregon) only implemented their sick pay mandates very recently and we do not have enough data points to evaluate their labor market outcomes (yet). Figure 1 also depicts the regions whose labor market dynamics are evaluated by this paper.

Second, we only formally evaluate the effects in Washington DC as an illustrative example of a case where the Synthetic Control Group Method is less appropriate due to a poor fit—at least for employment. This poor fit is due to many reasons, including *(a)* DC has a very unique employment structure with many non-profit, public sector, and lobbying jobs. Thus, finding appropriate control counties for DC is very challenging. *(b)* DC's original mandate had many exemptions that are difficult to model with our data (e.g. no health care or restaurant workers). Moreover, DC extended the mandate in September 2014, but retrospectively effective February 2014. *(c)* The first DC mandate became effective shortly after the Great Recession hit in October 2008 which makes it very difficult to disentangle labor market effects due to the mandate from the confounding effect of the recession. Because of *(a)*, the recession also affected DC differently than most other US counties.

The second column in Table B1 indicates the US counties that we formally evaluate. The case for San Francisco (SF) is clear given that the city boundaries equal the county boundaries. However, in the case of Seattle, Portland, Newark, Jersey City, and New York City the county boundaries are not identical with the city boundaries where their mandate formally applied. Portland almost entirely lies within Multnomah County, but small portions fall into Clackamas and Washington County which also include large(r) parts that do not belong to Portland. We focus on Multnomah County when evaluating the Portland mandate. As for Seattle, Newark and Jersey City: they all lie *within* the county that we formally evaluate. For example, in 2014, King County had 2,079,967 residents but Seattle only 668,342 (United States Census Bureau, 2016c). Essex County had 795,723 residents but Newark only 280,579. And Hudson County had 669,115 residents in 2014, but Jersey City only 262,146 (United States Census Bureau, 2016c). The fact that these three cities only formally count a third of the total county population simply means that we evaluate the entire county, not just the core city as in case of SF, NYC, or Portland. Comparing the results of these two groups of geographic units allows us to indirectly test whether firms re-located just outside the city boundaries to circumvent the mandate. This hypothesis would be reinforced, for example, if we found negative employment effects for the core cities but no impact when assessing employment in the entire county that surrounds the city.

Lastly, we do not separately evaluate the five counties of New York City (NYC) but aggregate them to one regional NYC unit for two reasons: The five regions together represent the entire area where the law formally applied. Employment ratios and wages in Manhattan are extremely high and they are relatively low in the other NYC counties. The reason for this disparity is that most people who work in NYC live in one of the four surrounding counties and commute to Manhattan. These unique features of the NYC labor market would make it difficult to find other US control counties of similar population, employment, and wage structure. NYC can be seen as one integrated labor market and not four separate ones. For these reasons, we treat NYC as one statistical unit.

Control Regions. As we describe in more detail in the next section, we employ the Synthetic Control Group Method (SCGM) to model an ideal hypothetical control county for each treatment county as listed in Table B1. For example, as for the county level evaluation, Table 1 lists the variables *county population*, *public sector employment*, *non-service* and *service sector employment*, and *private sector wages* which we use to find suitable control "donor" counties. In other words, in addition to having identical pre-reform outcome dynamics, we seek control counties with similar employment and population structures as the treatment counties. Tables A1 and A2 list all counties that are eventually used for each treatment county to model the pre-treatment employment and wage dynamics as closely as possible.

Sample Selection. In total, the county level dataset contains information on 3,069 individual counties. However, in order to proceed with the automated SCGM as described in the next section, we pre-select the total pool of counties. The main reason for this pre-selection is that running the SCGM with 3,069 donor counties would not technically be feasible due to multiple equilibria and too many degrees of freedom. We pre-select potential donor counties based on similarities in the dimensions: *county population*, *private sector employment* and *private sector wages*. To be specific, we *(a)* separately rank all 3,069 available counties along all three dimensions. Then, we *(b)* select all counties ranked above and below the treated county using a bandwidth of 500 ranks for the first dimension *county population*. Next, we *(c)* proceed with the same procedure on dimension two and three. Finally, we *(d)* use the counties that overlap on all three dimensions and fall within a ranking bandwidth of +/- 500 ranks on each dimension. This pre-selection procedure results in about 200 potential control counties for each treatment county (see the denominator in column (5) of Table 3 for the exact number), which are similar in terms of population and labor market structure.

In addition to this sample pre-selection for the county level dataset, to harmonize the analysis, we additionally restrict both datasets in Tables 1 and 2 as follows: *(i)* For each treatment region, we focus on 4 pre-treatment years (48 months or 16 quarters). *(ii)* We evaluate up to 3 post-treatment years (36 months or 12 quarters) but, depending on when the mandate was enacted, the length of the post-reform periods differ by treatment region. For example, in case of San Francisco and Connecticut we can evaluate the maximum post-reform period of three years. The other post-reform periods are about two years for King County and about one year for the mandates that became effective in 2014 (NYC, Portland, Newark, New Jersey City).

5 Empirical Approach: The Method of Synthetic Control Groups

To evaluate employment and wage dynamics after the implementation of sick pay mandates, we make use of Abadie and Gardeazabal (2003)'s Synthetic Control Group Method (SCGM). The basic idea is to use fractions of several control units to build an ideal—synthetic—control group whose pre-reform outcome dynamics are very similar to those of the treatment group (Abadie et al., 2010). The treated-control difference in post-reform outcome dynamics is then taken to assess the causal effect of the reform.

In our context, following Table B1, the treatment units are counties or states that implemented sick leave mandates, and all potential control units consist of the remaining US counties or states. We analyze the effects for each treatment unit separately. Thus, the notation below refers to one single treatment unit and J control units.

y_{it}^0 denotes the natural logarithm of the outcome ($y_{it}^0 = ln(Y_{it}^0)$ that would have been observed in country i at time t in the absence of the sick pay mandate. Moreover, y_{it}^1 denotes the natural logarithm of the outcome for the treated county i at time t, where sick pay mandates were implemented at time $T_0 + 1$. We assume $y_{it}^1 = y_{it}^0 \forall t = 1, \ldots, T_0, \forall i = 1, \ldots, J + 1$.

Abadie et al. (2010) suggest that the counterfactual y_{it}^0 can then be represented by a factor model:

$$y_{it}^0 = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \epsilon_{it} \tag{1}$$

where δ_t is a common time effect, θ_t is a vector of possibly time-dependent coefficients, λ_t is a vector of unobserved common factors, and μ_i is a vector of unknown factor loadings.

The SCGM allows for some degree of endogeneity in the treatment indicator—the treatment may be correlated with unobservables. First, applied to our case, one necessary assumption is that employment rates and wages of the control counties are not affected by the treatment, i.e., sick pay mandates. This implies the absence of spatial employment spillovers. Note that, in our case, the treated units are rather small and thus very unlikely to trigger large labor market spillover effects. More specifically, Tables A1 and A2 list all control counties used to build the synthetic control county for each treatment county. For example, the synthetic control counties to evaluate King County (WA) are the following, making the 'no spatial labor market spillover' assumption very reasonable: Fulton (GE), Denver (CO), San Mateo (CA), Santa Klara (CA), Durham (NC), Richmond City (VA), and Midland (TX).

Second, as in traditional parametric difference-in-differences (DiD) models, one also assumes away shocks affecting the outcome differently for treatment and control groups in post-reform years. In our case, shocks violating this assumption would be employment policies that are correlated with sick pay mandates. However, one could argue that the SCGM controls for such shocks (better than traditional methods) because the control units—the synthetic control groups—are by construction generated to produce identical outcomes to the treated unit, including unexpected exogenous shocks.

Third, and again similar to traditional DiD methods, treatment-induced geographic migration would lead to biases. When employment prospects worsen due to sick pay mandates and employees lose their jobs (or experience stagnant wages), they might migrate to more prosperous counties. Also, firms could leave the cities or states in response to the laws. For several reasons, economic migration is not a severe issue in our context. First, our data and outcome measures allow to directly test for such migration pattern. In fact it is precisely one objective of this paper to test for changes in *normalized* employment. Recall that we use official population and normalized employment data. In addition, we stratify the effects by the time since implementation and would thus identify negative wage or employment effects in the short-run. As above, it is extremely unlikely that the few control counties—chosen out of a total of 3,069 US counties—are contaminated by worker or firm migration from the treatment counties.

Lastly, in most SCGM settings, only one single treatment unit is evaluated. In our setting, we rely on seven different treatment units, counties of different size as well as one state. There is some probability that single unobserved shocks to single treatment units may confound the evaluation of one county. But, the probability that all seven treatment units will be all coincidentally affected by random unobserved labor market shocks is practically zero.

5.1 Implementation

SCGM requires the estimation of two matrices: V is the weighting matrix determining the relative predictive power of Z_i and of y^0_{it}. The vector W is a vector of non-negative weights given to the J control countries. The criterion to be minimized is:

$$\|\bar{X}_1 - \bar{X}_0 W\|_V = \sqrt{(\bar{X}_1 - \bar{X}_0 W)' \, V \, (\bar{X}_1 - \bar{X}_0 W)}, \tag{2}$$

where $\bar{X}_j$ is a vector of averages over the pre-treatment elements of Z_i and y_i, for both treated and control units. In our case, $\bar{X}_j$ includes the variables listed in Tables 1 and 2 (except for population). This means, for the county level analysis, $\bar{X}_j$ includes *private sector employment*, *private sector wages*, *public sector employment*, *service* and *non-service sector employment*. For the state level analysis, $\bar{X}_j$ includes *service sector employment* in large and small firms, *non-service sector employment* and *service sector wages* in large firms, as well as *public sector employment*.

As such, depending on the variable weight ($W^*(V)$), we obtain an optimal weight matrix among all diagonal positive definite matrices. The elements of V are chosen to minimize the distance to the outcome variable. In other words, an optimal weight matrix minimizes the root of the mean squared prediction error (RMSPE) for all pre-reform periods:

$$RMSPE = \sqrt{\frac{\sum_t (y^1_t - y^0_t W^*(V))^2}{T_0}}, \tag{3}$$

where T_0 represents the number of pre-reform time periods, i.e., in our case 48 months or 16 quarters.

5.2 Treatment Effects and Inference

In addition to calculating the RMSPE for the pre-treatment period, we also calculate the RMSPE for the post-reform period as well as the ratio of the two, as suggested by Abadie et al. (2010). Whereas the RMSPE for the pre-reform years can be used as an indicator to assess the fit of the synthetic control group, the ratio between post and pre RMSPE indicates the size of a possible treatment effect. Assuming that model fit is stable across pre and post-reform years, a ratio larger than 1 indicates that the average differences between treated and synthetic control group is larger (in absolute terms) post as compared to pre-reform, indicating a potential treatment effect.

One disadvantage of this *RMSPE Ratio* (=RMSPE post/RMSPE pre) is, however, that it only yields a relative measure of the treatment effect. Moreover, the sign of the treatment effect remains ambiguous. Hence, we calculate the *Percent Treatment Effect (PTE)* as

$$PTE = \frac{\sum_{T_0+1}^{T} (y_t^1 - y_t^0 W^*(V))}{T - T_0}, \tag{4}$$

and the *Level Treatment Effect (LTE)* as

$$LTE = \frac{\sum_{T_0+1}^{T} (Y_t^1 - Y_t^0 W^*(V))}{T - T_0}. \tag{5}$$

Note that, theoretically, the sign of the treatment effect could change over time. Then positive and negative effects would cancel each other out and bias the *PTE* and *LTE*. Still, in this case, both indicators would provide evidence on the cumulative sign and size of the long-run effect over all post-reform periods.

