# Econ 758 Homework 1

Ege Can, John Appert

February 23, 2019

## 1 Question 1

a Give a short description of the relevant aspects of the EITC expansion in 1993.(Hint: Have a look at Eissa and Hoynes, 2004.) Briefly discuss the theoretical predictions for the impact of the reform on the labor market participation of single women with children. You do not need to present a formal model!

The earned income tax credit began as a reform to traditional welfare programs in that it requires the receipient to earn income in order to receive benefits. This feature of the EITC was designed to counteract disincentives to work in traditional welfare programs.

There was an unintended effect from the program however. While the program provides increasing benefits based on the number of children in a family, it was focused on household income. This means that in a household where the mother is a secondary earner the choice to work is made after the husband's earnings are taken into account. This feature results in the decision for a women to join the labor market to occur only if the husband has not maximized the benefit of EITC.

In 1993 the EITC was modified to provide increased benefits to single parents. In this case we would expect the impact single mothers to be significantly positive as we are increasing the benefit for them to work.

b Would you expect the number of children to influence the size of the effect Why or why not? Explain.

Yes, we expect the number of children to impact the size of the effect. The more children someone has the greater the benefit for entering the workforce. Therefore, as the number of children increases we would expect a greater rise in the labor market participation rate.

c Generate a table with descriptive statistics (Table 1, structured as in Table I in Eissa and Liebman, 1996), which contains the sample means of the variables nonwhite age ed work earn for two groups: single women with and without children. You do not need to display the standard deviations. Briefly discuss the differences.

We included this table and our discussion of results as part of our answer to the next question.

d Now calculate the sample means separately for single women with one child and women with two or m?ore children (add the information to Table 1). How do they differ from each other?

Table 1 shows the mean of some of the variables for the single women in terms of the number of children they have. Firstly, single women with children are more likely to have less years of education by half years. Single women with children are more likely to join labor force by 11 percentage points. Single women without children are 12 percentage points less likely to be either black or Hispanic compared to single women with children. Additionally, women with no children has a higher earnings by about 5,700 dollars compared to single women with children. As for in group differences for women with children, single women with one child earn about 3,300 dollars on average more than single women with at least two children. Age and education level are similar but single women with single child are 12 percentage points more likely to join to the labor work force and 11 percentage less likely to have the race of either Black or Hispanic.

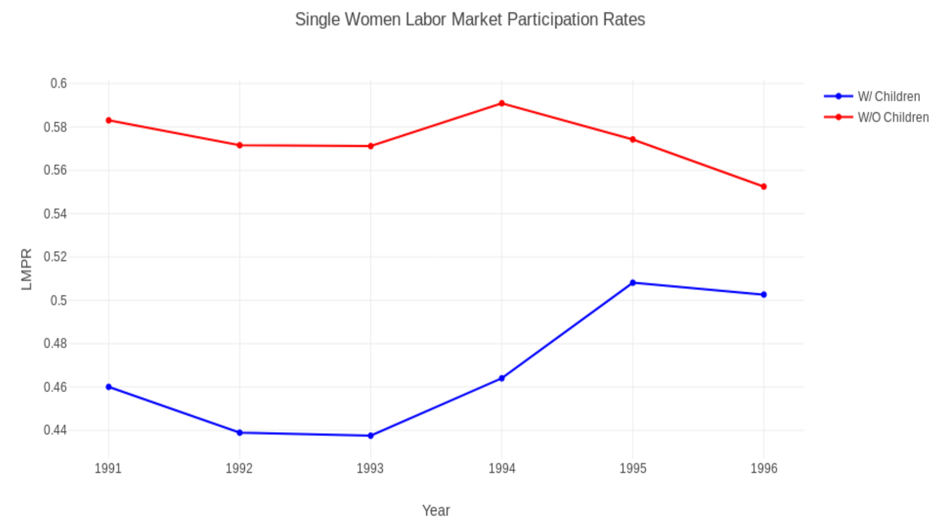
TABLE I
Summary Statistics

| Variable | Without Children (1) | With Children (2) | One Child (3) | 2 or more Children (4) |
|---|---|---|---|---|
| Nonwhite | 0.52 | 0.66 | 0.60 | 0.71 |
| Age | 38.50 | 32.72 | 33.76 | 32.05 |
| Earn | 13,760 | 7,910 | 9,928 | 6,614 |
| Work | 0.57 | 0.46 | 0.54 | 0.42 |
| Ed | 8.55 | 9.00 | 8.99 | 9.00 |
| Observations | 5,927 | 7,819 | 3,058 | 4,761 |

