UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

The impact of the Equal Pay Act of 2003 on gender equality in the Illinois labour market.

Jon-Samuel Andersson

Gustav Ekman

Spring 2018

Abstract:

Using data extracted through the Current Population Survey (CPS), this thesis evaluates the effects of the Illinois 2003 Equal Pay Act and its amendments. This is done through studying the effects that the implementation of this legislation had on gender specific labour market outcomes. The results from a quantitative approach, though not conclusive, indicate that the legislation may have had our anticipated effects on labour market outcomes in Illinois: The gender pay gap, as well as the female labour market participation, decreased. This suggests that the wage of a female employee compared to that of a male employee increased, whilst, contrary to the purpose of the law, the total number of hours worked by females compared to males decreased after the implementation of the protective legislation in Illinois. This thesis also addresses the need for further research of the appropriateness of labour legislation as a protective measure against labour market discrimination.

Key words: gender equality, Difference-in-Difference, Synthetic Control Method, Equal Pay Act, labour market, policy evaluation

Kandidatuppsats i nationalekonomi /
Bachelor's thesis in Economics (15hp)

Department of Economics,
School of Business, Economics and Law
University of Gothenburg

Supervisor: Maksym Khomenko

Acknowledgment

We would like to thank our supervisor Maksym Khomenko for valuable support throughout the procedure of writing this thesis.

Table of contents

1 Introduction

The gender pay gap has been heavily discussed for the last couple of decades, and at this point its existence is widely taken for granted. After examining some basic income data, certain individuals will probably report that the size of it adds up to a troublesome 20%, or maybe even 30 %, gap resulting from pure female discrimination. Others will proclaim that these numbers are perfectly reasonable, harmless even, as they are accounted for in terms of differences in education level, experience, hours spent working, and so on. The aim of this paper is not to take part in these types of discussions, but rather to focus on a very particular, and highly relevant policy change, the likes of which seeks to remedy this presumed injustice, namely *the Illinois equal pay act of 2003*.

Now, it might seem curious that we concern ourselves with such a particular law in one particular state of one particular country. The possibility of the gender pay gap is, of course, not restricted to one small geographical region. The law on the other hand, might only apply to a specific state, but nothing prevents policy makers in other regions from implementing similar types of legislation in aiming towards eliminating unjust inequalities. Before doing so, it would be expedient to evaluate the effect of such interventions, and one natural way of doing so is by examining previous examples thereof. Illinois offers an attractive way of doing so, partly because the labour force data in the USA is so easily accessible, but also because it was one of the first (the second) states to implement its own amendment to the 1963 federal equal pay act, yet, to our knowledge, its effectiveness has still not been thoroughly investigated. This provides justification for our research question: *Does the Illinois 2003 equal pay act affect gender equality in the Illinois labour market?*

In this paper we aim to examine the casual effect of the act and its amendments on gender equality in the Illinois labour market. In doing so we use pooled cross-sectional cps data over the years 2000-2016. We use a difference-in-difference (DiD) method, comparing the pay gap in Illinois, the treated group, with a similar group which was not treated, the control group. In order to find a sufficiently similar comparison group, we create a synthetic untreated control state which uses a set of predictors to follow the trending gender pay gap in Illinois in the years prior to the policy change. The synthetic state uses these predictors to estimate the 'had-been' gaps in Illinois absent the equal pay act, thus providing an appropriate comparison group. We measure the gender pay-gap by the difference in median hourly wage between men and women. We also evaluate whether the implementation of the law influenced the number of hours worked by female employees compared to the number of hours worked by male employees. This is done in order to control whether or not legislation of this sort has an impact on the employability of female workers.

Policy makers and general public alike, who regard discrimination in the workplace as problematic, should be interested in what effect anti-discrimination legislation has on such problems. Our thesis contributes to existing research on that topic in the following ways: we try to present a robust analysis of the possible outcomes incentivized by the legislation on the extensive, and intensive, margin of gender specific labour market outcomes. We use a relatively new methodology within policy evaluation; synthetic control model with additional placebo testing. In addition to previously mentioned aspects, we have not been able to come across any previous economic study on the effect the Illinois pay act on the gender pay gap.

2 Background

In January 2004 the state of Illinois implemented its first gender specific equal pay act, the equal pay act of 2003 (EPA). The law is an addition to the federal equal pay act of 1963 with the aim to take further action towards closing the gender pay gap. The law was later extended through two amendments, the first implemented in August 2009 and the second in August 2014.

2.1 The Equal Pay Act

The *federal* Equal Pay Act was signed in 1963 by president John F. Kennedy. This law prohibited wage discrimination based on gender (Wilkins, 2012). It was well known that initially, the law would not have a substantial impact. This was because the labour market segregation was substantial, which meant that it was hard to find male and female workers who performed similar work. The implementation of the law was however, seen as an initial progress towards closing the pay gap between the sexes, and changing attitudes within the labour market. According to the Obama administration, the equal pay act of 1963 and its amendments have been crucial for the decline in the gender pay gap that we have seen during the last decades. (National Equal Pay Enforcement Task Force, 2013)

The Illinois equal pay act was implemented in January 2004 and forbids employers from paying men and women different wages for the same, or substantially similar, work (ILCS, 2003). Employees who believe they have been receiving a lower wage than their opposite sex counterpart may file complaints to the Illinois Department of Labor (IDOL) either via a form available at the departments website, or by directly contacting the department, who will then investigate the complaint further. If the employer is found underpaying, then the employer is obliged to '…pay the entire amount of any underpayment together with interest and the costs and reasonable attorney's fees as may be allowed by the court and as necessary to make the employee whole.' (Illinois General Assembly, 2003).

The Equal Pay Act was amended in 2009. Originally, the complaint was due within 180 days from the underpayment. This time period was now extended to one year. Previously the period for which the underpayment charges was valid was limited by 'the date the employer learned of the underpayment' but now, the limitations period runs from 'the date of the underpayment', which is defined as 'each time wages are underpaid' (Mungerson, 2009).

In 2014, the law was once again amended, this time giving IDOL more power when it comes to investigating pay discrimination as each agreement, regardless of written or oral, may be seen as a contract and scrutinized as such (Halpern, 2015). This amendment allowed additional prosecution

through enabling the department to refer to the Department of Human Right, meaning that the allegations may go to trial under the Human Right Act in addition to the Equal Pay Act (National Conference of State Legislation, 2014).

2.2 Earlier studies

There has been extensive research within economics, psychology and sociology, for example, on the underlying causes of gender discrimination in the workplace. The research on the effectiveness of anti-discrimination legislation on the other hand is scarce. This may be due to the complexity of examining the effect of such legislation in isolation from other unexplained changes. John J. Donohue (2005), for instance, suggests that there is empirical evidence that imply that anti-discrimination laws may help those that already have jobs, whilst it may dampen the supply of new job openings. This is just one example that illustrates the multi-layered aspects of the effects of policy changes in the labour market associated with gender equality.

Evaluating the impact of a single law or policy change is also particularly problematic due to the unobservable nature of counterfactual outcomes. If such a change occurs researchers cannot with absolute certainty know what would have been the case absent the change, simply because the world in which they conduct their inquiries does not present that scenario. Nonetheless this field is not unexplored, here we summarize some of the more notable previous literature on the field.

The United States does, according to Blau and Kahn (1996), compare well to the other OECD countries in terms of human capital and occupational distribution. The U.S. had one of the longest history of anti-discrimination legislation. However, they did rank in the bottom of the countries with the largest pay gap between male and female workers. The combination of extensive anti-discrimination legislation with a substantial pay gap might sound contradictory, and it is investigated by both Waldfogel (1998), and Darity and Mason (1998). Waldfogel argues that legislation and policy changes are helpful, but not sufficient in closing the differences that persist in the labour market. Similarly, according to Darity and Mason the movement towards a closing wage gap that has been seen during the last decades is largely empowered by the Equal Pay Act of 1963 and its amendments, though they also conclude that there were strong demographic and social trends that drove this development as well.

In addition to the inquiries about the gender pay gap, researchers have also tried to evaluate whether anti-discrimination legislation has effects on the employability of the discriminated group in question. Interestingly, Neumark and Stock (2001), as well as Donohue (2005), found that anti sex-discrimination laws, such as equal pay laws, actually *reduced* the relative employment of women.

Nonetheless, they do not argue that this result would induce a net negative effect of antidiscrimination laws, but rather that there might be a trade-off between higher pay and lower employability.

3 Theoretical framework

In this section we present some of the most prevalent theories on labour market discrimination following Altonji and Blank (1999), who published a chapter in the *Handbook of Labor Economics; Race and gender in the labor market.* In this chapter the authors present an extensive discussion of current theories on the sources of race and gender differences in the labour market. The theories presented are: *the human capital model*, *taste-based models of discrimination*, *theories of crowding and occupational exclusion based on social norms*, *the added worker theory*, *the theory of alternate costs*, and finally, *the theory of statistical discrimination*. These theories provide the grounds for the discussion section and will help us analyse the outcomes that we will encounter.

The human capital model (Becker, 1964) states that wage discrimination is the residual difference in wages that exist among individuals after differences in, education, age, investment in human capital and preferences for market vs. non-market work etc. Naturally, if this residual is non-existent, implying that the difference in wages between men and women results entirely from differences in human capital, then the anti-discrimination law should have no effect on the pay gap. Our expectation on the effect of the law will therefore depend on the size of this residual.