In terms of inference, we follow Abadie et al. (2010) and run placebo estimates. Because we assess multiple treatments at different points in time, we first construct placebo estimates for each individual treatment unit. Then we rank the treated and all placebo estimates by their *RMSPE Ratio*. Following Abadie et al. (2010), the rank of the true treatment unit relative to the N placebo estimates then determines the p-value of the H_0 hypothesis of no treatment effect. As for the *RMSPE Ratio*, this means that the *RMSPE Ratio* of the treated unit is smaller or equal to the *RMPSE Ratio* of the placebo counties (H_0 : *RMPSE Ratio*$_{Treat}$ $\leq$ *RMPSE Ratio*$_{Placebo}$). Formally, we calculate the percentile rank $p = \hat{F}(RMSPERatio_e)$, where $\hat{F}$ stands for the empirical cumulative distribution of all *RMSPE*

Ratios, as obtained from the placebo estimates. For instance, if the true treatment county had the highest rank among 99 + 1 (placebo + treatment) counties, the p-value would be 1/100 = 0.01, one would reject the H_0, and the treatment effect would be highly significant. In the results section, we carry out this testing procedure for the *RMPSE Ratio*, the *LTE* and the *PTE*. Finally, we follow Dube and Zipperer (2015) and calculate joint p-values based on the sum of the previously obtained p-values using the Irwin-Hall distribution.

As in the standard parametric case, p-values could be statistically insignificant for two reasons: either there is no effect, or we do not have enough statistical power to identify an effect. To assess the statistical power of our estimates, we test the p-value of alternative hypotheses, thereby analyzing how broad or narrow the confidence intervals are—following Dube and Zipperer (2015): Calculating the *PTE* and *LTE*, we carry out all N placebo estimates as above, but now assume that the true effect was x percent or q changes of natural units. Then we assess the probability with which our treated unit originates from that distribution, thereby calculating p-values. Using the notation above this means that we calculate $p = \hat{F}(PTE - x)$ and $p = \hat{F}(LTE - q)$. To provide additional intuition: In the SCGM setting, placebos are usually run to see how the treated unit differs from the placebos. The placebos are, by definition, non-treated units and should thus have a treatment effect of zero. Using the distribution of placebo treatment effects, one can then derive the likelihood that the treated unit stems from this (non-treated) distribution. Here, we just slightly modify this basic idea and impose an artificial treatment effect of x (percent) or q (changes) on the placebos. Then, as in the standard case, we assess the likelihood that the treated unit comes from this modified distribution of placebos.

6 Results

6.1 Evaluating Labor Market Effects of City Sick Leave Mandates

We begin by evaluating the labor market dynamics of sick leave mandates at the city level. First we evaluate potential changes in employment and then move on to wages.

Employment Dynamics

Figure 2 shows the development of normalized county level employment in five different treatment counties as classified by Table B1. The composition of each synthetic control county—the weights W of the J control counties—-are displayed in Table A1.

Employment dynamics for five treated counties are in the left column of Figure 2. The solid lines represent the treatment counties and the dashed lines represent the synthetic control counties. The solid vertical lines at point zero on the x-axes represent the month when the sick pay mandates became officially law of the city, i.e., became effective and were enforced. The dotted lines to the left indicate the month when the bills were passed; they facilitate an assessment of whether there is evidence of anticipation effects. The dotted lines to the right of the vertical law effectiveness lines indicate when the accrual periods were over. Recall that most bills stipulate a three month accrual period during which sick days could be earned, but not taken. To be specific, during the three month accrual period, *paid* sick leave could not be taken but many employees gained legally guaranteed access to unpaid sick leave (Section 3).

We learn the following from the left column of Figure 2: First, the counties exhibit significantly different employment levels. Whereas San Francisco and King County have employment levels of around or above 50%, the level for Multnomah County is below 50%. The levels for New York City (NYC) and Essex County are even below 40%. The variation in baseline employment across treated counties suggests that the region of support of our results is broad and based on several heterogeneous counties, strengthening the external validity of our findings for other US counties.

Second, for all five treatments (and Hudson County in the Appendix, Figure B1), the employment dynamics of treated and synthetic controls are basically identical in the pre-reform period, suggesting that the synthetic control counties represent a valid counterfactual for the treated counties in the absence of sick pay mandates.

[Insert Figure 2 about here]

Third, one cannot visually identify sizable and systematic reform-related employment effects. In post-reform years, control and treatment county employment for King and Hudson County (Figure B1) follows literally an identical trajectory; post-mandate employment dynamics in San Francisco, NYC, and Multnomah County even seem to be on a slightly more robust trajectory than their synthetic control counterparts, whereas employment appears to be a little bit weaker in Essex County.

Fourth, to quantitatively evaluate the differences in employment dynamics and conduct inference, we follow Abadie and Gardeazabal (2003) and Abadie et al. (2010) and carry out the statistical tests as discussed in Section 5. All statistics for the treated counties and employment are shown in Table 3. The first column shows the level of the outcome measure, Y_{it}^1, the *Employment Ratio*—defined as private sector employment as a share of the total county population—averaged over the entire pre-reform period. As discussed, the region of support is broad and ranges from 30% for Hudson County to 57% for San Francisco (disregarding the 77.5% for DC, see below).

[Insert Table 3 about here]

Column (2) of Table 3 shows the RMSPE for pre-reform years as defined by equation (3). Note that we take the logarithm of the outcome variable before minimizing. Thus the values in column (2) can be interpreted as percentages of the outcome variable. With the exception of DC which we disregard due to a poor fit but show for completeness (see discussion in Section 4), all pre-mandate RMSPE's are very low—around 1% of the outcome measure. This implies that we could very successfully replicate the employment dynamics of the treatment counties by building a synthetic control county. As a comparison, evaluating the effects of a tobacco control program in California on cigarette consumption, Abadie et al. (2010) have a pre-reform RMSPE of 3 relative to a mean of about 100.

Column (3) shows the post-RMSPE for each treated county. Although it is slightly larger than the pre-reform RMSPE, there appears to be no substantial differences between pre and post RMSPEs as illustrated by column (4): The *RMPSE Ratio* [RMPSE-post/RMPSE-pre; column (3)/column(2)] lies between 0.5 for King County and 2.6 for NYC.

Next, we conduct inference using placebo methods as proposed by Abadie et al. (2010) and described in Section 5. Specifically, for each treatment county, we apply the SCGM method to non-treated placebo counties with similar labor market characteristics and population. In particular, in a first step, we rank all 3,069 potential donor counties on three dimensions, and then pre-select a set of potential control counties with a similar population and labor market structure as the treatment county (see Section 4 for a more detailed discussion). Then we replicate the standard SCGM procedure with each placebo county pretending it had been treated at the same time as the treated county.

Column (4) of Table 3 shows how we calculate the p-value for the hypothesis H_0 : *RMSPE Ratio*$_{Treat} \leq$ *RMSPE Ratio*$_{Placebo}$, as [Rank *RMSPE Ratio* Treated County/#Total Counties Assessed]. In other words, after calculating the *RMPSE Ratio* for each placebo evaluation and ranking all of them, we can assess the position of the *RMPSE Ratio* for the treated county in the test statistic distribution (Abadie et al., 2010). As seen in column (4), the total number of counties [placebo + 1] assessed for each treatment varies between 93 and 193, and the rank of the true treatment county falls typically somewhere in between: treatment counties do not systematically rank high in the distribution. Accordingly, six of the seven p-values are not even close to being considered statistically significant by conventional levels and are above 0.3. The only marginally significant p-value is NYC with 0.082. However, the direction of the employment effect is positive.

Finally, we calculate the sum of all p-values in the second last row (abstaining from DC due to a poor fit) and then evaluate the joint p-value—based on the Irwin-Hall distribution—as the sum of six independently drawn uniform random variables as suggested by Dube and Zipperer (2015). This yields an overall p-value of 0.34 for the joint test in column (5).[5] Consequently, overall, we clearly cannot reject the null of no employment effects as a result of US sick pay mandates.

The right column of Figure 2 shows the results of our placebo analyses and permutation inference graphically. Following the convention in the literature, the graphs for the treatment counties plot the difference in the logarithm of employment ratios along with the differences of all placebo evaluations with good fit ($\text{RMSPE}_{Placebo} \leq \text{RMSPE}_{Treat} \cdot 2$). The latter lines are in gray whereas the lines for the real treatment counties are solid black. The visual assessment follows the quantitative analysis in Table 3. For pre-reform periods, the solid black line fluctuates very closely around the horizontal zero line implying that the synthetic control counties very closely map the employment dynamics of the treatment counties. After the reform implementation, as indicated by the black solid vertical line, employment differentials between treated and control counties remain very small and straight flat for King County, Multnomah County, and Essex County. In line with the estimates in Table 3, for NYC and San Francisco, the differential even appears to be positive although this is not true in a statistical sense.

Column (5) of Table 3 shows the *Percent Treatment Effect (PTE)* for the post-reform period, whereas column (6) shows the *Level Treatment Effect (LTE)* in natural units, i.e., private sector employment as a share of the total population. As seen, the sign of the calculated treatment effect varies and is only negative for three out of the six treatment counties (abstaining from DC due to a poor fit). The size of the post-reform employment differences between treated and synthetic control counties varies between -1.7% (0.5ppt) for Hudson County and +2.8% (1.2ppt) for NYC.

Column (7) of Table 3 carries out the placebo testing procedure for the PTE, by ranking all counties by their PTE, starting with the most negative PTE. Thus here we test the hypothesis H_0 : $\text{PTE}_{Treat} \geq \text{PTE}_{Placebo}$. Note that the number of placebos in the denominator slightly changes (as compared to column (5)); column (7) only uses placebo counties with a good pre-treatment fit with $\text{RMSPE}_{Placebo} \leq \text{RMSPE}_{Treat} \cdot 2$.[6] The resulting p-values in column (7) reinforce what we found in

[5] Including DC would lead to a p-value of 0.29.

[6] The less restrictive $\text{RMSPE}_{Placebo} \leq \text{RMSPE}_{Treat} \cdot 5$ only slightly changes the p-values.

column (4) for the *RMSPE Ratio*: most p-values are not even close to conventional statistical levels and several of the treatment effects are rather positioned at the lower end of the treatment effect distribution. Also, the only two effects that could be assessed as marginally significant (0.12 for Essex and 0.1 for Hudson County) differ from the marginally significant RMSPE effect which was identified for NYC. In any case, the p-value of the joint test is 0.56 and thus far from exhibiting statistical significance.

[Insert Figure 3 about here]

Finally, columns (9) and (10) of Table 3 present p-values for alternative hypotheses, namely a treatment effect of -3% (column 9) and -2ppt (column 10). The null hypotheses are H_0 : $\text{PTE}_{Treat} \geq -0.03$ and H_0 : $\text{LTE}_{Treat} \geq -0.02$, respectively. Figure 3 illustrates how we calculated the p-values for these hypotheses. We first added this hypothetical treatment effect to each of the placebos, and then recalculated the p-value.[7] Figure 3 shows the resulting graph for Multnomah. After adding a pseudo treatment effect of -3% to the placebos, the placebos are centered around -3% instead of zero as in Figure 2. Multnomah, our exemplary treated county, is represented by the black line. It appears to be a positive outlier when assessing the distribution of treatment effects. In fact, Multnomah has the highest PTE and thus we can reject the hypothesis H_0 : $\text{PTE}_{Treat} \geq -0.03$ with a significance level of 1/181=0.0055 (column (9) in Table 3, only p-value shown).

As seen in columns (9) and (10), most p-values for single counties are very small implying that we are able to reject a true decrease in employment of 3% with probabilities of 99% for Multnomah County and NYC. Such a decrease in employment could be identified with 95% probability for San Francisco and 88% for King County. Only the two counties with slightly negative PTE and LTE point estimates and marginally significant p-values, Essex and Hudson County, have slightly higher p-values, but we would still be able to identify true negative employment effects with 86 and 82% probability. As column (10) shows, an employment decrease of 2ppt could be identified with probabilities between 94% and 99%. Finally, the p-values for the joint tests converge strictly toward zero. When we carry out the joint test for a hypothetical employment decrease of just 1%, we still would be able to identify this 1% decrease with 92% probability (number not shown in Table 3).