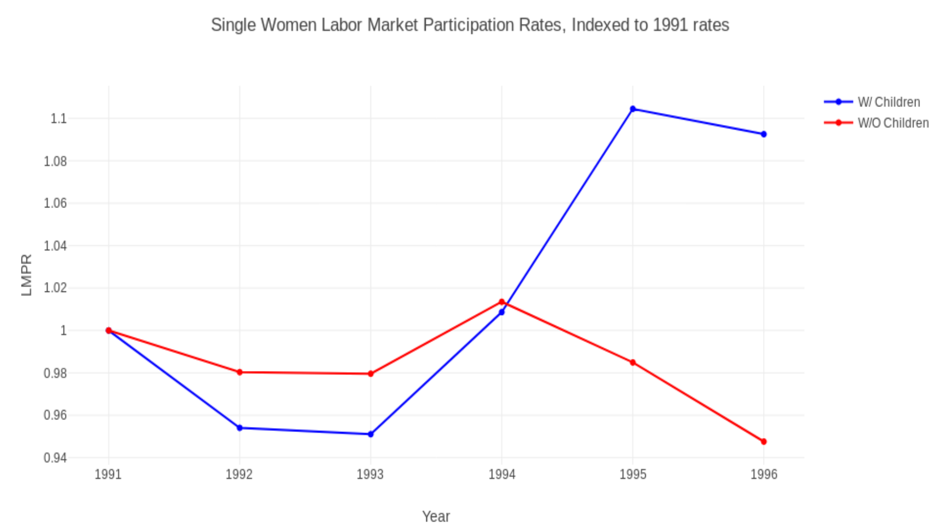
# 2 Question 2

For the following analysis you need to generate two dummy variables to identify the treatment group (single women with children) [call it child] and the post-treatment period (1994-1996) [call it post1993].

a Create a figure (Figure 1) that illustrates the annual mean labor market participation rates by year (1991-1996) for single women with children (treatment group) and single women without children (control group).Label the axes and include a title and a legend into the graph.




b Now normalize the value of the labor force participation rate for each of the two groups to group-specific 1991 values. That is, the mean of the labor market participation rates in 1991 become equal to 1. Plot a graph (as the one before, including labeling, title, and legend) in Figure 2.




c Based on Figures 1 and 2, discuss the validity of using single women without children as control group.

When looking at figure one it is difficult to determine whether or not the idea of using single women without children as a control group is valid. The levels of labor market participation are significantly different and both trends seem similar. However, once we index the labor market participation rate to 1991 and look at changes in the level with respect to 1991 we see that both groups track closely until

1994 when there is a divergence. This implies that we can use single woment without children as a control group.

d Calculate the sample means of labor force participation rates (work) of women with and without children for the pre-(average over 1991-1993) and post-reform (average over 1994-1996) period. Organize your table (Table 2) as in Table II in Eissa and Liebman (1996).

TABLE II
Labor Force Participation Rates of Single Women

| Group | Pre-1993 (1) | Post-1993 (2) | Diff (3) | Diff-in-Diff (4) |
|---|---|---|---|---|
| With Children | 0.446 (0.0076) | 0.491 (0.0084) | 0.0448 | |
| Without Children | 0.575 (0.0088) | 0.573 (0.0094) | -0.002 | 0.047 |

e Calculate the within-and between-group differences as well as the unconditional difference-in-differences estimate and add them to Table 2. Briefly comment on your results.

| TABLE II<br>Labor Force Participation Rates of Single Women | | | | |
|---|---|---|---|---|
| Group | Pre-1993<br>(1) | Post-1993<br>(2) | Diff<br>(3) | Diff-in-Diff<br>(4) |
| With Children | 0.446<br>(0.0076) | 0.491<br>(0.0084) | 0.0448 | |
| Without Children | 0.575<br>(0.0088) | 0.573<br>(0.0094) | -0.002 | 0.047 |
| Between Group Differences | -0.129 | -0.082 | | |

We find that there is an increase in the labor supply of single women with children from the pre-1993 period to the post-1993 period. This is inline with what we would expect from the theory.

f Repeat the comparison separately for women with one child and for women with at least two children for the years before and after the EITC expansion. Again compute the within-and between-group differences and the difference-in-differences estimates. Compare each of the two groups separately to single women without children (the control group). Display the results in Table 3 and discuss your findings. For which of the two groups do you find larger treatment effects? Is this consistent with the theoretical predictions?