Taste-based models of discrimination, assumes that agents act discriminating on the basis of 'taste'. The idea is that members of majority groups hold certain tastes against interacting with members of minority groups. According to Collard and Becker (1972), these 'tastes' are founded on prejudices held by one group towards another group, and these prejudices drive differences of labour market outcomes between majority and minority groups within markets. This is exactly the kind of discrimination that the act prohibits. Again, the effect of the law could vary with respects to *extensive* margin, which would be changes in the number of people who are employed, and *intensive* margins, which would be effects, such as hours worked and wages, for people already employed. Granted that the cost of discriminating would be high enough that it 'trumps' males distaste against working with females the pay gap with respect to the intensive margin would increase. Meanwhile, the law could make women even more unemployable since the presumed distaste against female workers held by certain men is now complemented with a higher cost to hiring women. Thus, the theory seems to imply that, given the right circumstances, the gap on the intensive margin would decrease, i.e. female wages increase, whereas the gap would increase on the extensive margin, meaning that the participation of women on the labour market decreases.

The theory of occupational crowding, (Bergmann, 1974) argues that gender gap in labour market outcomes can be explained through concentration of group specific occupations, e.g. male dominated occupations such as software development in contrast to female dominated occupations such as caregiving. Pay gaps stemming from these phenomena will not be affected by the policy,

this because the law requires that the jobs between which the wage discrepancies unfairly occur should be comparable and similar. This comparison is hard to perform when the occupations and tasks may differ substantially between industries.

Theories of statistical discrimination, introduced and emphasized by Phelps (1972), and Arrow (1973) assumes that discrimination will persist even tough economic agents are rational and non-prejudiced. Discrimination will be based on the average behaviour of a group e.g. gender specific characteristics, average behaviour of specific ethnic groups etc. As women are often expected to bare children, they might get discriminated against since the probability that they will have an extended parental leave is higher than that of comparable men, which makes them a less attractive employee candidates. Unless equal pay would prevent women from being expected to take parental leave to a larger extent than men, which is a question that unfortunately goes beyond the scope of this thesis, this theory of discrimination would suggest that the law in question might not be as effective as called for.

Bertrand (2011) published her chapter in the *Handbook of Labor Economics* `New perspectives on Gender´, the author presents psychological and socio-psychological perspectives as explanations to differences between male and females on the labour market. These includes attitudes towards competition, attitudes towards negotiations, and the impact of gender identity on outcomes on the labour market. Legislation may induce a decline in the labour market outcomes that is corresponding to these aspects, e.g. attitudes towards negotiations if a company is not allowed to pay an employee differently based on gender, the importance of this aspect may diminish.

We expect the effectiveness of the EPA to depend on the general economic situation within the country, and the states. This is supported by the theory of *Added worker* that suggests that in times of economic unrest the labour market participation of women will increase, as married women apply for work when their unemployed husbands stay at home. This would imply that the extensive margin would increase after the 2009 amendment, as a result of the financial crash of 2008 (Mincer, 1962; Bredtmann et al., 2017).

According to *The theory of alterative cost*, the act of choosing is performed through different options. In the labour market this can be whether or whether not to apply for a job. If the alternative cost is too high, suggesting that it is preferable to stay at home rather than working, this is what the rational agent will do. Likewise, if the alternative cost of not working is too high, suggesting that the pay is too high to resist, the rational agent will apply for a job. This is something that has an impact on all agents on the labour market, though not as relevant for men since it is usually women with small children who have the additional choice of paying for childcare or staying

at home themselves. According to this theory, we may expect that the labour force participation of women will increase since the alternative cost to not working will change as women are paid more, making it more desirable to work. We may also expect the pay gap to decrease since there will be more female employees performing jobs that are comparable to jobs performed by male employees (Streissler, 2008).

In light of these theories we have varying expectations of the effect of the EPA. The law might decrease the gender pay gap while having mixed effects on the difference in the amount worked between men and women. This is primarily according to *the taste based models of discrimination* together with *the theory of alternative cost*. The assumption is that there is a gender pay gap resulting from taste-based discrimination. If committing to such discrimination becomes costly for employers, the unjustly lower female wages would be expected to rise. Naturally, this would decrease the gender pay gap. This could also, in line with the theory of alternative costs, give positive effects on the extensive margin. If female wages increase, their leisure time becomes more expensive which causes them to increase their labour supply. Alternatively, the law yielding higher female wages would increase the cost of hiring women, for whom employers have a distaste, which would make women less hireable. This would result in a negative impact on the female labour extensive margin.

There are also theories in this section which would lead us to expect little or no effect of the intervention on the Illinois gender pay gap. We suspect discrimination to be existent in the labour market rather than the pay gap resulting entirely from differences in human capital. Even so, if *statistical* discrimination, which will not be affected by the EPA, is more prevalent than *taste-based* discrimination, which might be affected, the law should not have a substantial impact on the gender pay gap.

4 Method

To test the impact of the Equal Pay Act (EPA) on the gender equality in the Illinois labour market we use data extract from the Current Population Survey (CPS). The data extract which we use is called 'March CPS' and is freely available at the Census Bureau's website and covers data from march to march each year (Center for Economic and Policy Research, 2016). We first use a Difference-in-Difference fixed effects model. To handle problems with parallel trend assumptions violation, we then create a synthetic control state, using weighted averages of a set of outcome predictors from various control states. To control the validity of the synthetic estimates for the extensive and intensive margin of Illinois we perform a placebo test. This is done by creating synthetic controls for each state in the donor pool and comparing these outcomes with their respective factual outcome. These differences in turn are compared to the difference between Illinois and synthetic Illinois.

4.1 Data

We use pooled cross sectional randomized household CPS data covering roughly 200 000 households per year across the united states for the years 2000-2016. We use a Difference-in-Difference (DiD) methodology with state and year fixed effects. In the process, we turn the cross-sectional individual data into panels on state level. We choose to drop observations for individuals below the age of 25 and above the age of 55, this is done to get the most appropriate strata to evaluate the gender gap, the rationale being that outliers, such as individuals who are either unemployed, or have wages significantly above the averages are more likely outside this spectrum. Further, we only evaluate data of individuals that state that they are employed. This is done to reduce the skewness of data, since there are more men than women in the workforce.

4.2 Difference-in-Difference

We use a Difference-in-Difference method for estimating the effect of the EPA on the gender gap on labour market outcomes. The notions of the extensive and intensive margins are important to keep in mind while evaluating the results in order to mitigate the risk of misinterpretation and over/under valuing the possible effects of the legislation. As noted above, it is possible that the pay gap decreases, which would be a change in the *intensive* margin, while at the same time, the share of women in the work force decrease, which would be a change in the *extensive* margin. This of course would not be a desired effect of the law.

In performing the DiD we compare our outcome variables in Illinois after the implementation of the law to the outcome variables for the rest of the USA, which did not implement the policy, the idea being that differences between Illinois and the rest of the USA are results of the effect of the EPA. For the DiD method to be informative on the effect of the policy change, it is important that the treatment group (workers in Illinois after the treatment) would have had a similar development of the gender pay gap absent the implementation of the law to the one the control states had. If this is not the case, there would have been differences in pay gap between Illinois and the rest of the U.S. regardless of the law-enactment, but then the difference (in the outcome variables) between the two groups cannot be attributed to the policy change. As we will see, it is uncertain whether this requirement, typically referred to as *the parallel trends assumption*, is met in our case. In order to make the assumption more likely to hold, a synthetic control method will be used. It will be explained in the next section.

For the DiD estimates, we use the following model with data collapsed by state, and fixed effects for state and year:

$$y_{st} = \propto + \gamma_t + \delta_s + \beta_1 post_s + \beta_2 Illinois_t + \beta_3 (Illinois * post)_{st} + X_{st} + \varepsilon_{st}$$

where the subscripts *s* indicates state and *t* indicates year. *Post* is a dummy variable assigned the value 1 if state *s* is observed in the year after the policy and 0 if the state is observed prior to the policy change. Similarly, we set *Illinois* to be a dummy variable which takes the value 1 if state *s* is the state of Illinois, and 0 otherwise. We generate our variable of interest, *Illinois*post*, which is an interaction dummy for the treatment group, equal to 1 if state *s* is Illinois after the policy change, and 0 otherwise. The idea is that if the parallel trend assumption holds, the coefficient for this variable will capture the effect of the policy change[1]. Finally, we include a group of control variables, represented by the vector X, which consists of state median values for the variables *age* and *education level*[2], as well as the mean values for the racial indicators[3], dummy variable that indicates household members under 18 *fchild* and a control variable, *married,* for whether the individual is married or not. The control variables age and education, and the outcome variables are given in medians to make the results less influenced by outliers. We use means for the control dummies as this gives the percentage of the population in state *s* at year *t* with the property denoted by each dummy. Lastly, we use two different outcome variables (y_{st}). The first is the natural logarithm of the

[1] See appendix A. for full explanation.
[2] E*ducation level* is a categorical variable taking the value 1 (No high school), 2 (High school), 3 (Some college), 4 (College), or, 5 (Advanced).
[3] The racial dummies are Black, White, Asian, Hispanic and Other.

difference between median hourly wage for men and women representing the gender pay gap (*Hr_diff*). The second outcome variable is the share of total hours per week worked by females divided by the total hours worked per week by both sexes combined, *work_share*. The use of these outcome variables enables us to sort out changes in the extensive and the intensive margins. They are defined a bit more formally as:

$$Gender\ pay\ gap = Log(wage\ if\ male) - Log(wage\ if\ female)$$

$$Work\ share = \frac{hours\ worked\ by\ female\ employees}{hours\ worked\ by\ female\ employees + hours\ worked\ by\ female\ employees}$$

For the estimates of our variables of interest to capture the effects of the policy, the parallel trends assumption must hold.

4.3 Synthetic control model

As mentioned, the parallel trend assumption is a crucial element for any DiD estimation. For the assumption to hold, the outcomes for the treated and control groups must be parallel to each other. This would imply that '…without the intervention, outcomes for the treated and control groups would have followed parallel trajectories over time.' (Kreif et al., 2016). This means that the treated and control states are allowed to exert different absolute values in outcomes, but only so long as the change in these values are constant *relative to each other* over time.

