# Problem Set 4

This week we discussed to non-experimental means of generating a control group: difference-in-differences and propensity score matching. This problem set makes use of the National Supported Work (NSW) data. This data is used by LaLonde (1986) and Dehejia and Wahba (2002) to show how non-experimental estimates generated through standard econometric analysis compare to the experimental ideal.
Let us briefly recap why we need a control group. Suppose we are interested in the treatment effect on the treated (TOT). If $\tau_i = Y_{i1} - Y_{i0}$ is the treatment effect for individual $i$ then TOT is

$$E[\tau_i|T_i = 1] = E[Y_{i1}|T_i = 1] - E[Y_{i0}|T_i = 1]$$

The trouble is that we never observe $Y_{i0}|T_i = 1$! That is why we search for a group that *is not treated and behaves the same way as the treated group would have behaved in the absence of treatment*, a **control group**.
In this problem set we will walk through both difference-in-differences estimates and propensity score matching estimates. To do this, download and then read the experimental data in nswre74_treated.txt and nswre74_controls.txt using the command `infile` in STATA . You will need to specify all the variables in each file:

```
infile treatment age education black hispanic married nodegree re74 re75 re78 using
"path/nswre74_treated.txt"
```

[You will need to specify the path where the raw data has been downloaded.] Do the same thing for the control datasets (nswre74_controls). Then join the two files using the command:

```
append using filename
```

where filename is the dataset you are not currently using.
**All stata outputs are in the end of the solutions.**

# Question 1 - Estimating Causal Effects

Keep in mind the timing of the experiment. It happened between 1975 and 1977. So, data from 1974 and 1975 are pre-treatment, and data from 1978 are pos-treatment. The first part of the question uses **only** experimental data. Also, we use data from 1978 **only**.

A) **Question:** *Get the means of each variable. Test if these means differ between treatment and control group (HINT: you can use the command* `ttest varname, by(treatment)` *for the various descriptive statistics to do this.) Why are these tests helpful in establishing the credibility of the experiment?*

**Answer:** For two examples of an answer to this question check stata output 1. This and further tests show how for most observable characteristics which might influence the outcome as well there is no significant difference between the control and treatment group. It is precisely the point of random assignment that all observable and - more importantly - unobservable characteristics are in expectation the same in both the treatment and control group, so this is what we would expect under random assignment.

B) **Question:** *Estimate the treatment effect from the experiment using the outcome re78 (income in 1978). [Note: that we have perfect compliance in this case.] You can do this by estimating:*
`reg re78 treatment`
*or*
`reg re78 treatment age education black hispanic married nodegree`
*Should the treatment effects be significantly different between these two specifications? What happens to the R-squared in the second regression? Why does this matter?*

**Answer:** Using the experimental control group we estimate the treatment effect to be 886 dollars (Stata output 2).

The estimated effect when including further controls is 886 (but this difference to the previous estimate is not statistically significant), and it is significant (at least at 10%) - see Stata output 3. This is what we would expect since the controls should with random assignment be similar in both the treatment and control group and hence not matter.

The R-squared goes up in the second regression because we have reduced the residual variance, so including the covariates, however, increases efficiency of the estimation.

We now move to observational data, and see what we could have learned from it. We will use pre-treatment data, too. So we pretend we don't have experimental control data and, instead, we use observational data to construct the control group. How should we estimate the TE? If we have data from before and after the treatment (as we have), diff-in-diff might be a good option (or, sometimes, the only one).