  We find a treatment effect of approximately 3 percent for single women with one child verus a treatment effect of almost 6 percent for women with two children. This is in line with theoretical predictions as the incentive to work increases with the number of children.

TABLE III

Labor Force Participation Rates of Single Women

| | Pre-1993 (1) | Post-1993 (2) | Within Group Difference (3) | Difference-in-Differences (4) |
|---|---|---|---|---|
| A. Treatment and Treatment | | | | |
| Two or more Children | 0.396 | 0.449 | 0.031 | |
| One Child | 0.523 | 0.554 | 0.053 | 0.022 |
| Between Group Differences | -0.127 | -0.105 | | 0.022 |
| B. Treatment and Control | | | | |
| One Child | 0.523 | 0.524 | 0.031 | |
| No Children | 0.575 | 0.573 | -0.002 | 0.033 |
| Between Group Differences | -0.052 | -0.019 | | 0.033 |
| C. Treatment and Control | | | | |
| Two or More Children | 0.396 | 0.449 | 0.053 | |
| No Children | 0.575 | 0.573 | -0.002 | 0.055 |
| Between Group Differences | -0.179 | -0.124 | | 0.055 |

g Return to the comparison of women with and without children. Estimate the difference-in-differences effect from the EITC expansion by running OLS regressions. As dependent variable, use the dummy indicating labor market participation(work). First run a regression without controls ("unconditional diff-in-diff estimate"). Then add control variables (urate nonwhite age ed) to obtain the "conditional diff-in-diff estimate". Present your results (including standard errors) in Table 4 and interpret them. Compare the estimates and their statistical significance for the conditional and unconditional difference-in-differences estimates. Also comment on the estimated coefficients of child and post1993.

We ran a regression without control variables for labor rate participation. Post1993 indicates the data is after the expansion of EITC. The coefficient of this variable captures total changes in labor force participation between the two periods. The child coefficient captures the impact of having children on labor force participation. The coefficient of the interaction variable captures the impact of the expansion of EITC on the labor force participation rate for single women with at least one child.

After running the unconditional regression we obtain the results in table 4. Probability of labor participation for women with children is 12.9 percent less than women without children without considering any reform. Due to the 1993 reform, the probability of labor participation for women with children increased by 4.7 percentage compared to the women without children. The coefficient of the variable for post1993 is not significant. This variable allows us to control for trends in the data and we find no significant trend impact based on time period alone.

Controlling for factors like race, education and overall unemployment rate allow us to improve the model by removing effects that may bias our estimate. For example, as years of education increase by a year the probability of labor participation increases by 1.7 percent. Women of African-American or Hispanic heritage are 4.4 percent less likely to participate in the labor market. Including these controls improves the specification, improves the significance of the findings, and reduces the potential for unobserved variables bias.

TABLE IV
OLS Results for post 1993 Treatment

| Dependent variable: work | | |
|---|---|---|
| Variable | Unconditional (1) | Conditional (2) |
| constant | 0.575***(0.009) | 0.496 |
| child | -0.129***(0.012) | -0.118***(0.012) |
| post1993 | -0.002(0.013) | -0.0234*(0.014) |
| interactionterm | 0.046***(0.017) | 0.049***(0.017) |
| urate | | -0.016***(0.003) |
| nonwhite | | -0.044***(0.009) |
| education | | 0.017***(0.002) |
| age | | 0.002***(0.001) |
| Observations | 13476 | 13476 |
| R-Squared | 0.0126 | 0.0273 |

Standard error in parentheses
***p<0.01, **p<0.05, **p<0.1

h Estimate a conditional (i.e., including urate nonwhite age ed), “placebo” treatment model on the pre-treatment period. For this purpose, take data from the years 1991-1993 only and leave the treatment and control groups unchanged. Assume for the analysis that the placebo reform would have taken place on January 1st, 1992 (generate a dummy variable postplacebo that is one for year 1992 and after and an interaction with child) and present your results (including standard errors) in Table 5. What do you find?