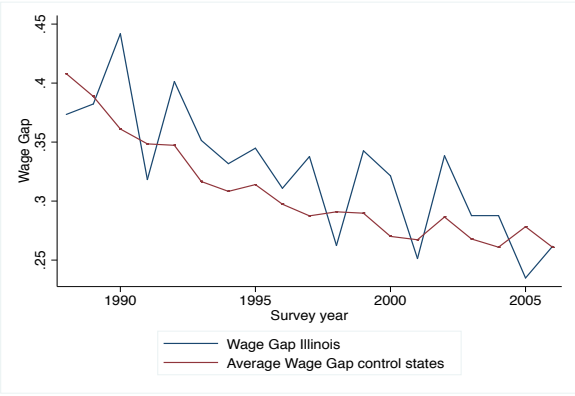


Figure 1 Time trends of female labour market participation for Illinois (blue line) and average for the control states (pink line).

Figure 2 Time trends Wage Gap between female and male employees in Illinois (blue line) and the average for the control states (pink line).

In figures 1 and 2, we present mean values of our outcome variables for Illinois compared to the mean values of the comparison group prior to the intervention. The reasoning goes as follows: trends that were parallel before the intervention would have stayed parallel absent the intervention, and, conversely, it is unlikely that non-parallel pre-trends would have been parallel after the intervention (absent the intervention). The treated and comparison units in the graphs seem to be trending in the same direction, but to claim that they are parallel might be too strong. This does not mean that it is impossible for the parallel trends assumption to holds in our DiD, but it seems questionable. This is where the Synthetic control method, which is designed to make the assumption more likely to hold, comes in to play. In what follows we will give a slightly more technical exposition of the synthetic control method.

Following Abadie and Gardeazabal (2003), we create a control group for which the parallel trend assumption is more likely to hold by combining weighted averages of our predictors for the states in the USA which best mimic the values of our outcome predictors. The states from which these values can be taken will be referred to as the donor pool. This gives us our 'Synthetic control state'. The weight W is assigned so as to minimize the expression:

$$\sum_{m=1}^{k} v_m (X_{1m} - X_{0m}W)^2 \quad \text{(Abadie et al., 2015, p.497)}$$

Intuitively, the W, is the weight that minimizes the difference between a given predictor in the treated state (Illinois) and said predictor in the synthetic state of Illinois. X_{1m} then, represents predictor m at state 1 (which in our case would be Illinois) and X_{0m} represents the values of the predictor m from the states in the donor pool. The W is the weight and it is assigned as to minimize the expression.[4] Intuitively we want the difference prior to the intervention between Illinois and synthetic Illinois to be as small possible, this very naturally translates into the synthetic state being given by the W (that sums to 1) that minimizes the above expression.

Next the estimated effect following Abadie et al. (2015, p.498), much like in the DiD case, is given by:

$$Y_{1t} - \Sigma_{j=2}^{J+1} w_j^* Y_{jt} \text{ (ibid, p.498)}$$

[4] For illustrative purpose we ignore the 'v' in the equation, which is another weight used for ranking the importance of predictor m relative to the other predictors, for now. We will come back to this operator in the discussion.

Now, Y_{1t} is the post intervention value for the outcome variable at time t. In our case, this would be Illinois after the implementation of the EPA. The subtrahend is the sum of all the donor states, j=2 through J+1, multiplied with each state´s respective weight w_j^*, where Y_{jt} is the output level of donor state j at time t. Again, in our case this would be, for example (output for Wisconsin * 0.03) + (output for Idaho * 0.005)+…+(output J+1 * weight for state J+1), where the sum of all the weights equals to one. So, Synthetic Illinois is given by the sum of each state j weighted by $w_j^* \in [0,1]$, where J+1 is the total number of possible donor states, j=1 is the treated state which is why j=2 is the first donor pool state. Intuitively, this expresses the difference in post intervention outcome values between Illinois(treatment) and synthetic Illinois (which is, of course, untreated) i.e. the ATT (average treatment effect on treated).

We have now seen how the synthetic state provides a reasonable comparison state for which problems with parallel trend assumptions are resolved. This does not mean that running a simple Synthetic Control Method (SCM) gives completely unbiased estimates. There are certain restraints and problems to this method which we will hold off with until the discussion part of the paper.

Using the technique described above we generate the synthetic state of Illinois. Table 1 shows the mean values of the predictors for actual Illinois (Treated), counterfactual Illinois (Synthetic) and the control states (Whole donor pool). The table demonstrates how well the synthetic control state compares to taking the average of all the states combined, which is what the DiD does. From table 1 we see that for the first set of predictors, the synthetic state is similar to both Illinois as well as the whole donor pool. This is not the case for difference in wages for each respective year. Here, the synthetic state is more similar to actual Illinois than the rest of the control states.

Predictors	Treated	Synthetic	Whole donor pool
Age	39.75	40.2229	40.3727
Education	3	2.9962	2.9984
Black	.1304	.0682	.099
Asian	.0332	.019	.0401
Other	.0035	.0067	.0191
Hispanic	.1578	.0602	.0961
Married	.6623	.7181	.6874
Fchild[5]	.5546	.5793	.5739
Wage gap 2000	.3214	.3218	.2720
Wage gap 2001	.2513	.2515	.2709
Wage gap 2002	.3385	.3387	.2872
Wage gap 2003	.2877	.2880	.2679

Table 1 shows the values of the predictors for the Treated group (factual value) the Synthetic group (estimated values of counterfactual) and the values for the whole donor pool (mean values of the complete group of control states).

It is generally not a good idea to include every available control group in the donor pool, and there are several reasons for why this is the case. First, as the weights are only set as to generate the best fit in state averages of predictor variables between treated state and synthetic state, states with a very small population may give a disproportionally large contribution to the synthetic state. Second, states that are in general substantially different, when it comes to things that would affect the impact of the policy change, e.g. culture and GDP, should be omitted. This is because, naturally, we are not interested in what would happen in states where the policy effect is due to factors that are not prevalent in Illinois. However, this issue will probably not cause any major concerns here since our states are all in the same country after all. Lastly, one should discard states that recently underwent similar policy changes. For these reasons, we omit Alabama, California, Maine and New Jersey[5]. The resulting control state (which we will henceforth be referred to as synthetic Illinois is a weighted average of predictors for which the trend in outcome variables prior to the EPA implementation matches that of Illinois. This is a way of creating an observable counterfactual control unit, which can then be compared to the actual values of the outcome variables in Illinois after the policy change. The difference between these two will be interpreted as the effect of the policy change. For illustration purposes, we present an example of the results below (Abadie et al., 2010; Abadie et al., 2015).

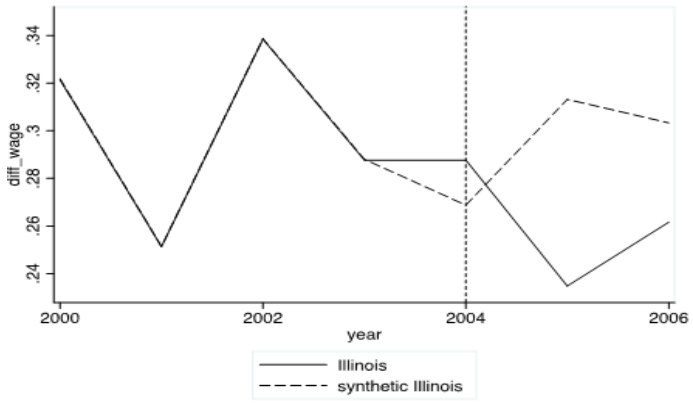


Figure 3. Difference in median hourly wages between men and women in Illinois(factual) and synthetic Illinois (counterfactual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while the trajectories go in different direction post implementation.

[5] Alabama because of its disproportionally substantial wage gap. California, New Jersey and Maine because of similar already implemented legislations on gender equality.

4.4 Placebo-test

In order to test the robustness of the results, as well as calculating the significance level of the effects, estimated by the synthetic control model, of the policy change, we perform a placebo test as developed by . We treat every single control state as if it was subject to the law enactment and create synthetic control states for all the states, just as we previously did for Illinois. We then compare the outcome for each actual state to its respective synthetic state. This is done by calculating the root mean square prediction error (RMSPE) post/pre-EPA ratio for each state. The RSMPE measures how well the trend in a variable of a synthetic state matches the trend of that variable in the actual state. We will need two RSMPE figures for each state, one pre and one post, the ratio will be the post RSMPE divided by the pre RSMPE. A large ratio means that the gap between synthetic and actual outcomes are large, indicating that the treatment had effect. But for the placebo states there were no treatment, thus unless something unexplained by our model happened during the treatment year, we should expect low RMSPE ratios for the placebos states.

To get the p-value, we divide the number of states for which the RMSPE ratios are equal to or greater than that of Illinois, with the number of states in the donor pool plus the treated state. This will give the probability that a randomly chosen state presents an effect at least as high as the one Illinois presents. As usual, if this probability is higher than some prespecified significance level, the estimated effect can no longer be attributed to the implementation of the equal pay act.

5. Results

5. 1 Difference-in-Difference

The DiD of hour wage suggests that the implementation of the 2003 law had a positive effect on the gender pay gap. The DiD of the share of hours worked by women compared to men indicates that there was a negative effect in Illinois during 2004 and 2005 (see Table 2.)