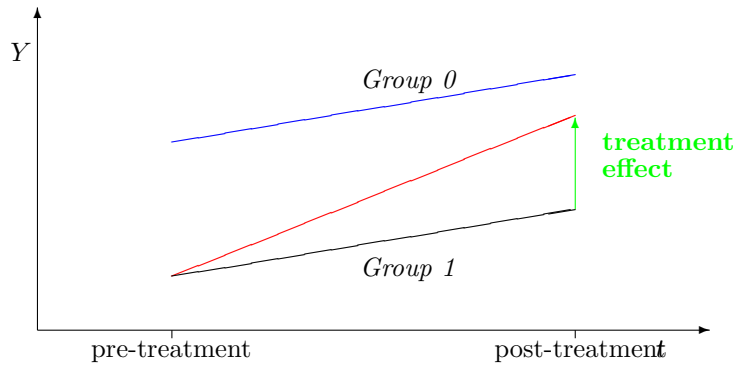
C) **Question:** *Now instead of using the true experimental controls we will use the non-experimental ones from the PSID. To do this, you must once again infile the data, this time from cps3_controls.txt.*
`infile treatment age education black hispanic married nodegree re74 re75 re78 using cps3_controls.txt`
*Append the treatment group data on once again. Now you can construct a difference-in-differences estimate. To do this, construct a before after difference for your treatment group and you control group. You can do this by typing:*
`gen ba_diff = re78 - re75`
*Then you test the significance of the difference. Do this by typing:*
`ttest ba_diff, by(treatment)`
*How do these results compare to the experimental results?*

**Answer:** Suppose we have data on the outcome we are interested in for two groups both before and after some treatment occurred: *Group* 1 was actually treated. And for some untreated group, *group* 0, we believe that the **trend of its outcome data** is the same as the trend for *group* 1 would have been if *group* 1 had not been treated.

Then we can take the actual development of the outcome for *group* 0 and together with the initial value (pre-treatment) for *group* 1 to calculate where *group* 1 would have been, if it had not been treated, i.e. $E[Y_{i0}|T_i = 1]$.

The difference between what the outcome actually was and what it would have been in the absence of treatment is our **difference-in-difference estimate** of the treatment effect.

This can be understood in the following graph, where the blue and red line show the observed data and the black line shows how we expect *group* 1 would have developed in the absence of treatment and under the common trend assumption.



Using the diff-in-diff method we estimate a treatment effect of 299 dollars of the training programme under study (Stata output 4). The estimated treatment effect is substantially smaller than the effect estimated by the experimental method. What is going on?

D) **Question:** *To construct this sample from the CPS, Lalonde tried to pick a comparable group of individuals. What would he do to test that? If you compare the various characteristics like in part A, you find many significant differences. The difference-in-difference framework allows for this however. What is the assumption that must be made to interpret the difference in difference estimate as a causal effect? Why is this important?*

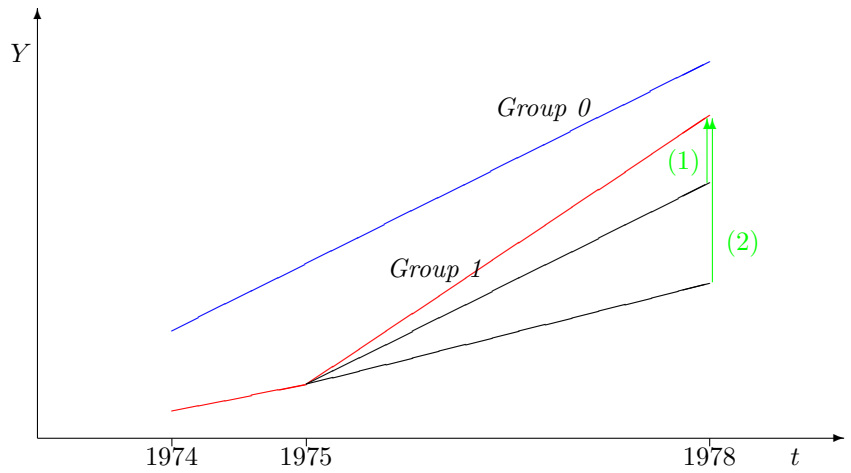
**Answer:** For the estimation of the treatment effect to work we must believe that the parallel trend assumption holds, so we must believe is that the differences between the treatment and control group are fixed over time. But we do not need to believe that the absolute values are the same. Hence the fact that many observable characteristics are different between the treatment group and the control group picked by LaLonde (1986) is as such not a problem for our diff-in-diff estimation.