We find that when building a ”placebo” model there are no significant effects on labor supply of single women with children in the period preceding the change in the EITC. This supports the argument that the change we see in the later period is due to the changes in the EITC and not some other unobserved variable. In this case we find the coefficient of the interaction variable is -0.0127 with a standard error of 0.024.

TABLE V

OLS Results for placebo Treatment

| Dependent variable: work | | |
|---|---|---|
| Variable | Unconditional | Conditional |
| | (1) | (2) |
| constant | 0.583***(0.015) | 0.540***(0.048) |
| child | -0.123***(0.020) | -0.111***(0.020) |
| postPlacebo | -0.012(0.018) | -0.000(0.018) |
| interactionterm | -0.010(0.024) | 0.049(0.024) |
| urate | | -0.021***(0.004) |
| nonwhite | | -0.039***(0.012) |
| education | | 0.016***(0.002) |
| age | | 0.002***(0.001) |
| Observations | 7,401 | 7,401 |
| R-Squared | 0.0167 | 0.0312 |

Standard error in parentheses

***p<0.01, **p<0.05, **p<0.1

# 3 Code

## 3.1 Regression Outputs from Python

| | | | |
|---|---|---|---|
| **Dep. Variable:** | work | **R-squared:** | 0.013 |
| **Model:** | OLS | **Adj. R-squared:** | 0.012 |
| **Method:** | Least Squares | **F-statistic:** | 58.45 |
| **Date:** | Wed, 20 Feb 2019 | **Prob (F-statistic):** | 1.54e-37 |
| **Time:** | 08:01:06 | **Log-Likelihood:** | -9884.9 |
| **No. Observations:** | 13746 | **AIC:** | 1.978e+04 |
| **Df Residuals:** | 13742 | **BIC:** | 1.981e+04 |
| **Df Model:** | 3 | | |

| | **coef** | **std err** | **t** | **P>\|t\|** | **[0.025** | **0.975]** |
|---|---|---|---|---|---|---|
| **const** | 0.5755 | 0.009 | 65.060 | 0.000 | 0.558 | 0.593 |
| **parent** | -0.1295 | 0.012 | -11.091 | 0.000 | -0.152 | -0.107 |
| **Post1993** | -0.0021 | 0.013 | -0.160 | 0.873 | -0.027 | 0.023 |
| **interact** | 0.0469 | 0.017 | 2.732 | 0.006 | 0.013 | 0.081 |

| | | | |
|---|---|---|---|
| **Omnibus:** | 5.965 | **Durbin-Watson:** | 1.934 |
| **Prob(Omnibus):** | 0.051 | **Jarque-Bera (JB):** | 2175.929 |
| **Skew:** | -0.051 | **Prob(JB):** | 0.00 |
| **Kurtosis:** | 1.054 | **Cond. No.** | 7.14 |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

| | | | |
|---|---|---|---|
| **Dep. Variable:** | work | **R-squared:** | 0.027 |
| **Model:** | OLS | **Adj. R-squared:** | 0.027 |
| **Method:** | Least Squares | **F-statistic:** | 55.09 |
| **Date:** | Wed, 20 Feb 2019 | **Prob (F-statistic):** | 3.84e-78 |
| **Time:** | 08:04:28 | **Log-Likelihood:** | -9781.8 |
| **No. Observations:** | 13746 | **AIC:** | 1.958e+04 |
| **Df Residuals:** | 13738 | **BIC:** | 1.964e+04 |
| **Df Model:** | 7 | | |

| | **coef** | **std err** | **t** | **P>\|t\|** | **[0.025** | **0.975]** |
|---|---|---|---|---|---|---|
| **const** | 0.4959 | 0.036 | 13.960 | 0.000 | 0.426 | 0.565 |
| **parent** | -0.1179 | 0.012 | -9.891 | 0.000 | -0.141 | -0.095 |
| **Post1993** | -0.0234 | 0.014 | -1.730 | 0.084 | -0.050 | 0.003 |
| **urate** | -0.0164 | 0.003 | -4.962 | 0.000 | -0.023 | -0.010 |
| **nonwhite** | -0.0445 | 0.009 | -4.945 | 0.000 | -0.062 | -0.027 |
| **age** | 0.0020 | 0.000 | 4.466 | 0.000 | 0.001 | 0.003 |
| **ed** | 0.0171 | 0.002 | 10.477 | 0.000 | 0.014 | 0.020 |
| **interact** | 0.0495 | 0.017 | 2.905 | 0.004 | 0.016 | 0.083 |