Dependent variable:	(1) Wage Gap	(2) Wage Gap	(3) Work Share	(4) Wage Gap	(5) Wage Gap	(6) Work Share	(7) Wage Gap	(8) Wage Gap	(9) Work Share
Post	-0.0149**	-0.0164	-0.00138	0.00428	-0.00284	-0.00102	-0.00999	-0.0121	-0.00717
	(0.00676)	(0.0215)	(0.00489)	(0.00877)	(0.0209)	(0.00482)	(0.00861)	(0.0248)	(0.00551)
Post · Illinois	0.00285	0.00245	-0.00696***	-0.0692***	-0.0735***	-0.0142***	-0.0280***	-0.0427**	0.00354
	(0.00676)	(0.00747)	(0.00187)	(0.00877)	(0.0123)	(0.00283)	(0.00861)	(0.0165)	(0.00285)
Constant	0.274***	-0.121	0.401***	0.274***	-0.139	0.426***	0.274***	-0.107	0.399***
	(0.00132)	(0.319)	(0.0568)	(0.00172)	(0.302)	(0.0542)	(0.00168)	(0.326)	(0.0643)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year	2004	2004	2004	2005	2005	2005	2006	2006	2006
Observations	230	230	230	230	230	230	230	230	230
R-squared	0.019	0.111	0.165	0.009	0.073	0.177	0.010	0.108	0.133
Number of states	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2. Output DiD fixed effects 2004 through 2006. The outcome variables are Wage Gap, which is logged and can thereby be interpreted as a percentage indicating the change on the intensive margin, and Workshare, which gives an indication of the change on the extensive margin. All regressions are controlled for fixed effects with regards to year and state effects.

Table 2 DiD estimates for the 2003 intervention, year 2004 – 2006 separated by year, this means that the estimate shown in each column is the effect from the intervention for each year after the treatment. During 2004 there is no significant effect of the intervention on the gender pay gap, there is however a significant and negative effect with regards to work share. This means that given that all our assumptions are met the intervention had a negative impact of roughly 0.7% on the share of hours worked by women in Illinois post-treatment. The DiD-estimate for 2005 shows a significant and negative effect of the intervention on the gender pay gap by roughly 7.3% which is interpreted as the intervention having reduced the gender pay gap in Illinois by 7.3% compared to what it would have been absent of the intervention. There was a significant and negative effect of the intervention during 2005 by roughly 1.42%. During 2006 the DiD-estimate of the gender

pay gap shows a significant and negative coefficient of roughly 4.3%, there is no significant effect in terms of the work share.

Table 3 shows the DiD-estimates of the 2009 amendment. The DiD-estimates of the gender pay gap is negative and significant straight through. This suggests that the intervention had the anticipated effect on the gender pay gap. The DiD-estimate of the work share shows negative coefficients for 2010 and 2011 (not significant), and a positive and significant coefficient for year 2012.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable:	Wage Gap	Wage Gap	Workshare	Wage Gap	Wage Gap	Workshare	Wage Gap	Wage Gap	Workshare
Post	-0.00182	-0.00272	0.0164***	-0.0133*	-0.0165	0.0136***	-0.0219**	-0.0333*	0.00746**
	(0.00626)	(0.0108)	(0.00264)	(0.00681)	(0.0152)	(0.00310)	(0.00937)	(0.0180)	(0.00358)
Post · Illinois	-0.0276***	-0.0196*	-0.00231	-0.0693***	-0.0668***	-0.00176	-0.0356***	-0.0316**	0.0128***
	(0.00626)	(0.0103)	(0.00338)	(0.00681)	(0.0125)	(0.00400)	(0.00937)	(0.0146)	(0.00328)
Constant	0.264***	0.334	0.516***	0.264***	0.325	0.354***	0.264***	0.309	0.389***
	(0.00122)	(0.339)	(0.0983)	(0.00133)	(0.284)	(0.0808)	(0.00183)	(0.261)	(0.0911)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year	2010	2010	2010	2011	2011	2011	2012	2012	2012
Observations	230	230	230	230	230	230	230	230	230
R-squared	0.003	0.046	0.322	0.035	0.075	0.249	0.048	0.092	0.196
Number of state	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. Output DiD fixed effects 2010 through 2012 The outcome variables are the same as in table 2.

Table 4 Shows the DiD-estimates of the 2013 amendment. The DiD-estimate for year 2014 is significant and positive, suggesting that the gender pay gap increased by roughly 5.5% compared to the comparison group, that is, the estimated counterfactual gender pay gap. The work share DiD-estimate of 2014 shows a significant and negative effect of roughly 1.5% compared to the estimated counterfactual. During 2015 the DiD-estimate of the gender pay gap is significant and negative standing for roughly 6.1% compared to the estimated counterfactual. The work share DiD-estimate shows a significant and negative coefficient of roughly 1.2% compared to the estimated counterfactual. The work share DiD-estimate of 2016 shows a value of roughly negative 0.8%.

VARIABLES	(1) Wage Gap	(2) Wage Gap	(3) Workshare	(4) Wage Gap	(5) Wage Gap	(6) Workshare	(7) Wage Gap	(8) Wage Gap	(9) Workshare
Post	-0.0306*** (0.00908)	-0.0830*** (0.0190)	-0.00194 (0.00397)	-0.0416*** (0.00806)	-0.0731*** (0.0190)	0.00137 (0.00474)	-0.0232** (0.00914)	-0.0591*** (0.0181)	0.00497 (0.00495)
Post · Illinois	0.0412*** (0.00908)	0.0547*** (0.0120)	-0.0150*** (0.00281)	-0.0614*** (0.00806)	-0.0614*** (0.0159)	-0.0120*** (0.00323)	-0.00123 (0.00914)	-0.0157 (0.0140)	-0.00798** (0.00348)
Constant	0.251*** (0.00148)	0.862*** (0.235)	0.335*** (0.0767)	0.251*** (0.00131)	0.295 (0.270)	0.353*** (0.0868)	0.251*** (0.00149)	0.486 (0.311)	0.240** (0.0927)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year	2014	2014	2014	2015	2015	2015	2016	2016	2016
Observations	276	276	276	276	276	276	276	276	276
R-squared	0.061	0.228	0.137	0.125	0.231	0.144	0.037	0.204	0.159
Number of state	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Output DiD fixed effects 2014 through 2016 The outcome variables are the same as for table 2 and 3.

5.2 SCM Results

5.2.1 2003 SCM Results

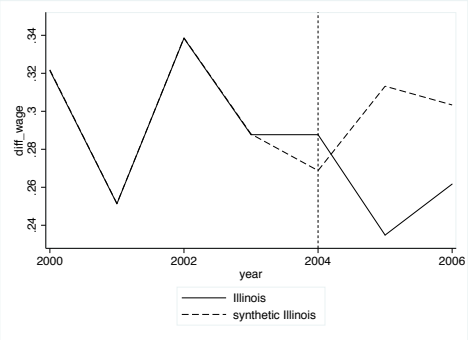


Figure 1. Synth estimation intensive margin 2003 EPA. Difference in median hourly wages between men and women in Illinois(factual) and synthetic Illinois(counter factual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while the trajectories go in different direction post implementation.

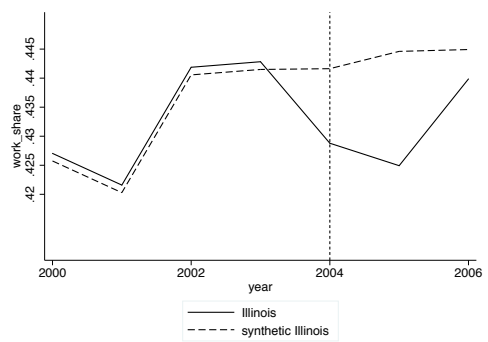


Figure 2. Synth estimation extensive margin 2003 EPA. Difference in median share of hours worked by female employees as a share of the total hours worked in the labour market in Illinois(factual) and synthetic Illinois(counter factual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while the trajectories go in different direction post implementation.

There is a clear difference with respect to the gender pay gap 'synthetic Illinois' compared to 'treated Illinois' the two lines follow the same trend before the treatment, after the treatment (2004) there is a clear split between the two lines, there is however a trend of convergence apparent from 2005 to 2006 (see fig. 2). There is an apparent gap between the work share between synthetic and treated Illinois, the synthetic control says that the work share would have continued on a steady level after the treatment, however the treated Illinois shows a clear drop in work share (see fig. 3).

5.2.2 SCM First Amendment Results

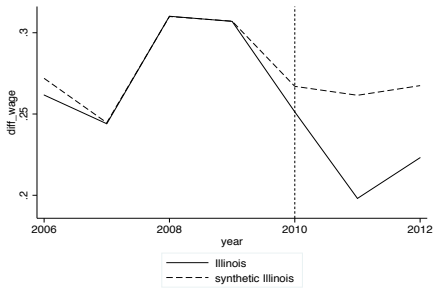


Figure 4. Synth estimation intensive margin 2009 amendment. Difference in median hourly wages between men and women in Illinois(factual) and synthetic Illinois (counterfactual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while there seems to be a difference between the trajectories post implementation

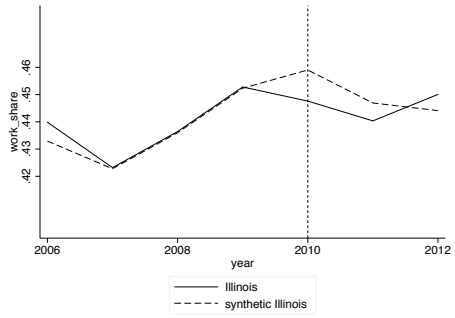


Figure 5. Synth estimation extensive margin 2009 amendment. Difference in median share of hours worked by female employees as a share of the total hours worked in the labour market in Illinois (factual) and synthetic Illinois (counterfactual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while the trajectories differ slightly around the amendment.

The two lines seem to be well fitted with respect to the gender pay gap, they follow the same trend pre-treatment. There is a gap between the two lines after the treatment, this suggest that the treatment had an effect on the gender pay gap (see fig. 4). The lines for the work share well fitted pre-treatment. There is a gap between the synthetic and the treated post-treatment, this suggests that the treatment had an effect on the work share, suggesting that female workers worked less hours compared to what they would have done in the counterfactual (see fig. 5).