[7] Dube and Zipperer (2015) suggests adding the negative treatment effect to the treated county, which gives exactly the same result.

Wage Dynamics

Next we analyze the wage effects. As discussed in Section 4, we use quarterly nominal county level wages as provided by the Quarterly Census of Employment and Wages (QCEW) (Bureau of Labor Statistics (BLS), 2016). The left column of Figure 4 shows the wage dynamics for five of the treatment regions. We make the following observations: First, there exists a clear cyclical pattern in nominal wage payments. This is not surprising, given that the wage measure includes bonus payments and fringe benefits, which are not equally spread out over all four quarters of the year. Second, we observe a positive trend in the wage dynamics representing natural nominal wage increases over time. However, interestingly, not only do the wage levels differ substantially between local labor markets, so does the slope representing wage growth. This is one of the reasons why we decided against further manipulation of the raw data, e.g., de-trending or correcting for the consumer price index. In fact, there are two reasons why we abstain from such manipulation. First, the SCGM is able to precisely replicate local and time-variant differences in wage dynamics. Actually it is a method that is very well suited for such purposes. Second, because no monthly (or quarterly) county level CPI measure is available, one would have to convert nominal wages into presumable 'real' wages using a common rigid discount rate which may not capture the properties of the local labor market appropriately.

[Insert Figure 4 about here]

Third, in terms of baseline levels, weekly pre-reform wages in the private sector vary between \$916 in Multnomah and \$1617 in New York City (column (1) of Table 4). In terms of post-treatment wage differences, all values in columns (6) and (7) are close to zero with the exception of Hudson County, where we observe a 12% wage differential between Hudson County and the synthetic control county. However, column (2) reveals that the difference between Hudson County and its synthetic control county was already quite large before the mandate (RMSPE=0.13), i.e., the SCGM fit is poor. The reason is that the wage in Hudson County is much higher than the wages in donor counties with a comparable employment ratio and population (see Table A2 for the list of donor counties). Therefore the SCGM minimization procedure is unable to find a suitable synthetic control group. A similar argument applies to King County although the fit is slightly better here (RMSPE=0.0772).

Fourth, with the exception of King County and Hudson County (in Figure B2), all pre-mandate treatment group wage changes could be relatively accurately replicated with SCGM. Column (2) of Table 4 shows that the largest pre-reform RMSPE is 0.13 for Hudson County—relative to a pre-reform weekly wage level of $1343 (Column (1)). The other relative RMSPEs are substantially smaller, e.g. the one for San Francisco is 0.023 relative to $1263.

[Insert Table 4 about here]

Fifth, there is not much evidence for significant wage decreases (or increases) as a result of mandating sick leave. Visually, it is hard to detect substantial and systematic effects (Figure B2), and the RMPSE Ratios in column (4) of Table 4 are mostly below 1 (except for King County: 1.27). When running placebo models using non-treated counties with a similar labor market structure and population, one obtains the RMSPE p-values in column (5). None of these p-values are even close to being statistically significant, neither separately nor in a joint test. The same holds for the PTE p-values in column (8), which are all above 0.4, except for King County which is significant at the 3% level with a treatment effect of -$110. However, as discussed, the quality of the SCGM fit as assessed by the pre-reform RMSPE (column (2)) is poor for King County, which is why we have to be cautious with the interpretation of this single significant effect. In any case, the joint p-value for all treatment units is large and 0.63.

The graphical representation of the estimated treated-control wage differential for real treatment groups and their placebos is in the right column of Figure 4. Again, the placebo estimates are in gray whereas the black solid line represents the wage differential of the real treatment county. As seen, in line with the complementary graphs on the left and the statistics in Table 4, there is not much evidence that wages decreased significant in the aftermath of the sick pay mandate.

The power tests in columns (9) and (10) of Table 4 follow the structure of Table 3. With the exception of King County, which ranks #4 out of 150+1 SCGM runs, for 4 out of the remaining 5 counties with good pre-RMSPE fit, we would be able to identify wage decreases of 3% with more than 90% statistical probability. The joint test p-value is 0.003 implying that we could identify joint wage decreases of 3% with 99.7% probability. Using joint tests, the SCGM method would identify wage decreases of more than 1% with 90% statistical probability (number not shown in Table 4).

Recall that we derived a static wage decrease of about 3% as an upper bound in the Introduction. This static calculation assumes that *(i)* all employees fully exhaust their paid sick leave credit, *(ii)* all firms were affected by the law and went from zero paid sick leave to fully complying with the law. Further, this static view assumes *(iii)* the absence of any decrease in work productivity when working sick and zero spillover effects from sick workers, and *(iv)* that 100% of work productivity is lost during times of sick leave (no compensation for lost labor after returning to work). Hence, the absence of negative wage effects may not be surprising.

6.2 Evaluating Labor Market Effects of the Sick Pay Mandate in Connecticut

Lastly, we evaluate the labor market effects in Connecticut following the implementation of a sick pay mandate for full-time service sector workers in non-small firms. As discussed in Section 3, the law only applied to about 20% of the workforce and was not considered very stringent. This section makes use of QCEW data by firm size and industry (Bureau of Labor Statistics (BLS), 2016), see Section 4. We generate treatment groups that are defined as establishments in the service sector with more than 49 employees. We also assess labor market effects for small firms with less than 50 employees in the service sector because one might expect employment shifts to smaller firms either through conversion of firms or real shifts in employment.

Figure 5 shows the results for employment and Figure 6 shows the results for wages. The SCGM test statistics are in Table 5.

Starting with employment and Figure 5, one visualizes a pretty good synthetic control group fit for pre-reform years. This impression is confirmed by column (2) of Table 5 and the very low RMSPE values of 0.005 and 0.007, respectively. There is some evidence of lower employment dynamics in the treatment group (big service sector firms). Column (6) yields a post-reform employment reduction of around 2.6% and the RMPSE Ratio in column (4) is large, at 5.45. The placebo analysis using the other 51 US states confirms this impression. As column (5) shows, there is only one other placebo state with a larger RMSPE ratio. The p-value is statistically significant at the 4% level. The right column of Figure 5 is in line with the suggestive evidence of employment losses of about 2.6%. It seems like small employment losses kicked in just after the reform became effective in 2012. The weaker job growth has been increasing smoothly but steadily over time.

As seen in the second row of Table 5, where we test for employment effects of small service firms, there also exists suggestive evidence that employment slightly increased for smaller firms. One plausible explanation could be that big service sector firms converted to several smaller ones to circumvent the law. However, the employment increase of 1.3% in smaller firms has only a RMSPE Ratio of 2 (column (4)) and ranks #14 among 51 states when conducting placebo tests. Consequently, the employment increase is not statistically significant at conventional levels.

In contrast, there is not much evidence that wages among the employed decreased post-reform. The synthetic control group match in the upper left corner of Figure 6 seems to be very good, which is confirmed by a pre-RMPSE of just 0.016 (column (2) of Table 5). Connecticut's wage dynamics for big service sector firms ranked 29/51 in the three post-reform years between 2012 and 2015, and the p-value is accordingly a large 0.57. The upper right corner of Figure 6 visualizes the wage differential along with the placebo tests, and it does not yield not much evidence for wage changes. We consider the wage fit for smaller service sector firms in Connecticut not good enough to conduct any statistical inference (lower row of Figure 6, column (2) of Table 6).

7 Discussion and Conclusion

This paper systematically evaluates the labor market consequences of all existing US sick pay mandates—whenever the data are already available and whenever we obtain a reasonably good pre-reform fit for wage and employment dynamics using the Synthetic Control Group Method (SCGM). To date, we are able to carry out sick pay mandate evaluations for San Francisco (2007), Connecticut (2012), Seattle (2012), New York City (2014), Portland (2014), Jersey City (2014) and Newark (2014).

The setting is ideally suited for the SCGM. First, especially when evaluating counties, we have a very rich pool of donor counties—in fact thousands of them—which we can exploit to build convincing synthetic control counties that map the labor market dynamics of the treated counties very closely. We also rely on many pre-treatment observations, and matching treated-control labor market dynamics over a long pre-reform time period strengthens the identifying assumptions. Because our treated units are very small and geographically dispersed, we can also convincingly assume that general equilibrium and spillover effects from treated to control counties do not exist. Additionally, because we rely on six different treatment counties and one treatment state—all of which have very diverse labor markets—our findings have a broad range of common support. This allows to strengthen the external validity for other US counties. Finally, seven treatment regions reduce the likelihood that unobserved shocks confounded post-reform labor market dynamics, especially because the treatments have been spread out over almost one decade. Having half a dozen treatment units also increases the statistical power of our estimates.

We do not find much evidence that employment and wage growth have been substantially dampened by mandating employers to allow employees to earn paid sick leave. This may be a function of how the US laws are designed. In fact, they seem to be more incentive-compatible than their European counterparts and minimize shirking behavior, a main concern of opponents. The reason for this incentive-compatibility is that paid sick days are personalized and employees 'earn' them. For every 30-40 hours worked—i.e., for every week a full-time employee works—employees earn one hour of paid sick leave. This means that employees earn about one day of paid sick leave for every two months worked, up to typically seven days per year. Unused sick days roll over to the next year. Because earned sick days represent a personalized insurance credit—similar to health savings accounts—against future exogenous health shocks that are likely to occur (e.g. flu or disease of child), we expect shirking to play a minimal role for most employees. Static calculations suggest that the US version of earning sick days equals a fringe benefit that is worth up to 3.3% of the wage. The static calculation is an upper bound because it assumes that employees fully exhaust their sick days, could maintain 100% work productivity when working sick, and would not compensate for any of the foregone work productivity during their sick leave. However, it may be possible that employment growth remains weaker subsequent such employer mandates because employers overestimate their real work impact on labor costs, or as a function of other psychologically triggers that may affect the business climate.

In line with this reasoning, our findings suggest that neither employment nor wage growth has been significantly affected by the sick pay mandates. We can exclude with high statistical precision that employment or wages have decreased by more than 1% when assessing all treatment counties jointly. One exception could be Connecticut where the law was the least comprehensive and only applied to 20% of the workforce—that is, full-time service sector employees in firms with at least 50 employees. Here we find some suggestive evidence that, as compared to the same sector in the other US states, employment growth has been lagging behind as a result of the mandate.

The US needs more economic research on sick leave.

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.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the Basque Country. *The American Economic Review 93*(1), 113–132.

Ahn, T. and A. Yelowitz (2015). The short-run impacts of Connecticuts paid sick leave legislation. *Applied Economics Letters 22*(15), 1267–1272.

Ahn, T. and A. Yelowitz (2016). Paid sick leave and absenteeism: The first evidence from the U.S. `https://sites.google.com/site/tomsyahn/`, retrieved March 17, 2016.

Andrèn, D. (2007). Long-term absenteeism due to sickness in Sweden: How long does it take and what happens after? *The European Journal of Health Economics 8*, 41–50.

Appelbaum, E. and R. Milkman (2011). Leaves that pay: Employer and worker experiences with paid family leave in California. report, Center for Economic and Policy Research (CEPR).

Aronsson, G., K. Gustafsson, and M. Dallner (2000). Sick but yet at work: An empirical study of sickness presenteeism. *Journal of Epidemiology & Community Health 54*(7), 502–509.

Askildsen, J. E., E. Bratberg, and Ø. A. Nilsen (2005). Unemployment, labor force composition and sickness absence: A panel study. *Health Economics 14*(11), 1087–1101.

Autor, D. H. and M. G. Duggan (2006). The growth in the Social Security Disability Rolls: A fiscal crisis unfolding. *Journal of Economic Perspectives 20*(3), 71–96.

Barmby, T. and M. Larguem (2009). Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. *Journal of Health Economics 28*(5), 1012–1017.

Barmby, T., J. Sessions, and J. G. Treble (1994). Absenteeism, efficiency wages and shirking. *Scandinavian Journal of Economics 96*(4), 561–566.