E) **Question:** *Compare the difference in the pre-training incomes by constructing a difference between re74 and re75. Do the same comparison of means that you did in C, what do you find?*

**Answer:** The idea behind this question is the following: Obviously we never know how the treated would have developed in the absence of treatment, so we cannot test the crucial part of the **common-trend assumption**. However, if we can show that for some periods prior to the treatment the control and treatment groups were developing similarly, this would be a convincing argument for our control group to be a good control group.

In our case we have data for 1974 and 1975 (the treatment happened between 1975 and 1978). It turns out the trends prior to treatment are in fact not at all similar! (Stata output 5)

This explains why the diff-in-diff estimate is so much lower than the experimental estimate (which is probably close to the truth). To understand this consider the following graphs which show the diff-in-diff estimate (1) obtained under the false common trend assumption and the true treatment effect (2).[1]




## Question 2 - Propensity Score Matching

*We discussed a different way to construct a control group in which we get a propensity score for every individual in both the control and treatment group and then match based on that propensity score. For*

[1] The slopes depicted do not correspond directly to what we find in the data, but the idea is the same.

*this question, assume the propensity score for individuals with some given characteristics* $\boldsymbol{X}$ *is known (that is, for every individual, you know their propensity and that is the true value not an estimate).*
To understand the propensity score estimator as different method to create a control group think first a simple matching estimator. There we *match* treated observations to non-treated observations with similar observable characteristics $\mathbf{X}$. Then, **if** the treatment conditional on $\mathbf{X}$ is as good as random, comparing the outcomes for both groups gives an unbiased estimator. This assumption is called 'conditional independence assumption' (CIA).
However, if $\mathbf{X}$ is multidimensional this would require a lot of data - and is hence often not possible. **Propensity-score matching** allows us to reduce the 'dimensionality of the problem': It has been proven that if the CIA indeed holds, meaning $(Y_{1i}, Y_{0i})$ is independent of $T_i$ conditional on $x_i$, then in fact $(Y_{1i}, Y_{0i})$ is as well independent of $T_i$ conditional on $P(x_i)$. Hence, rather than matching on all kind of combination of the multidimensional $x_i$, we can just match on $P(x_i) = E[T_i|x_i]$, which is one-dimensional.
Note that it is important, that the CIA indeed holds!

A) **Question:** *Suppose we decided to begin by simply matching individuals on propensity scores, p(X). We then estimate a regression for each* 0.1 *interval of the propensity score. How might we do this in a single regression? What parameters would we be interested in?*

**Answer:** Suppose we know the true propensity scores (which we never do). Then we can define

- a set of dummy variables $p_m$ where $p_m = 1$ if an individuals propensity score $p(X)$ satisfies $m < p(X) \leq m + 0.1$ and $m \in M = 0, 0.1, 0.2, ..., 0.9$.
- a set of interaction terms $T_i \cdot p_{m,i}$ which is 1 if the individual is in the treatment group **and** the propensity score is between $m$ and $m + 0.1$.

We can then estimate

$$Y_i = \delta T_i + \sum_{m \in M} \gamma_m p_{m,i} + \sum_{m \in M} \beta_m (T_i \cdot p_{m,i}) + \epsilon_i$$

where we would be interested in the $\beta_m$'s.

B) **Question:** *What assumption must we make for the specification you suggested in part A to recover the cause effect of training on income?*

**Answer:** This will only work if the Conditional Independence Assumptions (CIA) does hold! So we must think that conditional on the $x_i$'s, $D_i$ is as if randomly assigned. Put another way, we must believe that the $x_i$'s fully characterize the selection into treatment and no additional variables which would as well influence the outcome influence this selection.

C) **Question:** *Some people argue that propensity scores are not very flexible because while they allow non-linearity and multiple interactions in deriving the propensity score, they are not flexible when estimating the differences. How does your answer in part A address this criticism?*

**Answer:** The answer in part A allows the treatment effect to vary across propensity scores, hence increasing the flexibility. But the fewer bins you have, the lower is this flexibility.