5.2.3 Second Amendment Results

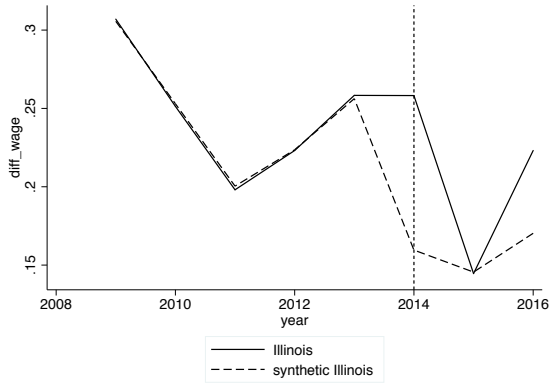


Figure 6 Synth estimation intensive margin 2013 amendment. Difference in median hourly wages between men and women in Illinois(factual) and synthetic Illinois (counterfactual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while there seems to be a difference between the trajectories post implementation.

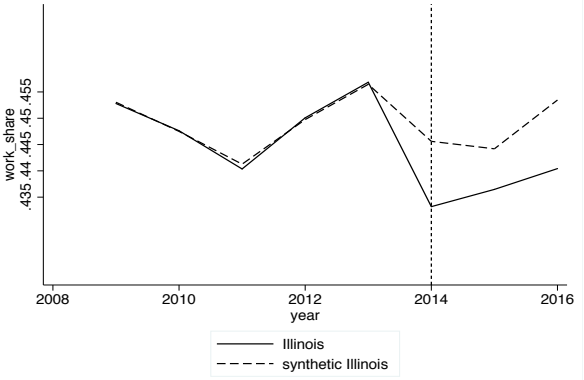


Figure 7 Synth estimation extensive margin 2013 amendment. Synth estimation extensive margin 2003 EPA. Difference in median share of hours worked by female employees as a share of the total hours worked in the labour market in Illinois(factual) and synthetic Illinois (counterfactual). The dotted vertical line indicates the time of the implantation of the EPA. The graph presents a good fit between actual and synthetic state prior to the intervention while the trajectories go in different direction post implementation.

The lines representing the synthetic and the treated are relatively well fitted pre-treatment. There is a gap between the synthetic and treated group around the treatment, the pay gap with respect to hour wage decline more for the synthetic than the treated, however it seems to be the same trend but lagged (see fig. 6). With regards to work share there is a gap between the two groups, there is a clear drop in treated Illinois compared to synthetic Illinois, suggesting that women worked less in the treated than in the counterfactual (see fig. 7).

We carried out a difference-in-difference estimation with fixed effects (see results, Table 2-4), these estimations indicate that the implementation of the law had a significant impact upon the gender pay gap in Illinois. As mentioned previously a crucial assumption for difference-in-difference estimations is the assumption of parallel trends. To amend the possible breach of this assumption, we decided to implement the synthetic control model as well. By observing the graphs (see fig. 2-3), one can see that there is a clear difference in the outcome for 'treated' and 'synthetic' Illinois post intervention. This seem to support the outcome from the difference-in-difference estimations, but the possibility that these outcomes are random and not due to the EPA still remains. This is controlled for through a placebo test (see fig. 8-9).

5.3 Placebo Results

5.3.1 2003 Placebo Results

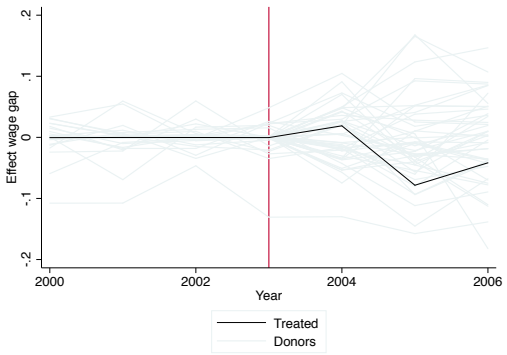


Figure 8 Placebo test; Gender pay gap 2000 – 2006. Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

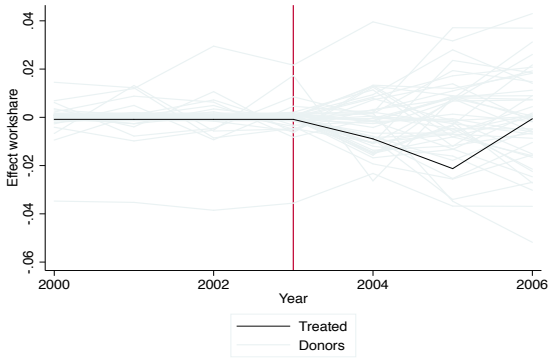


Figure 9 Placebo test; Workshare 2000 – 2006 Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

The black line represents the difference between treated Illinois and synthetic Illinois, the grey lines represent the difference between each donor state and its hypothetical treated, synthetic, state of world. The placebo test of the gender pay gap shows large variation (see fig. 8). There are several states, represented by the grey lines, that present larger negative effects than Illinois, which makes it hard to tell whether the Illinois effect is due to the EPA or just random. The placebo illustration for work share shows the same tendency (see fig. 9).

5.3.2 First Amendment Placebo Results

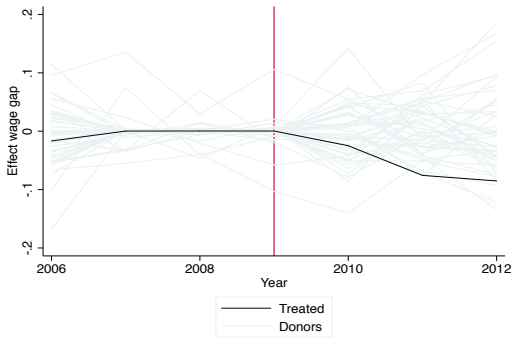


Figure 10 Placebo test; Gender pay gap 2006 – 2012 Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

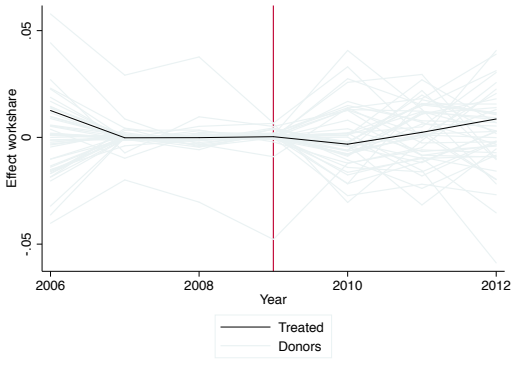


Figure 11 Placebo test; Work share 2006 – 2012 Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

The placebo illustration of the gender pay gap shows that Illinois is one of the states that differs most on the negative scale (see fig. 10), there is however a large variation that complicates the graphical analysis. The illustration of the placebo test of work share shows that Illinois follows the synthetic trend quite well (see fig. 11) indicating that the treatment did not have a large effect on the work share.

5.3.3 Second Amendment Results

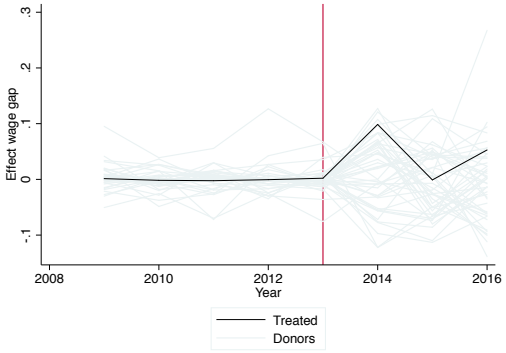


Figure 12 Placebo test; Gender pay gap 2009 – 2016 Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

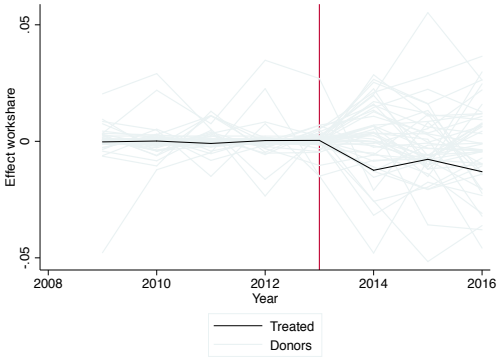


Figure 13 Placebo test; Work share 2009 – 2016 Dark line represents the difference between treated Illinois and synthetic Illinois and Grey lines represents the corresponding values for the donor states.

It is hard to tell the effects on Illinois apart from the placebos of the donor states, suggesting that the treatment effect, again, may have been random.

5.4 P-values from the placebo tests in comparison to DiD estimates

Wage gap	(1) Standard DiD ATT (p − value[a])	(2) Synthetic control ATT (p − value[b])
2004	.0024 (0.744)	.0189 (0.756)
2005	-.0735 (<0.01)	-.0783 (0.244)
2006	-0.427 (0.013)	-0.416 (0.533)
2010	-.0196 (0.065)	-.0249 (0.622)
2011	-.0668 (>0.01)	-.0757 (0.044)
2012	-.0316 (0.035)	-.0852 (0.2)
2014	.0547 (>0.01)	.0099 (0.133)
2015	-.0614 (>0.01)	-.001 (1)
2016	-.0157 (0.269)	.0528 (0.422)

Work share	(1) Standard DiD ATT (p − value[a])	(2) Synthetic control ATT (p − value[b])
2004	-0.007 (0.001)	-.0089 (0.511)
2005	-.0142 (<0.01)	-.0213 (0.2)
2006	.0035 (0.220)	-.0006 (0.933)
2010	-.0023 (0.498)	-.0032 (0.778)
2011	-.0018 (0.662)	.0024 (0.844)
2012	.0128 (>0.01)	.0086 (0.622)
2014	-.015 (>0.01)	-.0124 (0.467)
2015	-.012 (>0.01)	-.0077 (0.622)
2016	-.008 (0.026)	-.0130 (0.422)

Table 5 ATT-estimates on the gender wage gap (left) and female workshare (right) with respective p-values for corresponding estimate. P-values in parentheses. p-values[a] are from robust standard errors and p-values[b] from the placebo method.

