The Effect of Universal Pre-kindergarten Policy on Female Labor Force Participation — A Synthetic Control Approach*

Hao Li†

Southern Methodist University

Abstract

A complete evaluation of universal pre-k requires the understanding of such programs on not just child welfare, but also maternal welfare. To that end, this paper examines the effect of universal pre-k on labor force participation of fertility age women in Oklahoma. I investigate the policy effect from both theoretical and empirical perspectives. The theoretical model suggests that the maternal labor force participation rate is likely to increase with a price reduction or a quality improvement of the universal pre-k programs, but the effect on mothers' working hours is ambiguous. I apply the synthetic control method (SCM) to the Current Population Survey (CPS) data to identify the causal relationship between universal pre-k and female labor outcomes. I find that universal pre-k increases labor supply of women aged 25 to 45 in Oklahoma. The empirical results also suggest that universal pre-k has heterogenous effects on subsamples stratified by education level, marital status, poverty status and number of children in the household.

Key words: Child Care, Universal Pre-k, Female Labor Force Participation, Synthetic Control

JEL: I20, J13, J21

*I am thankful to Daniel Millimet and Elira Kuka for their valuable advice and continued feedback. I am also grateful to Alberto Abadie, Elizabeth Cascio, Raj Deb and the participants in 15th Texas Econometics Camp for helpful comments.

†Hao Li, Department of Economics, Box 0496, Southern Methodist University, Dallas, TX, 75206-0496. Tel: (214)768-4337. Fax:(214)768-1821. E-mail: hhli@smu.edu.

1 Introduction

In an effort to increase the human capital development of children, universal pre-kindergarten (pre-k) has become an increasingly popular policy response. While a substantial literature shows that universal pre-k improves child outcomes (Gormley and Gayer, 2005; Berlinski et al., 2009; Drange and Telle, 2015; Chor, Andresen and Kalil, 2016; Herbst, 2017), there is little analysis of the policy's causal impacts on other household members. A well-known fact is that time-consuming home child care and costly day care are barriers to maternal labor force participation. Cohany and Sok (2007) show that the labor force participation rate of married women with young children declined during a time of increase in the overall female labor force participation rate. High-quality child care can provide young children a good start in education while allowing their parents to work through an increase of pre-k enrollment rate. It thus potentially breaks the cycle of intergenerational poverty and improves welfare for two generations simultaneously. Understanding the effects of universal pre-k on maternal labor market participation is also essential to uncover the potential mechanisms through which universal pre-k affects child outcomes. There might be a secondary effect of universal pre-k on child outcomes through the increase of maternal labor supply, resulting in an increase in total household income and parental investment in their children. Therefore, the study of causal relationship between universal pre-k and maternal labor supply is vital for complete welfare analysis of pre-k programs.

In this paper, I examine the impact of universal pre-k on maternal labor market behavior. This paper focuses on the high quality universal pre-k in Oklahoma as a case study. The expansion of universal pre-k should not only focus on the free assess to child care, but also on the quality of child care offered. However, the quality of universal pre-k programs is implicit and thus difficult to quantify. However, the pre-k program of Oklahoma is generally considered to be a high quality program because of its government mandated small classroom size and highly educated teachers.[1] To further investigate the quality effect of universal pre-k, I compare the empirical results to Georgia where the program is typically considered to be of relatively lower quality.

To show the importance of pre-k program quality, I first propose a theoretical model to investigate how quality and price changes affect maternal labor outcomes. The static labor supply model suggests that price reduction and quality improvement both increase maternal labor force participation at the extensive margin. While quality improvement unambiguously increases mothers' hours of work, the policy effect along the intensive margin is ambiguous due to opposing income and substitution effects. The empirical study investigates the effects of universal pre-k on both extensive and intensive margin, including labor force participation rate, employment rate, the percentage change in full-time job participation and working hours.

[1] Section 2 briefly introduces the history and major characteristics of Oklahoma universal pre-k.

The contribution of this paper is to use the synthetic control method (SCM) to examine the causal impact of universal pre-k on multiple labor outcomes of potential mothers. The major challenge in analyzing causal-relationship between universal pre-k and maternal labor force participation is to construct a credible counterfactual. This paper employs the SCM to construct a well-fitted control group. Before this paper, three identification strategies have been applied to link the expansion of child care programs to labor supply of mothers: an instrumental variable approach (Gelbach, 2002), a regression discontinuity approach (Fitzpatrick, 2010) and a difference-in-differences approach (Cascio and Schanzenbach, 2013).

The most commonly used empirical method in the universal pre-k literature is the difference-in-differences (DID) method. Comparing to DID, which constructs a control group with equally weighted untreated, the SCM is advantageous because it allows for the construction of a linear combination of unaffected states that minimizes the distance between Oklahoma and its synthetic control group in terms of observed pre-intervention characteristics. I find that the synthetic control group has a seemingly better fit than the equally weighted control group. Moreover, since Oklahoma is a small US state with more noisy data than larger US states, using the traditional SCM developed by Abadie and Gardeazabal (2003) and Abadie et al. (2010) does not provide a good counterfactual for Oklahoma. Doudchenko and Imbens (2017) suggest that allowing for negative weights may well improve the out-of-sample prediction. Therefore, I improve the SCM by allowing for negative weights on the unaffected units in the donor pool. I also include all lags of the dependent variables as control variables in the regression procedure to emphasize the importance of lags. In this setting, the improved SCM produces a much better counterfactual than the traditional SCM. Another advantage of the SCM is that the estimation results show the deviation between Oklahoma and "Synthetic Oklahoma" in each post-intervention year, thus providing a dynamic view of the treatment effect.

The empirical analysis uses the Current Population Survey (CPS) from 1980 to 2007. Oklahoma's universal pre-k program begun in 1998. The results of this paper are as follows. Oklahoma universal pre-k increases the labor force participation rate, employment rate, and working hours of women of childbearing age. However, universal pre-k has little effect on full-time labor force participation of women of childbearing age. I also stratify the sample of potential mothers by four socioeconomic characteristics—education level, marital status, poverty status and numbers of children in the household. I find that married mothers, mothers with lower than high school and higher than college education, and mothers with no more than two children are more likely to increase their labor supply after the implementation of the universal pre-k policy. These findings suggest that a high-quality universal pre-k impacts on the labor market decisions of mothers from disadvantaged as well as non-disadvantaged backgrounds. The synthetic control analysis yields no effect of the 1998 Oklahoma universal pre-k policy on labor outcomes of mothers with younger than 5 years old children and mothers with 4-year-olds only. My finding shows that pre-k has positive effects on labor supply for all women of fertility age, but insignificant effects when focusing only on mothers with younger children. It suggests that women may be more inclined to work before having children since they expect to continue to work after having kids, but end up with staying at home once they actually have kids. The empirical analysis on Georgia universal pre-k presents no evidence of significant effects of Georgia universal pre-k on maternal labor outcomes. It is possible that the difference between the effects of Oklahoma and Georgia universal pre-k is caused by quality variation of universal pre-k programs.

Broadly previous studies on the relationship between child care and maternal labor supply can be classified into two categories. The first strand of literature examines the labor supply elasticity of market prices of child care. Anderson and Levine (1999) find that the response of the female labor force participation rate to the price of child care is decreasing in education levels for women with children under age 13. Connelly and Kimmel (2003) find a significant negative effect of child care price on the employment of single mothers. Lundin et al. (2008) show that the maternal labor supply did not respond to a child care reform in Sweden which set a cap on the price of child care. Another strand of literature studies the labor market effects of child care policies, including child care subsidies and pre-kindergarten programs. Michalopoulos et al. (1992) show that a refundable child care tax credit would increase the labor supply of mothers. Tekin (2005) shows that single mothers are highly responsive to child care subsidies by increasing their employment while moving away from parental and relative care toward center care in the process. Blau and Tekin (2007) find that the subsidy recipients were about 13 percentage points more likely to be employed than nonrecipients.

Since the emergence of nationwide and statewide universal pre-k programs, there has been an increasing interest in the estimation of the effects of universal pre-k on maternal labor market decisions. The empirical evidence from non-US countries generally indicates a remarkable increase of maternal labor supply with a nationwide expansion to universal pre-k. Schlosser (2005) finds that introducing free public preschool to 3- or 4-year-old children in Israel increased the labor supply of mothers by about seven percentage points. Goux and Maurin (2010) also find that early school availability to 2- and 3-year-old children in France had a significant employment effect on single mothers. Baker et al. (2008) show that the labor supply of married women rose about 14.5 percent with the highly subsidized and universally accessible child care program, "$5 per day child care" in Quebec, Canada. Simonsen (2010) shows that a price increase of €1 per month decreases employment by 0.08%, which corresponds to a price elasticity of -0.17 using Danish data. Koebel and Schirle (2016) show that the Canadian Universal Child Care Benefit has significant negative effects on the labor supply of legally married mothers but positively affects the labor supply of single mothers.

The US evidence, however, shows mixed results. Cascio and Schanzenbach (2013) find that universal pre-k programs had a positive impact on the employment rate of less educated mothers with 4-year-olds only. Fitzpatrick (2010) shows that the universal pre-k increased pre-k enrollment by 12 to 15 percent, but had no robust impact on maternal labor supply, she also finds increased labor supply in rural families. Fitzpatrick (2010) uses the regression discontinuity (RD) method to address the causal relationship. This method, however requires information on birth months of Oklahoma children and the sample size is much smaller than Cascio and Schanzenbach (2013). Herbst (2017) finds that maternal employment increased substantially after the US Lanham Act of 1940, a heavily subsidized and universal childcare program during World War II.

The rest of the paper is organized as follows. Section 2 describes the background of US pre-k programs and the Oklahoma universal pre-k policy. Section 3 provides a static labor supply model to examine how price and quality changes affect the consumption of market child care and the maternal labor supply. Section 4 describes the data used for this analysis. In section 5, I describe the synthetic control method. Section 6 presents and discusses the results from the synthetic control analysis. Section 7 concludes.

2 Background on US Pre-kindergarten Education Programs and Oklahoma Universal Pre-k

Since the President's 2013 State of the Union address, 34 states have increased funding for their preschool programs, amounting to over $1 billion in new state resources dedicated to early education.[2] Nowadays, expanding access and improving the quality of pre-k programs become major concerns of policymakers in early education. Before the universal pre-k, several federal and statewide pre-k programs have been established to help children from low-income families. Since 1935, the federal government has supported early care and education for poor children to promote their healthy development and give them a better chance to succeed. But the improvement in early education of the past 80 years is slow due to "a haphazard array of uncoordinated programs, shaped by outdated science and entrenched political interests, and long driven by addressing unintended consequences of previous policies rather than core goals".[3]

The federal government now funds dozens of small programs providing services to children from birth through age five, but the preponderance of federal funds—$17.2 billion—is spent on three major programs: Head Start at $9.2 billion, the Child Care Development Fund (CCDF) at $5.4 billion, and child care expenditures from Temporary Assistance for Needy Families (TANF) at $2.6 billion annually.[4] All three programs fund poor children's participation in early care and education. Since those early child care programs are almost all highly targeted on poor families, nowadays, policymakers and economists are interested in expanding traditional public pre-k to a universal level, so that every pre-k age child (3- or 4-years-old) would

[2] White House Website, https://www.whitehouse.gov/the-press-office/2013/02/13/fact-sheet-president-obama-s-plan-early-education-all-americans

[3] Besharov et al.(2017). A Safety Net that Works.

[4] Besharov et al.(2017). A Safety Net that Works.

have access to free public pre-k.

In the state level, several states have been on the path to funding universal pre-k during the past two decades. Georgia first established a universal pre-k program in 1995, followed by Oklahoma in 1998, Florida in 2005 and then Illinois in 2006. California and New York also started to establish universal pre-k in the 1990s, but it has not yet been implemented statewide due to budget and political issues. In 2014, of the 41 states with state-funded pre-k programs (including the District of Columbia), only nine served more than half of all 4-year-olds statewide, and eleven states served less than ten percent of all 4-year-old children. Only three states—Georgia, Oklahoma and Florida—are believed to "truly" have universal pre-k programs in terms of their pre-k enrollment rates.[5]

The Oklahoma universal pre-k program is in high quality and is believed to be a successful example. In the spring of 1998, House member Joe Eddins and state senator Penny Williams secured approval to amend the school formula so that four-year-olds would be included in the school funding formula. Since 1998, Oklahoma has provided universal access to public pre-kindergarten. Children in Oklahoma who turn 4 years old on or before September 1st are eligible for the public pre-kindergarten program. The Oklahoma universal pre-k program offers two options: full day or half day, both of which provide high standard curriculum for young children. With 74 percent of all 4-year-olds enrolled in pre-k programs in 2014, Oklahoma maintains a high enrollment rate.

Besides the free access for every pre-k age child and the high enrollment rate, the quality of Oklahoma universal pre-k is also a remarkable feature. This is primarily based on the high quality of teachers and the small group size. All teachers must have college degrees and certificates in early childhood education and receive the same compensation as teachers in public elementary schools. In regard to the classes, the group size is set to not exceed 20 and the maximum child to staff ratio is set at 10 to 1.[6] Additionally, National Institute for Early Education (NIEER) reports that the Oklahoma universal pre-k has a generous expenditure of almost \$7,427 per child, while the spending per child in Georgia is \$3,490 and the average annual pre-k expenditure per child in the US is no more than \$5,000.

3 The Model

In this section, I present a static labor supply model to investigate how the price and quality changes of a pre-k program affect maternal labor market decisions. Assume a mother can always find a job and her working hours are perfectly flexible, the utility maximization problem is given by

$$\max_{X,H} \; U(X, Q, L) \tag{1}$$

[5] Sara Mead, The Building Blocks of Success, June 26, 2015, http://www.usnews.com/opinion/knowledge-bank/2015/06/26/setting-the-record-straight-on-state-pre-k-programs

[6] Oklahoma State Department of Education, http://sde.ok.gov/sde/early-childhood-and-family-education

$$s.t.\ \ X + P_m H = WH + Y \tag{2}$$

$$H + L = T \tag{3}$$

$$Q = \frac{H}{T} Q_m + \frac{T-H}{T} Q_h \tag{4}$$

In this setup, a mother's utility $U(X, Q, L)$ is derived from three sources: consumption of numeraire good X, average quality of child care Q, which is defined as the time weighted average of the market child care (day care) and the home child care following Michalopoulos et al. (1992), and leisure time $L = T - H$, where T is the time endowment and H is number of hours worked per day.

Equation (2) is the budget constraint. W is the hourly wage rate of working mothers and P_m is the price of the market pre-k per hour. Assume that the utility of other family members is exogenous to the utility function of mothers. I also take the income of other family members as part of mother's exogenous non-labor income Y. Further, I assume that working mothers have to purchase market pre-k or participate in a public pre-k program during their work time, i.e., no one else in the family will take care of their children when the mothers are at work. Rewrite equation (2), we have

$$X = (W - P_m)H + Y = \tilde{W}H + Y \tag{5}$$

where $\tilde{W}$ is the hourly wage rate net of market pre-k cost.

Equation (3) is the time constraint. Assume that mothers spend all their leisure time on child care at home, so that the time spent on home child care is $T - H$. Then the average quality is given by Equation (4), where Q_m is the quality of market child care and Q_h is the home child care quality provided by mothers. Finally, assume no market child care is better in quality than home child care provided by mothers, such that $Q_m = \alpha Q_h$ and $0 < \alpha < 1$. Therefore, we can rewrite equation (4) as

$$\begin{aligned} Q &= \frac{H}{T}\alpha Q_h + \frac{T-H}{T} Q_h \\ &= [1 - (1-\alpha)\frac{H}{T}]Q_h \end{aligned} \tag{6}$$

We can observe that mothers are affected by the universal pre-k policy through price and quality changes on both the extensive and the intensive margins.