| | | | |
|---|---|---|---|
| **Omnibus:** | 4.872 | **Durbin-Watson:** | 1.939 |
| **Prob(Omnibus):** | 0.088 | **Jarque-Bera (JB):** | 2046.360 |
| **Skew:** | -0.046 | **Prob(JB):** | 0.00 |
| **Kurtosis:** | 1.112 | **Cond. No.** | 330. |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

| | | | |
|---|---|---|---|
| **Dep. Variable:** | work | **R-squared:** | 0.031 |
| **Model:** | OLS | **Adj. R-squared:** | 0.030 |
| **Method:** | Least Squares | **F-statistic:** | 34.06 |
| **Date:** | Wed, 20 Feb 2019 | **Prob (F-statistic):** | 4.84e-47 |
| **Time:** | 08:05:32 | **Log-Likelihood:** | -5254.1 |
| **No. Observations:** | 7401 | **AIC:** | 1.052e+04 |
| **Df Residuals:** | 7393 | **BIC:** | 1.058e+04 |
| **Df Model:** | 7 | | |

| | **coef** | **std err** | **t** | **P>\|t\|** | **[0.025** | **0.975]** |
|---|---|---|---|---|---|---|
| **const** | 0.5403 | 0.048 | 11.281 | 0.000 | 0.446 | 0.634 |
| **parent** | -0.1092 | 0.020 | -5.490 | 0.000 | -0.148 | -0.070 |
| **Post1992** | -0.0002 | 0.018 | -0.009 | 0.993 | -0.036 | 0.036 |
| **urate** | -0.0210 | 0.004 | -4.750 | 0.000 | -0.030 | -0.012 |
| **nonwhite** | -0.0394 | 0.012 | -3.265 | 0.001 | -0.063 | -0.016 |
| **age** | 0.0019 | 0.001 | 3.237 | 0.001 | 0.001 | 0.003 |
| **ed** | 0.0157 | 0.002 | 7.103 | 0.000 | 0.011 | 0.020 |
| **interact** | -0.0127 | 0.024 | -0.525 | 0.599 | -0.060 | 0.035 |

| | | | |
|---|---|---|---|
| **Omnibus:** | 0.010 | **Durbin-Watson:** | 1.968 |
| **Prob(Omnibus):** | 0.995 | **Jarque-Bera (JB):** | 1083.431 |
| **Skew:** | 0.003 | **Prob(JB):** | 5.44e-236 |
| **Kurtosis:** | 1.126 | **Cond. No.** | 328. |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