Table 5 shows the estimated ATT for the extensive and intensive margin of each year through the simple DID (1) and for the synthetic control (2). Their respective p-values are presented within the parenthesis. The p-values for the synthetic control are based on the placebo-test, illustrating the probability that the estimate may be the result of chance. The p-values for the DiD are based on the standard errors of the model. The p-values indicate that, besides the gender gap coefficient for year 2011, we cannot reject that the EPA had no effect for any additional years through the synthetic control model. This suggests that we cannot say that the effect was incentivised by the treatment with respect to neither the intensive margin nor the extensive margin.

6 Discussion

We have attempted to estimate the effect of equal pay legislation on the gender pay gap in Illinois. Our simple DiD estimates suggest that the legislation and its amendments had effects on both the intensive margin (gender pay gap) and the extensive margin (work share of women compared to men). The estimated effect on the gender pay gap (i.e. difference in hourly wage between men and women) was negative, indicating that the legislation led to a decrease in the gender pay gap, which our theories predicted. It is also the desired effect of the law. Further, the overall DiD estimates showed a negative difference in the outcome variable work share, suggesting that legislation led to a lower female work force participation compared to pre-treatment.

In contrast to these findings, the placebo-tests on the SCM showed no significant impacts on either of our outcome variables. This is contradictory to some of our predictions, which calls for discussion of these results.

6.1 DiD

The results from the DiD suggest a negative effect of the legislation and its amendments on the intensive margin, whilst there are some positive and some negative effects on the extensive margin. Even though the possibility that the results do not capture the causal effect of the EPA, due to the possible failing of the parallel trends assumption, interpreting the results relative to our theoretical framework will still inform on interesting possible impacts of anti-discrimination laws in general.

The negative effect on the intensive margin, that is a decreasing gender pay gap, is in line with *the theory of taste based discrimination*. The reasoning is straightforward: if the act of underpaying females, for whom certain employees share a distaste, becomes costly, it seems natural that these employees would mitigate such underpayment, thereby decreasing the wage gap. Further the positive effect on the extensive margin is in line with the same theory together with the *theory of alternative cost.* If female workers get offered a higher wage, their leisure time gets more expensive. According to the theory, females will then increase their labour supply which yields positive effects on the extensive margin.

The negative effect on the extensive margin coupled with the positive effect on the intensive margin are particularly interesting. These highlight an important risk with anti-discrimination legislation, which might not be taken into consideration beforehand: equal pay legislation can have negative effects on the very same gender equality issues in the labour market they were intended to cure. Indeed, in line with the findings of Neumark and Stock (2001), as well as the findings of Donohue (2005), we expected that the higher cost of hiring women might reduce women´s

employability, and as we have seen, the DiD results are compatible with such an interpretation of the effect of the EPA.

6.2 Synthetic control & placebos

The first step of the synthetic control suggested that there was a difference between the factual outcome of Illinois and the counterfactual outcome of synthetic Illinois, seemingly confirming the findings of the DiD, that is, that legislation had impacts on both the intensive and extensive margins. These findings, then, are also consistent with the above explanation of what our theoretical framework predicted.

Now, the theories just discussed predicted that we would observe a negative effect on the intensive margin, and some suggested that there would be a negative effect on the extensive margin as well. In the theoretical framework section, some theories lead us to expect no effect of the law. In line with these, the placebo tests suggest that there were no significant results on the outcome variables. Though, from table 5 and 6, it *looks* as if the law affected the female intensive and extensive margin in the Illinois labour market, this might not actually be the case. The placebo test, carried out to control whether the results were reliable, indicated that the synthetic control method might have picked up on state level occurrences which were unexplained by our models, rather than any effect of the equal pay act.

There are several reasons for why this could be the case, besides the theories discussed in the above section simply being wrong, to be discussed here. As mentioned, we anticipated the possibility that there were little or no taste-based discrimination, but rather statistical discrimination, in the Illinois labour market. This would mean that employers discriminate based on actual, on average, labour force gender differences. Females are more likely to take time off from work to raise children for instance, which over time would make them less hireable and less experienced workers. Inequalities resulting from such discrimination would remain unaffected by the policy change.

Moreover, there are explanations for how there could be no effect of the law which are compatible with all the theories described in this paper. The first, and perhaps the simplest explanation is that, it is possible that employers did not believe that they would suffer any consequences for discriminating against women. This could be the case because it was not thought that the law would be enforced in practise. Laziness, fear of reappraisal, a general disbelief in the judicial system, and so on could prevent reports from being filed. Alternatively, it could be that the

law was not known enough to be considered by employers when setting wages. There is actually some support for the latter explanation. For instance, here it is referred to by the law firm Wessels Sherman as 'this little known law' (Joerg, 2013). These reasons would be compatible with all the theories described above as their predictive power, of course, stems from people taking the law seriously.

Another possible explanation is that even if the lack of impact of the intervention was not (entirely) due to lack of knowledge about, or belief in, the law, the absence of the effect could be explained by the current economic situation. This would be in line with the added worker theory, as described by Bredtmann et al (2014). The second amendment of the intervention, for instance, was implemented close to the financial crisis of 2008, which may have disturbed the outcomes as this is a strong exogenous shock that may have large impacts on both the intensive margin and the extensive margin. According to Bredtmann's theory of added worker, this would have a positive effect on the female extensive margin. This is because more women would apply for jobs when their husbands lose theirs, which naturally would result in an increase in the share of hours worked by women. Interestingly, we do observe a positive effect on female workshare on the year fixed effects in the vicinity preceding this crisis[6].

Lastly, there remains of course the possibility that the law had no effect on the pay gap simply because there was no discrimination at all. This is not likely to be the case though, (see for example Blau & Kahn 2016).

6.3 SCM revisited

We have seen that the promising results given by the DiD were contrasted by the placebo tests from the synthetic control method. Nonetheless, the synthetic control method itself is not without its flaws. This has already been hinted at when the method was first explained in the method-section. These issues will now be discussed a bit more thoroughly.

One potential threat to the method is that the similarity between treated unit and units in the donor pool might not be similar enough. The fundamental idea of the SCM is to construct a counterfactual state which accurately represents the value of the outcome variable as it would have been if the intervention under investigation never happened. Abadie et al. (2015, p. 498) mention that this method is vulnerable particularly when it comes to… 'the presence of unmeasured factors affecting the outcome variable as well as by heterogeneity in the effects of observed and unobserved factors.'. The general response to this kind of critique is that, if the pre-treatment

[6] See appendix C.

period is long enough, unobservable factors that might interfere with the results should cancel out, otherwise the matching between actual and synthetic pre-trends would not fit. If they do, that is evidence that the problem is taken care of (ibid). Furthermore, Pfeifer et al. (2015) pointed out that SCM´s in which the outcome variable, for every year preceding the treatment, is included (as has been standard procedure) the other predictors get dismissed in order to produce the best fit between actual and synthetic state. To the extent that one thinks these predictors are important when it comes to accurately forecasting the dependent variable, disregarding these predictors is of course problematic. In fact, Pfeifer et al (ibid) show that, as a result of this, the achieved estimates might be biased. To address these issues, we run a new SCM where we extend our pre-treatment and exclude the outcome variables for all years except the last one before the intervention. Indeed, when this is done, the fit becomes slightly worse[7]. This indicates that there could be problems with unobservable factors, and/or relevant predictors not contributing to the synthetic state. Nonetheless, our results were already insignificant, and they remain so after these issues are considered. So, although it is likely that we have not constructed the optimal comparison group required for validating the results of the SMC, accounting for these problems does not significantly change our findings.

6.4 External validity

The data set used for this research is to be considered large, and constitutes a well randomized cross-sectional sample. This suggests that it is plausible that the findings of this research on our data-sample to be generalized to the labour market in Illinois. Nevertheless, the Illinois EPA is specific for Illinois. This specificity makes it unreliable to generalize the results of this evaluation beyond the U.S.A., and maybe even beyond Illinois. This means that it is not possible to draw conclusions as to whether a similar EPA in another state or country would have a similar impact or not, unless they are substantially similar to Illinois. The heterogeneous nature of states and countries weakens the external validity of this research.

[7] See appendix B. for an example of this

7 Conclusion

The aim of the present research, was to examine whether legislation was effective or not in the task of closing the gender gap. The study used the 2003 Equal Pay Act of Illinois as study subject, as this was found to be an appropriate experimental setting for a natural experiment.

Our research found conflicting results from the different methods used. For defenders of the DiD method, it might look as though the legislation does have effects on both the extensive as well as the intensive margin of labour market outcomes for women. These effects are in line with the theories of taste-based discrimination together with the theory of alternative costs, and the added worker effect, but only if the parallel assumption holds. Proponents of The SCM on the other hand, would view the effect of the policy in a more pessimistic manner, since these results came out insignificant. Regardless of the results, or lack thereof of this study, there is room for further research of the appropriateness of gender equality legislation as a policy tool to close the gender gap. First, due to scope restrictions, we have only used the RSMPE-method for p-values, and one particular form of placebo test, but there are many different ways of calculating p-values using SCM as well as other robustness checks[8]. Carrying out such tests would help evaluation of the EPA further. Second, research on the effects that equal pay legislation has on psychological and psycho-sociological attitudes which may affect labour market outcomes would help properly investigating the success of such laws. Lastly, taking the insignificant findings of the SCM carried out in this paper at face value could have multiple implications. Either it could be taken as a reason for investigating other means for reaching the desired gender equality in the labour market, such as gender quotas or gender specific subsidies. Also, as mentioned, these results may be specific for Illinois, in which case the effects of implementing similar laws in other areas should be properly examined before ruled out.

[8] See for instance, the leave-one-out robustness check from Albadie et al., (2015), and DiD combined with matching (Kreifer et al., 2015)

Bibliography

Abadie, A., Diamond, A. & Hainmueller, J. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association,* 105**,** 493-505.

Abadie, A., Diamond, A. & Hainmueller, J. 2015. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science,* 59**,** 495-510.

Abadie, A. & Gardeazabal, J. 2003. The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review,* 93**,** 113-132

Altonji, J. G. & Blank, R. M. 1999. Chapter 48 Race and gender in the labor market. *Handbook of Labor Economics.* Elsevier.