The working decision of a mother at the extensive margin depends on whether the utility from working exceeds the utility from staying at home and taking care of her child (children). Let D denote the working decision of the mother. For the working decisions of mothers on the extensive margin, $D = 1$ denotes a mother chooses to work and $D = 0$ demotes the mother chooses to stay at home. Therefore we have,

$$D = \begin{cases} 1 & \text{if } U|_{H=H^*>0, X=X^*, Q=Q^*} - U|_{H=0, X=Y, Q=Q_h} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $U|_{H=0,X=Y,Q=Q_h}$ is the non-working utility and H^*, X^* and Q^* are the utility-maximizing values of each variable conditional on working.

Similarly, for the working decisions of mothers on the intensive margin, $D = 1$ denotes a mother chooses to work more hours and $D = 0$ demotes the mother chooses to work as long as before the policy implementation. Therefore we have,

$$D = \begin{cases} 1 & \text{if } U|_{H=H^*=H'+\epsilon, X=X^*, Q=Q^*} - U|_{H=H', X=Y, Q=Q_h} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $U|_{H=H', X=Y, Q=Q_h}$, $U|_{H=H^*=H'+\epsilon, X=X^*, Q=Q^*}$ are the working utilities before and after the universal pre-k policy and ϵ is a positive number of working hours.

The threshold of entering the labor market is quantified by a reservation wage. Equation (2) and (5) indicate that the price of market pre-k P_m affects a mother's working decision through the net wage $\tilde{W}$. When the implementation of universal pre-k leads to a price reduction, the net wage $\tilde{W}$ increases and it is more likely to exceed the mother's reservation wage. Either $\tilde{W}$ is higher than the reservation wage and mothers shift from staying at home to working, or $\tilde{W}$ is still below the reservation wage and mothers remain at home.

Besides the large price reduction, quality improvement of a pre-k program would also have an impact on mother's working decisions. Equation (4) suggests that the quality of market pre-k Q_m affects a mother's utility function through the allocation of working hours and leisure. For a given market pre-k price, when the quality of market pre-k or public pre-k improves, the parameter α in equation (6) increases, so that the utility loss from lower-quality child care is smaller and the gap between mothers' net wage and reservation wage decreases. Hence, for the extensive margin, quality improvement increases the likelihood of working.

However, for working mothers, a rise in the net wage $\tilde{W}$ is well known to have a theoretically ambiguous effect on working hours due to the trade-off between the negative income effect and positive substitution effect. For the effect of quality improvement, as α increases, the average quality of childcare Q is higher, the marginal utility from home child care decreases, and mothers will allocate less time to home child care and correspondingly allocate more time to work. Thus, quality improvement increases mothers' working hours.

In summary, universal pre-k should unambiguously increase the maternal labor force participation rate, especially when the program is in high quality. However, the model does not make a decisive prediction on mothers' working hours, although the effect is more likely to be positive when the program quality is higher.

4 Empirical Strategy

The empirical method employed in this paper is the synthetic control approach, first introduced by Abadie and Gardeazabal (2003) and developed by Abadie et al. (2010). The SCM allows for estimation in settings where a single unit is exposed to an event. It provides a data-driven procedure to construct a synthetic control unit that approximates the characteristics of the treated unit. In this section, I will briefly introduce the SCM used to analyze the effect of the 1998 Oklahoma universal pre-k on female labor market decisions.

Suppose we observe $S+1$ states, one of which is our treated state, Oklahoma, which we call state 1. Let Y_{it}^{N} denote the outcomes of interest that would be observed for state i at time t in the absence of the intervention, where $i = 1, \ldots, S+1$, $t = 1, \ldots, T$. Let T_0 be the number of pre-intervention periods and $1 \leq T_0 < T$. Let Y_{it}^{I} be the outcome that would be observed for state i at year t if state i is exposed to the universal pre-k policy in period $T_0 + 1$ to T. Two assumptions are needed for the synthetic control method.

Assumption 1: No anticipation effects

The intervention has no effect on the outcomes before the implementation period $t \in \{1, \ldots T_0\}$. Under Assumption 1, for $t \in \{1, \ldots T_0\}$ and all $i \in \{1, \ldots S+1\}$, we have $Y_{it}^{N} = Y_{it}^{I}$.

Assumption 2: No interference on untreated units

There is no interference between treated and untreated states, the outcomes of the untreated states are not affected by the intervention implemented in the treatment state.

This estimated treatment effect is $\alpha_{1t} = Y_{1t}^{I} - Y_{1t}^{N}$. Since Y_{1t}^{N} is unobserved, we need to estimate the counterfactual Y_{1t}^{N} with the following factor model

$$Y_{1t}^{N} = \delta_t + \theta_{\mathbf{t}} \mathbf{Z_1} + \lambda_{\mathbf{t}} \mu_{\mathbf{1}} + \varepsilon_{1t} \tag{7}$$

where δ_1 is an unknown state fixed effect, Z_1 is a $(r \times 1)$ vector of observed covariates (not affected by the intervention), θ_t is a $(1 \times r)$ vector of unknown parameters, $\lambda_t = (\lambda_{t1}, \lambda_{t2}, \ldots, \lambda_{tF)}$ is a $(1 \times F)$ vector of unobserved time fixed effect for $t = 1, \ldots F$, $\mu_1 = (\mu_{11}, \mu_{12}, \ldots \mu_{1F})$ is an $(F \times 1)$ vector of unobserved factor loading for the treated state Oklahoma, and the error terms ε_{it} are unobserved transitory shocks at the state level with zero mean. Note that we will obtain the difference-in-differences (fixed effect) model from equation (7) if λ_t is constant over time.

The synthetic control group is obtained by assigning weights $\boldsymbol{\omega} = (\omega_2, ..., \omega_{S+1})$ to each untreated unit in the donor pool. The value of the outcome variable for each synthetic control indexed by $\boldsymbol{\omega}$ is $\sum_{s=2}^{S+1} \omega_s Y_{st} = \delta_t + \theta_{\mathbf{t}} \sum_{s=2}^{S+1} \omega_s \mathbf{Z_s} + \lambda_{\mathbf{t}} \sum_{s=2}^{S+1} \omega_s \mu_{\mathbf{s}} + \sum_{s=2}^{S+1} \omega_s \varepsilon_{st}$. The optimal weight vector $\boldsymbol{\omega}^*$ is the one that minimizes the distance between the pre-intervention observed characteristic vector $\mathbf{X_i}$ of the treated state and the selected control group, $||\mathbf{X_1} - \mathbf{X_0} W||$, where $\mathbf{X_i}$ includes both covariates $\mathbf{Z_i}$ and outcome variables Y_i, $\mathbf{X_1}$ is the pre-intervention observed characteristics of the treated unit and the $\mathbf{X_0}$ is the same observed variable set of the untreated states.[7]

Abadie et al. (2010) impose three restrictions in analyzing the effect of California's tobacco control program—no intercept, the sum of weights add up to 1, and all weights are non-negative. In this paper,

[7]See details on optimal weight vector selection on Abadie et al. (2010)

to obtain a better counterfactual for the treated state, I relax the non-negativity assumption ($\omega_s \geq 0$, $s = 2, 3, ..., n+1$). Doudchenko and Imbens (2017) claim that allowing for negative weights may well improve the out-of-sample prediction. First, allowing for weights on the observed characteristics of the untreated unites may be able to better fit those of the treated unit. Doudchenko and Imbens (2017) show an example when there is one treated state and two control states and the key characteristic is the share of young people. If the share of young people in the treated state is 2/3, and the shares of the control states are 1/2 and 1/3, assigning weights 2 and -1 respectively to the two control states can produce the identical share of young people to the treated states. The second reason is negative weights can help with bias-reduction as the bias goes to zero in a faster rate in settings with many covariates to be matched on.[8] I still keep the no intercept and sum up to 1 restrictions in the minimization procedure because imposing these two restrictions helps produce a unique weight matrix, though Doudchenko and Imbens (2017) claims no restrictions are necessary for the SCM. The intuition behind the negative weights is that the control group is constructed by some virtual states, which have characteristics opposite to some real US states.

For inference, Abadie et al. (2010) suggest using placebo tests to measure the significance of estimates. The basic idea behind the placebo tests is to apply the synthetic control method to all the control units in the donor pool as if they were also exposed to the universal pre-k policy and test whether the treated unit behaves in a significantly different way from unexposed units. Under the null hypothesis that policy intervention has no impact on the treated unit, the estimate for the treated unit is expected to lie within the distribution of the placebo estimates. I also apply the pre/post rooted mean squared prediction error (RMSPE) ratio test to measure the significance of treatment effects. The pre/post RMSPE ratio test is an extension of the placebo tests, it measures the closeness between the observed variables of the treated unit and the synthetic control group before and after the policy intervention. Therefore, we can compare the pre/post RMSPE ratio of the treated unit to that of the untreated unit to examine whether there is a relative larger increase in post RMSPE.

5 Data

The primary data set is the March Current Population Survey (CPS), which provides detailed labor statistics and demographic characteristics for individuals and households in the US annually. The entire study period is from 1980 to 2007 to cover a long enough pre-intervention period before the intervention year 1998 and the sample ends in 2007 to avoid the effect of the financial crisis and more recent expansion of pre-k in other states. Since the enrollment of Oklahoma universal pre-k starts at the end of February each year, in case the labor outcomes of women in Oklahoma is immediately affected by the new policy in 1998,

[8]See Abadie and Imbens (2011).

the post-intervention period is set to start from 1998.

Table 1. Demographic and Economic Characteristics Before and After 1998 in Oklahoma

	Female (25-45)	
	Before 1998	After 1998
Labor force participation rate	0.71	0.74
Working hours	28.67	28.94
Family total income	32716.80	55539.84
Mother's age: 25-30	0.25	0.21
Mother's age: 30-35	0.24	0.23
Mother's age: 35-40	0.25	0.25
Mother's age: 40-45	0.26	0.31
Mother's education level: below high school	0.11	0.10
High school graduates	0.40	0.34
College graduates	0.44	0.50
Above college degree	0.05	0.06
White	0.86	0.76
Black	0.07	0.09
Other race	0.07	0.15
Family size	3.45	3.45
Food Stamp Recipients	0.10	0.10
Children enrolled in free lunch program	0.35	0.43
Below poverty line	0.14	0.14
Number of observations	6060	3331

The major outcomes of interest are labor outcomes of potential mothers (women aged 25 to 45),[9] mothers with younger than 5 year olds and mothers with 4 year old children only after the universal pre-k policy. There are three reasons to focus on the sample of potential mothers, defined as women between 25 to 45 years old, as the primary sample of interest. First, the provision of free and high-quality pre-k shortens the duration of utilization of day care and/or home care, resulting in a higher possibility of females to stay in their current jobs before their children turn 3 or 4 years old. Therefore, there might be a spillover effect on mothers whose children is younger than 4 years old. In addition, universal pre-k may also have an effect on fertility decisions and it will thus affect the labor market outcomes of all women of childbearing age. Prior studies have shown that the increase of child care subsidies or the expansion of child care programs would increase fertility rates (Blau and Robins, 1989; Baughman and Dickert-Conlin, 2003; Haan and Worhlich, 2011; Bauernschuster et al., 2015). Third, the sample size of mothers with only 4-year-olds is small in the Current Population Survey (CPS), especially for a small US state like Oklahoma, and the sample size issue would be more severe in the study of heterogeneous effects by subgroups.

The labor outcomes includes labor force participation rate, percentage change of full-time labor force

[9]In Appendix A, I present the synthetic control method analysis on potential mothers in different age ranges, including women aged 25-35, 25-40, 20-45, 25-50, the results are robust for slightly narrowing or expanding age groups, I choose women aged 25 to 45 mainly because this group is the commonly used sample of fertility age women.

participation, hours-of-work for working mothers and employment rate. The major labor outcomes following the model in Section 2 are labor force participation rate as the outcome on the extensive margin and working hours as the outcomes on the intensive margin. Universal pre-k is expected to increase the likelihood that a mother is employed since mothers who have free full-time child care options are better able to work. Employment is also restricted by job availability though. It is possible that mothers who are willing to work cannot find jobs, especially after years of unemployment. Therefore, the empirical analysis is needed to investigate the effect on employment rate. The percentage change in full-time working mothers, defined as the ratio of full-time working mothers to the whole working sample of mothers, is a proxy of change in job type. When a high-quality universal pre-k policy is implemented, the marginal rate of substitution between the utility of working and home child care is smaller and thus working mothers are more likely to switch from part-time jobs to full-time jobs. Unlike the change of working hours, percentage change in full-time working females could also indicate a change in type of jobs.

The CPS defines an individual to be 'in the labor force' if she is employed or unemployed, so that those who are at school, retired and staying at home are not in the labor force. The employment rate is the ratio of those who are employed to the population of females in the labor force. The 'working hours' variable in this paper is defined as how many hours an individual work (not including zero) the week before the survey and percentage of full-time labor force participation is directly from the indicator variable of full-time or part-time employment. The studied outcomes of interest are state level labor statistics obtained from aggregating the individual level data in CPS.

Since the dependent variables are state level labor outcomes and the universal pre-k policy takes place at the state level, it is convenient to estimate the effect in aggregate level. Meanwhile, the synthetic control method is good for the comparative case study in aggregate entities or administrative areas. Thus I use the individual CPS data to form state level controls such as the fraction of population by race and education level. CPS personal sample weights are used in the data aggregation. The sample size of the women aged 25 to 45 in the CPS is 1,478,181 for all states from 1980 to 2007. The state level data used for the synthetic analysis has a sample size of 1,428, which comes from the multiplication of the number of states and districts in the CPS (51) and the number of study years (28). The corresponding set of explanatory variables $\{\mathbf{Z}_i\}$ consists of demographic characteristics such as age, race and economic characteristics including total family income, personal income, family size and spouse's working hours.[10] To capture state level shocks to female labor outcomes and to better construct the synthetic control group, I also take the labor outcomes of women aged 45 to 60 as an additional control variable. In addition, I add state gross domestic product (GDP) into the set of explanatory variables $\{\mathbf{Z}_i\}$, obtained from another data source—Bureau of Economics Analysis (BEA).

[10] The full set of the explanatory variables is listed in Appendix C.

Table 2. Summary Statistics of Selected Variables at State Level—OK vs. rest of US

	Before 1998		After 1998	
	Oklahoma	rest of US	Oklahoma	rest of US
Labor force participation rate	0.72	0.74	0.75	0.78
Working hours	28.84	28.59	29.22	30.25
Family income	33164.04	35674.11	54624.93	61654.86
Mother's age: 25-30	0.07	0.08	0.07	0.07
Mother's age: 30-35	0.21	0.21	0.20	0.20
Mother's age: 35-40	0.59	0.60	0.60	0.61
Mother's age : 40-45	0.13	0.12	0.14	0.12
Mother's education level: below high school	0.04	0.04	0.09	0.09
High school graduates	0.15	0.13	0.34	0.30
Some college, no degree	0.08	0.07	0.22	0.20
Associate Degree,	0.03	0.04	0.09	0.11
Bachelors degree	0.05	0.07	0.19	0.22
Above college degree	0.02	0.02	0.06	0.08
White	0.85	0.85	0.79	0.82
Black	0.07	0.11	0.08	0.11
Other race	0.07	0.04	0.13	0.07
Family size	3.43	3.40	3.32	3.26
Food Stamp Recipients	0.10	0.10	0.09	0.07
Children enrolled in free lunch program	0.30	0.26	0.33	0.26
Missing in free lunch data	0.24	0.30	0.26	0.29
Below poverty line	0.16	0.14	0.14	0.12
GDP	58584.26	107485.4	111232.7	233695
Number of observations	19	950	9	450

Sources: State GDP data is from the Bureau of Economics Analysis (BEA), all other statistics are from the Current Population Survey (CPS).
Notes: The table presents the summary statistics of Oklahoma and the rest of all other U.S. states. In the state level, there is one observation each year for each state. There are 50 states and 1 district (District of Columbia) in the date set. Excluding Oklahoma, the number of the rest of the US states is 50.