### 3.2 Python Code from Jupyter Notebook

```
import pandas as pd
import numpy as np

#import data visualization library
import plotly.plotly as py
import plotly.graph_objs as go
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot

#suppress warnings
import warnings
warnings.simplefilter("ignore")

#regression
import statsmodels.api as sm

#Read in the data
df=pd.read_stata('ps1.dta')

#Initial data munging

df['employed']=np.where(df['work']==1,1,0)
df['unemployed']=np.where(df['work']==0,1,0)
df['parent']=np.where(df['children']!=0,1,0)

#pivot by year and parent and then reset the index
df1=df.groupby(['year', 'parent']).sum()
df1=df1.reset_index()

#calculate the lfpr for both parents and no parents
df1['urate']=(df1['employed'])/(df1['employed']+df1['unemployed'])
parent=df1[df1['parent']==1]
nparent=df1[df1['parent']==0]

#Generate figure 1
# Add data
year = parent['year']
parentLMPR= parent['urate']
nparentLMPR = nparent['urate']


# Create and style traces
trace0 = go.Scatter(
    x = year,
    y = parentLMPR,
    name = 'W/ Children',
    line = dict(
        color = ('blue'),
        width = 2)
)
trace1 = go.Scatter(
    x = year,
    y = nparentLMPR,
    name = 'W/O Children',
    line = dict(
        color = ('red'),
        width = 2,)
)


data = [trace0, trace1]

# Edit the layout
layout = dict(title = 'Single Women Labor Market Participation Rates',
              xaxis = dict(title = 'Year'),
              yaxis = dict(title = 'LMPR'),
              )
```

```

fig = dict(data=data, layout=layout)
py.iplot(fig, filename='raw-plot')

#Reindexing to 1991
pBaseLevel=parent.iloc[0,3]
nBaseLevel=nparent.iloc[0,3]
parent['index']=parent['urate']/pBaseLevel
nparent['index']=nparent['urate']/nBaseLevel

#Generate figure 2
# Add data
year = parent['year']
piLMPR= parent['index']
niLMPR = nparent['index']


# Create and style traces
trace0 = go.Scatter(
    x = year,
    y = piLMPR,
    name = 'W/ Children',
    line = dict(
        color = ('blue'),
        width = 2)
)
trace1 = go.Scatter(
    x = year,
    y = niLMPR,
    name = 'W/O Children',
    line = dict(
        color = ('red'),
        width = 2,)
)


data = [trace0, trace1]

# Edit the layout
layout = dict(title = 'Single Women Labor Market
         Participation Rates, Indexed to 1991 rates',
              xaxis = dict(title = 'Year'),
              yaxis = dict(title = 'LMPR'),
              )

fig = dict(data=data, layout=layout)
py.iplot(fig, filename='index-plot')

#Calculating diff-in-diff
parent=df[df['parent']==1]
nparent=df[df['parent']!=1]

#calculate the average of the treatment group pre-1994
tc1=parent[parent['year']<1994]
tc1_empl=tc1['work'].sum()
tc1_mean=tc1_empl/len(tc1)

#calculate the average of the treatment group post-1994
tc2=parent[parent['year']>1993]
tc2_empl=tc2['work'].sum()
tc2_mean=tc2_empl/len(tc2)

#calculate the average of the control group pre-1994
cg1=nparent[nparent['year']<1994]
cg1_empl=cg1['work'].sum()
cg1_mean=cg1_empl/len(cg1)

#calculate the average of the control group post-1994
```

```python
cg2=nparent[nparent['year']>1993]
cg2_empl=cg2['work'].sum()
cg2_mean=cg2_empl/len(cg2)

#calculate diffs
dif1=tc2_mean-tc1_mean
dif2=cg2_mean-cg1_mean
dif_dif=dif1-dif2

#print (tc1_mean, tc2_mean, cg1_mean, cg2_mean)

l1=["Treatment Group", len(parent), tc1_mean, tc2_mean, dif1, '']
l2=["Control Group", len(nparent), cg1_mean, cg2_mean, dif2, dif_dif]

table=[l1, l2]

headers=['Group', 'Sample Size', 'Pre-1993',
    'Post-1993', 'Difference', 'Difference-in-differences']

table2=pd.DataFrame(table, columns=headers)

#table2

#diff-in-diff w/ one child and two children

one_child=df[df['children']==1]
two_child=df[df['children']>1]

#calculate the average of the treatment group with one child pre-1994
tg1c1=one_child[one_child['year']<1994]
tg1c1_empl=tg1c1['work'].sum()
tg1c1_mean=tg1c1_empl/len(tg1c1)

#calculate the average of the treatment group with one child post-1994
tg2c1=one_child[one_child['year']>1993]
tg2c1_empl=tg2c1['work'].sum()
tg2c1_mean=tg2c1_empl/len(tg2c1)

#calculate the average of the treatment group with two children pre-1994
tg1c2=two_child[two_child['year']<1994]
tg1c2_empl=tg1c2['work'].sum()
tg1c2_mean=tg1c2_empl/len(tg1c2)

#calculate the average of the treatment group with two child post-1994
tg2c2=two_child[two_child['year']>1993]
tg2c2_empl=tg2c2['work'].sum()
tg2c2_mean=tg2c2_empl/len(tg2c2)

#calculate diffs
dif3=tg1c2_mean-tg1c1_mean
dif4=tg2c2_mean-tg1c2_mean
dif_dif3=dif3-dif2
dif_dif4=dif4-dif2

l3=["One Child", len(one_child), tg1c1_mean, tg2c1_mean, dif3, '']
l4=["Control Group", len(nparent), cg1_mean, cg2_mean, dif2, dif_dif3]
l5=["Two Child", len(two_child), tg1c2_mean, tg2c2_mean, dif4, '']
l6=["Control Group", len(nparent), cg1_mean, cg2_mean, dif2, dif_dif4]

table=[l1, l2, l3, l4, l5, l6]

headers=['Group', 'Sample Size', 'Pre-1993', 'Post-1993',
    'Difference', 'Difference-in-differences']

table2=pd.DataFrame(table, columns=headers)

#table2