Arrow, K. 1973. The theory of discrimination. *Discrimination in labor markets,* 3**,** 3-33.

Becker, G. S. 1964. *Human capital; a theoretical and empirical analysis, with special reference to education*, New York : National Bureau of Economic Research; distributed by Columbia University Press, 1964.

Bergmann, B. 1974. Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex. *Eastern Economic Journal,* 1**,** 103-110.

Bertrand, M. 2011. Chapter 17 - New Perspectives on Gender. *In:* CARD, D. & ASHENFELTER, O. (eds.) *Handbook of Labor Economics.* Elsevier.

Blau, F. D. & Kahn, L. M. 1996. Wage Structure and Gender Earnings Differentials: An International Comparison. *Economica,* 63**,** S29-S62.

Bredtmann, J., Otten, S. & Rulff, C. 2017. Husband's Unemployment and Wife's Labor Supply: The Added Worker Effect across Europe. *ILR Review.*

Center for Economic and Policy Research 2016. March CPS Uniform Extracts, Version 1.0. Washington, DC.

Collard, D. & Becker, G. S. 1972. The Economics of Discrimination. *The Economic Journal,* 82**,** 788.

Darity, W. A. & Mason, P. L. 1998. Evidence on Discrimination in Employment: Codes of Color, Codes of Gender. *The Journal of Economic Perspectives,* 12**,** 63-90.

Donohue, J. J., Iii 2005. The Law and Economics of Antidiscrimination Law. *National Bureau of Economic Research Working Paper Series,* No. 11631.

Halpern, J. S. 2015. *Amendments to Illinois Wage Payment and Collection Act Regulations: Big Changes with Little Notice* [Online]. Available: https://www.muchshelist.com/insights/article/amendments-illinois-wage-payment-and-collection-act-regulations-big-changes-little [Accessed May 5 2018].

Ilcs 2003. Equal Pay Act of 2003. *In:* ILCS (ed.) *820 112/1.* Illinois General Assembly

Illinois General Assembly. 2003. Equal Pay Act of 2003. Available: http://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=2501&ChapterID=68 [Accessed April 27, 2018].

Joerg, E. N. 2013. What Does an Illinois Employer Need to Understand About the Illinois Equal Pay Act? Available: https://www.wesselssherman.com/Articles/What-Does-an-Illinois-Employer-Need-to-Understand-About-the-Illinois-Equal-Pay-Act.shtml [Accessed May 5, 2018].

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S. & Sutton, M. 2016. Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units. *Health Econ,* 25**,** 1514-1528.

Mincer, J. 1962. Labor Force Participation of Married Women: A Study of Labor Supply. *Aspects of Labor Economics*. National Bureau of Economic Research, Inc.

Mungerson, M. D. 2009. Illinois Amends State Equal Pay Act to Extend Filing Periods and Adopt Ledbetter Paycheck Standard. Available: https://www.littler.com/files/press/pdf/2009_09_Midwest_ILAmends_StateEqualPayAct_LedbetterPaycheckStandard.pdf [Accessed 2018, April 27].

National Conference of State Legislation. 2014. *2014 EQUAL PAY LEGISLATION* [Online]. Available: http://www.ncsl.org/research/labor-and-employment/2014-equal-pay-legislation.aspx [Accessed May 16 2017].

National Equal Pay Enforcement Task Force 2013. "Fifty years after the equal pay act". *The White House*.

Neumark, D. & Stock, W. 2001. The Effects of Race and Sex Discrimination Laws. *NBER Working Paper Series***,** 8215.

Pfeifer, G., Klößner, S., Kaul, A. & Schieler, M. 2015. *Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates*.

Phelps, E. S. 1972. The Statistical Theory of Racism and Sexism. *The American Economic Review,* 62**,** 659-661.

Streissler, E. 2008. *Wieser, Friedrich Freiherr, (Baron) von (1851–1926)*.

Waldfogel, J. 1998. Understanding the "Family Gap" in Pay for Women with Children. *The Journal of Economic Perspectives,* 12**,** 137-156.

Wilkins, E. R. 2012. Equal Pay Act of 1963. *In:* ROTHWELL, W. J. & PRESCOTT, R. K. (eds.) *Encyclopedia of Human Resource Management : Key Topics and Issues*. New York, UNITED STATES: Center for Creative Leadership.

Appendix

A. Derivation of the interaction term as the causal effect of the policy change:

Treated state post intervention:

$$y_{11} = \propto + \gamma_t + \delta_s + \beta_1 post_s + \beta_2 Illinois_t + \beta_3 (Illinois * post)_{st} + X_{st} + \varepsilon_{st}$$

Control state post intervention:

$$y_{01} = \propto + \gamma_t + \delta_s + \beta_1 post_s + X_{st} + \varepsilon_{st}$$

Treated state pre intervention

$$y_{10} = \propto + \gamma_t + \delta_s + \beta_2 Illinois_t + X_{st} + \varepsilon_{st}$$

Control state pre intervention

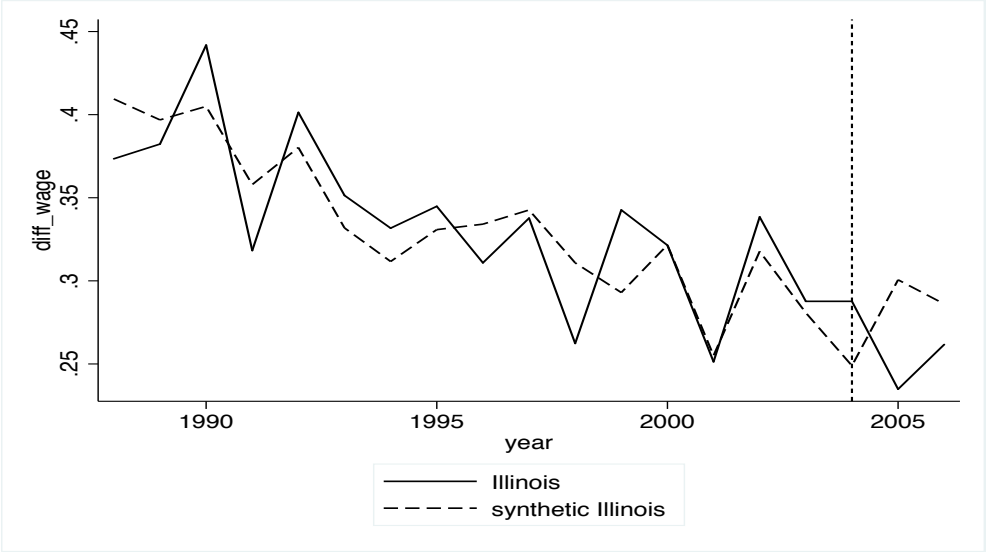
$$y_{00} = \propto + \gamma_t + \delta_s + X_{st} + \varepsilon_{st}$$

Now, we take the difference between treated and control after the intervention minus the difference between treated and control state before intervention

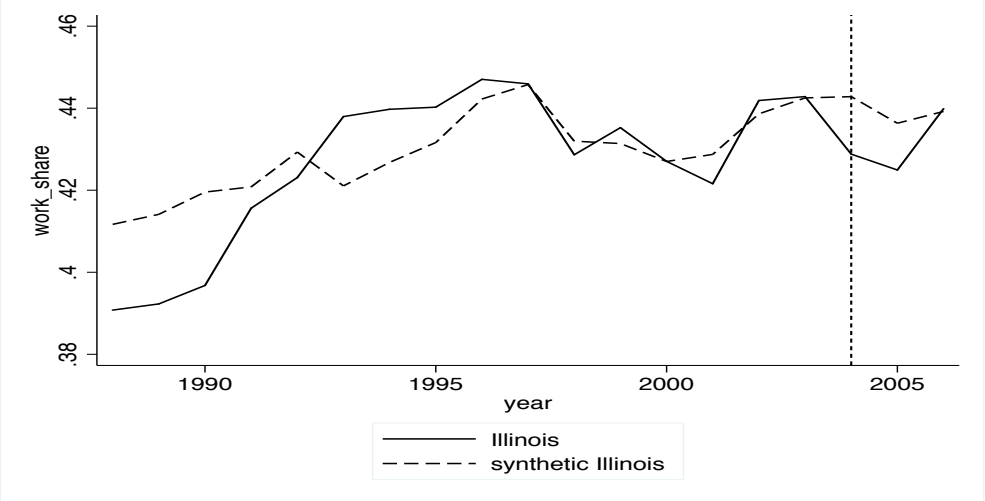
$$(y_{11} - y_{01}) - (y_{10} - y_{00}) = \beta_3 (Illinois * post)_{st}$$