Bassok, D., M. Fitzpatrick, and S. Loeb (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. *Journal of Urban Economics 83*(C), 18–33.

Bauhoff, S. (2014). The effect of school district nutrition policies on dietary intake and overweight: A synthetic control approach. *Economics & Human Biology 12*(C), 45–55.

Baum, C. L. and C. J. Ruhm (2016). The effects of paid family leave in California on labor market outcomes. *Journal of Policy Analysis and Management 35*(2), 333–356.

Billmeier, A. and T. Nannicini (2013). Assessing economic liberalization episodes: A synthetic control approach. *The Review of Economics and Statistics 95*(3), 983–1001.

Böckerman, P. and E. Laukkanen (2010). What makes you work while you are sick? Evidence from a survey of workers. *The European Journal of Public Health 20*(1), 43–46.

Bohn, S., M. Lofstrom, and S. Raphael (2014). Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrant population? *The Review of Economics and Statistics 96*(2), 258–269.

Boots, S. W., K. Martinson, and A. Danziger (2009). Employers' perspectives on San Francisco's paid sick leave policy. Technical report, The Urban Institute.

Borghans, L., A. C. Gielen, and E. F. P. Luttmer (2014). Social support substitution and the earnings rebound: Evidence from a regression discontinuity in disability insurance reform. *American Economic Journal: Economic Policy 6*(4), 34–70.

Brown, S. and J. G. Sessions (2004). Absenteeism, presenteeism, and shirking. *Economic Issues 9*(1), 15–23.

Bureau of Labor Statistics (BLS) (2016). *Quarterly Census of Employment and Wages (QCEW)*. `http://www.bls.gov/cew/datatoc.htm#NAICS_BASED`, last accessed on February 28, 2016.

Burkhauser, R. V. and M. C. Daly (2012). Social Security Disability Insurance: Time for fundamental change. *Journal of Policy Analysis and Management 31*(2), 454–461.

Burkhauser, R. V., M. C. Daly, and N. Ziebarth (2015). Protecting working-age people with disabilities: Experiences of four industrialized nations. Working Paper Series 2015-8, Federal Reserve Bank of San Francisco.

Busse, R. and A. Riesberg (2004). *Health care systems in transition: Germany* (1 ed.). WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.

Campolieti, M. and D. Hyatt (2006). Further evidence on the Monday effect in Workers' Compensation. *Industrial and Labor Relations Review 59*(3), 438–450.

Chatterji, M. and C. J. Tilley (2002). Sickness, absenteeism, presenteeism, and sick pay. *Oxford Economic Papers 54*, 669–687.

Colla, C. H., W. H. Dow, A. Dube, and V. Lovell (2014). Early effects of the San Francisco paid sick leave policy. *American Journal of Public Health 104*(12), 2453–2460.

Connecticut Department of Labor (2015). *Connecticut General Statue 31-57r—Paid Sick Leave*. `http://www.ctdol.state.ct.us/wgwkstnd/sickleave.htm`, last accessed on May 28, 2015.

Dahl, G. B., K. V. Løken, M. Mogstad, and K. V. Salvanes (2016). What is the case for paid maternity leave? *Review of Economics and Statistics*. forthcoming.

Dale-Olsen, H. (2013). Absenteeism, efficiency wages, and marginal taxes. *Scandinavian Journal of Economics 115*(4), 1158–1185.

Dale-Olsen, H. (2014). Sickness absence, sick leave pay, and pay schemes. *Labour 28*(1), 40–63.

De Paola, M., V. Scoppa, and V. Pupo (2014). Absenteeism in the Italian public sector: The effects of changes in sick leave policy. *Journal of Labor Economics 32*(2), 337–360.

Dew, K., V. Keefe, and K. Small (2005). Choosing to work when sick: Workplace presenteeism. *Social Science & Medicine 60*(10), 2273–2282.

Dube, A. and B. Zipperer (2015). Pooling multiple case studies using synthetic controls: An application to minimum wage policies. IZA Discussion Papers 8944, Institute for the Study of Labor (IZA).

Fevang, E., S. Markussen, and K. Røed (2014). The sick pay trap. *Journal of Labor Economics 32*(2), 305–336.

Gilleskie, D. (2010). Work absences and doctor visits during an illness episode: The differential role of preferences, production, and policies among men and women. *Journal of Econometrics 156*(1), 148–163.

Gilleskie, D. B. (1998). A dynamic stochastic model of medical care use and work absence. *Econometrica 66*(1), 1–45.

Goerke, L. and M. Pannenberg (2015). Trade union membership and sickness absence: Evidence from a sick pay reform. *Labour Economics 33*(C), 13–25.

Gruber, J. (1994). The incidence of mandated maternity benefits. *American Economic Review 84*(3), 622–641.

Hansen, B. (2016). California's 2004 Workers' Compensation reform: Costs, claims, and contingent workers. *Industrial and Labor Relations Review forthcoming*.

Hansen, C. D. and J. H. Andersen (2008). Going ill to work what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine 67*(6), 956 – 964.

Herrmann, M. A. and J. E. Rockoff (2012). Does menstruation explain gender gaps in work absenteeism? *Journal of Human Resources 47*(2), 493–508.

Heymann, J., H. J. Rho, J. Schmitt, and A. Earle (2009). Contagion nation: A comparison of paid sick day policies in 22 countries. Technical Report 2009-19, Center for Economic and Policy Research (CEPR).

Ichino, A. and G. Maggi (2000). Work environment and individual background: Explaining regional shirking differentials in a large Italian firm. *The Quarterly Journal of Economics 115*(3), 1057–1090.

Ichino, A. and E. Moretti (2009). Biological gender differences, absenteeism, and the earnings gap. *American Economic Journal: Applied Economics 1*(1), 183–218.

Ichino, A. and R. T. Riphahn (2005). The effect of employment protection on worker effort. A comparison of absenteeism during and after probation. *Journal of the European Economic Association 3*(1), 120–143.

Johansson, P. and M. Palme (1996). Do economic incentives affect work absence? Empirical evidence using Swedish micro data. *Journal of Public Economics 59*(1), 195–218.

Johansson, P. and M. Palme (2005). Moral hazard and sickness insurance. *Journal of Public Economics 89*(9-10), 1879–1890.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior 31*(4), 519–542.

Karlsson, M. and S. Pichler (2015). Demographic consequences of HIV. *Journal of Population Economics 28*(4), 1097–1135.

Kostol, A. R. and M. Mogstad (2014). How financial incentives induce disability insurance recipients to return to work. *American Economic Review 104*(2), 624–655.

Lalive, R., A. Schlosser, A. Steinhauer, and J. Zweimüller (2014). Parental leave and mothers' careers: The relative importance of job protection and cash benefits. *The Review of Economic Studies 81*(1), 219–265.

Lovell, V. (2003). No time to be sick: Why everyone suffers when workers don't have paid sick leave. Policy report, Institute for Women's Policy Research.

Markussen, S. (2012). The individual cost of sick leave. *Journal of Population Economics 25*(4), 1287–1306.

Markussen, S., A. Mykletun, and K. Røed (2012). The case for presenteeism: Evidence from Norway's sickness insurance program. *Journal of Public Economics 96*(11), 959–972.

Markussen, S., K. Røed, O. J. Røgeberg, and S. Gaure (2011). The anatomy of absenteeism. *Journal of Health Economics 30*(2), 277–292.

McInerney, M. and E. Bronchetti (2012). Revisiting incentive effects in Workers' Compensation: Do higher benefits really induce more claims? *Industrial and Labor Relations Review 65*(2), 288–315.

Meyer, B. D., W. K. Viscusi, and D. L. Durbin (1995). Workers' compensation and injury duration: Evidence from a natural experiment. *American Economic Review 85*(3), 322–340.

Miller, K. and C. Williams (2015). Valuing good health in Connecticut: The costs and benefits of paid sick days. report, Institute for Women's Policy Research. `http://www.iwpr.org/publications/pubs/valuing-good-health-in-connecticut-the-costs-and-benefits-of-paid-sick-days`, last accessed on May 28, 2015.

Nordberg, M. and K. Røed (2009). Economic incentives, business cycles, and long-term sickness absence. *Industrial Relations 48*(2), 203–230.

Pauly, M. V., S. Nicholson, D. Polsky, M. L. Berger, and C. Sharda (2008). Valuing reductions in on-the-job illness: 'Presenteeism' from managerial and economic perspectives. *Health Economics 17*(4), 469–485.

Peipins, L., A. Soman, Z. Berkowitz, and M. White (2012). The lack of paid sick leave as a barrier to cancer screening and medical care-seeking: Results from the National Health Interview Survey. *BMC Public Health 12*(1), 520.

Pichler, S. (2015). Sickness absence, moral hazard, and the business cycle. *Health Economics 24*(6), 692–710.

Pichler, S. and N. R. Ziebarth (2015). The pros and cons of sick pay schemes: Testing for contagious presenteeism and shirking behavior. Upjohn institute working paper 15-239. http://research.upjohn.org/up_workingpapers/239/, retrieved March 14, 2016.

Restrepo, B. J. and M. Rieger (2016). Trans fat and cardiovascular disease mortality: Evidence from bans in restaurants in New York. *Journal of Health Economics*, –. forthcoming.

Riphahn, R. T. (2004). Employment protection and effort among German employees. *Economics Letters 85*, 353–357.

Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (2013). The effects of California's Paid Family Leave Program on mothers leave? Taking and subsequent labor market outcomes. *Journal of Policy Analysis and Management 32*(2), 224–245.

Ruhm, C. J. (1998). The economic consequences of parental leave mandates: Lessons from Europe. *The Quarterly Journal of Economics 113*(1), 285–317.

Sandy, R. and R. F. Elliott (2005). Long-term illness and wages: The impact of the risk of occupationally related long-term illness on earnings. *Journal of Human Resources 40*(3), 744–768.

Schultz, A. B. and D. W. Edington (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation 17*(3), 547–579.

Susser, P. and N. R. Ziebarth (2016). Profiling the us sick leave landscape: Presenteeim among females. *Health Services Research forthcoming*. www.nicolasziebarth.com, retrieved March 15, 2016.

Thomas, M. (2015). The impact of mandated maternity benefits on the gender differential in promotions: Examining the role of adverse selection. Technical report. mimeo.

United States Census Bureau (2016a). *Population Estimates—County Totals: Vintage 2013*. http://www.census.gov/popest/data/historical/2010s/vintage_2013/county.html, last accessed on February 28, 2016.

United States Census Bureau (2016b). *Population Estimates: 1990s County Tables*. http://www.census.gov/popest/data/historical/1990s/county.html, last accessed on February 28, 2016.

United States Census Bureau (2016c). *QuickFacts*. http://www.census.gov/quickfacts/table/PST045215/00, last accessed on February 28, 2016.

US Congress (2015). *H.R.1286 - Healthy Families Act*. https://www.congress.gov/bill/113th-congress/house-bill/1286, last accessed on May 28, 2015.

van Kammen, B. (2013). Sick leave mandates and employment. mimeo, University of Wisconsin-Milwaukee.