The CPS data also shows high serial correlations in female labor outcomes. For example, the correlation between the female labor force participation rate and its one-year lagged variable is as high as .8877. To deal with this issue, I take the lags of outcome variables in each pre-intervention period into the control variable set. Therefore, lags of dependent variables are given more importance in the synthetic control analysis relative to other controls in $\{\mathbf{Z}_i\}$.

Table 1 summarizes the individual level data to show the before and after differences in the demographic and economic characteristics in Oklahoma. There are a number of things that are worth pointing out from the table. First, the statistics show that the labor force participation rate increased after 1998, but working hours seemed to remain the same. Second, family total income rose dramatically from about 32k to 55k on average.[11] Table 1 also indicates that a larger percentage of females are pursuing higher level of education. Next, as expected, the demographic characteristics in the summary statistics were stable over time; whereas

[11]In the CPS data, the family income is not adjusted by CPI, for reference, $32k in 1990 is about $42k in 2000.

only the racial composition had a notable change. Finally, although the number of free lunch recipients increased after 1998, the fractions of food stamp recipients and households below the poverty line did not change much in Oklahoma.

Table 2 summarizes the state level statistics for Oklahoma and the rest of US before and after 1998. It suggests that, compared to the rest of US states, women aged 25 to 45 in Oklahoma were less likely to participate in the labor market and worked fewer hours per week. The GDP difference shows that the economy of Oklahoma was below the average level of other US states. Hence, it should not be surprising that the labor force participation rate of Oklahoma was lower than average. Table 2 also shows that Oklahoma had a higher fraction of lower-educated women, food stamp and free lunch recipients, and high-poverty households. As for the racial composition, Oklahoma had more non-whites than other states. Ultimately, when controlling for time difference, the summary statistics in Table 2 appear to show a non-substantial change in the labor force participation rate of fertility age women in Oklahoma relative to the simple average of the rest of US.

6 Empirical Estimates

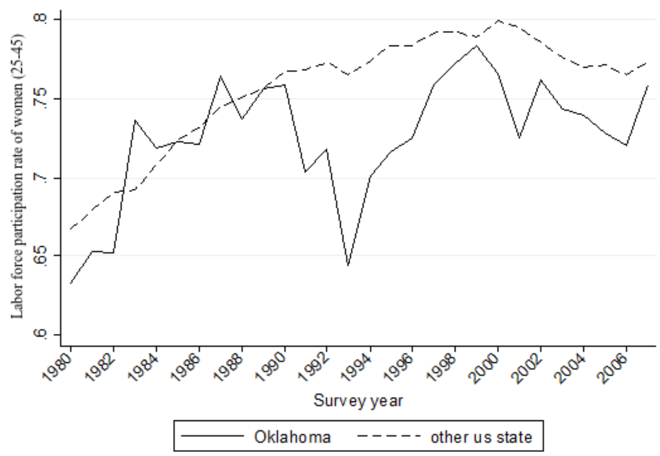
6.1 Main Results

Now I estimate the impact of the 1998 universal pre-k policy on Oklahoma female labor force participation. The main results presented below investigate the labor outcomes of childbearing age women (women aged 25-45) and mothers with young children. In this section, I first present the effects of universal pre-k on the labor force participation rate of fertility age women to further illustrate the empirical strategy. Then I show the estimation effects on three other labor outcomes of potential mothers in Oklahoma—percentage of full-time labor force participation, average weekly working hours and employment rate. Lastly, I repeat the empirical analysis on mothers with younger than 5-years-olds and mothers with 4-year-olds only to compare with previous evidence.

Before applying the synthetic control method, Figure 1 plots the labor force participation rates of potential mothers in Oklahoma compared to a naive control constructed by assigning equal weights to all unaffected US states. Note that before 1990, the naive control group appears to be a good control, however, after 1990 and up to 1998, there is a significant gap. Therefore, using equally weighted untreated US states would not be a good strategy to investigate the policy of interest.

Figure 2 presents the synthetic control without allowing for negative weights on the untreated units. Comparing to the simple average control in Figure 1, which is normally used in the difference-in-differences approach, the traditional SCM provides a better counterfactual in the pre-intervention (pre-1998) period.

Figure 1. Maternal Labor Force Participation Rate—Oklahoma vs. the Rest of US States (excluding Georgia)

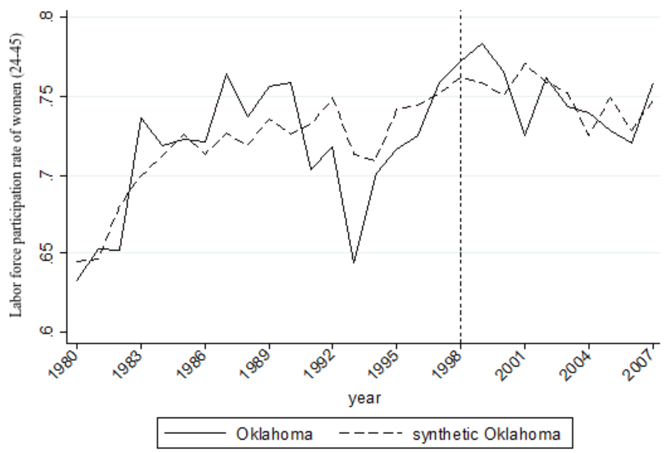


Notes: The graph presents the labor force participation rate each year. The solid black line stands for the trend of labor force participation rate of Oklahoma and the dash line represents the trend of labor force participation rate of the rest of US states in average from 1980 to 2007 excluding Georgia, who had established universal pre-k policy in 1995.

The better fit from employing the SCM is due to the fact that the SCM does not assume equal weight to each untreated unit in the control group and it takes the lags of outcome variables in minimizing the distance between the observed characteristics of treatment and control. Except for the period 1992 to 1994, labor force participation rates of the synthetic control group in most years before 1998 were generally close to labor force participation rates of Oklahoma. One possible explanation for this big drop is that the oil price crisis in the early 1990s had a stronger effect on Oklahoma's economy. Historical statistics show that in each oil price decline (defined as inflation-adjusted oil prices falling nearly continually by more than 30 percent and more than $20 a barrel), the employment rate in the oil and gas sector in Oklahoma alone decreased by at least seven percent.[12] The huge drop in labor force participation rate may also result from measurement error and small sample bias.

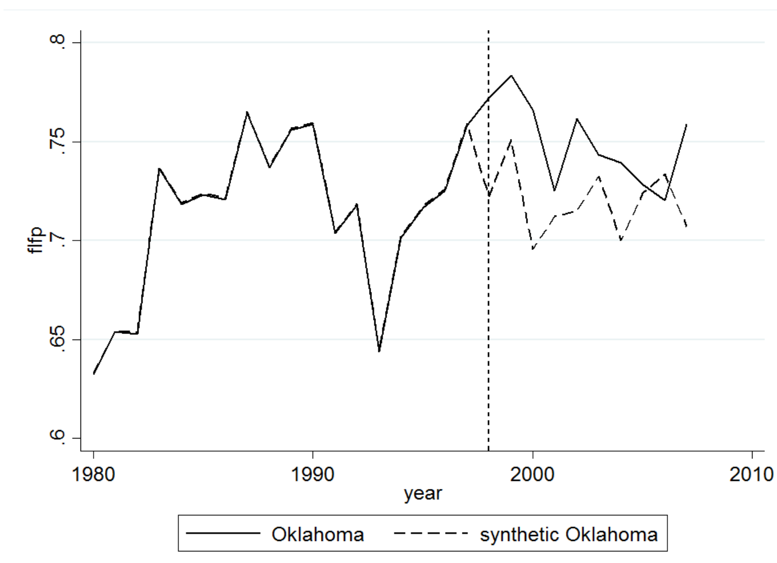
[12] Chad Wilkerson, How will Oklahoma be affected be decline of oil price? March 11, 2015. Federal Reserve Bank of Kansas City, Denver, Oklahoma City, Omaha. https://www.kansascityfed.org/publications/research/oke/articles/03-11-2015/oke-03-11-15.

Figure 2. Maternal Labor Force Participation Rate (25-45 sample)—Oklahoma vs. Synthetic Oklahoma (Traditional SCM)



Note: The vertical dash line indicates the policy intervention year 1998, the starting date of the universal pre-k policy was September 1st, 1998. I treat 1998 as the first year of post-intervention period, the results are robust if 1998 is taken as a pre-intervention year.

Figure 3. Maternal Labor Force Participation Rate (25-45 sample)—Oklahoma vs. Synthetic Oklahoma (SCM allowing for negative weights)



Note: Same as Figure 2

Figure 2 shows that even though the traditional SCM provides a better counterfactual than the difference-in-differences method, it may not be able to provide a close counterfactual to a small state with noisy data like Oklahoma. In order to obtain a better counterfactual, I allow for negative weights in constructing the control group and add labor force participation rates from 1980 to 1997 into the control variable set to emphasize on the strong correlation between the outcome variable and its lags. Figure 3 shows the synthetic control allowing for negative weights. Obviously, the synthetic control with negative weights provides a better counterfactual than the traditional synthetic control with a non-negativity restriction. Figure 3 suggests that the labor force participation rates of potential mothers in Oklahoma are higher than the labor force participation rates of potential mothers in the synthetic control group in the post-intervention period.

Table 3. Weights in the Synthetic Control Group

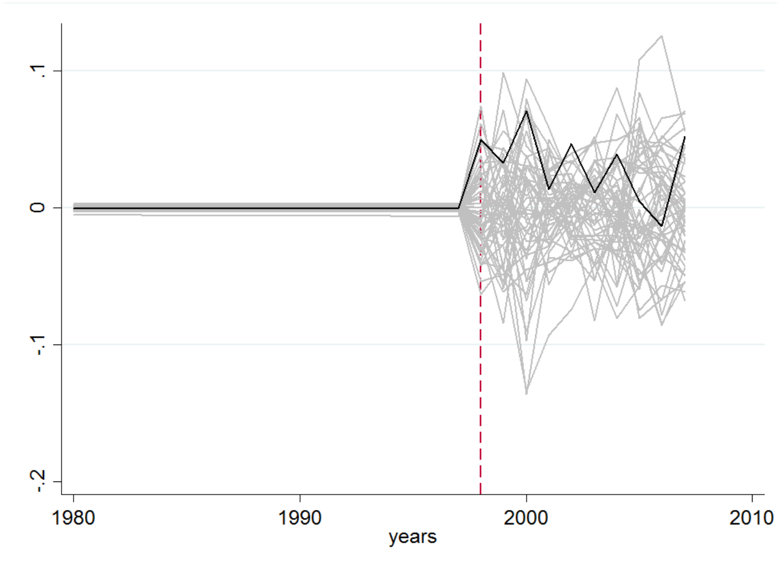
State	Weight	State	Weight
Alabama	0.313	Montana	-0.111
Alaska	-0.108	Nebraska	-0.403
Arizona	-0.104	Nevada	0.431
Arkansas	-0.126	New Hampshire	0.228
California	0.005	New Jersey	0.028
Colorado	0.195	New Mexico	0.116
Connecticut	0.219	New York	-0.155
Delaware	-0.166	North Carolina	-0.036
District of Columbia	0.048	North Dakota	-0.045
Florida	-0.167	Ohio	0.178
Hawaii	0.134	Oregon	-0.088
Idaho	-0.081	Pennsylvania	0.011
Illinois	-0.034	Rhode Island	-0.209
Indiana	-0.045	South Carolina	0.352
Iowa	0.158	South Dakota	-0.049
Kansas	0.239	Tennessee	-0.09
Kentucky	0.177	Texas	0.15
Louisiana	0.152	Utah	0.203
Maine	-0.151	Vermont	0.199
Maryland	-0.036	Virginia	-0.146
Massachusetts	0.091	Washington	0.216
Michigan	-0.129	West Virginia	0.059
Minnesota	-0.025	Wisconsin	-0.23
Mississippi	-0.192	Wyoming	-0.026
Missouri	0.051		

Note: Georgia is not included in this table of weights. Since Georgia universal pre-k starts in 1995, including Georgia in this analysis violates Assumption 2 of the synthetic control framework.

Table 3 provides the weights of each state used to construct the "Synthetic Oklahoma". A state will be assigned zero weight if it is not chosen to construct the synthetic control group. Note that Georgia is eliminated from the donor pool of the unaffected states, hence the synthetic control group is ensured to be untreated before the policy intervention year 1998.

Figure 3 shows that the trend in female labor force participation rate of "Synthetic Oklahoma" well-matches that of Oklahoma in the pre-intervention period. Table 4 further displays the closeness of the observable characteristics between Oklahoma and "Synthetic Oklahoma", these variables are used to construct the synthetic control. I also list the differences in observable characteristics between Oklahoma and the simple average of unaffected US states in Table 4 for comparison. It shows that the observed charateristics of the

Figure 4. Placebo Tests on Synthetic Control Results of Oklahoma



Note: All lines represent the distribution of estimated gaps between the treated unit and non-intervened control states. The black line stands for the estimated gaps of Oklahoma and the estimated gaps of the placebos are indicated by grey lines.

synthetic control group closely match the observed charateristics of Oklahoma.

Now regarding the post-intervention period, the female labor force participation rate in Oklahoma shows no break from the pre-intervention period. However, Figure 3 suggests a large distance between the labor force participation rates of Oklahoma and "Synthetic Oklahoma" after the policy implementation. Moreover, the labor force participation rate of Oklahoma is larger than that of the synthetic control group in almost every post-intervention year. Abadie et al. (2010) suggest to use placebo tests to investigate the significance of empirical results obtained from the SCM. The nature of these tests is to conduct a series of placebo studies by iteratively applying the SCM to states other than Oklahoma. Figure 4 shows the placebo tests of the synthetic control estimation. The lines in Figure 4 represent the distribution of estimated gaps between the treated unit and their synthetic control group in labor force participation rate. The gap line for Oklahoma is in black and the gap lines for other states are in grey. As expected, it is found that the estimated gap between Oklahoma and "Synthetic Oklahoma" generally lies above most of the placebos. Therefore, the placebo test shows a statistically significant and positive impact of the Oklahoma universal pre-k policy on the labor force participation rate of women of childbearing age.