B. Graphs of extended pre-treatment SCM

B.1 SCM Gender Pay Gap



B.2 SCM Female labour market participation



C. DiD Results

C.1 First implementation of the EPA

VARIABLES	(1) WageGap	(2) WageGap	(3) Workshare	(4) WageGap	(5) WageGap	(6) Workshare	(7) WageGap	(8) WageGap	(9) Workshare
Post	-0.0149**	-0.0164	-0.00138	0.00428	-0.00284	-0.00102	-0.00999	-0.0121	-0.00717
	(0.00676)	(0.0215)	(0.00489)	(0.00877)	(0.0209)	(0.00482)	(0.00861)	(0.0248)	(0.00551)
Post · Illinois	0.00285	0.00245	-0.00696***	-0.0692***	-0.0735***	-0.0142***	-0.0280***	-0.0427**	0.00354
	(0.00676)	(0.00747)	(0.00187)	(0.00877)	(0.0123)	(0.00283)	(0.00861)	(0.0165)	(0.00285)
Education		0.0191	0.00545		0.00835	0.00557		0.0360*	0.0125
		(0.0323)	(0.00526)		(0.0217)	(0.00388)		(0.0202)	(0.00841)
Age		0.00461	0.00246**		0.00370	0.00141		0.00179	0.00167
		(0.00694)	(0.00109)		(0.00697)	(0.00122)		(0.00668)	(0.00137)
Married		0.105	-0.175***		0.221	-0.165***		0.198	-0.179***
		(0.202)	(0.0452)		(0.197)	(0.0496)		(0.211)	(0.0455)
Fchild		0.161	0.0849**		0.117	0.0963**		0.156	0.123***
		(0.207)	(0.0362)		(0.190)	(0.0430)		(0.240)	(0.0379)
Black		-0.00372	0.0295		0.369	0.0930*		0.000474	-0.00869
		(0.204)	(0.0435)		(0.300)	(0.0543)		(0.247)	(0.0611)
Asian		-0.288	0.167*		-0.281	0.126		-0.00171	0.176*
		(0.287)	(0.0971)		(0.401)	(0.125)		(0.428)	(0.0997)
Hispanic		0.276	-0.138**		0.190	-0.137*		0.0639	-0.142**
		(0.213)	(0.0622)		(0.297)	(0.0726)		(0.279)	(0.0616)
Other		-1.364***	-0.114		-0.872*	-0.207*		-1.423***	-0.169*
		(0.358)	(0.129)		(0.463)	(0.113)		(0.302)	(0.0935)
2001 FE		-0.00353	0.000360		-0.00138	0.00115		-0.00218	0.000610
		(0.0125)	(0.00233)		(0.0121)	(0.00241)		(0.0122)	(0.00248)
2002 FE		0.00578	-0.00392		-0.000233	-0.00515		-0.000582	-0.00673
		(0.0234)	(0.00435)		(0.0212)	(0.00464)		(0.0272)	(0.00472)
2003 FE		-0.00716	-0.00209		-0.0157	-0.00242		-0.0107	-0.00419
		(0.0215)	(0.00425)		(0.0201)	(0.00434)		(0.0259)	(0.00480)
2004 FE		-	-						
2005 FE					-	-			
2006 FE								-	-
Constant	0.274***	-0.121	0.401***	0.274***	-0.139	0.426***	0.274***	-0.107	0.399***
	(0.00132)	(0.319)	(0.0568)	(0.00172)	(0.302)	(0.0542)	(0.00168)	(0.326)	(0.0643)
Year	2004	2004	2004	2005	2005	2005	2006	2006	2006
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	230	230	230	230	230	230	230	230	230
R-squared	0.019	0.111	0.165	0.009	0.073	0.177	0.010	0.108	0.133
Number of state	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

C.2 First Amendment of the EPA

VARIABLES	(1) WageGap	(2) WageGap	(3) Workshare	(4) WageGap	(5) WageGap	(6) Workshare	(7) WageGap	(8) WageGap	(9) Workshare
Post	-0.00182	-0.00272	0.0164***	-0.0133*	-0.0165	0.0136***	-0.0219**	-0.0333*	0.00746**
	(0.00626)	(0.0108)	(0.00264)	(0.00681)	(0.0152)	(0.00310)	(0.00937)	(0.0180)	(0.00358)
Post · Illinois	-0.0276***	-0.0196*	-0.00231	-0.0693***	-0.0668***	-0.00176	-0.0356***	-0.0316**	0.0128***
	(0.00626)	(0.0103)	(0.00338)	(0.00681)	(0.0125)	(0.00400)	(0.00937)	(0.0146)	(0.00328)
Education		-0.0284	-0.0113		-0.0262	-0.00486		0.000328	-0.00801
		(0.0188)	(0.0120)		(0.0249)	(0.00888)		(0.0305)	(0.0122)
Age		-0.00337	-0.000757		-0.00432	0.000664		-0.00225	0.000753
		(0.00543)	(0.00144)		(0.00499)	(0.00144)		(0.00609)	(0.00148)
Married		0.298	-0.0536		0.295	0.0346		0.116	0.00496
		(0.229)	(0.0555)		(0.223)	(0.0701)		(0.219)	(0.0596)
Fchild		-0.148	0.0521		-0.0995	0.0802		-0.0917	0.0513
		(0.175)	(0.0634)		(0.193)	(0.0523)		(0.183)	(0.0503)
Black		0.00287	0.0366		0.229	0.0427		-0.0146	0.0754
		(0.215)	(0.0801)		(0.230)	(0.0843)		(0.221)	(0.0771)
Asian		0.540	0.0287		0.503	0.130		0.992	0.277**
		(0.681)	(0.139)		(0.663)	(0.146)		(0.766)	(0.114)
Hispanic		0.0438	-0.160*		0.0692	-0.0962		-0.117	-0.109
		(0.268)	(0.0863)		(0.271)	(0.0757)		(0.369)	(0.115)
Other		0.306	0.112		-0.155	-0.0959		-0.465	-0.0930
		(0.551)	(0.176)		(0.823)	(0.196)		(1.157)	(0.178)
2007 FE		-0.000487	0.00243		-0.00157	0.00168		-0.00215	0.00139
		(0.0103)	(0.00200)		(0.0105)	(0.00223)		(0.0110)	(0.00207)
2008 FE		-0.00673	0.00542**		-0.00761	0.00551**		-0.0111	0.00451*
		(0.00863)	(0.00232)		(0.00891)	(0.00236)		(0.00938)	(0.00242)
2009 FE		0.00749	0.0128***		0.00734	0.0120***		0.00282	0.0112***
		(0.0100)	(0.00251)		(0.0101)	(0.00254)		(0.0108)	(0.00251)
2010 FE		-	-						
2011 FE					-	-			
2012 FE								-	-
Constant	0.264***	0.334	0.516***	0.264***	0.325	0.354***	0.264***	0.309	0.389***
	(0.00122)	(0.339)	(0.0983)	(0.00133)	(0.284)	(0.0808)	(0.00183)	(0.261)	(0.0911)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year	2010	2010	2010	2011	2011	2011	2012	2012	2012
Observations	230	230	230	230	230	230	230	230	230
R-squared	0.003	0.046	0.322	0.035	0.075	0.249	0.048	0.092	0.196
Number of state	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

C.3 Second Amendment of the EPA

VARIABLES	(1) WageGap	(2) WageGap	(3) Workshare	(4) WageGap	(5) WageGap	(6) Workshare	(7) WageGap	(8) WageGap	(9) Workshare
Post	-0.0306***	-0.0830***	-0.00194	-0.0416***	-0.0731***	0.00137	-0.0232**	-0.0591***	0.00497
	(0.00908)	(0.0190)	(0.00397)	(0.00806)	(0.0190)	(0.00474)	(0.00914)	(0.0181)	(0.00495)
Post · Illinois	0.0412***	0.0547***	-0.0150***	-0.0614***	-0.0614***	-0.0120***	-0.00123	-0.0157	-0.00798**
	(0.00908)	(0.0120)	(0.00281)	(0.00806)	(0.0159)	(0.00323)	(0.00914)	(0.0140)	(0.00348)
Education		0.0150	0.0110***		0.0423***	-0.00178		0.0350***	0.0114***
		(0.0113)	(0.00223)		(0.0109)	(0.00928)		(0.00899)	(0.00173)
Age		-0.00266	0.000717		-9.85e-05	0.000539		0.000667	0.00244
		(0.00653)	(0.00140)		(0.00599)	(0.00157)		(0.00666)	(0.00159)
Married		-0.776***	0.0386		-0.238	0.0835		-0.508**	0.0147
		(0.253)	(0.0536)		(0.217)	(0.0498)		(0.213)	(0.0555)
Fchild		0.106	0.0409		0.0589	0.0402		0.116	0.0947*
		(0.161)	(0.0453)		(0.150)	(0.0539)		(0.163)	(0.0550)
Black		-0.347	-0.0388		-0.0425	-0.0416		-0.196	0.000576
		(0.286)	(0.0562)		(0.272)	(0.0484)		(0.303)	(0.0715)
Asian		0.432	0.0825		0.371	0.0478		0.294	0.102
		(0.439)	(0.117)		(0.459)	(0.113)		(0.437)	(0.118)
Hispanic		-0.107	-0.00735		-0.192	-0.0266		-0.285	-0.0223
		(0.261)	(0.0652)		(0.313)	(0.0746)		(0.273)	(0.0753)
Other		-1.580***	0.0404		-0.639	0.132		-1.504***	0.242**
		(0.522)	(0.108)		(0.487)	(0.0961)		(0.389)	(0.106)
2010 FE		-0.0197**	0.00475**		-0.0125	0.00519**		-0.0145*	0.00502**
		(0.00859)	(0.00191)		(0.00777)	(0.00205)		(0.00811)	(0.00208)
2011 FE		-0.0395***	0.000895		-0.0278**	0.00205		-0.0310***	0.00198
		(0.0110)	(0.00278)		(0.0115)	(0.00297)		(0.0110)	(0.00278)
2012 FE		-0.0545***	-0.00209		-0.0378***	-0.000278		-0.0431***	-0.00107
		(0.0117)	(0.00362)		(0.0121)	(0.00382)		(0.0107)	(0.00347)
2013 FE		-0.0672***	-0.00592*		-0.0486***	-0.00389		-0.0549***	-0.00522
		(0.0150)	(0.00343)		(0.0141)	(0.00350)		(0.0135)	(0.00329)
2014 FE		-	-						
2015 FE					-	-			
2016 FE								-	-
Constant	0.251***	0.862***	0.335***	0.251***	0.295	0.353***	0.251***	0.486	0.240**
	(0.00148)	(0.235)	(0.0767)	(0.00131)	(0.270)	(0.0868)	(0.00149)	(0.311)	(0.0927)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year	2014	2014	2014	2015	2015	2015	2016	2016	2016
Observations	276	276	276	276	276	276	276	276	276
R-squared	0.061	0.228	0.137	0.125	0.231	0.144	0.037	0.204	0.159
Number of state	46	46	46	46	46	46	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1