## Question 3 - Propensity Score Matching with Data

*Return to Stata, now to do a propensity score match. To do this, we will use a new data set so you must infile it:*

```
infile treatment age education black hispanic married nodegree re74 re75 re78 using
"cps_controls.txt"
```

*and once again append the nswre74_treated sample. Having done this you may also need to install the* `pscore` *program from Stata. To do this, simply search for the command* `pscore` *and install the relevant programs that come up.*

A) **Question:** *Begin to estimate a propensity score. We will limit our estimates to the common support (comsup) and simply estimate the propensity score in blocks (so that the mean propensity score in a block is the same). To do this, you type*

```
pscore treatment varlist , pscore(p) blockid(b) comsup
```

*You need to come up with the* `varlist` *. You can begin by including all the descriptive variables available and then progressively dropping some until the balancing property is satisfied. Recall that the balancing property requires that the X's for individuals in the control and treatment group with the same propensity score must have the same distribution of X's. What variables did you need to drop? Why might this happen? What does this imply for our interpretation of propensity scores?*

**Answer:** As we noted before, we don't know the true propensity score - we do know whether somebody was treated or not, but not what was the a priori probability that he will be treated. So it has to be estimated from the data. Typically, this is done by using a technique to handle a dummy variable on the LHS of you regression - probit or logit (more on this in a couple of weeks). `pscore` does it with probit.

If we use the **probit model** we have to drop the *nodegree* variable for the balancing property to be satisfied. This means that the distribution of the variable nodegree is quite different between the control and treatment group within some blocks of $P(X)$. Thus if certain descriptive characteristics are very skewed towards one group, they are usually excluded from a p-score estimation. If there are some characteristics that are very different for the control and treatment groups and we believe those characteristics are correlated with the outcome and the treatment probability then this will suggest that the p-score method is invalid.

B) **Question:** *Compare your p-scores between the control and treatment group. To make it easier, reduce your p-score variable to only two decimal places. You can do that quickly by typing:*

```
replace p = (int(p*100))/100
```

*Then compare your control and treatment p-scores by using the* `tab` *command and typing* `tab p treatment` *. What do you find?*

**Answer:** We see that a lot of observations in the control group are estimated to have 0 probability p-scores (Stata output 6). This highlights how many of the observations in the control group might not be comparable to the treated observations would hence not be a good control. No p-scores are estimated to be above 0.4.

C) **Question:** *Try instead, to just estimate the probability of treatment from a linear regression. To do this, type:*

```
reg treatment age education black hispanic married
predict p
replace p = (int(p*100))/100
tab p treatment
```

*Compare the* `p` *in the treatment and control group. Why is this OLS specification helpful in interpreting the predicted probability of treatment? What is a problem with the predicted values from OLS?*

**Answer:** If we use a linear probability model instead, we can see what the effect of different variables will be in changing the predicted probability of treatment. For example, a change in 1 year of schooling, reduces the probability of treatment by 0.1 percentage point (Stata output 7). In contrast, it is not transparent how a change in education affects the p-score estimate. The problem is that OLS sometimes predicts negative probability of treatment or a probability greater than 1 which do not make sense.

D) **Question:** *Estimate a regression with different dummy variables for different p-score values. To do this you can type:*

```
for num 0(0.1)0.3 \ num 0(1)3 : gen p_Y = (p > X) & (p <= X + 0.1)
```

*you will also need to define interaction terms for all these variables*

```
for num 0/3: gen p_treatX = p_X * treatment
```

*Then simply regress these dummy variables on the income in 1978 or*

```
reg re78 p_1 p_2 p_3 treatment p_treat1 p_treat2 p_treat3
```

*What do the results suggest about the significance of treatment? Is this effect constant over all values of the pscore?*

**Answer:** Finally, we use the estimated propensity scores and match the treated and untreated individuals in 4 bins of propensity scores (0.0-0.1, 0.1-0.2,...). The treatment effect in each bin of propensity score can then be estimated using the specification presented earlier. The results are in Stata output 8. The treatment effect for bins 2-4 is the estimated coefficient on the interaction term plus the effect of treatment in bin 1 (which is the baseline). The results suggest that for individuals with a propensity score of 0-0.1, the estimated difference between the treatment and control groups is $-7105.64$ dollars. The treatment effect is negative for bins 1-3 and only for bin 4 it is positive. The reason might be some unobserved characteristic which makes some individuals have a higher income in the future and hence a low propensity to be treated. For those individuals treatment might actually be bad because of the negative signal it sends or the time it needs.

