Problemt Set 2

Global Poverty and Economic Development

Wenjie Tu

Fall Semester 2021

1 Theory Question: Intertemporal Choices

Given information:

$$\begin{cases} u(c) = \ln(c) & \text{utility function} \\ f(x) = \frac{1}{\rho}x^{\rho} & \text{production function} \\ \pi_L = 1 - \pi_H = \frac{1}{2} & \text{probabilities of the state of the world} \\ p = 1 & \text{output price} \end{cases}$$

Perfect Insurance Markets

1.1

Utility maximization:

$$\begin{aligned} \max_{b,x_r,x_h} \quad & u(c_1) + \delta \mathbf{E}[u(c_2)] \\ \text{s.t.} \quad & \begin{cases} c_1 = Y - x_r - x_h - b \\ c_2 = R \cdot b + \pi_H A_H f(x_r) + (1 - \pi_H) A_L f(x_h) \end{cases} \end{aligned}$$

Rewrite maximization:

$$\mathcal{L} = u(\underbrace{Y - x_r - x_h - b}_{c_1}) + \delta \mathbf{E}[u(\underbrace{R \cdot b + \pi_H A_H f(x_r) + (1 - \pi_H) A_L f(x_h)}_{c_2})]$$

FOC w.r.t. b:

$$\frac{\partial \mathcal{L}}{\partial b} = -u'(c_1) + \delta R u'(c_2) = 0$$

$$u'(c_1) = \delta R u'(c_2) \tag{1}$$

We know that $u(c) = \ln(c)$

$$\frac{1}{c_1} = \delta R \frac{1}{c_2} \implies c_2 = \delta R c_1 \quad \text{(Euler Equation)}$$

1.2

FOC w.r.t. x_r:

$$\frac{\partial \mathcal{L}}{\partial x_r} = -u'(c_1) + \delta \pi_H A_H u'(c_2) \frac{\partial f(x_r)}{\partial x_r} = 0$$

$$u'(c_1) = \delta \pi_H A_H u'(c_2) \frac{\partial f(x_r)}{\partial x_r} \tag{2}$$

FOC w.r.t. x_h:

$$\frac{\partial \mathcal{L}}{\partial x_h} = -u'(c_1) + \delta (1-\pi_H) A_L u'(c_2) \frac{\partial f(x_h)}{\partial x_h} = 0$$

$$u'(c_1) = \delta (1-\pi_H) A_L u'(c_2) \frac{\partial f(x_h)}{\partial x_h} \tag{3}$$

Combine equations (1), (2), (3)

$$\begin{cases} u'(c_1) = \delta R u'(c_2) \\ u'(c_1) = \delta \pi_H A_H u'(c_2) \frac{\partial f(x_r)}{\partial x_r} \\ u'(c_1) = \delta (1-\pi_H) A_L u'(c_2) \frac{\partial f(x_h)}{\partial x_h} \end{cases}$$

$$R = \pi_H A_H \frac{\partial f(x_r)}{\partial x_r} = (1-\pi_H) A_L \frac{\partial f(x_h)}{\partial x_h} \tag{4}$$

$$f(x) = \frac{1}{\rho} x^\rho \implies f'(x) = x^{\rho-1}$$

Rewrite equation (4):

$$R = \pi_H A_H x_r^{\rho-1} = (1-\pi_H) A_L x_h^{\rho-1}$$

$$\begin{cases} x_r = \left(\frac{\pi_H A_H}{R}\right)^{\frac{1}{1-\rho}} \\ x_h = \left(\frac{(1-\pi_H) A_L}{R}\right)^{\frac{1}{1-\rho}} \end{cases}$$

With complete markets, the production and the consumption decisions are separable (from equation (4)). Wealth or impatience and other utility parameters do not enter the investment choice.

1.3

With Euler Equation $c_2 = \delta R c_1$, we can rewrite the second constraint $c_2 = R \cdot b + \pi_H A_H f(x_r) + (1 - \pi_H) A_L f(x_h)$:

$$\delta R c_1 = R \cdot b + \pi_H A_H f(x_r) + (1-\pi_H) A_L f(x_h)$$

Combine with the first constraint:

$$\begin{cases} c_1 = Y - x_r - x_h - b \\ \delta R c_1 = R \cdot b + \pi_H A_H f(x_r) + (1 - \pi_H) A_L f(x_h) \end{cases}$$

All the parameters are given, so we solve two equations with two unknowns (b and c_1).