```

```
#1st regression
df['Post1993']=np.where(df['year']<1994,0,1)
df['interact']=df['Post1993']*df['parent']
X=df[['parent', 'Post1993', 'interact']]
y=df['work']
mod=sm.OLS(y, sm.add_constant(X))
res=mod.fit()
print (res.summary())

#second regression
X=df[['parent', 'Post1993','urate',
  'nonwhite', 'age', 'ed', 'interact']]
y=df['work']
mod=sm.OLS(y, sm.add_constant(X))
res=mod.fit()
print (res.summary())

#placebo regression

df2=df[df['year']<1994]
df2['Post1992']=np.where(df2['year']<1992,0,1)
df2['interact']=df2['Post1992']*df2['parent']
X=df2[['parent', 'Post1992','urate',
  'nonwhite', 'age', 'ed', 'interact']]
y=df2['work']
mod=sm.OLS(y, sm.add_constant(X))
res=mod.fit()
print (res.summary())
\end{lstlisting}
```

### 3.3 Stata Code

```
*Question 1 c
use "\\files\users\ecan\Desktop\ps1.dta", clear

*mean for education - years of education
sum ed if child>0
sum ed if child==0
*mean for nonwhite - if hispanic or black it is 1
sum nonwhite if child>0
sum nonwhite if child==0
*mean for age - age of the women
sum age if child>0
sum age if child==0
*mean for work - if worked last year
sum work if child>0
sum work if child==0
*mean for earn - annual earning
sum earn if child>0
sum earn if child==0

*Question 1 d
*mean for education - years of education
sum ed if child==1
sum ed if child>1
*mean for nonwhite - if hispanic or black it is 1
sum nonwhite if child==1
sum nonwhite if child>1
*mean for age - age of the women
sum age if child==1
sum age if child>1
*mean for work - if worked last year
sum work if child==1
sum work if child>1
*mean for earn - annual earnings
sum earn if child==1
sum earn if child>1

```

```
*Question 2
generate child=1 if children>0
replace child=0 if children==0
generate post1993=1 if year>1993
replace post1993=0 if year<1994

*Question 2 part a
**Obtaning the means for annual labor participation for women with children
sum work if child==1 & year==1991
sum work if child==1 & year==1992
sum work if child==1 & year==1993
sum work if child==1 & year==1994
sum work if child==1 & year==1995
sum work if child==1 & year==1996

collapse (mean) work, by(post1993 child)

**Obtaning the means for annual labor participation for women without children
sum work if child==0 & year==1991
sum work if child==0 & year==1992
sum work if child==0 & year==1993
sum work if child==0 & year==1994
sum work if child==0 & year==1995
sum work if child==0 & year==1996

*Question 2 d & e
**Obtaining the means for post 1993 and pre 1993 for control and treatment groups
sum work if child==1 & post1993==1
sum work if child==1 & post1993==0
sum work if child==0 & post1993==1
sum work if child==0 & post1993==0

*Question 2 f
sum work if children==1 & post1993==1
sum work if children==1 & post1993==0
sum work if children>1 & post1993==1
sum work if children>1 & post1993==0

*Question 2 g
generate interactionterm=post1993*child
reg work child post1993 interactionterm
reg work child post1993 interactionterm ed urate nonwhite age

*Question 2 h
use ”\\files\users\ecan\Desktop\ps1.dta”, clear
drop if year>1993
generate child=1 if children>0
replace child=0 if children==0
generate postplacebo=1 if year>1991
replace postplace=0 if year==1991
generate interactionterm2=postplacebo*child
reg work postplacebo child interactionterm2
reg work postplacebo child interactionterm2 urate ed nonwhite age
```