E) **Question:** *Lastly, estimate a regression adjusted effect. To do this, use the regression in part D but add the control variables used in constructing the propensity score. To do this type:*

```
reg re78 p_1 p_2 p_3 treatment p_treat1 p_treat2 p_treat3 age education black
hispanic married
```

*Do your results differ from part E? Why might this be?*

**Answer:** If we now include the covariates $X$, nothing should change if the CIA indeed holds. However, the results do differ and are substantially bigger in all cases (Stata output 9) - in particular for bin 1. This suggests that the $X$'s are correlated with the outcome above and beyond the information contained in the propensity score. This should make us seriously doubt that the propensity score method is valid here. Further evidence for this is that the results are far away from the experimental benchmark.

# Stata Outputs

## Stata Output 1

```
-> ttest age, by(treatment)

Two-sample t test with equal variances

   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     425    24.44706    .3196754    6.590276    23.81871    25.0754
       1 |     297    24.62626    .3879837    6.686391    23.86271    25.38982
---------+--------------------------------------------------------------------
combined |     722    24.52078    .2465922    6.625947    24.03665    25.0049
---------+--------------------------------------------------------------------
    diff |           -.1792038    .5014259               -1.163635    .8052277
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.3574
Ho: diff = 0                     degrees of freedom =      720

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.3605         Pr(|T| > |t|) = 0.7209          Pr(T > t) = 0.6395

-> ttest education, by(treatment)

Two-sample t test with equal variances

   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     425    10.18824    .0785178    1.618686     10.0339    10.34257
       1 |     297    10.38047    .1054743    1.817712     10.1729    10.58805
---------+--------------------------------------------------------------------
combined |     722    10.26731    .0634451    1.704774    10.14275    10.39187
---------+--------------------------------------------------------------------
    diff |           -.1922361     .128823               -.4451497    .0606775
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.4922
Ho: diff = 0                     degrees of freedom =      720

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0680         Pr(|T| > |t|) = 0.1361          Pr(T > t) = 0.9320
```

## Stata Output 2

```
. reg re78 treatment

      Source |       SS       df       MS              Number of obs =     722
-------------+------------------------------           F(  1,   720) =    3.52
       Model |  137332501      1  137332501            Prob > F      =  0.0609
    Residual |  2.8053e+10   720  38962866.3           R-squared     =  0.0049
-------------+------------------------------           Adj R-squared =  0.0035
       Total |  2.8191e+10   721  39099301.3           Root MSE      =    6242

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   treatment |   886.3037   472.0863     1.88   0.061    -40.52635    1813.134
       _cons |   5090.048   302.7826    16.81   0.000     4495.606    5684.491
------------------------------------------------------------------------------
```

## Stata Output 3

```
. reg re78 treatment age education black hispanic married nodegree

      Source |       SS       df       MS              Number of obs =     722
-------------+------------------------------           F(  7,   714) =    2.58
       Model |  696533107      7  99504729.6            Prob > F      =  0.0123
    Residual |  2.7494e+10   714  38507091.2           R-squared     =  0.0247
-------------+------------------------------           Adj R-squared =  0.0151
       Total |  2.8191e+10   721  39099301.3           Root MSE      =  6205.4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   treatment |   793.6092   471.8952     1.68   0.093    -132.8588    1720.077
         age |   20.10478    36.4909     0.55   0.582    -51.53753    91.74708
   education |   205.8794   180.9277     1.14   0.256    -149.3346    561.0933
       black |  -1765.638   803.4878    -2.20   0.028     -3343.12   -188.1571
    hispanic |  -133.9468   1053.144    -0.13   0.899    -2201.575    1933.682
     married |   540.9907   644.9783     0.84   0.402    -725.2901    1807.272
   nodegree |  -522.3149   749.1767    -0.70   0.486    -1993.168    948.5378
       _cons |   4268.577   2624.619     1.63   0.104    -884.3171    9421.472
------------------------------------------------------------------------------
```