Incomplete Insurance Markets

The individual has different consumption in period 2 depending on the state of the world (H or L).

$$\begin{cases} c_{2H} = R \cdot b + A_H f(x_r) & \text{with probability } \pi_H \\ c_{2L} = R \cdot b + A_L f(x_h) & \text{with probability } \pi_L \end{cases}$$

1.5

Utility maximization:

$$\max_{b, x_r, x_h} \quad u(c_1) + \delta \left(\pi_H \mathbf{E}[u(c_{2H})] + (1 - \pi_H) \mathbf{E}[u(c_{2L})] \right)$$

$$\text{s.t.} \quad \begin{cases} c_1 = Y - x_r - x_h - b \\ c_{2H} = R \cdot b + A_H f(x_r) \\ c_{2L} = R \cdot b + A_L f(x_h) \end{cases}$$

$$\mathcal{L} = u(\underbrace{Y - x_r - x_h - b}_{c_1}) + \delta \left(\pi_H \mathbf{E}[u(\underbrace{R \cdot b + A_H f(x_r)}_{c_{2H}})] + (1 - \pi_H) \mathbf{E}[u(\underbrace{R \cdot b + A_L f(x_h)}_{c_{2L}})] \right)$$

FOC w.r.t. b:

$$\frac{\partial \mathcal{L}}{\partial b} = -u'(c_1) + \delta \left(\pi_H R u'(c_{2H}) + (1 - \pi_H) R u'(c_{2L}) \right) = 0$$

Euler Equation:

$$u'(c_1) = \delta R \left(\pi_H u'(c_{2H}) + (1 - \pi_H) u'(c_{2L}) \right) \tag{5}$$

Consumption: marginal utility in period 1 is equal to the expected marginal utility in period 2.

1.6

FOC w.r.t. x_r:

$$\frac{\partial \mathcal{L}}{\partial x_r} = -u'(c_1) + \delta \pi_H A_H \frac{\partial f(x_r)}{\partial x_r} u'(c_{2H}) = 0$$

$$u'(c_1) = \delta \pi_H A_H \frac{\partial f(x_r)}{\partial x_r} u'(c_{2H}) \tag{6}$$

FOC w.r.t. x_h:

$$\frac{\partial \mathcal{L}}{\partial x_h} = -u'(c_1) + \delta(1-\pi_H)A_L \frac{\partial f(x_h)}{\partial x_h} u'(c_{2L}) = 0$$

$$u'(c_1) = \delta(1-\pi_H)A_L \frac{\partial f(x_h)}{\partial x_h} u'(c_{2L}) \tag{7}$$

$$A_H f(x_r) > A_L f(x_h) \implies c_{2H} > c_{2L} \implies u'(c_{2H}) < u'(c_{2L})$$

Combine equations (5) and (6):

$$\begin{cases} u'(c_1) = \delta R\left(\pi_H u'(c_{2H}) + (1-\pi_H)u'(c_{2L})\right) \\ u'(c_1) = \delta \pi_H A_H \frac{\partial f(x_r)}{\partial x_r} u'(c_{2H}) \end{cases}$$

Investment: underinvestment in risky input

$$\pi_H A_H \frac{\partial f(x_r)}{\partial x_r} = R\left(\pi_H + (1-\pi_H)\frac{u'(c_{2L})}{u'(c_{2H})}\right) > R$$

Combine equations (5) and (7):

$$\begin{cases} u'(c_1) = \delta R\left(\pi_H u'(c_{2H}) + (1-\pi_H)u'(c_{2L})\right) \\ u'(c_1) = \delta(1-\pi_H) A_L \frac{\partial f(x_h)}{\partial x_h} u'(c_{2L}) \end{cases}$$

Investment: overinvestment in hedging input

$$(1-\pi_H)A_L \frac{\partial f(x_h)}{\partial x_h} = R\left(\pi_H \frac{u'(c_{2H})}{u'(c_{2L})} + (1-\pi_H)\right) < R$$

Separation fails: investment now depends on wealth and utility parameters.

2 Experiments

2.1

Research question: the impact of a new type of insecticide-diffuser on malaria rates in a region of Tanzania.