Besides the placebo tests, I also apply the pre/post RMSPE ratio test for inference. The pre/post RMSPE ratio test is an extension of the placebo tests, but using numbers rather than a graph to compare the pre-post difference. In the synthetic control method, the rooted mean squared prediction error (RMSPE) is used to measure the closeness between the observed variables of the treated unit and the syn-

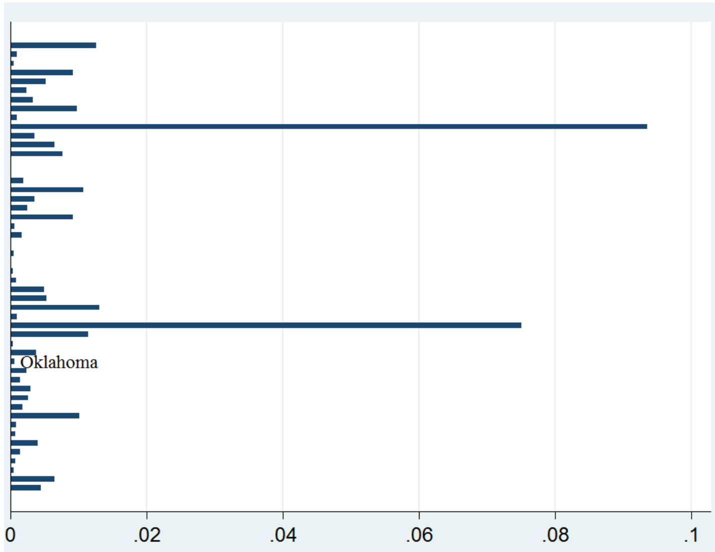
Table 4. Mean Differences between Oklahoma and "Synthetic Oklahoma" and Oklahoma and the Rest of US States in Demographic Economic Characteristics (Partially)

	OK vs. Synthetic OK	OK vs. Rest of US
female education: less than high school	-0.001	0.063
female education: high school	0.036	0.265
female education: some college	-0.001	0.145
female education: associate degree	0.036	0.048
female education: college	-0.003	0.090
female education: master or doctoral	0.005	0.020
family total income	-5814.66	5829.97
married	0.038	0.090
separated, divorced, or widow	-0.009	0.024
never married	-0.050	-0.114
food stamp	0.015	0.003
free lunch	0.020	0.052
below poverty line	0.033	0.024
white	0.072	0.005
black	-0.049	-0.037
other race	0.018	0.031

thetic control group, it is defined as $\sqrt{\frac{1}{T}\sum_{t=0}^{t=T} e_t^2}$, where $e = |X_{treated} - X_{synth}|$ is the distance between the treated unit and the synthetic control group in the value of variable X. The pre/post RMSPE ratio $\sqrt{\frac{1}{T_0}\sum_{t=0}^{t=T_0} e_t^2}/\sqrt{\frac{1}{T-T_0+1}\sum_{t=T_0+1}^{t=T} e_t^2}$ is the relative measure of the pre-intervention and the post-intervention difference in observed characteristics. The pre/post RMSPE ratio of the treated unit is compared with the ratios of the placebos. A relatively smaller pre/post RMSPE means the post-intervention difference between the treated unit and the synthetic control group is larger than pre-intervention difference, thus it is likely to show a significant treatment effect. Figure 5 shows the pre/post RMSPE ratio of labor force participation rates of all US states. Oklahoma has a small pre/post RMSPE ratio, and the pre/post RMSPE ratio of Oklahoma is one of the smallest among all 50 US states and the District of Columbia. It suggests that the difference in the female labor force participation rate between Oklahoma and the synthetic control group before the implementation of the universal pre-k policy is smaller than the post-intervention difference. In other words, the treatment effect relative to pre-intervention difference is large.

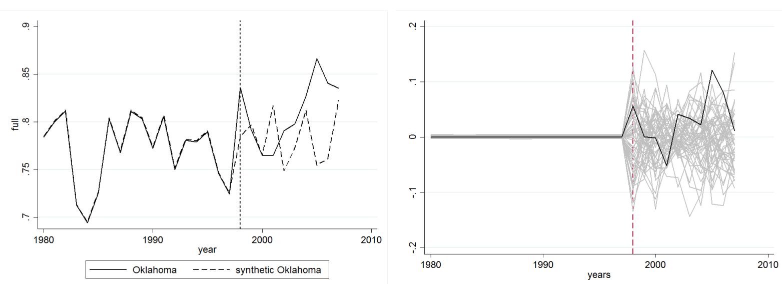
Next, I examine the response of other three labor outcomes of women aged 25 to 45 to the universal pre-k policy: percentage change of full-time labor market participation, employment rate and mean weekly working hours. I first examine the policy effect on the percentage change of full-time job participation, which indicates not only a change in labor force participation but also a change in job type. Figure 6 shows that the effect of universal pre-k on the percentage of full-time female workers aged 25 to 45 in Oklahoma is statistically insignificant. There is no great difference between the full-time job participation rates of

Figure 5. Pre/post RMSPE Ratio of Female Labor Force Participation Rate by States



potential mothers in Oklahoma and the synthetic control group after 1998, and the placebo tests also show the gap in full-time job participation rate of fertility age women in Oklahoma is generally inside the range of placebos.

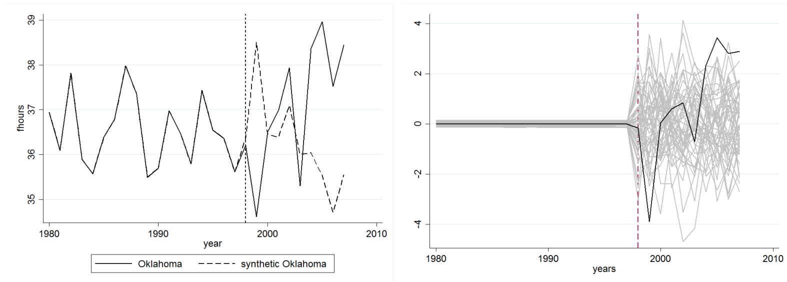
Figure 6. Synthetic control analysis on Percentage of Full-time Working Women Aged 25-45



Then I examine the effect on hours of work of working mothers. Since the theoretical model has an ambiguous solution on evaluating the effect of universal pre-k on the intensive margin, it is especially useful to utilize empirics. The synthetic control estimation in Figure 7 shows that the post-trend of average working hours in Oklahoma is opposite to the post-trend of average working hours in the synthetic control group. The working hours of women aged 25-45 appear to decrease right after the universal pre-k policy and then increase to be much higher than the average working hours of the synthetic control group. This suggests that the universal pre-k increases working hours of women aged 25-45 in the long run.

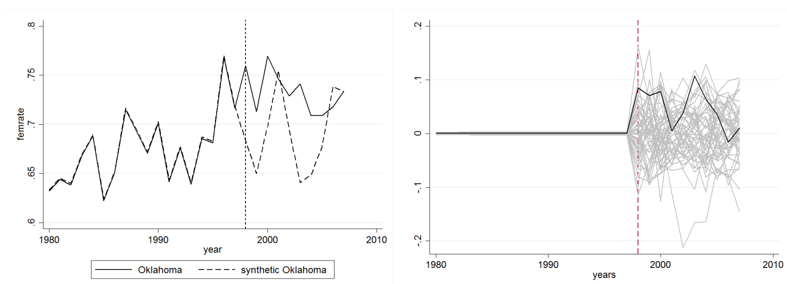
The last labor outcome to analyze is state level employment rate of women aged 25 to 45. Figure 8 shows a higher employment rate of potential mothers in Oklahoma than that in the synthetic control group for most of the post-intervention years. The placebo tests also suggest the gap between the employment rate

Figure 7. Synthetic Control Analysis on Working Hours of Women Aged 25-45



in Oklahoma and the synthetic control is generally large and above zero, but the differences in employment rates of potential mothers between Oklahoma and the synthetic control group in 2001 and 2006 are not significantly different from zero.

Figure 8. Synthetic Control Analysis on Employment Rate of Women Aged 25-45



At last, to compare with previous studies, I apply the synthetic control method on mothers with younger than 5 years old children and the direct policy-affected sample—mothers with 4-year-olds only. In this section, I will only show the effects of universal pre-k on the labor force participation rate and mean weekly working hours of the two samples, the estimation results of the other two labor outcomes are presented in Appendix D. Figure 9 and Figure 10 show the effects of universal pre-k on the labor force participation rate and weekly working hours of mothers with 5 years old children, respectively. Surprisingly, the results suggest no statistically significant effects of the universal pre-k policy on the extensive and intensive margin labor outcomes of mothers with younger than 5-year-olds. Figure 9 shows that the post-intervention labor force participation rates of mothers in Oklahoma are higher than the synthetic control group, but the distance is not significantly different from zero. The labor force participation rate of mothers with younger than 5 years old children increased in the first two years after the policy implementation and then dropped back to level almost identical to the synthetic control group. Figure 10 shows an increasing post-trend of mean weekly working hours of mothers with younger than 5 years old children in Oklahoma, however, the synthetic control group also follows an increasing trend close to Oklahoma. The placebo tests also show that the gap between mean weekly working hours of mothers with younger than 5 years old children in Oklahoma and its synthetic control group is not significantly different from zero.

Figure 9. Synthetic Control Analysis on Labor Force Participation Rate of Mothers with Children Younger than 5 Years Old

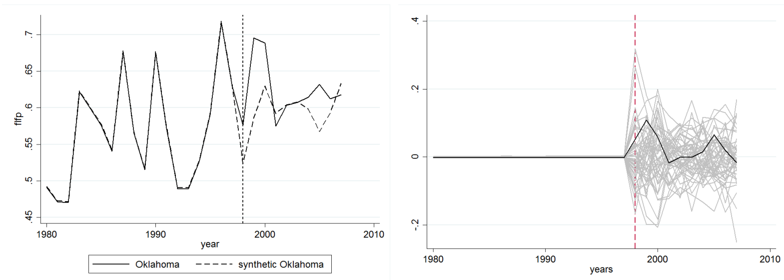
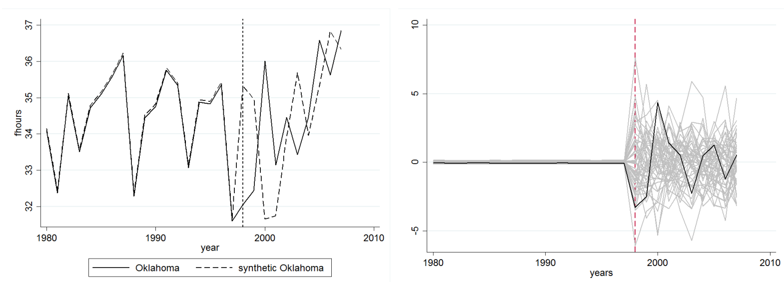


Figure 10. Synthetic Control Analysis on Weekly Working Hours of Mothers with Children Younger than 5 Years Old



For the sample of mothers with 4 years old children only, first note that sample size of mothers with 4 years old children is small, especially for small states. Also note that the study time period of this sample is different from the sample of women of childbearing age. I match mothers with 4-year-olds using their household ID, which is not available in IPUMS-CPS prior to 1988. Figure 11 presents the synthetic control analysis on the labor force participation rate of mothers with 4-year-olds only. Except for a large decrease in 2005, the labor force participation rate in Oklahoma is not significantly different from that of the synthetic control group in the post-trend. This is also confirmed by the placebo tests in the right panel. In regard to the intensive margin effect, Figure 12 shows the weekly working hours of mothers with only 4 years old children in Oklahoma does not follow a clear increasing and decreasing pattern. In fact, the mean weekly working hours of mothers with only 4 years old children in Oklahoma moves up and down cyclically. Therefore, there is no clear break from the pre-trend, and there is no evidence of significant effect of universal pre-k on mean weekly working hours of mothers with 4-year-olds only. Thus, different from previous literature, this paper suggests that the effect of universal pre-k on the labor outcomes of the policy-targeted sample is statistically insignificant.

To summarize, the main results show that the universal pre-k policy has positive effects on labor force participation rate, employment rate and weekly working hours of women of childbearing age, but the effect

Figure 11. Synthetic Control Analysis on Labor Force Participation Rate of Mothers with 4-year-olds Only

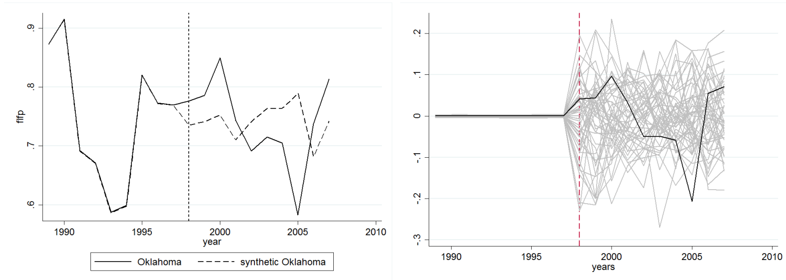
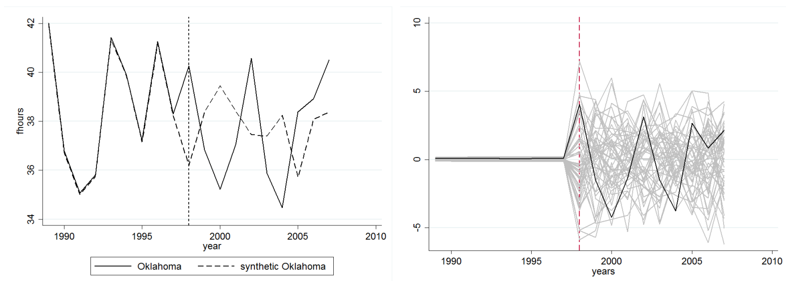


Figure 12. Synthetic Control Analysis on Weekly Working Hours of Mothers with 4-year-olds Only



is statistically insignificant on full-time labor force participation of women in the fertility age. The empirical result agrees with the predictions of the theoretical model in Section 3. It suggests that the universal pre-k policy increases extensive margin labor outcomes of potential mothers. Though the policy effect on the intensive margins is ambiguous in theory, the empirical result shows universal pre-k positively affect mothers' working hours. However, the positive effects of universal pre-k are neither consistent for mothers with younger than 5 years old children nor mothers with only 4 years old children. Although the labor force participation rate of mothers with younger than 5 years old children in Oklahoma is also positively affected by the universal pre-k policy, it is not statistically significant, and there is no evidence of significant and positive effect on mean working hours of mothers with younger than 5 years old children. If the study sample is further specified to mothers with 4 years old children only, the positive effect of universal pre-k vanishes. The null effect of universal pre-k on the labor outcomes of mothers with young children may be due to several possibilities.

First, the universal pre-k policy affects maternal labor supply through actual enrollment in child care programs. Albeit the total enrollment rate of Oklahoma is increasing in the post intervention period[13], mothers are not required to work to qualify for the universal pre-k program. It is possible that marginal utility from leisure and/or home child care for young children exceeds marginal utility of working, thus mothers remain to stay at home.

[13]See Appendix C.

Second, since the universal pre-k policy may have a positive effect on the fertility decisions of women, it is also possible that the decision to have additional children would reduce a mother's incentive to work. Thus the positive effect of the universal pre-k policy might be canceled out by the indirect effect of childbearing decisions.[14]

Third, the universal pre-k policy may simply crowd out existing private day care and mothers transfer their children from private pre-k to public pre-k without changing their labor market decisions. The fourth explanation is that female labor force participation rates of highly developed countries, such as the US and Sweden (Lundin et al., 2008), are already high before a further expansion on child care policy. Moreover, studies on the wage elasticity of female labor supply show that women are no longer as responsive to wage changes as before (Blau and Kahn, 2007; Heim, 2007). Hence universal pre-k may have limited effect on the maternal labor supply.

At last, the theoretical labor supply model suggests mothers with lower family income are more likely to enter the labor market with low-cost or free child care programs, however, the existing child care programs for poor families, such as the Head Start, may already enable mothers from disadvantaged background to participate in the labor market.

6.2 Treatment Effect Heterogeneity

So far the analysis has primarily focused on all potential mothers. However, the impact of universal pre-k may vary in education level, marital status, family income and number of children mothers. Anderson and Levine (1999) find that the response of female labor force participation to the price of child care decreases in education levels for women with children under age 13. Koebel and Schirle (2016) show that the Canadian Universal Child Care Benefit has significantly negative effects on the labor supply of legally married mothers but has significantly positive effects on the labor supply of single mothers. In this section, I focus on the labor outcomes of women of childbearing age to avoid small sample size problem that maybe produced from further sample restrictions. And I will only show the effects of universal pre-k on labor force participation rate and working hours of potential mothers. [15]

I first investigate the effects of universal pre-k on the labor force participation rate of women with family income below and above the poverty line. The theoretical model shows that mothers with lower non-labor incomes have stronger incentives to work for a given consumption level since their expected wage rates are more likely to be higher than their reservation wages. Moreover, as universal pre-k provides free child care to all pre-k age children, I also expect mothers with higher family income to respond to the policy. The results

[14]In Appendix D, I show annual fertility rate in Oklahoma and estimate the effect of universal pre-k on fertility rate using the SCM.