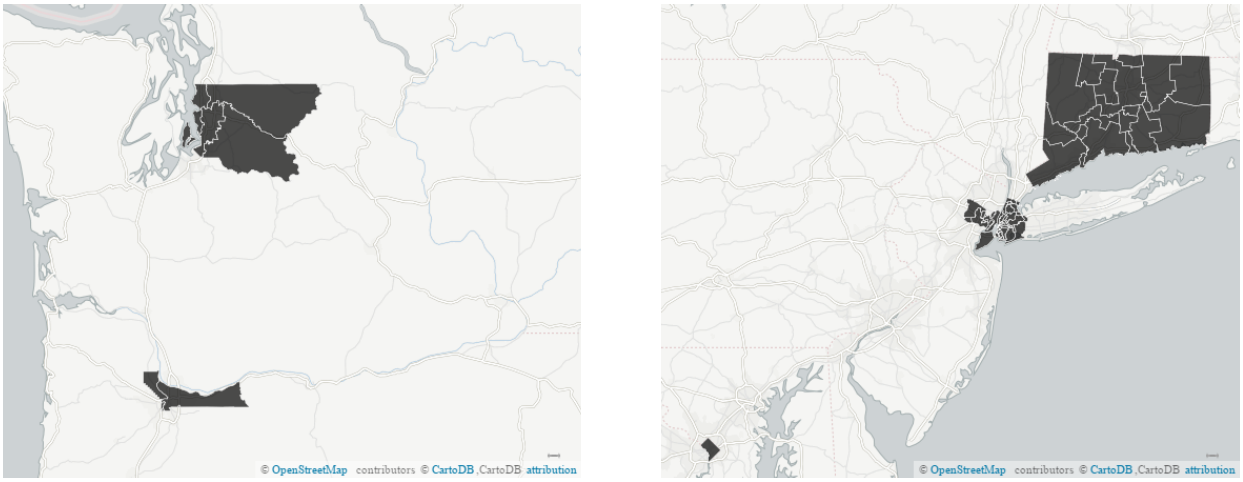
Waldfogel, J. (1998). Understanding the "family gap" in pay for women with children. *Journal of Economic Perspectives 12*(1), 137–156.

Ziebarth, N. R. and M. Karlsson (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics 94*(11-12), 1108–1122.

Ziebarth, N. R. and M. Karlsson (2014). The effects of expanding the generosity of the statutory sickness insurance system. *Journal of Applied Econometrics 29*(2), 208–230.

Figures and Tables

Figure 1: Evaluated Sick Pay Mandates in the (a) Northwest, (b) Northeast, (c) Southwest of the US



The upper left graph (Figure 1a) shows counties that are evaluated in this paper and located in the Northwest of the USA. Specifically, the northern county is King County in Washington, and the southern county is Multnomah County in Oregon. The upper right graph (Figure 1b) shows counties that are evaluated in this paper and located in the Northeast of the USA. Specifically, all nine counties in the state of Connecticut are highlighted (top of graph). The county in the bottom of Figure 1b is the District of Columbia. In the middle of Figure 1b, the four counties belonging to NYC are highlighted as well as Hudson and Essex County (New Jersey). The bottom middle graph (Figure 1c) highlights San Francisco County. For more information about the sick pay reforms, see Table B1.

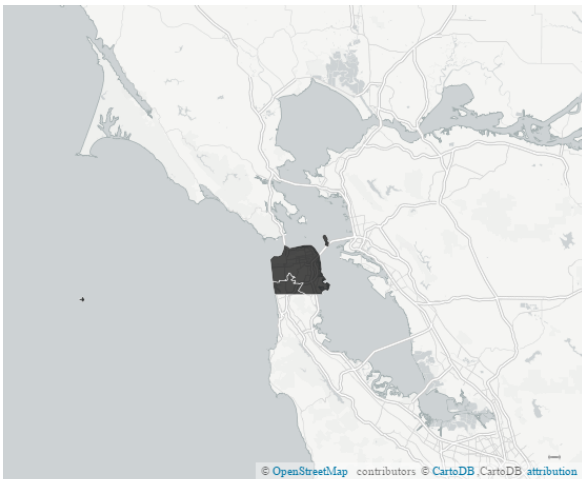
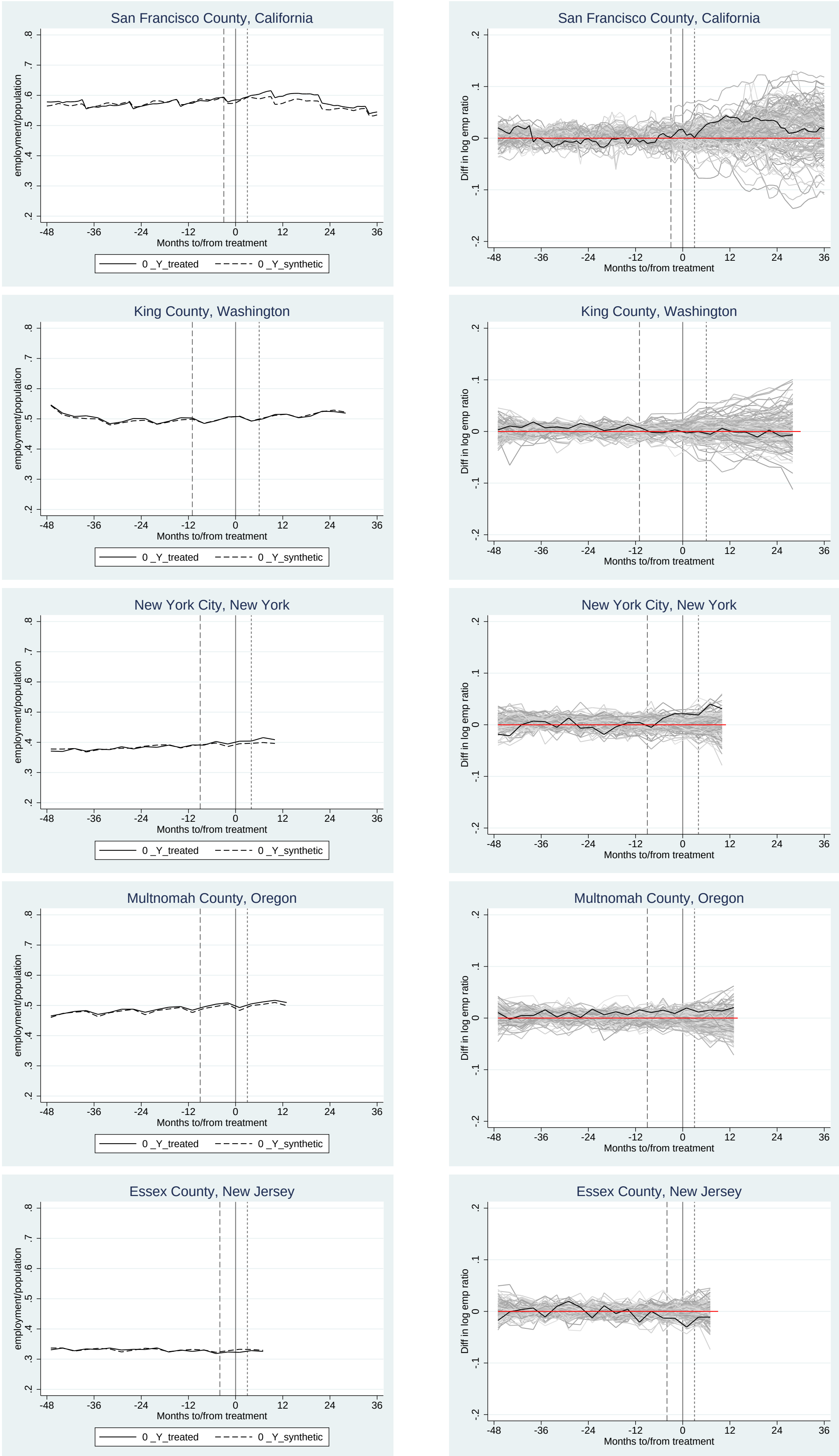
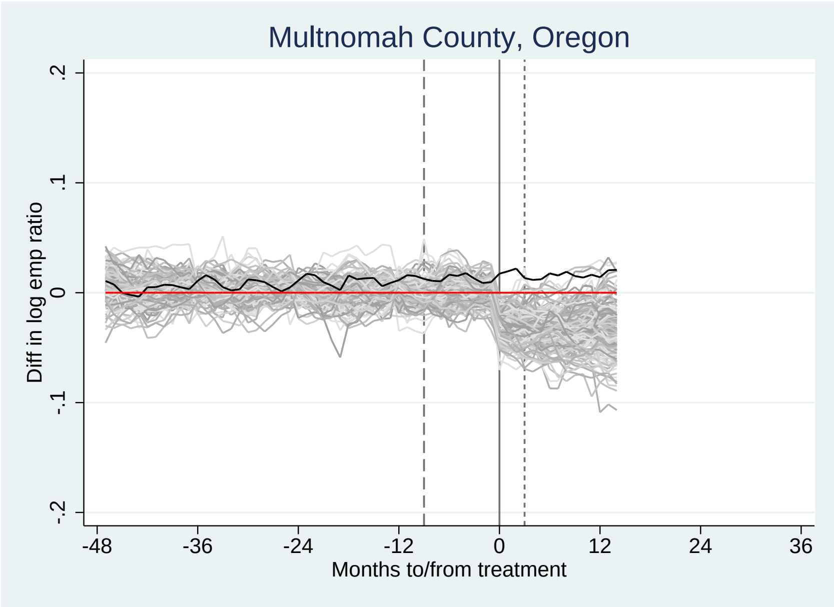


Figure 2: Development of Employment Ratios in Treated vs. Synthetic Control Counties



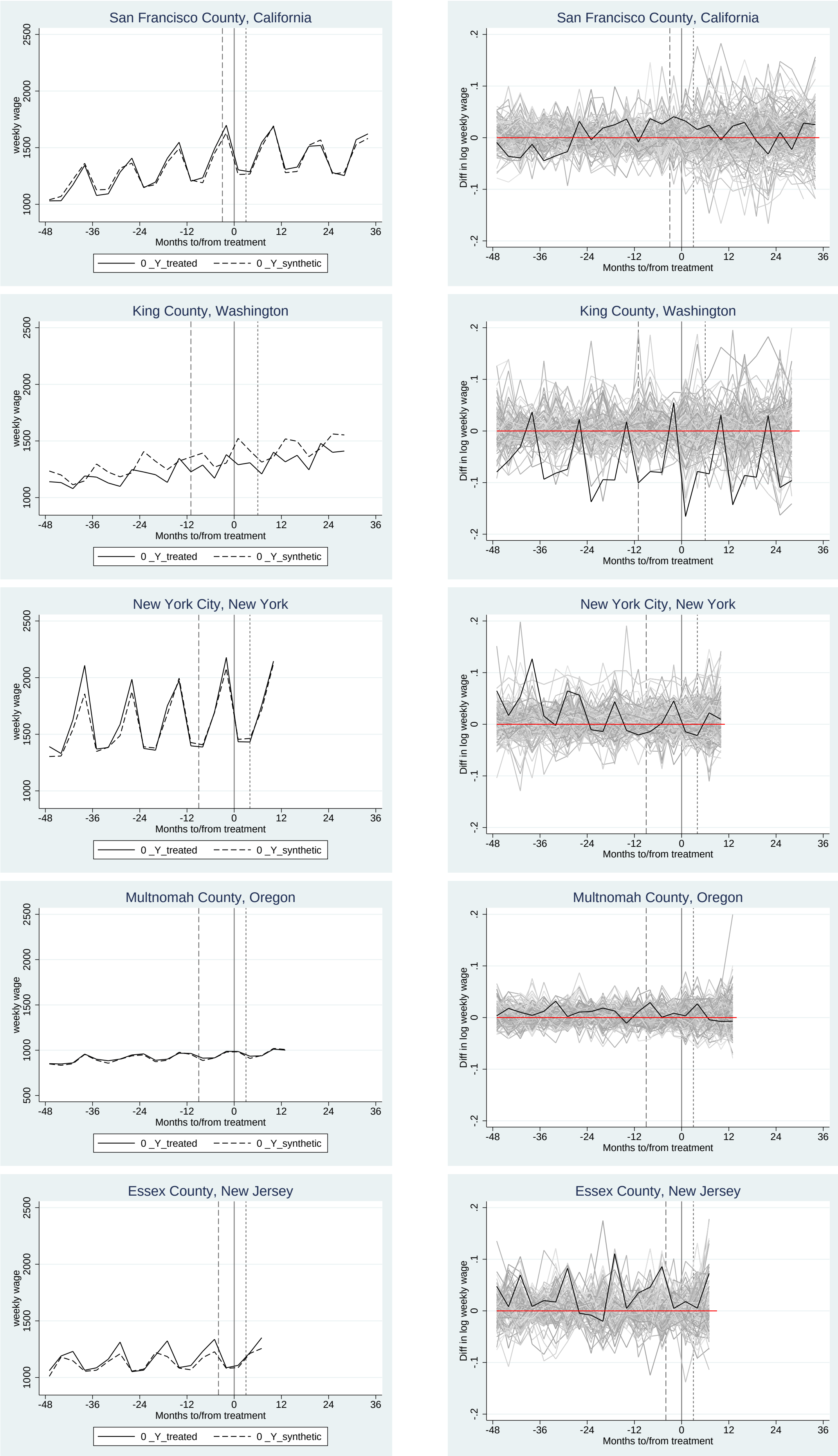
The left column shows the development for the treated counties (solid line) vs. the synthetic control group counties (dashed line). The composition of the synthetic control counties is in Table A1. The right column shows the difference of the logarithm of the employment ratios between treated and control counties (treated-controls) along with placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). For more information about the sick pay reforms, see Table B1.