We can run a natural experiment in the region of Tanzania.

Experimental strategy: randomized controlled trial

- We randomize villages and assign to 50 villages to the treatment group and the other 50 to the control group.
- In the treatment group: households use the new type of insecticide-diffuser.
- In the control group: households do not use the original type of insecticide-diffuser.

Outcome measures:

We can measure the outcome in the following way:

$$\text{Malaria Rate} = \frac{\text{The Number of Infected Individuals in Village i in One Month}}{\text{Total Population in Village i in One Month}}$$

- Malaria rate can be measured as the number of infected individuals divided by the population in a village in one month. The incubation period in most cases varies from 7 to 30 days. One month would be a reasonable period to measure the malaria infection rate
- Most malaria cases occur during rainy season. Ideally, we can conduct experiments during summer.

2.2

Spillovers:

- Spillovers within the village: there is a positive externality within the village if households use the new type of insecticide-diffuser. Even if some hoseholds in the treatment group do not use the new type of insecticide-diffuser, they can still benefit from the positive externality generated by other households that use the new insecticide-diffuser.
- Geographic spillovers between C and T villages: if the new type of insecticide-diffuser is very effective in repelling mosquitoes, these mosquitoes will migrate to the neighboring control villages, which leads to more intense transmission in villages where the new type of insecticide-diffuser is not used. With this spillover effect, the treatment effect is more likely to be overestimated.

Spillover effect is stronger within the village than across villages.

The spillovers would most likely be positive, for example if the diffuser is killing mosquitos carrying the disease, thereby lowering the chance of others being infected. In order to account for spillover effects, you would want to randomize at the village level as well as the individual level. This would allow you to capture the effects of the spillovers, by controlling the outcomes of individuals in control villages with those untreated individuals in treated villages.

2.3

2.3 (a)

I would cluster regressions at a village level to address spillovers within the village.

2.3 (b)

I will do a intention-to-treat analysis to test if attrition is a big problem. Households in C villages may use the new type of insecticide-diffuser intentionally or accidentally. I can also instrument the treatment variable with the intent-to-treat variable and estimate the local average treatment effect on compliers.

Collect data on baseline characteristics between groups (those who dropped out and those who stayed), to ensure that there was no selection issues. You would also want to test that your control and treatment groups are still comparable. Would want to ensure that the attrition is random.

2.3 (c)

If the randomization is conducted without any compromise, I would expect estimates to be the same more or less after including baseline covariates. Randomized controlled trials always produce an unbiased estimate for the treatment effect even without adjusting for any baseline covariate.

If the randomization was properly done, it should not affect the estimates significantly. It will increase precision if the control variables are correlated with the potential outcomes. These variables are not of interest to us in estimating treatment effects directly, but we want to control for any variation they introduce into the dependent variable, thereby increasing precision, but not affecting treatment.

However, this does not hold if you have a sample that is so small that adding too many covariates would decrease precision slightly, since it reduces your degrees of freedom. This is mostly true if your covariates are independent of your outcome.

Finally, if the randomization was compromised, and you observe unbalance between treatment and control groups, then controlling for covariates is necessary to remove biases.

3 Data Exercise: Difference-in-Differences (DiD)

```
library(stargazer)
library(plm)
library(ggplot2)
library(dplyr)
library(did)
```

```
# read in data
d.did <- read.csv("./data_DiD.csv")
```

3.1

```
pre.treat <- mean(d.did[d.did$treatment == 1 & d.did$after == 0, "y"])

pre.control <- mean(d.did[d.did$treatment == 0 & d.did$after == 0, "y"])

post.treat <- mean(d.did[d.did$treatment == 1 & d.did$after == 1, "y"])

post.control <- mean(d.did[d.did$treatment == 0 & d.did$after == 1, "y"])
```

```
df <- data.frame(
  matrix(data=c(pre.treat, pre.control, post.treat, post.control),
         ncol=2, nrow=2)
)
rownames(df) <- c("Treatment Group", "Control Group")
colnames(df) <- c("Pre-Program", "Post-Program")

knitr::kable(df)
```

	Pre-Program	Post-Program
Treatment Group	2.548334	3.206209
Control Group	1.172755	1.172812

```
cat(sprintf("The DiD estimator is %.3f",
            (post.treat - pre.treat) - (post.control - pre.control)))
```

```
## The DiD estimator is 0.658
```