## Stata Output 4

```
. gen ba_diff =  re78 - re75

. ttest ba_diff, by(treatment)

Two-sample t test with equal variances
```

| Group | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| 0 | 429 | 4517.685 | 333.9821 | 6917.543 | 3861.236 | 5174.135 |
| 1 | 185 | 4817.088 | 608.4202 | 8275.409 | 3616.711 | 6017.465 |
| combined | 614 | 4607.896 | 296.5206 | 7347.49 | 4025.577 | 5190.216 |
| diff | | -299.4029 | 646.6766 | | -1569.377 | 970.5715 |

```
    diff = mean(0) - mean(1)                                      t =  -0.4630
Ho: diff = 0                     degrees of freedom =      612

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.3218         Pr(|T| > |t|) = 0.6435          Pr(T > t) = 0.6782
```

## Stata Output 5

```
. ttest bb_diff, by(treatment)

Two-sample t test with equal variances
```

| Group | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| 0 | 429 | -3152.752 | 280.3635 | 5806.977 | -3703.813 | -2601.691 |
| 1 | 185 | -563.5184 | 278.2071 | 3784.026 | -1112.404 | -14.63228 |
| combined | 614 | -2372.608 | 218.284 | 5408.865 | -2801.284 | -1943.933 |
| diff | | -2589.234 | 464.4893 | | -3501.42 | -1677.047 |

```
    diff = mean(0) - mean(1)                                      t =  -5.5744
Ho: diff = 0                     degrees of freedom =      612

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

## Stata Output 6

| p2 | treatment 0 | 1 | Total |
|---|---|---|---|
| 0 | 14,333 | 20 | 14,353 |
| .01 | 346 | 2 | 348 |
| .02 | 215 | 4 | 219 |
| .03 | 234 | 7 | 241 |
| .04 | 189 | 7 | 196 |
| .05 | 118 | 8 | 126 |
| .06 | 54 | 5 | 59 |
| .07 | 37 | 5 | 42 |
| .08 | 16 | 0 | 16 |
| .09 | 13 | 2 | 15 |
| .1 | 21 | 0 | 21 |
| .11 | 21 | 1 | 22 |
| .12 | 13 | 2 | 15 |
| .13 | 20 | 2 | 22 |
| .14 | 20 | 3 | 23 |
| .15 | 16 | 3 | 19 |
| .16 | 12 | 3 | 15 |
| .17 | 15 | 6 | 21 |
| .18 | 18 | 4 | 22 |
| .19 | 24 | 8 | 32 |
| .2 | 30 | 13 | 43 |
| .21 | 22 | 8 | 30 |
| .22 | 40 | 13 | 53 |
| .23 | 28 | 7 | 35 |
| .24 | 9 | 6 | 15 |
| .25 | 27 | 7 | 34 |
| .26 | 12 | 2 | 14 |
| .27 | 15 | 11 | 26 |
| .28 | 10 | 6 | 16 |
| .29 | 18 | 2 | 20 |
| .3 | 14 | 6 | 20 |
| .31 | 14 | 3 | 17 |
| .32 | 9 | 4 | 13 |
| .33 | 1 | 3 | 4 |
| .34 | 1 | 0 | 1 |
| .35 | 3 | 1 | 4 |
| .36 | 1 | 1 | 2 |
| .38 | 1 | 0 | 1 |
| .39 | 1 | 0 | 1 |
| .4 | 1 | 0 | 1 |
| Total | 15,992 | 185 | 16,177 |

## Stata Output 7