Figure 3: Calculation of P-value of a Potential Employment Decrease >3%



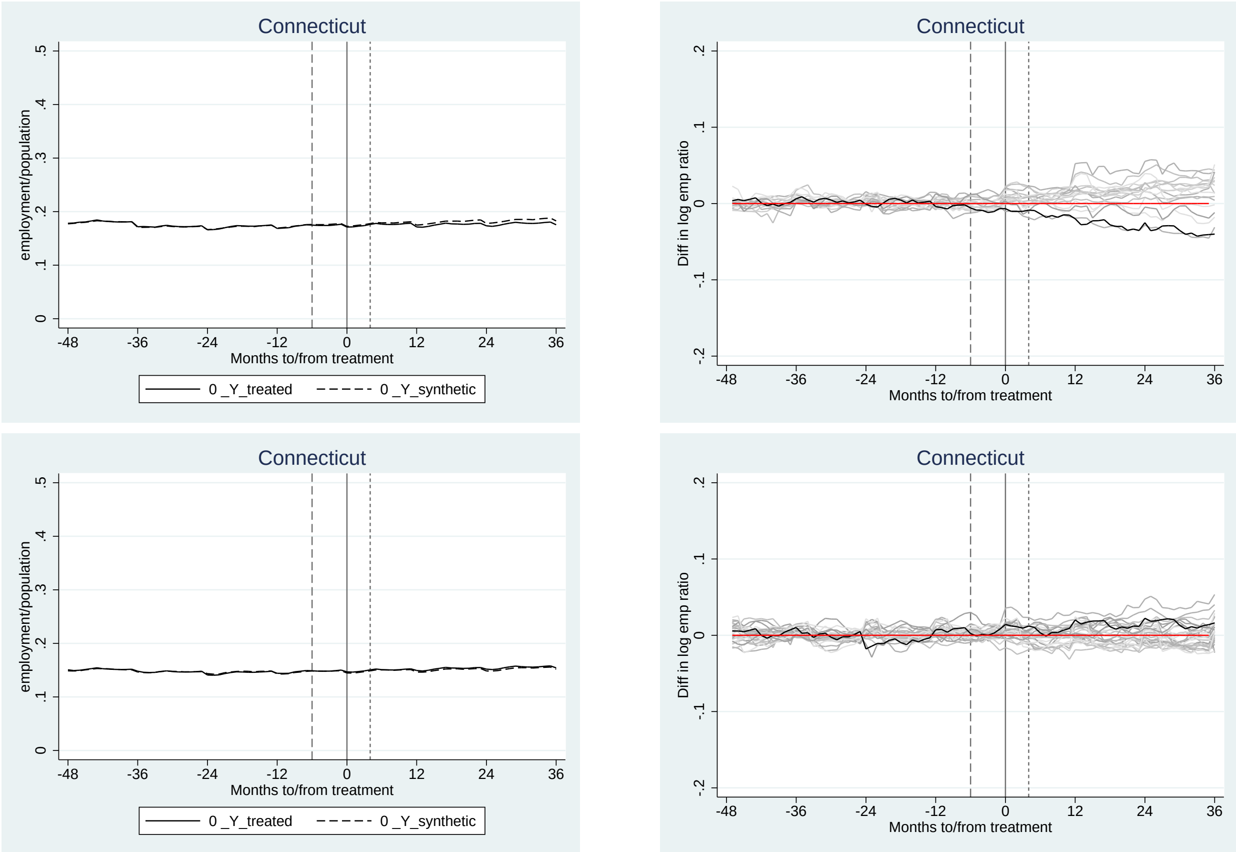
This graph shows shows the difference of the logarithm of the employment ratios between treated and control counties (treated-controls) along with placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). The graph illustrates how the p-values in columns (9) and (10) of Tables 3, 4, and 5 were constructed. Also see Section 5 for a discussion. The method employed here alters the placebo estimates by adding a pre-defined hypothetical treatment effect (-3% in this case) and then calculates the likelihood (p-value) that the treated county/state stems from this distribution. For more information see the discussion of the results in the main text.

Figure 4: Development of Weekly Wages in Treated vs. Synthetic Control Counties



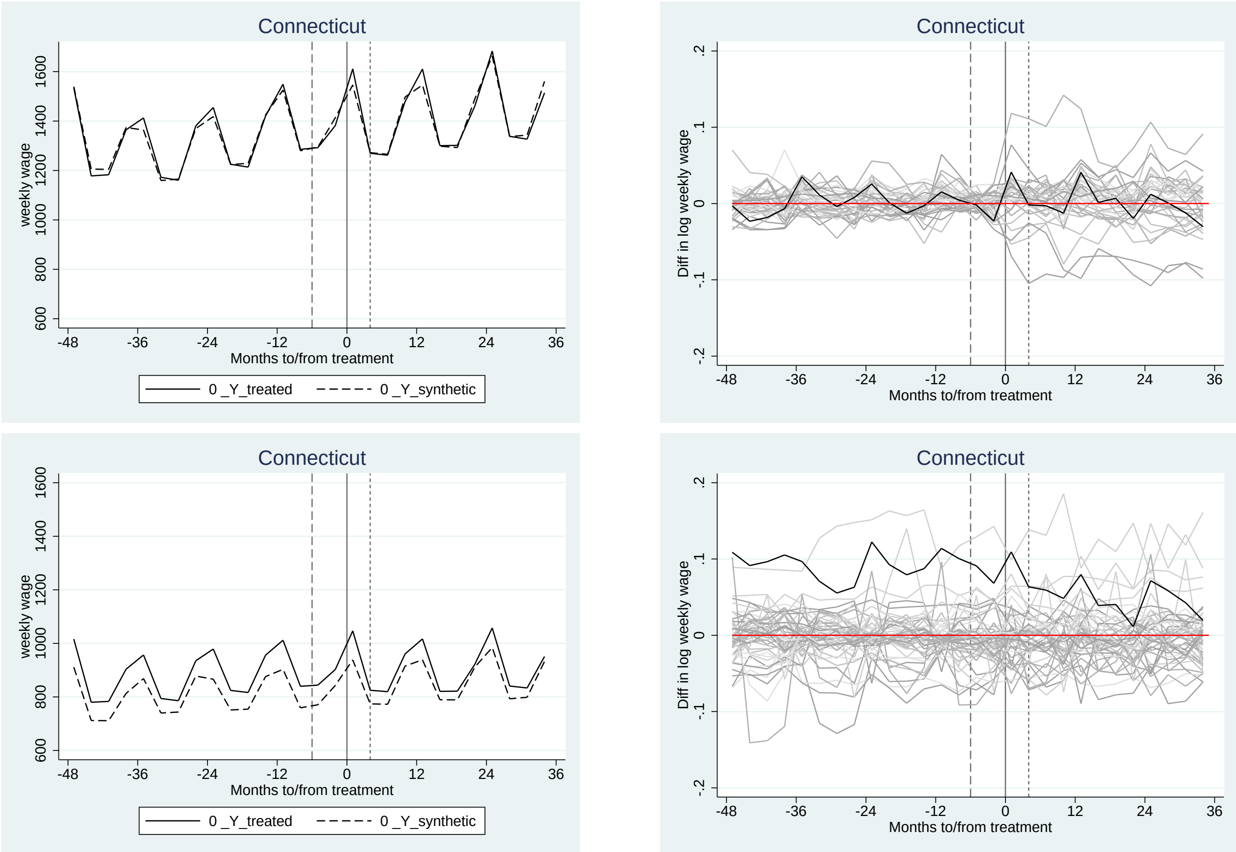
The left column shows the development for the treated counties (solid line) vs. the synthetic control group counties (dashed line). The composition of the synthetic control counties is in Table A2. The right column shows the difference of the logarithm of weekly wages (reported quarterly) between treated and control counties (treated-controls) along with placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). For more information about the sick pay reforms, see Table B1.

Figure 5: Development of Employment Ratios in Connecticut vs. Synthetic Control States: (a) Service Sector Firms >49 Employees, (b) Service Sector Firms <50 Employees



The left column shows the development for Connecticut (solid line) vs. the synthetic control group state (dashed line). The composition of the synthetic control state is in Table A3. The right column shows the difference of the logarithm of the employment ratios along with placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). The first row shows the result for firms with more than 49 employees (which were affected by the law), and the second row shows results for firms with less than 50 employees. For more information about the sick pay reforms, see Table B1.

Figure 6: Development of Weekly Wages in Connecticut vs. Synthetic Control States: (a) Service Sector Firms >49 Employees, (b) Service Sector Firms <50 Employees



The left column shows the development for Connecticut (solid line) vs. the synthetic control group state (dashed line). The composition of the synthetic control state is in Table A3. The right column shows the difference of the logarithm of weekly wages (reported quarterly) along with placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). The first row shows the result for firms with more than 49 employees (which were affected by the law), and the second row shows results for firms with less than 50 employees. For more information about the sick pay reforms, see Table B1.

Table 1: Quarterly Census of Employment and Wages (QCEW), County Level: 2001-2015

Variable	Mean	Std. Dev.	Min.	Max.	N
Private sector employment	0.271	0.134	0.011	4.019	534,006
Public sector employment	0.077	0.036	0.012	0.496	534,006
Service sector employment	0.192	0.109	0	2.996	534,006
Non-service sector(=production) employment	0.078	0.064	0	3.569	534,006
Private sector wages	603	175	158	5,124	178,002
Population	98,035	338,036	258	10,116,705	42,966

Source: QCEW (Bureau of Labor Statistics (BLS), 2016), own calculation and illustration. County level population data are taken from (United States Census Bureau, 2016b).

Table 2: Quarterly Census of Employment and Wages (QCEW), State Level: 2001-2015

Variable	Mean	Std. Dev.	Min.	Max.	N
Private service sector employment, >49 employees	0.153	0.054	0.047	0.547	8,721
Private service sector employment, <50 employees	0.146	0.028	0.081	0.355	8,721
Private non-service sector employment, >49 employees	0.045	0.016	0.006	0.098	8,721
Public sector employment	0.081	0.046	0.044	0.414	8,721
Private service sector wages, >49 employees	821.49	226.92	423.17	2030.24	2,907

Source: QCEW (Bureau of Labor Statistics (BLS), 2016), own calculation and illustration. State level population data are taken from (United States Census Bureau, 2016b).