[15]The synthetic control estimation and placebo tests on the labor outcomes of subsamples are presented in Appendix E.

suggest that the labor outcomes of income disadvantage mothers are not so responsive to universal pre-k. The mean weekly working hours of potential mothers with family income below the poverty line in Oklahoma is higher than in the synthetic control group for a few years after the policy intervention and then drop to be close to the synthetic control group. Though the theoretical model suggests that the labor market decisions of mothers from poorer families (with a lower value of Y) are more likely to be affected by a child care policy, several free public pre-k and child care credit programs for poor families existed before universal pre-k. It's possible that mothers from high-poverty families have already enrolled their children in other public pre-k programs and participated in the labor force before the availability of universal pre-k. The CPS data shows that labor force participation rate of women whose family income is below the poverty line is about 87% before 1998. Consequently, the maternal labor supply of poorer mothers would not be significantly affected by this new child care policy. The empirical results show that the labor force participation rate of women from higher income family is also not responsive to universal pre-k. However, the mean weekly working hours of potential mothers with family income above the poverty line in Oklahoma increase after 1998 and are higher than the synthetic control group. Therefore, the results suggest that universal pre-k programs, though in high quality, may not affect the labor force participation rate of women with higher family income. This is probably because they care more about taking care of their children than working to support their families. However, for mothers who work before the implementation of universal pre-k, they are able to increase their working hours by sending their children to a free and all-day pre-k program.

The second subsample analysis is on mothers with differential marital status. The estimation results show that the treatment effect on the labor outcomes of married women is similar to the effect on all women of childbearing age. The effect of universal pre-k on the labor force participation rate of married women is positive and statistically significant. The mean working hours of the married sample present an increasing trend after 1998, and the post-trend of mean working hours of Oklahoma gradually exceed that of the synthetic control group. In the analysis of unmarried mothers, the working hours of unmarried mothers in Oklahoma are larger than the working hours of unmarried mothers in the synthetic control group during the post-intervention period, which indicates a positive effect of universal pre-k on weekly working hours of unmarried mothers. However, the labor force participation rate of unmarried mothers in Oklahoma decreases after 1998 and it is largely below the the labor force participation rate in the synthetic control group after 2001. One possible explanation is that unmarried mothers only receive financial support from the government if their income is below a threshold level. For example, chilren are not eligible for the free lunch if their family income is higher than the federal poverty line. The universal pre-k increases total income of poor family by reducing their child care payment, therefore, losses in labor income being compensated by welfare transfers from the government disincentivizes work. Further justification could be that the pre-intervention labor force participation rate of unmarried mothers in Oklahoma is already high. Since a great majority of unmarried mothers (82%) were already working, the total effect of the policy on maternal labor supply is limited.

Third, I analyze the the labor outcomes of women of childbearing age with differential education levels. I categorize education levels of potential mothers into lower than high school, high school, and college and higher than college. The empirical results show that universal pre-k has a positive impact on women that are either low or high educated, especially those with a college or higher degree. This may be explained by the quality of Oklahoma universal pre-k because mothers with higher education level are more willing to send their children to high-quality pre-k programs. There is no evidence of statistically significant effect of universal pre-k on the working hours of women with either high or low education levels. This results may not be surprising because the effect of universal pre-k on working hours are intensive margin effect, which focuses on the sample of potential mothers who already work before the universal pre-k policy. Since education level is a proxy of working skills, it is highly possible that mothers' working hours after the policy implementation are determined by their unchanged working abilities.

I also investigate the effect of universal pre-k on mothers with different numbers of children in the households because taking care of more children are more time-consuming. As expected, no significant effects exist on the labor force participation rate of mothers with more than two children. However, the empirical result shows that mothers with fewer than two children increase their labor force participation after universal pre-k becomes available. This suggests that the universal pre-k policy is more likely to increase the labor force participation rate of mothers with fewer children. The effect of universal pre-k on the working hours of mothers with either fewer or more than 2 children are statistically insignificant.

In summary, universal pre-k has differential effects on mothers from different socioeconomic groups, and universal pre-k, which provides free pre-kindergarten to all pre-k age children, also has an impact on mothers from non-disadvantaged backgrounds.

6.3 Expansion of Case Study to Georgia Universal Pre-k

In this section, I apply the synthetic control method to the case of 1995 Georgia universal pre-k. Georgia is the very first state that established universal pre-k statewide. Unlike Oklahoma, Georgia universal pre-k program is available for both 3- and 4-year-old children. As a result, the per-child expenditure on Georgia universal pre-k is much less, and Georgia universal pre-k does not require a high standard of teacher quality in state legislation. The purpose of studying Georgia universal pre-k is to see the effect of universal pre-k on the labor outcomes of a relatively low-quality pre-k state.

Figure 13 and Figure 14 show the effects of Georgia universal pre-k on female labor force participation

Figure 13. Synthetic Control Analysis on Labor Force Participation Rate

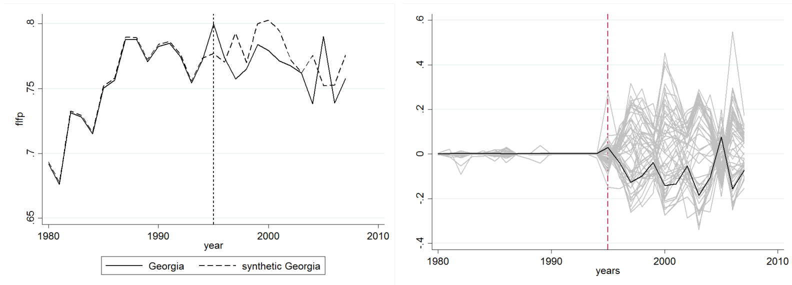
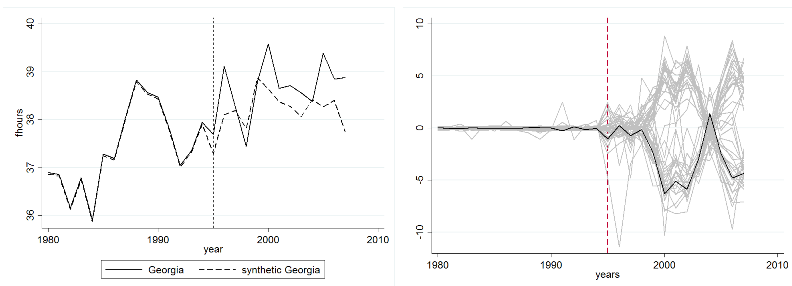


Figure 14. Synthetic Control Analysis on Weekly Working Hours



rate and working hours respectively, as well as their corresponding placebo tests. The pre-intervention period shows a good match between Georgia and its synthetic group. In the post-intervention period, both labor force participation rate and working hours of women aged between 25 to 45 are not significantly affected by Georgia universal pre-k policy. This might be explained by the possibility that maternal labor supply are more responsive to high quality universal pre-k in making labor market decisions.

7 Conclusion

This paper studies the impact of a high-quality universal pre-k program on the labor outcomes of mothers. I first present a theoretical labor supply model to predict the effect of universal pre-k on maternal labor supply. It suggests that price reduction and quality improvement in a child care program may increase the probability of a mother working but yields an ambiguous prediction regarding the working hours. In the empirical analysis, this paper chooses the high-quality Oklahoma universal pre-k program as the special case of interest. I apply a newly developed method—the synthetic control approach—to state level Current Population Survey data. To construct a better counterfactual for the treated state, Oklahoma, which is a small state with noisy data, I allow negative weights on the untreated state and add lags of the outcome variables in all pre-intervention years into the control variable set in the synthetic control analysis. I also use the placebo tests and pre/post RMSPE ratio test to investigate the significance of the treatment effect.

This paper examines the effects of universal pre-k on four labor market outcomes: labor force participation rates, employment rates, percentage of full-time labor force participation and working hours. The primary sample of interest is potential mothers, defined as women aged 25 to 45. The empirical findings suggest that the 1998 Oklahoma universal pre-k policy has a positive effect on the labor force participation rate, employment rate and weekly working hours of potential mothers in Oklahoma. The empirical results agree with the theoretical model predictions. And the empirics further show that the working hours of women of childbearing age are also increased by the universal pre-k policy. The empirical evidence also shows that there is little effect of universal pre-k on the labor outcomes of mothers with 4-year-olds only, though Oklahoma pre-k enrollment rate has been increased since the implementation of universal pre-k.

The analysis on heterogeneous treatment effects shows that universal pre-k has differential effects on mothers with different socioeconomic backgrounds. The universal pre-k policy increases the labor force participation rate for both low-educated (lower than high school education level) and high-educated (college or above college education level) mothers, but has no effect on high school graduates. Married women are more responsive to the universal pre-k policy than unmarried women in labor force participation, though the working hours of both married and unmarried mothers are increased by universal pre-k policy. The working hours of potential mothers whose family income is above the poverty line is increased by universal pre-k, and mothers with fewer children (no more than 2) increase their labor force participation after the policy implementation.

At the end of the empirical study, I expand the synthetic control analysis to another pre-k state—Georgia, which is believed to have a relatively lower-quality universal pre-k program compared to Oklahoma. Georgia universal pre-k is found to insignificantly affect female labor outcomes. This may suggest that mothers are not responsive to lower-quality universal pre-k if other characteristics related to maternal labor market behavior are not significantly different between Oklahoma and Georgia.

In conclusion, the universal pre-k policy not only increases the pre-k enrollment rate and school performance of pre-k age children, but also positively affect the labor outcomes of fertility age women who live in a high-quality universal pre-k state. Hence, the universal pre-k policy may help reduce the inter-generational education and income gap by providing children good starts as well as providing mothers chances to work. The empirical results of this paper shows no discernible effect of universal pre-k on the labor market decisions of mothers with young children. Future research examining the effects of pre-k, and child policies in general, should not necessarily restrict the sample to households currently with young children. Moreover, future studies on efficiency or benefits to costs of universal pre-k programs should consider the welfare of all family members, and policy makers should pay more attention on the quality of pre-k programs.

References

[1] Abadie, A. & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country, *American Economic Review*, 93(1), 112-132.

[2] Abadie, A., Diamond, A., & Hainmueller, J. (2012). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association.*

[3] Anderson, P. M., & Levine, P. B. (1999). Child care and mothers' employment decisions (No. w7058). National bureau of economic research.

[4] Baker, M., Gruber, J., & Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of political Economy*, 116(4), 709-745.

[5] Bauernschuster, S., Hener, T., & Rainer, H. (2015). Children of a (Policy) revolution: the introduction of universal child care and its effect on fertility. *Journal of the European Economic Association.*

[6] Baughman, R., & Dickert-Conlin, S. (2003). Did expanding the EITC promote motherhood?. *American Economic Review*, 93(2), 247-251.

[7] Berlinski, S., & Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3), 665-680.

[8] Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of public Economics*, 93(1), 219-234.

[9] Blau, D. M., & Robins, P. K. (1989). Fertility, employment, and child-care costs. *Demography*, 26(2), 287-299.

[10] Blau, D., & Tekin, E. (2007). The determinants and consequences of child care subsidies for single mothers in the USA. *Journal of Population Economics*, 20(4), 719-741.

[11] Cascio, E. U., & Schanzenbach, D. W. (2013). The impacts of expanding access to high-quality preschool education (No. w19735). National Bureau of Economic Research.

[12] Chor, E., Andresen, M. E., & Kalil, A. (2016). The impact of universal prekindergarten on family behavior and child outcomes. *Economics of Education Review*, 55, 168-181.

[13] Connelly, R., & Kimmel, J. (2003). The effect of child care costs on the employment and welfare recipiency of single mothers. *Southern Economic Journal*, 69(3), 498-519.

[14] Drange, N., & Telle, K. (2015). Promoting integration of immigrants: Effects of free child care on child enrollment and parental employment. *Labour Economics*, 34, 26-38.

[15] Fitzpatrick, M. D. (2008). Preschoolers enrolled and mothers at work? The effects of universal pre-kindergarten. US Census Bureau Center for Economic Studies Paper No. CES-WP-08-04.

[16] Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, 92(1), 307-322.

[17] Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma an evaluation of tulsa's pre-k program. *Journal of Human resources*, 40(3), 533-558.

[18] Doudchenko, N., & Imbens, G. W. (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis (No. w22791). National Bureau of Economic Research.

[19] Goux, D., & Maurin, E. (2010). Public school availability for two-year olds and mothers' labour supply. *Labour Economics*, 17(6), 951-962.

[20] Haan, P., & Wrohlich, K. (2011). Can child care policy encourage employment and fertility?: Evidence from a structural model. *Labour Economics*, 18(4), 498-512.

[21] Heim, B. T. (2007). The incredible shrinking elasticities married female labor supply, 1978-2002. *Journal of Human resources*, 42(4), 881-918.

[22] Herbst, C. (2017). Universal child care, maternal employment, and children's long-run outcomes: Evidence from the US Lanham Act of 1940. *Journal of Labor Economics*, 35(2), 519-564.

[23] Koebel, K., & Schirle, T. (2016). The Differential Impact of Universal Child Benefits on the Labour Supply of Married and Single Mothers. *Canadian Public Policy*, 42(1), 49-64.

[24] Lundin, D., Mörk, E., & Öckert, B. (2008). How far can reduced childcare prices push female labour supply?. *Labour Economics*, 15(4), 647-659.

[25] Michalopoulos, C., Robins, P. K., & Garfinkel, I. (1992). A structural model of labor supply and child care demand. *Journal of Human Resources*, 166-203.

[26] Schlosser, A. (2005). Public preschool and the labor supply of Arab mothers: Evidence from a natural experiment. Manuscript, The Hebrew University of Jerusalem.

[27] Sarah Flood, Miriam King, Steven Ruggles, & J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

[28] Simonsen, M. (2010). Price of High-quality Daycare and Female Employment. *The Scandinavian Journal of Economics*, 112(3), 570-594.

[29] Tekin, E. (2005). Child care subsidy receipt, employment, and child care choices of single mothers. *Economics Letters*, 89(1), 1-6.