```
. reg treatment age education black hispanic married

      Source |       SS       df       MS              Number of obs =   16177
-------------+------------------------------           F(  5, 16171) =  357.10
       Model |  18.1851766     5  3.63703532            Prob > F      =  0.0000
    Residual |  164.699165 16171  .010184847           R-squared     =  0.0994
-------------+------------------------------           Adj R-squared =  0.0992
       Total |  182.884342 16176  .011305906           Root MSE      =  .10092

------------------------------------------------------------------------------
   treatment |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   -.0002708    .000081    -3.34   0.001    -.0004296   -.0001119
   education |   -.0010997   .0002873    -3.83   0.000    -.0016628   -.0005366
       black |    .1110172   .0029347    37.83   0.000     .1052648    .1167696
    hispanic |    .0057716   .0031484     1.83   0.067    -.0003996    .0119428
     married |   -.0195236   .0019524   -10.00   0.000    -.0233506   -.0156966
       _cons |     .037838   .0046064     8.21   0.000     .028809    .046867
------------------------------------------------------------------------------
```

## Stata Output 8

```
. reg re78 p_1 p_2 p_3 treatment p_treat1 p_treat2 p_treat3

      Source |       SS       df       MS              Number of obs =   16177
-------------+------------------------------           F(  7, 16169) =   46.36
       Model |  2.9769e+10     7  4.2527e+09            Prob > F      =  0.0000
    Residual |  1.4831e+12 16169    91727540           R-squared     =  0.0197
-------------+------------------------------           Adj R-squared =  0.0193
       Total |  1.5129e+12 16176  93528158.7           Root MSE      =  9577.4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         p_1 |   -4101.933   700.8708    -5.85   0.000    -5475.717   -2728.148
         p_2 |   -6940.625   690.1353   -10.06   0.000    -8293.367   -5587.884
         p_3 |   -10853.74   1694.808    -6.40   0.000    -14175.75   -7531.728
   treatment |   -7105.638   1238.823    -5.74   0.000    -9533.868   -4677.409
    p_treat1 |    4166.125   2014.548     2.07   0.039     217.3883    8114.861
    p_treat2 |    3273.189   1831.397     1.79   0.074    -316.5515    6862.929
    p_treat3 |    7550.119   3470.609     2.18   0.030     747.3407    14352.9
       _cons |    15001.49   76.74002   195.48   0.000     14851.07    15151.91
------------------------------------------------------------------------------
```

## Stata Output 9

```
. reg re78 p_1 p_2 p_3 treatment p_treat1 p_treat2 p_treat3 age education black hispanic married

      Source |       SS       df       MS              Number of obs =   16177
-------------+------------------------------           F( 12, 16164) =  145.63
       Model |  1.4761e+11    12  1.2301e+10            Prob > F      =  0.0000
    Residual |  1.3653e+12 16164  84465418.8           R-squared     =  0.0976
-------------+------------------------------           Adj R-squared =  0.0969
       Total |  1.5129e+12 16176  93528158.7           Root MSE      =  9190.5

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         p_1 |    1632.492    763.606     2.14   0.033     135.7398    3129.245
         p_2 |    191.3484      755.1     0.25   0.800    -1288.731    1671.428
         p_3 |   -2109.405    1670.47    -1.26   0.207    -5383.711    1164.902
   treatment |   -4191.879   1201.372    -3.49   0.000    -6546.701   -1837.058
    p_treat1 |    1563.486   1940.636     0.81   0.420    -2240.376    5367.349
    p_treat2 |     464.183   1765.967     0.26   0.793    -2997.308    3925.674
    p_treat3 |    4733.499   3334.938     1.42   0.156    -1803.348    11270.35
         age |    64.86673   7.426033     8.74   0.000     50.31088    79.42258
   education |    446.6367   26.30531    16.98   0.000     395.0754     498.198
       black |   -2381.292   342.9629    -6.94   0.000    -3053.537   -1709.047
    hispanic |   -858.5486   286.9083    -2.99   0.003    -1420.921   -296.1765
     married |    4637.042   184.4558    25.14   0.000     4275.488    4998.596
       _cons |    4238.746   425.3136     9.97   0.000     3405.084    5072.407
------------------------------------------------------------------------------
```