Table 3: Key Statistics—Using the Synthetic Control Group Method to Assess Changes in Employment after the Implementation of Sick Pay Mandates

	$\bar{Y}^1_{i,pre}$ (1)	RMSPE pre (2)	RMPSE post (3)	RMPSE Ratio (4)	Rank RMSPE/ #Placebos= P-Value (5)	PTE (6)	LTE (7)	Rank PTE/ #Placebos= P-Value (8)	P-value PTE<-3% (9)	P-value LTE<-2ppt (10)
San Francisco,CA	0.5742	0.0117	0.0272	2.3195	82/167=0.4910	0.0245	0.0144	94/151=0.6225	0.0464	0.0132
King County, WA	0.5041	0.0102	0.0053	0.5214	151/153=0.9869	-0.0022	-0.0012	57/140=0.4071	0.1214	0.0571
New York City,NY	0.3835	0.0113	0.0292	2.5915	15/183=0.082	0.0282	0.0116	164/169=0.9704	0.0059	0.0059
Multnomah, OR	0.4854	0.0104	0.0169	1.6200	79/193=0.4093	0.0166	0.0083	161/181=0.8895	0.0055	0.0055
Essex County, NJ	0.3316	0.0106	0.0185	1.7409	57/163=0.3497	-0.0163	-0.0055	18/148=0.1216	0.1351	0.0135
Hudson County, NJ	0.2965	0.0141	0.0196	1.3859	36/93=0.3871	-0.0167	-0.0051	8/79=0.1013	0.1772	0.0506
District of Columbia	0.7752	0.0453	0.1076	2.3722	58/156=0.3718	0.1072	0.0761	156/156=1	0.0064	0.0064
Sum of P-Values					2.7060			3.1125	0.4916	0.1460
P-Value Irwin Hall					0.3424			0.5616	0.00002	$1.34389*10^{-8}$

Source: QCEW (Bureau of Labor Statistics (BLS), 2016), own calculation and illustration. All statistics displayed here are discussed in Section 5. Column (1) displays the outcome measure in levels for each treated county averaged over all pre-reform years. Columns (2) and (3) display the *RMSPE* as in equation (3) for pre and post-reform years, respectively. Column (4) displays the *RMSPE Ratio* [RMSPE post/RMSPE pre]. Column (5) calculates the p-value of the *RMSPE Ratio* for all treated counties using the indicated number of placebo estimates. Columns (6) and (7) show the *PTE* and *LTE* as in equations (4) and (5). Column (8) calculates the p-value of the *PTE* for all treated counties using the indicated number of placebo estimates. Columns (9) and (10) display the p-values of hypothetical employment decreases of 3% and 2ppt respectively (see main text and Figure 3 for more details). As for the joint tests and sum of all p-values per county, we exclude the District of Columbia due to a poor pre RMSPE fit. For more information, see the discussion on treatment regions in Section 4. For more information about the sick pay reforms, see Table B1.

Table 4: Key Statistics—Using the Synthetic Control Group Method to Assess Changes in Weekly Wages after the Implementation of Sick Pay Mandates

	$\bar{Y}^1_{i,pre}$ (1)	**RMSPE pre (2)**	**RMPSE post (3)**	**RMPSE Ratio (4)**	**Rank RMSPE/ #Placebos= P-Value (5)**	**PTE (6)**	**LTE (7)**	**Rank PTE/ #Placebos= P-Value (8)**	**P-value PTE$<$-3% (9)**	**P-value LTE$<$-€40 (10)**
San Francisco,CA	1263.27	0.0293	0.0237	0.8092	156/167=0.9341	0.0110	14.05	97/158=0.6139	0.0886	0.0506
King County, WA	1194.19	0.0772	0.0986	1.2769	97/153=0.6339	-0.0747	-110.45	4/151=0.0265	0.9669	0.9735
New York City,NY	1618.65	0.0475	0.0179	0.3763	180/183=0.9836	-0.0012	3.63	71/177=0.4011	0.1525	0.0847
Multnomah, OR	915.76	0.0148	0.0130	0.8770	146/193=0.7564	0.0023	1.76	63/141=0.4468	0.0709	0.0284
Essex County, NJ	1160.52	0.0489	0.0411	0.8397	129/163=0.7913	0.0292	40.26	135/158=0.8544	0.0253	0.0190
Hudson County, NJ	1343.03	0.1319	0.1260	0.9552	61/93=0.6559	0.1159	148.60	93/93=1	0.0108	0.0108
District of Columbia	1281.16	0.0271	0.0241	0.8902	118/156=0.7564	-0.0030	-6.04	59/144=0.4097	0.1875	0.1458
Sum of P-Values					5.5119			3.7525	1.5025	1.3128
P-Value Irwin Hall					0.9968			0.6269	0.0034	0.0013

Source: QCEW (Bureau of Labor Statistics (BLS), 2016), own calculation and illustration. All statistics displayed here are discussed in Section 5. Column (1) displays the outcome measure in levels for each treated county averaged over all pre-reform years. Columns (2) and (3) display the *RMSPE* as in equation (3) for pre and post-reform years, respectively. Column (4) displays the *RMSPE Ratio* [RMSPE post/RMSPE pre]. Column (5) calculates the p-value of the *RMSPE Ratio* for all treated counties using the indicated number of placebo estimates. Columns (6) and (7) show the *PTE* and *LTE* as in equations (4) and (5). Column (8) calculates the p-value of the *PTE* for all treated counties using the indicated number of placebo estimates. Columns (9) and (10) display the p-values of hypothetical weekly wage decreases of 3% and €40 respectively (see main text and Figure 3 for more details). As for the joint tests and sum of all p-values per county, we exclude the District of Columbia due to a poor pre RMSPE fit. For more information, see the discussion on treatment regions in Section 4. For more information about the sick pay reforms, see Table B1.

Table 5: Key Statistics—Using the Synthetic Control Group Method to Assess Labor Market Effects in Connecticut

	$\bar{Y}^1_{i,pre}$ (1)	RMSPE pre (2)	RMPSE post (3)	RMPSE Ratio (4)	Rank RMSPE/ #Placebos= P-Value (5)	PTE (6)	LTE (7)	Rank PTE/ #Placebos= P-Value (8)	P-value PTE<-3% (9)	P-value LTE <-2ppt/-€40 (10)
Panel A: Employment										
Firms >49 employees	0.1746	0.0051	0.0277	5.4588	2/51=0.0392	-0.0256	-0.0046	2/17=0.1176	0.7647	0.0588
Firms <50 employees	0.1478	0.0069	0.0139	2.0087	14/51=0.2745	0.0128	0.0019	26/30=0.8667	0.0333	0.0333
Panel B: Wages										
Firms >49 employees	1325.61	0.0157	0.0210	1.3362	29/51=0.5686	0.0030	5.4196	14/38=0.3684	0.0789474	0.05263
Firms <50 employees	883.08	0.0920	0.0598	0.6497	46/51=0.9019	0.0542	48.9511	47/50=0.94	0.06	0.04

Source: QCEW (Bureau of Labor Statistics (BLS), 2016), own calculation and illustration. All statistics displayed here are discussed in Section 5. Column (1) displays the outcome measure in levels averaged over all pre-reform years. Columns (2) and (3) display the *RMSPE* as in equation (3) for pre and post-reform years, respectively. Column (4) displays the *RMSPE Ratio* [RMSPE post/RMSPE pre]. Column (5) calculates the p-value of the *RMSPE Ratio* using the indicated number of placebo estimates. Columns (6) and (7) show the *PTE* and *LTE* as in equations (4) and (5). Column (8) calculates the p-value of the *PTE* using the indicated number of placebo estimates. Columns (9) and (10) display the p-values of hypothetical employment and weekly wage decreases (see main text and Figure 3 for more details). For more information, see the discussion on treatment regions in Section 4. For more information about the sick pay reforms, see Table B1.

Appendix A

Table A1: Counties for Synthetic Control Group: Employment

	San Francisco	King (WA)	NYC (NY)	Multnomah (OR)	Essex (NJ)	Hudson (NJ)	DC (DC)
Fulton, GE	0	0.027	0	0.011	a	a	0.72
Suffolk, MA	0.309	0	0	0	a	a	0.28
Montgomery, MD	0	0	0	0	0.537	a	0
Arlington, VA	0.244	0	0.211	0	a	a	0
Contra Costa, CA	a	a	a	a	a	0.447	a
Somerset, NJ	0.446	0	0	0	0	a	a
Westchester, NY	0	0	0	0	0	0.425	0
Marin, CA	0	0	0.403	0	0	0	0
Denver, CO	0	0.295	0	0	0	a	a
San Mateo, CA	0	0.285	0	0	0	a	0
Brevard, FL	a	a	a	a	0.281	0	a
Clayton, GE	a	a	0.23	0	0	0	a
Fayette, KY	0	0	a	0.199	a	a	0
Polk, IA	0	0	0	0.181	a	a	0
Santa Clara, CA	0	0.174	0	0	0	a	0
DeKalb, GA	a	a	0.156	a	0	0	a
Washoe, NV	0	0	0	0.14	0	a	0
Durham, NC	0	0.131	0	0	0	a	a
St. Mary's, MD	a	a	a	a	a	0.128	a
Travis, TX	0	0	0	0.112	0	a	0
Washtenaw, MI	0	a	0	0	0.106	0	a
St. Louis City, MO	0	0	0	0.104	a	a	0
Williamson, TN	0	0	0	0.085	0	a	0
Sacramento, CA	a	a	a	a	0.076	0	a
Richmond City, VA	0	0.07	0	0	0	a	0
Cass, ND	a	0	0	0.067	a	a	a
Albany, NY	0	0	0	0.065	a	a	0
Elkhart, IN	0	a	a	0.029	a	a	0
Midland, TX	0	0.018	0	0	a	a	0
Washington, OR	0	0	0	0.005	0	a	0
Kent, MI	0	0	0	0.002	a	a	0

Sources: QCEW, own calculation and illustration. 'a' indicates that the variables for employment, wages, and county population do not lie within the region of support of the treatment county. Thus these counties are not considered as potential "donors." '0' indicates that the county is a potential control county donor but has not actually been used as a donor. All counties with positive fractions indicate the donor share employed by the synthetic control group method for the treatment county in the column header. Thus, all fractions in one column add to 100%.

Table A2: Counties for Synthetic Control Groups: Weekly Wages

	San Francisco	King (WA)	NYC (NY)	Multnomah (OR)	Essex (NJ)	Hudson (NJ)	DC (DC)
Suffolk, MA	0.609	0.051	0.245	0.000	a	a	0.103
Arlington, VA	0.132	0.000	0.000	0.000	a	a	0.787
Westchester, NY	0.110	0.000	0.000	0.000	0.173	0.517	0.000
Somerset, NJ	0.149	0.000	0.650	0.000	a	a	0.000
Contra Costa, CA	a	a	a	a	a	0.483	a
Montgomery, MY	0.000	0.000	0.000	0.000	0.479	a	0.000
Cass, ND	a	0.235	0.000	0.200	a	a	a
Santa Clara, CA	0.000	0.380	0.000	0.000	0.000	a	0.000
Sacramento, CA	a	a	a	a	0.274	0.000	a
Travis, TX	0.000	0.000	0.000	0.226	0.000	a	0.000
San Mateo, CA	0.000	0.087	0.105	0.000	0.000	a	0.000
Fulton, GA	0.000	0.040	0.000	0.033	a	a	0.111
Alexandria City, VA	0.000	0.167	0.000	0.000	a	a	0.000
Polk, IA	0.000	0.000	0.000	0.163	a	a	0.000
Fayette, KY	0.000	0.000	a	0.123	a	a	0.000
Anne Arundel, MY	a	0.000	0.000	0.104	0.000	0.000	a
Douglas, CO	a	a	a	a	0.074	0.000	a
Washoe, NV	0.000	0.000	0.000	0.056	0.000	a	0.000
Kent, MI	0.000	0.000	0.000	0.040	a	a	0.000
Madison, AL	0.000	0.040	0.000	0.000	0.000	a	0.000
Albany, NY	0.000	0.000	0.000	0.033	a	a	0.000
Washtenaw, MI	0.000	a	0.000	0.012	0.000	0.000	a
St. Louis City, MO	0.000	0.000	0.000	0.006	a	a	0.000
Williamson, TN	0.000	0.000	0.000	0.003	0.000	a	0.000

Sources: **QCEW**, own calculation and illustration. 'a' indicates that the variables for employment, wages, and county population do not lie within the region of support of the treatment county. Thus these counties are not considered as potential "donors." '0 'indicates that the county is a potential control county donor but has not actually been used as a donor. All counties with positive fractions indicate the donor share employed by the synthetic control group method for the treatment county in the column header. Thus, all fractions in one column add to 100%.