Appendix A

Robustness Check on Different Age Groups of Childbearing Age Women

Figure A1. Synthetic Control Analysis Result for Women Aged 20-45

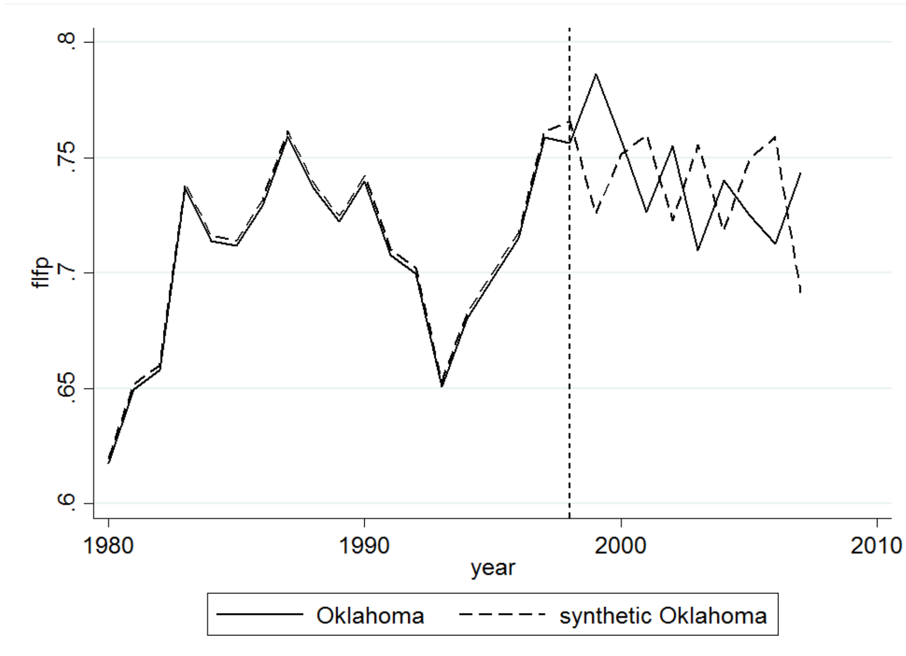


Figure A2. Synthetic Control Analysis Result for Women Aged 25-35

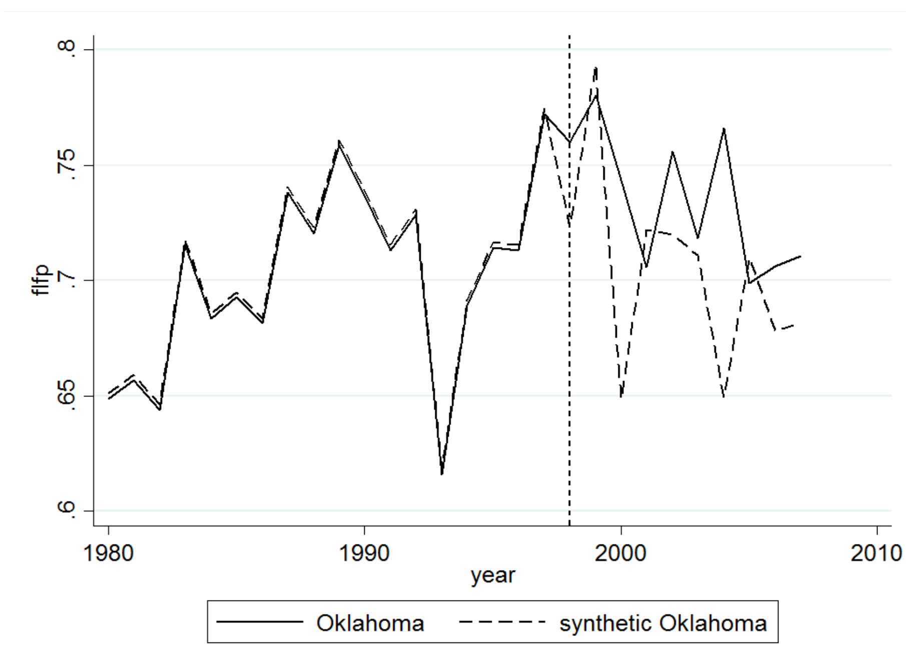


Figure A3. Synthetic Control Analysis Result for Women Aged 25-40

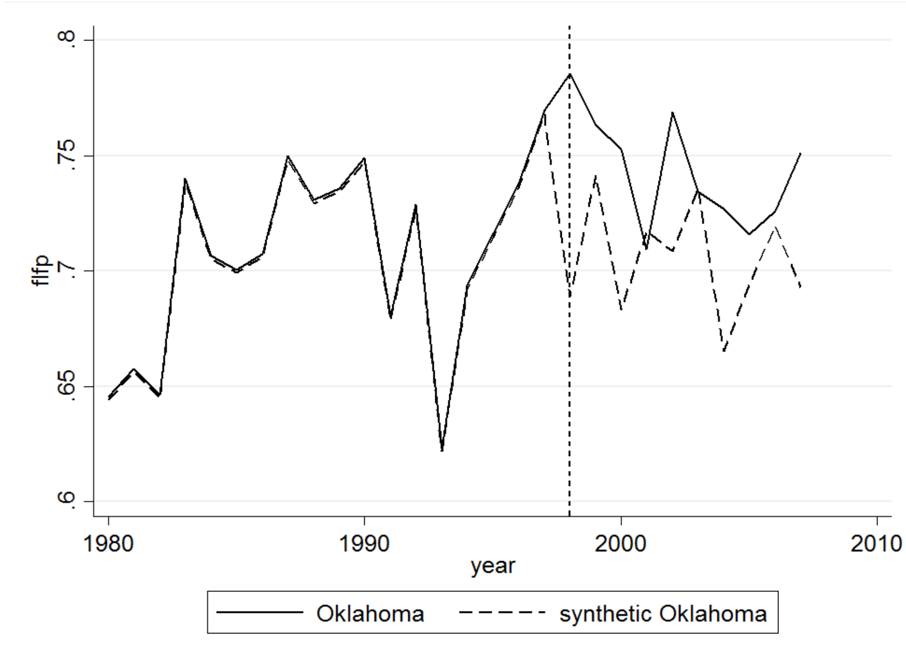
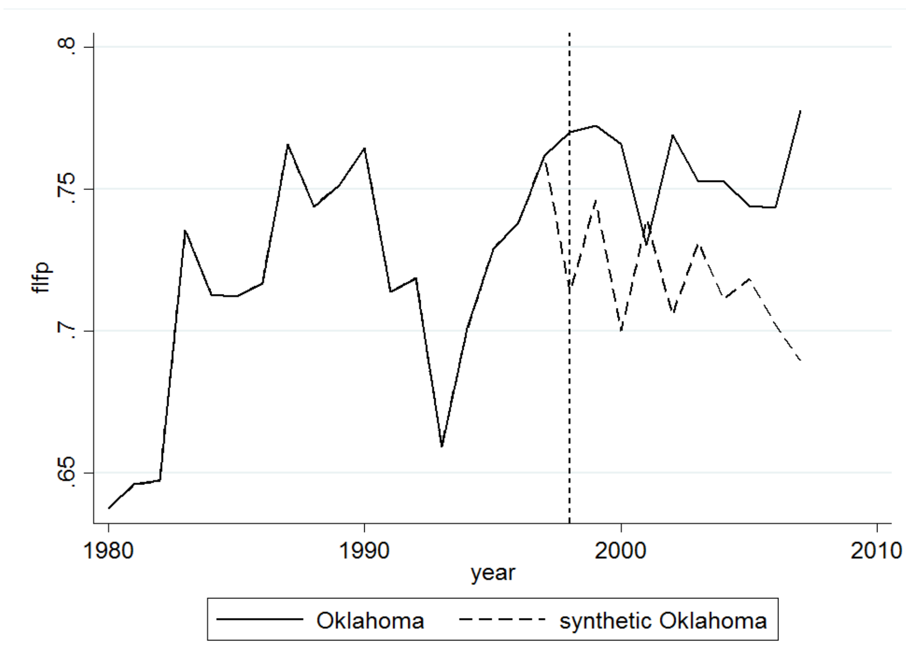


Figure A4. Synthetic Control Analysis Result for Women Aged 25-50



Appendix B

Full Set of Control Variables in the Synthetic Control Analysis on Labor Force Participation Rate

Table C1. Full Set of Explanatory Variables and Measure of Closeness

	Treated	Synthetic		Treated	Synthetic
average male education (1980-1991)			famsize1	0.133	0.158
less than high school	0.141	0.169	famsize-1 member	0.133	0.158
high school	0.344	0.356	famsize-2 members	0.167	0.179
college unfinished	0.228	0.211	famsize-3 members	0.217	0.219
college and above	0.254	0.232	famsize-4 members	0.293	0.254
average male education (1992-1997)			famsize-5 members	0.134	0.1205
less than high school	0.115	0.155	famsize-6 members	0.039	0.0438
high school	0.378	0.367	famsize-7 members	0.017	0.027
some college	0.210	0.201	married	0.733	0.678
2 years in college	0.055	0.063	separated, divorced and widowed	0.185	0.189
3 years in college	0.176	0.154	never married	0.082	0.134
college degree	0.066	0.061	not in good health(1996-2007)	0.322	0.322
average female education (1980-1991)			good health condition(1996-2007)	0.368	0.356
less than high school	0.147	0.161	very good health condition(1996-2007)	0.215	0.232
high school	0.416	0.423	excellent health condition(1996-2007)	0.095	0.092
college unfinished	0.218	0.208	male personal income	17922.22	18958.45
college and above	0.181	0.178	male working hours	40	40.04
average female education (1992-1997)			age(1980-1984)	34.14	33.79
less than high school	0.109	0.144	age(1990-1994)	34.81	34.82
high school	0.392	0.362	age(1995-1997)	35.25	35.07
some college	0.223	0.223	female lfpr(1980)	0.633	0.646
2 years in college	0.075	0.075	female lfpr(1981)	0.654	0.652
3 years in college	0.151	0.150	female lfpr(1982)	0.653	0.688
college degree	0.050	0.047	female lfpr(1981)	0.654	0.652
family total income	32673.94	33921.59	female lfpr(1982)	0.653	0.688
number of children			female lfpr(1983)	0.736	0.705
no children	0.309	0.357	female lfpr(1984)	0.719	0.718
1 child	0.211	0.214	female lfpr(1985)	0.723	0.723
2 children	0.211	0.214	female lfpr(1986)	0.721	0.711
3 children	0.300	0.260	female lfpr(1987)	0.764	0.732
4 children	0.130	0.115	female lfpr(1988)	0.737	0.726
5 children	0.036	0.038	female lfpr(1989)	0.756	0.737
more than 6	0.013	0.017	female lfpr(1990)	0.759	
number of children younger than 5			female lfpr(1991)	0.704	0.733

	Treated	Synthetic		Treated	Synthetic
1 child	0.745	0.745	female lfpr(1992)	0.718	0.749
2 children	0.192	0.191	female lfpr(1993)	0.644	0.717
3 children	0.058	0.058	female lfpr(1994)	0.701	0.712
more than 4	0.005	0.007	female lfpr(1995)	0.717	0.736
female lfpr(1996)	0.725	0.738	male lfpr(1980-1984)	0.953	0.943
female lfpr(1997)	0.759	0.749	male lfpr(1985-1989)	0.927	0.936
female lfpr 45-65 years old(1985-1989)	0.569	0.512	male lfpr(1990-1994)	0.918	0.916
female lfpr 45-65 years old(1990-1994)	0.536	0.546	male lfpr(1995-1997)	0.908	0.908
female lfpr 45-65 years old(1995-1997)	0.602	0.610	male lfpr 45-65 years old(1980-1984)	0.772	0.769
above poverty line	0.160	0.159	male lfpr 45-65 years old(1985-1989)	0.753	0.736
food stamp	0.099	0.108	male lfpr 45-65 years old(1990-1994)	0.793	0.740
free lunch	0.139	0.124	male lfpr 45-65 years old(1995-1997)	0.800	0.748
age<4	0.075	0.078	GDP	57309.33	94868.06
age4-18	0.214	0.212	white	0.853	0.823
age18-45	0.585	0.606	black	0.073	0.146
age>45	0.126	0.105	other race	0.073	0.032

Sources: IPUMS-CPS.
Notes: lfpr stands for labor force participation rates.

Appendix C

Fertility Rate of Women Aged 25-45 and its Synthetic Control Analysis

Figure C1. Pre-k Enrollment Rate in Oklahoma

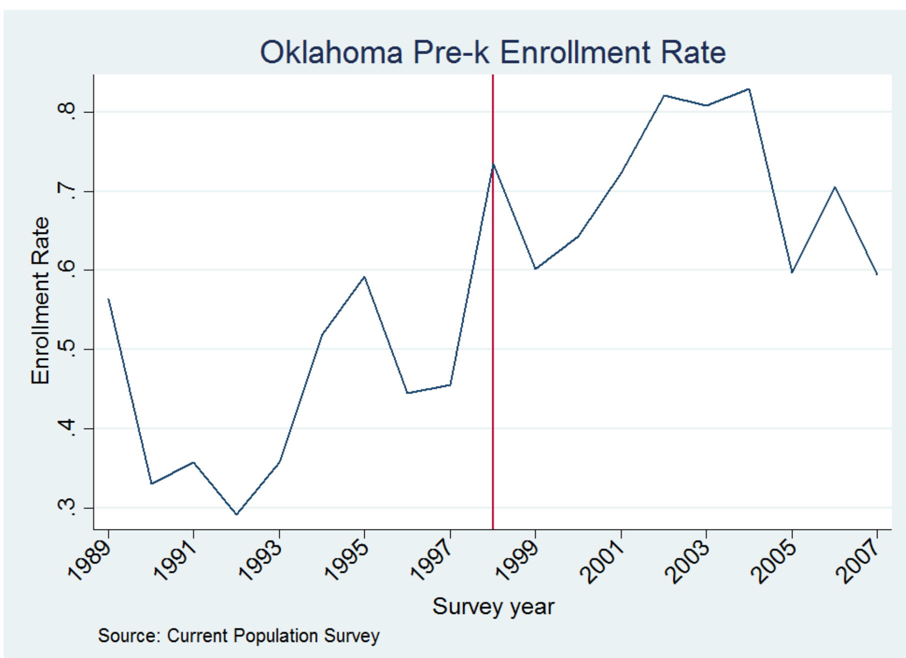


Figure C2. Synthetic Control Analysis on Oklahoma Pre-k Enrollment Rate

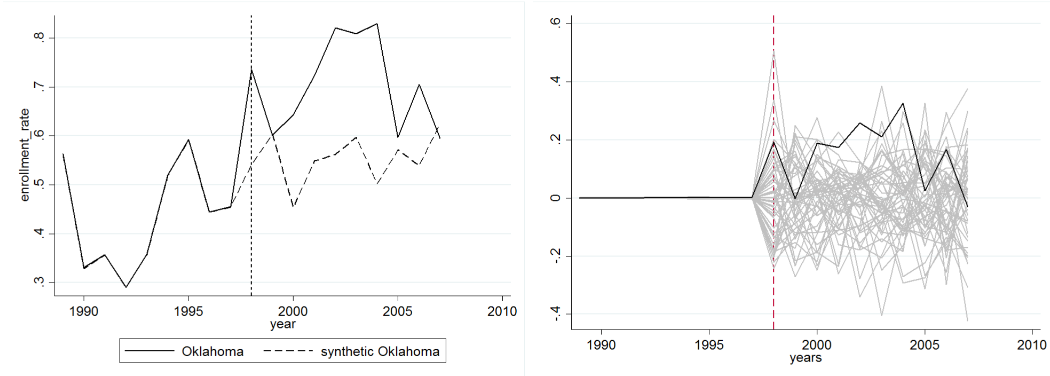
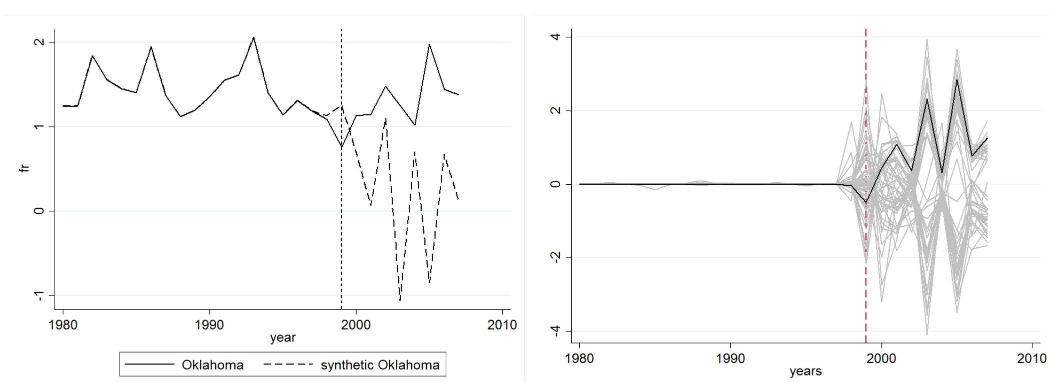


Figure C3. Fertility Rate of Oklahoma Women Aged 25-45

Figure C4. Synthetic Control Analysis on Fertility Rate of Women Aged 25-45



Appendix D

Main Results of the Effect of Oklahoma Universal Pre-k on Other Samples

Figure D1. Synthetic Control Analysis on Full-time Labor Force Participation Rate of Mothers with Children Younger than 5 Years Old