Table A3: States Used to build Synthetic Control State for Connecticut

	Private Service Sector Employment		Private Service Sector Wages	
	>49 employees	**<50 employees**	**>49 employees**	**<50 employees**
Massachusetts	0.226	0.433	0.478	0.523
New York	0.44	0.101	0.403	0.436
California	0.24	0.223	0.003	0
Iowa	0	0.175	0	0.016
Wisconsin	0.091	0	0	0.025
Delaware	0	0	0.116	0
Oregon	0.003	0.068	0	0
Wyoming	0	0	0	0

Sources: QCEW, own calculation and illustration. '0' indicates that the state is a potential donor state but has not actually been used. All states with positive fractions indicate the donor share employed by the synthetic control group method for the treatment state in the column header. Thus, all fractions in one column add to 100%.

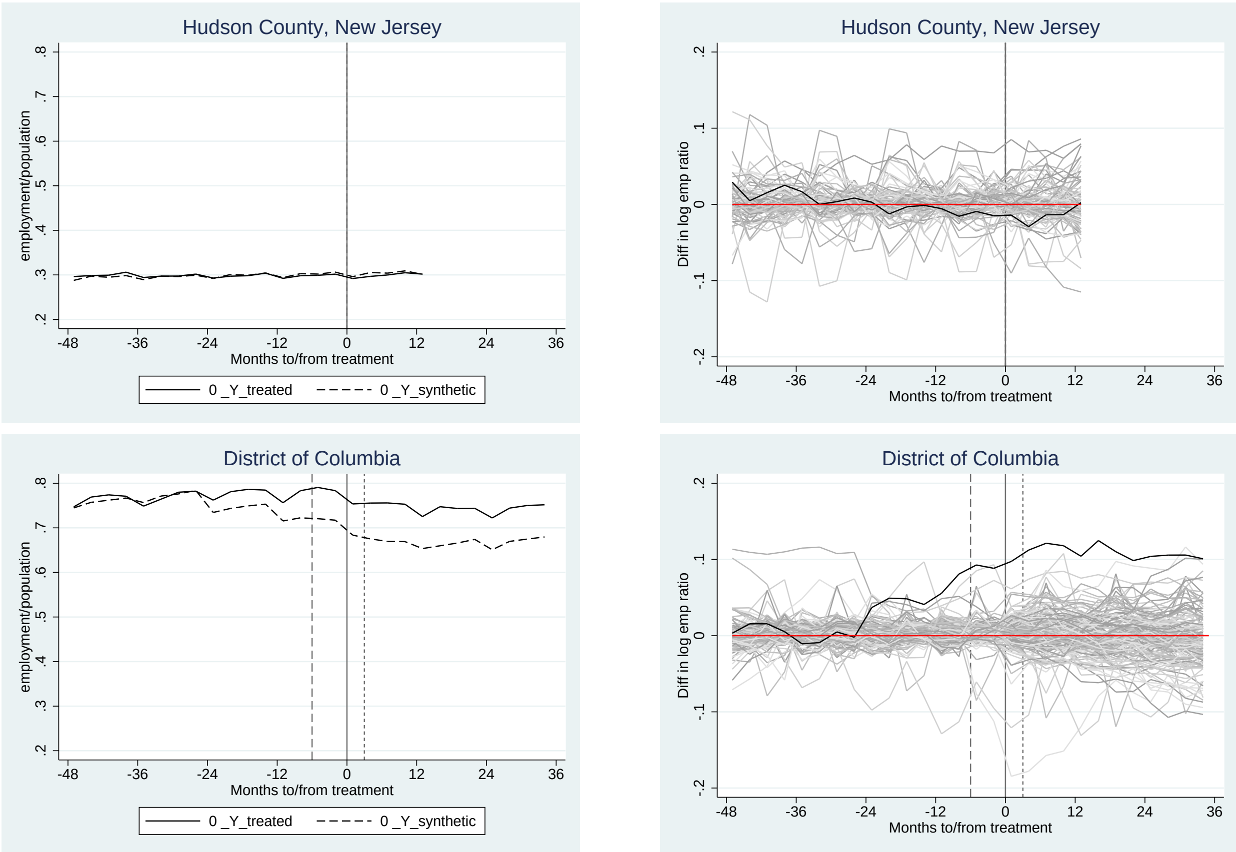
Appendix B

Table B1: Overview of Employer Sick Pay Mandates in the US

Region (1)	County (2)	Law Passed (3)	Law Effective (4)	Content (5)
San Francisco, CA	SF	Nov 7, 2006	Feb 5, 2007	all employees including part-time and temporary; 1 hour of paid sick leave for every 30 hours worked; up to 5 to 9 days depending on firm size; for own sickness or family member; 90 days accrual period
Washington, DC	DC	May 13, 2008	Nov 13, 2008	'qualified employees'; 1 hour of paid sick leave for every 43 hours, 90 days accrual period; up to 3 to 9 days depend. on firm size; own sickness or family; no health care or restaurant workers
		Dec 18, 2013	Feb 22, 2014 (retrosp. in Sep 2014)	extension to 20,000 temporary workers and tipped employees
Connecticut		July 1, 2011	Jan 1, 2012	full-time service sector employees in firms>49 employees (20% of workforce); 1 hour for every 40 hours; up to 5 days; own sickness or family member, 680 hours accrual period (4 months)
Seattle, WA	King	Sep 12, 2011	Sep 1, 2012	all employees in firms with >4 full-time employees; 1 hour for every 30 or 40 hours worked; up to 5 to 13 days depending on firm size, for own sickness or family member; 180 days accrual period
New York, NY	Bronx, Kings, New York, Queens, Richmond	June 26, 2013 Jan 17, 2014 extended	April 1, 2014	employees w >80 hours p.a in firms >4 employees or 1 domestic worker; 1 hour for every 30 hours; up to 40 hours; own sickness or family member; 120 days accrual period
Portland, OR	Multnomah	March 13, 2013	Jan 1 2014	employees w >250 hours p.a. in firms >5 employees; 1 hour for every 30 hours; up to 40 hours; own sickness or family member
Jersey City, NJ	Hudson	Sep 26, 2013 Oct 28, 2015 extended	Jan 22, 2014	all employees in private firms with >9 employees; 1 hour for every 30 hours; up to 40 hours; own sickness or family; 90 days accrual period
Newark, NJ	Essex	Jan 29, 2014	May 29, 2014	all employees in private companies; 1 hour for every 30 hours; 90 days accrual period; up to 24 to 40 hours depending on size; own sickness or family
Philadelphia, PA		Feb 12, 2015	May 13, 2015	all employees in firms >9 employees; 1 hour for every 40 hours; up to 40 hours; own sickness or family member; 90 days accrual period
California		September 19, 2014	July 1, 2015	all employees; 1 hour of paid sick leave for every 30 hours; minimum 24 hours; own sickness or family member; 90 days accrual period
Massachusetts		Nov 4, 2014	July 1, 2015	all employees in firms >10 employees; 1 hour for every 40 hours; up to 40 hours; own sickness or family member; 90 days accrual period
Oakland, CA		Nov 4, 2014	March 2, 2015	all employees in firms >9 employees; 1 hour for every 30 hours; 90 days accrual period; up to 40 to 72 hours depending on firm size; own sickness or family member
Oregon		June 22, 2015	Jan 1, 2016	all employees in firms >9 employees; 1 hour for every 30 hours; 90 days accrual period; up to 40 hours; own sickness or family member

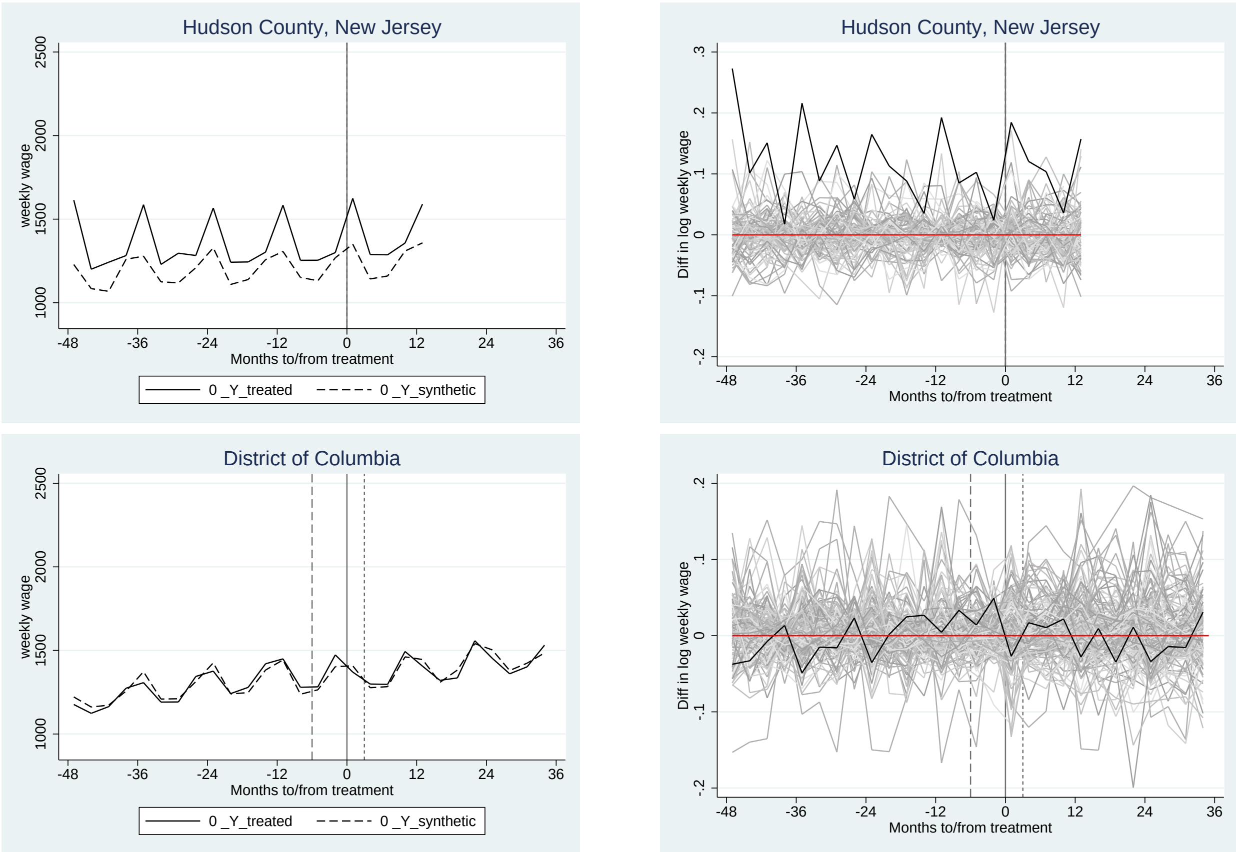
Source: several sources, own collection, own illustration. Labor market outcomes of cities in black are evaluated in this paper.

Figure B1: Employment in Hudson County, DC and Their Synthetic Control Counties



The left column shows the development for the two treated counties (sold lines) vs. the synthetic control group counties (dashed lines). The composition of the synthetic control counties is in Table A1. The right column shows the difference of the logarithm of the employment ratios between treated and control counties (treated-controls) along with all placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). For more information about the sick pay reforms, see Table B1.

Figure B2: Wages in Hudson County, DC and Their Synthetic Control Counties



The left column shows the development for the two treated counties (sold lines) vs. the synthetic control group counties (dashed lines). The composition of the synthetic control counties is in Table A2. The right column shows the difference of the logarithm of weekly wages between treated and control counties (treated-controls) along with all placebo estimates for counties with an RMSPE smaller than 2 times the RMSPE of the treated county (gray lines). For more information about the sick pay reforms, see Table B1.