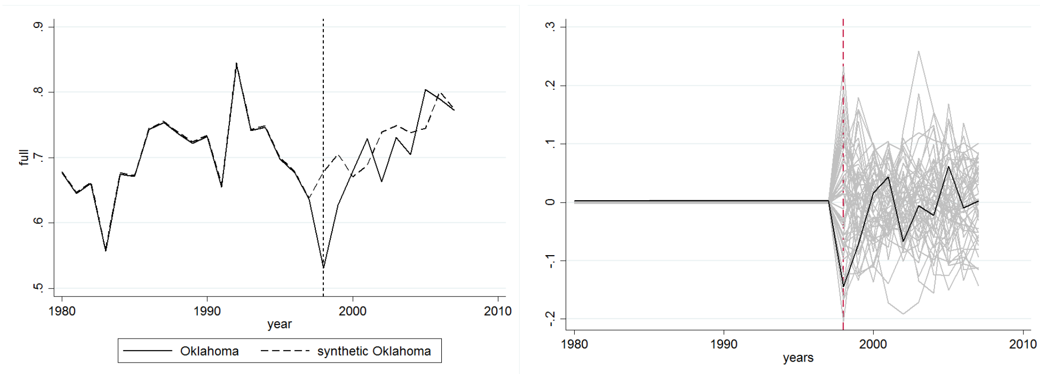


Figure D2. Synthetic Control Analysis on Employment Rate of Mothers with Children Younger than 5 Years Old

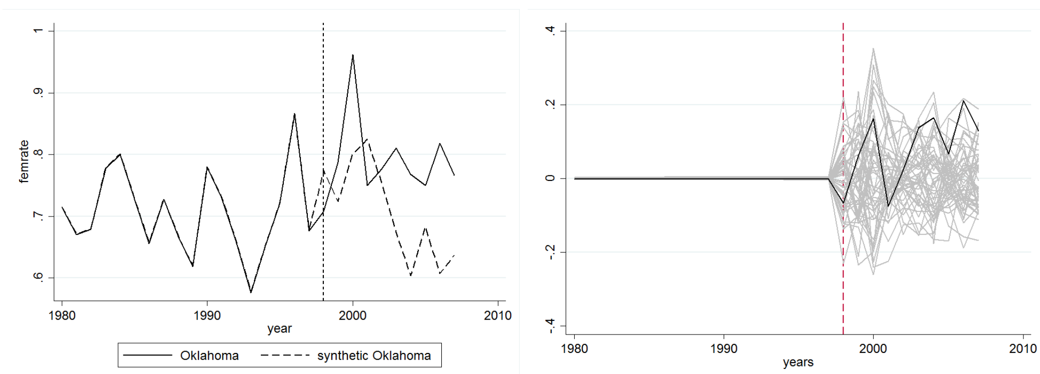


Figure D3. Synthetic Control Analysis on Employment Rate of Mothers with 4-year-olds Only

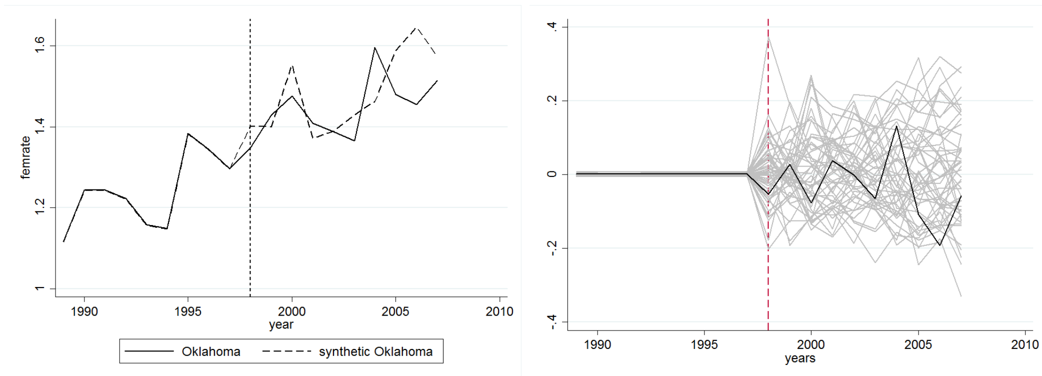


Figure D4. Synthetic Control Analysis on Full-time Labor Force Participation Rate (Georgia)

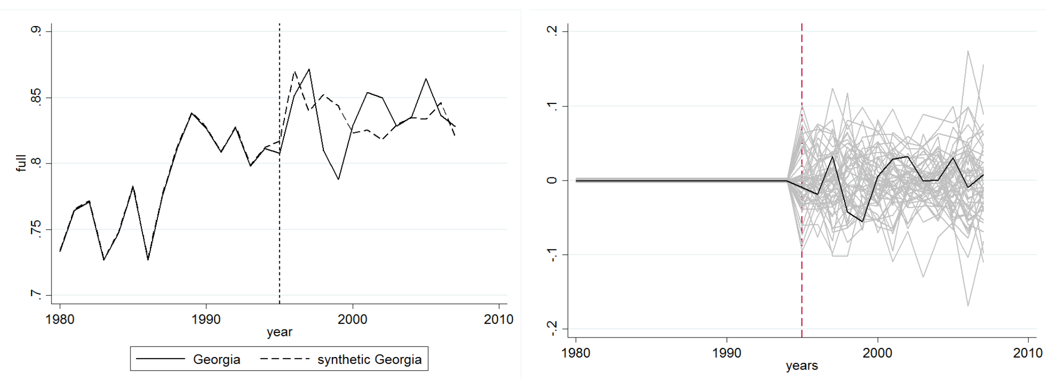
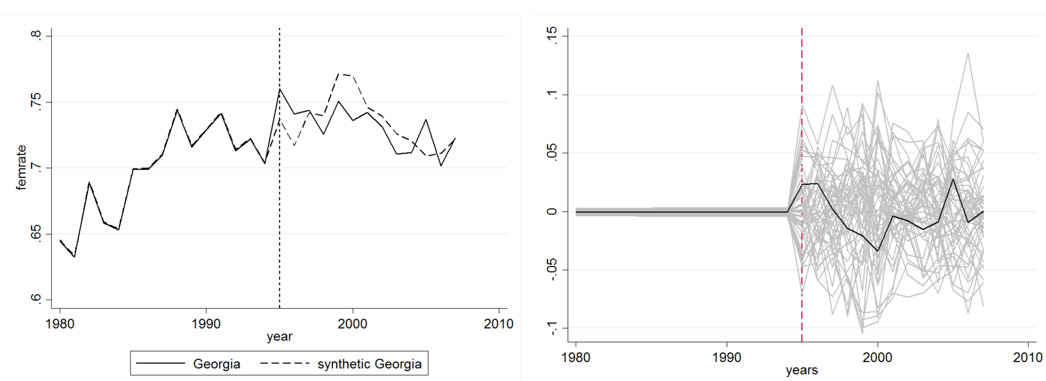


Figure D5. Synthetic Control Analysis on Employment Rate (Georgia)



Appendix E

Subsample Results of the Effect of Oklahoma Universal Pre-k

Figure E1. Synthetic Control Analysis on Labor Force Participation Rate (Below Poverty Line)

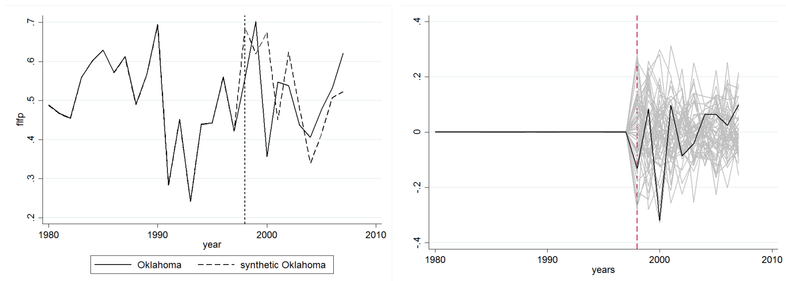


Figure E2. Synthetic Control Analysis on Weekly Working Hours (Below Poverty Line)

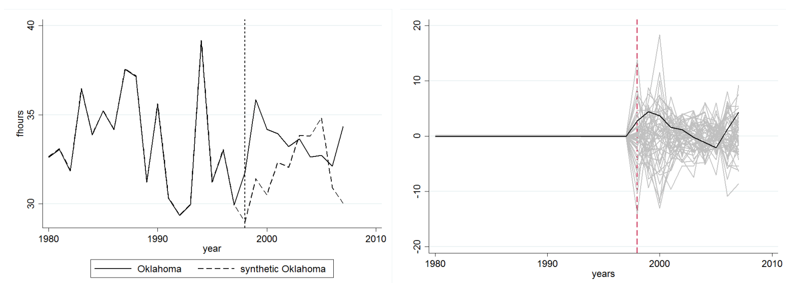


Figure E3. Synthetic Control Analysis on Labor Force Participation Rate (Above Poverty Line)

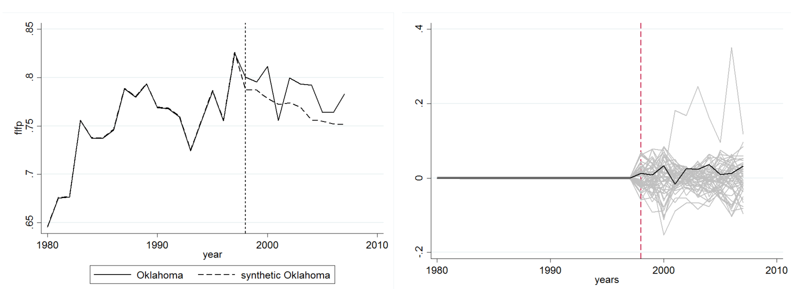


Figure E4. Synthetic Control Analysis on Weekly Working Hours (Above Poverty Line)

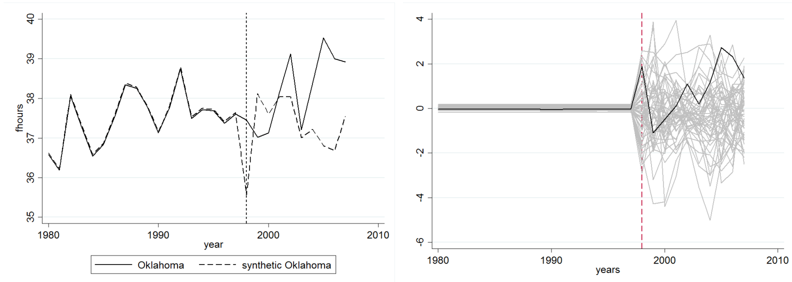


Figure E5. Synthetic Control Analysis on Labor Force Participation Rate (Married)

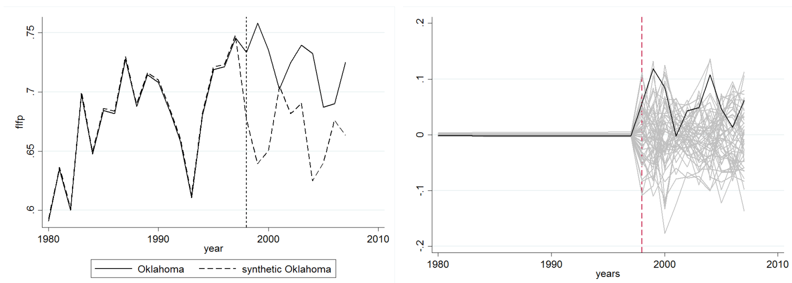


Figure E6. Synthetic Control Analysis on Weekly Working Hours (Married)

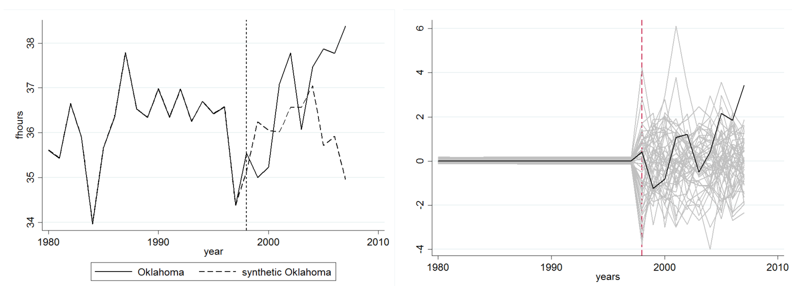


Figure E7. Synthetic Control Analysis on Labor Force Participation Rate (Unmarried)

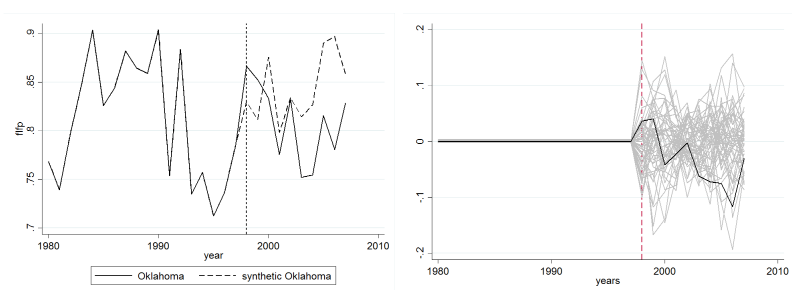


Figure E8. Synthetic Control Analysis on Weekly Working Hours (Unmarried)

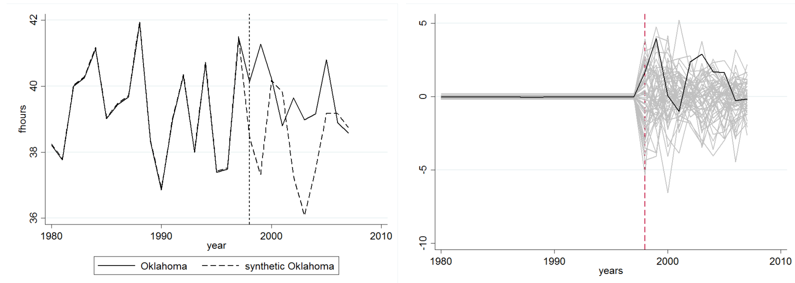


Figure E9. Synthetic Control Analysis on Labor Force Participation Rate (Differential Education Level—Lower than Highschool, High School, College and Above College by Order)

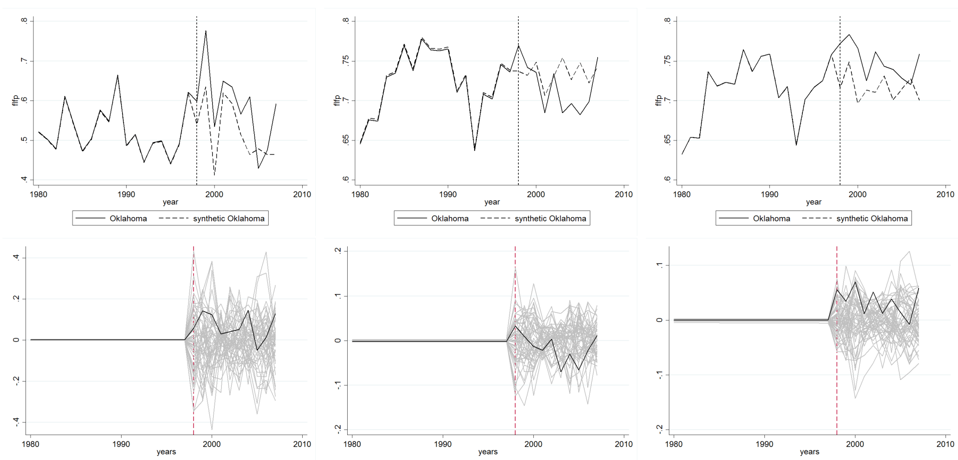


Figure E10. Synthetic Control Analysis on Weekly Working Hours (Differential Education Level—Lower than Highschool, High School, College and Above College by Order)

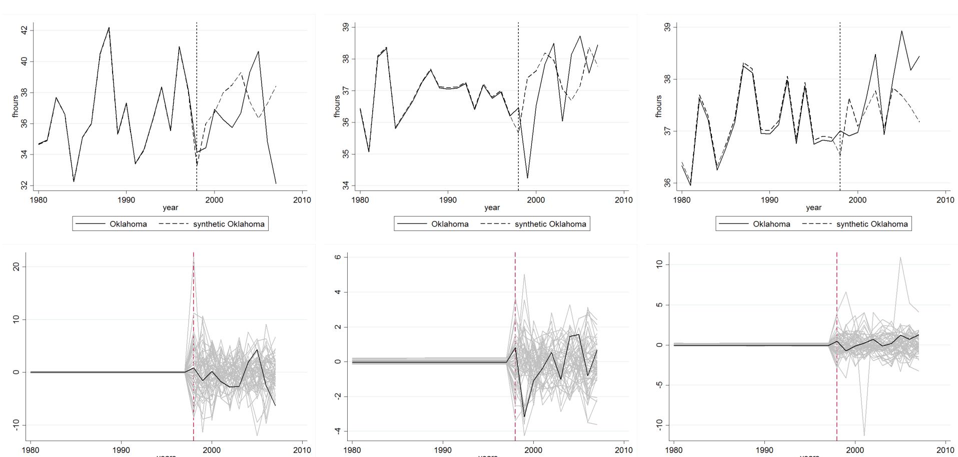


Figure E11. Synthetic Control Analysis on Labor Force Participation Rate (Fewer than 2 Children)

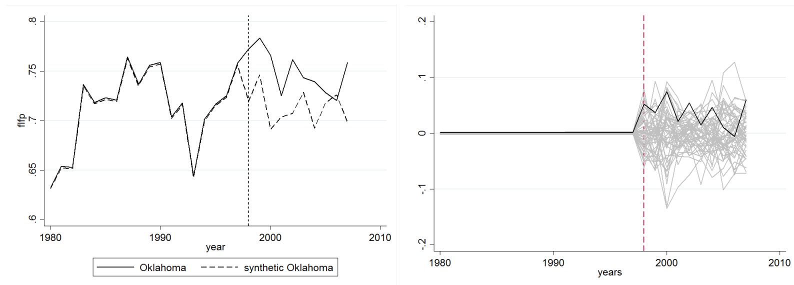


Figure E12. Synthetic Control Analysis on Weekly Working Hours (Fewer than 2 Children)

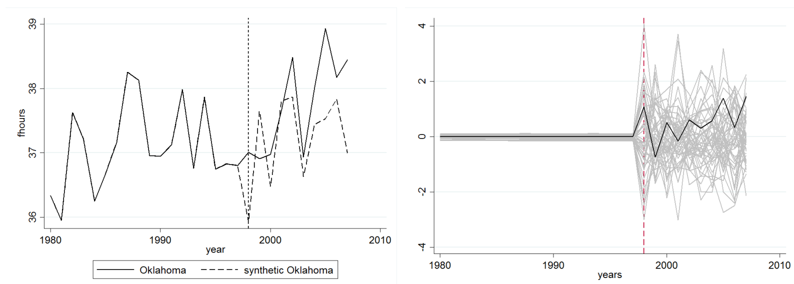


Figure E13. Synthetic Control Analysis on Labor Force Participation Rate (More than 2 Children)

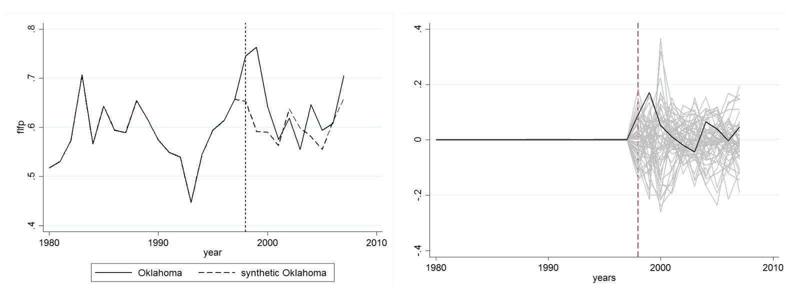


Figure E14. Synthetic Control Analysis on Weekly Working Hours (More than 2 Children)

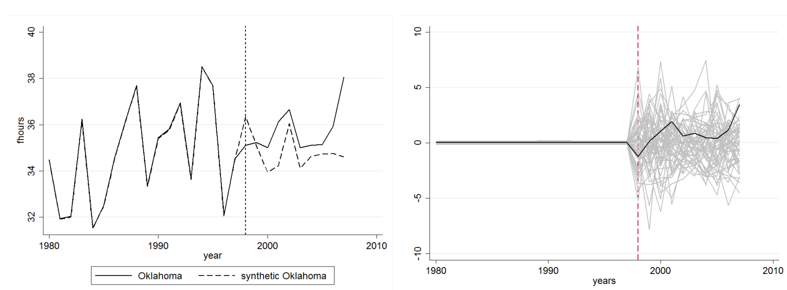


Figure E15. Synthetic control analysis on women (25-45) with less than high school education level

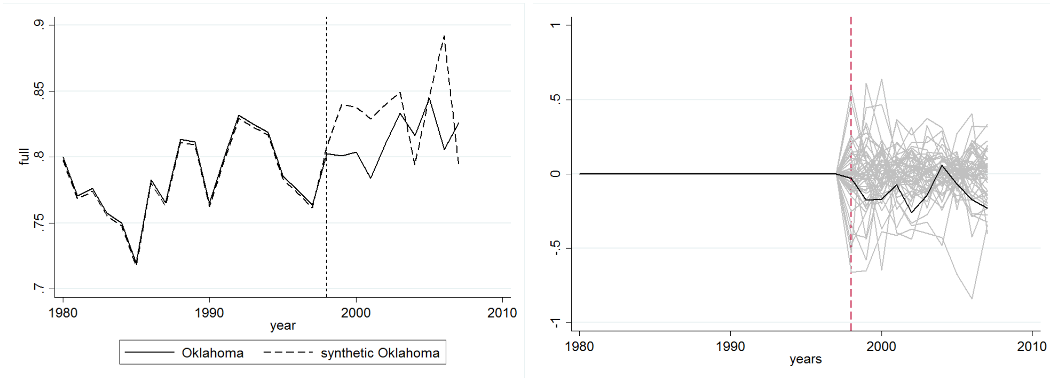


Figure E16. Synthetic control analysis on women (25-45) with high school education level

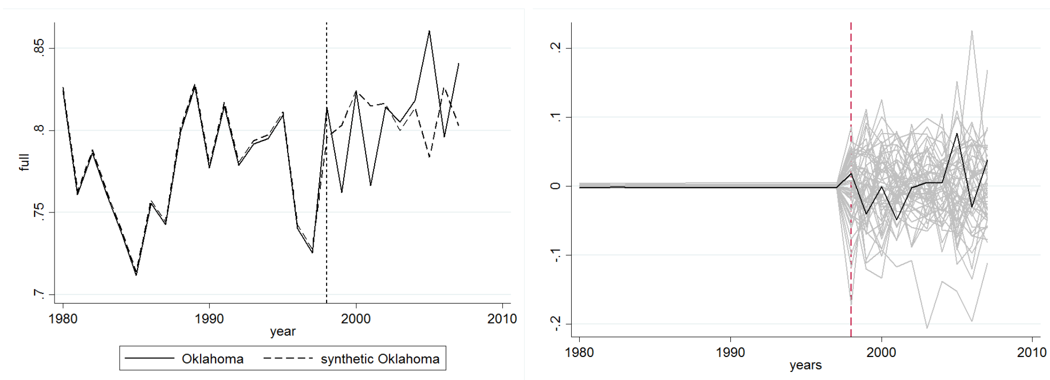


Figure E17. Synthetic control analysis on women (25-45) with college and above education level

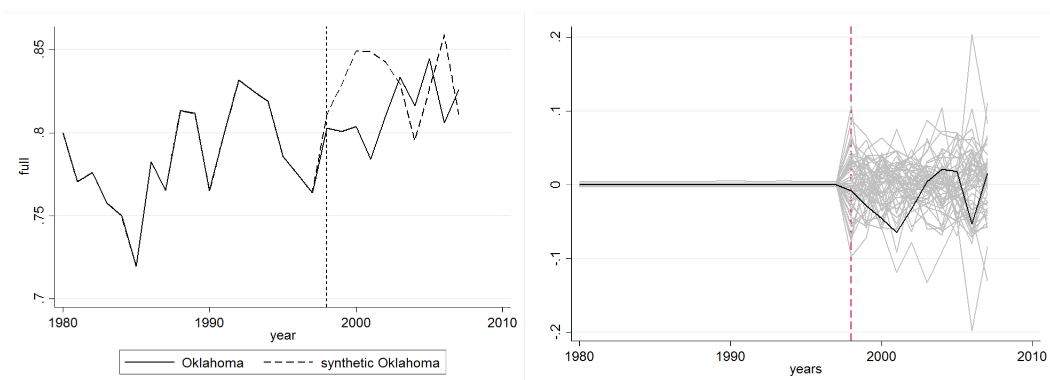


Figure E18. Synthetic control analysis on women (25-45) with family income above poverty line

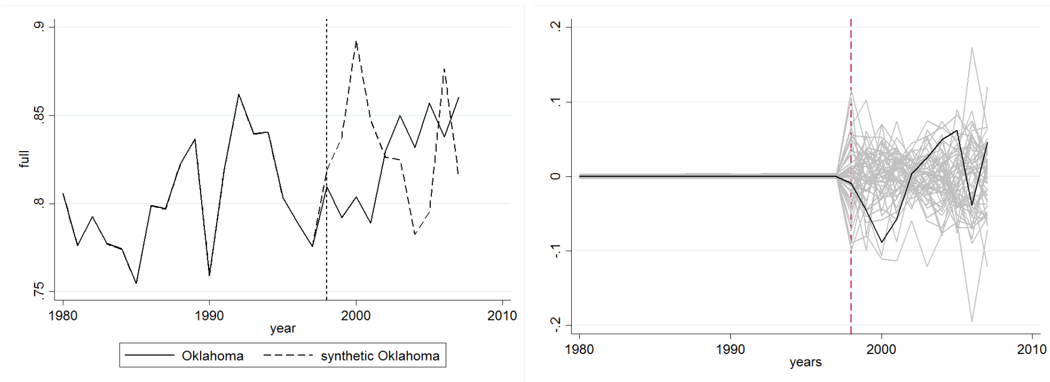


Figure E19. Synthetic control analysis on women (25-45) with family income below poverty line

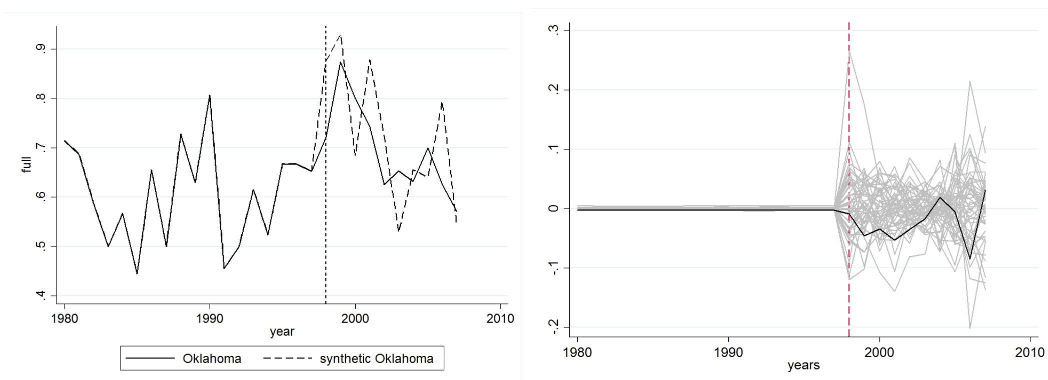


Figure E20. Synthetic control analysis on married mothers

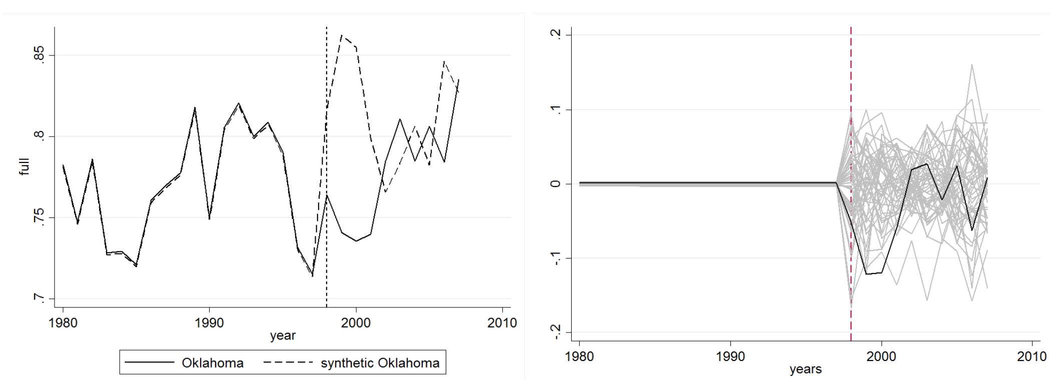


Figure E21. Synthetic control analysis on unmarried mothers

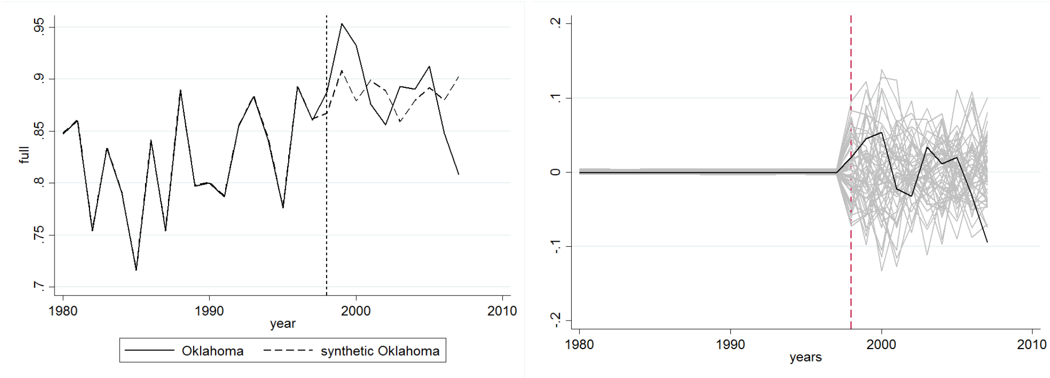


Figure E22. Synthetic control analysis on women (25-45) with more than 2 children

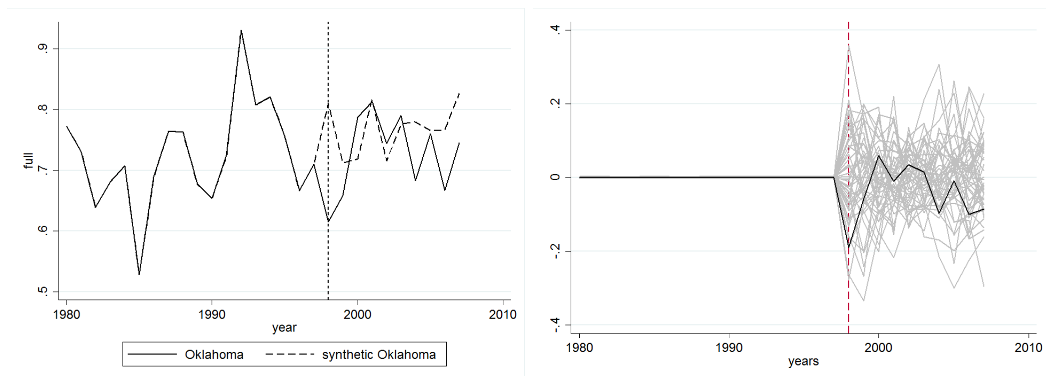


Figure E23. Synthetic control analysis on women (25-45) with no more than 2 children