3.2

```
did.ols <- plm(y ~ after*treatment, data=d.did, effect="individual",
               model="pooling", index=c("town", "t"))
vcov.ols <- vcovHC(did.ols, type="sss", cluster="group")
se.ols <- sqrt(diag(vcov.ols))
stargazer(did.ols, se=list(se.ols), single.row=T, keep.stat=c("rsq", "n"),
          title="OLS with SE clustered by town", header=F)
```

Table 2: OLS with SE clustered by town

	Dependent variable:
	y
after	0.0001 (0.022)
treatment	1.376*** (0.023)
after:treatment	0.658*** (0.033)
Constant	1.173*** (0.015)
Observations	5,000
R^2	0.698
Note:	*p<0.1; **p<0.05; ***p<0.01

The estimate from regression is the same as the estimate from the table in (1).

3.3

```
cat(sprintf(
"Mean value of x1 in treatment group is %.3f\n
Mean value of x1 in control group is %.3f\n
Mean value of x2 in treatment group is %.3f\n
Mean value of x2 in control group is %.3f\n",
mean(d.did[d.did$treatment==1, "x1"]),
mean(d.did[d.did$treatment==0, "x1"]),
mean(d.did[d.did$treatment==1, "x2"]),
mean(d.did[d.did$treatment==0, "x2"]))
)
```

```
## Mean value of x1 in treatment group is 0.495
##
## Mean value of x1 in control group is 0.504
##
## Mean value of x2 in treatment group is 8.006
##
## Mean value of x2 in control group is 3.008
```

The coefficient for x_1 is not significant for the first regression, indicating that x_1 is not different between treatment and control towns. On the contrary, x_2 is significantly different between the two groups. This is not necessarily a concern for the DiD approach, since x_1 and x_2 are controls, and therefore should not

directly effect our estimates for the treatment effect. What matters is that x1 and x2 are independent from the treatment assignment, and are not affected by the treatment. It doesn't affect the validity of the DiD approach as long as we control for it and as long a the parallel trend assumption holds. The coefficient is almost the same, controlling for covariates only increased precision.

```
did.ols1 <- plm(y ~ after*treatment + x1 + x2, data=d.did,
                effect="individual", model="pooling", index=c("town", "t"))
vcov.ols1 <- vcovHC(did.ols1, type="sss", cluster="group")
se.ols1 <- sqrt(diag(vcov.ols1))
stargazer(did.ols, did.ols1, se=list(se.ols, se.ols1), single.row=T, header=F,
          keep.stat=c("n", "rsq"), title="OLS with SE clustered by town")
```

Table 3: OLS with SE clustered by town

	Dependent variable:	
	y	
	(1)	(2)
after	0.0001 (0.022)	0.002 (0.017)
treatment	1.376*** (0.023)	−0.289*** (0.037)
x1		−0.562*** (0.021)
x2		0.332*** (0.006)
after:treatment	0.658*** (0.033)	0.657*** (0.024)
Constant	1.173*** (0.015)	0.456*** (0.024)
Observations	5,000	5,000
R^2	0.698	0.816
Note:		*p<0.1; **p<0.05; ***p<0.01

In table 3, we see a decrease both in the coefficient of interest and clustered standard error from column (1) to column (2) but the difference between the coefficients of interest is trivial. After controlling for covariates, the standard error would be smaller (stronger statistical power).

3.4

```
did.fe <- plm(y ~ x1 + x2 + after*treatment, data=d.did,
              effect="twoways", model="within", index=c("town", "t"))
vcov.fe <- vcovHC(did.fe, type="sss", cluster="group")
se.fe <- sqrt(diag(vcov.fe))
stargazer(did.ols, did.ols1, did.fe, se=list(se.ols, se.ols1, se.fe),
          single.row=T, keep.stat=c("n", "rsq"), header=F,
          title="FE regression with SE clustered by town")
```

The coefficient of interest from FE regression is similar to the one from OLS (without control variables) and OLS (with control variables) because we use clustered standard errors in all regression. Fixed-effects estimation takes into account unobserved unit-sepcific and time-specific confounders at the same time while cluster-adjusted standard errors account for within-cluster correlation or heteroscedasticity which the fixed-effects estimator does not take into account.

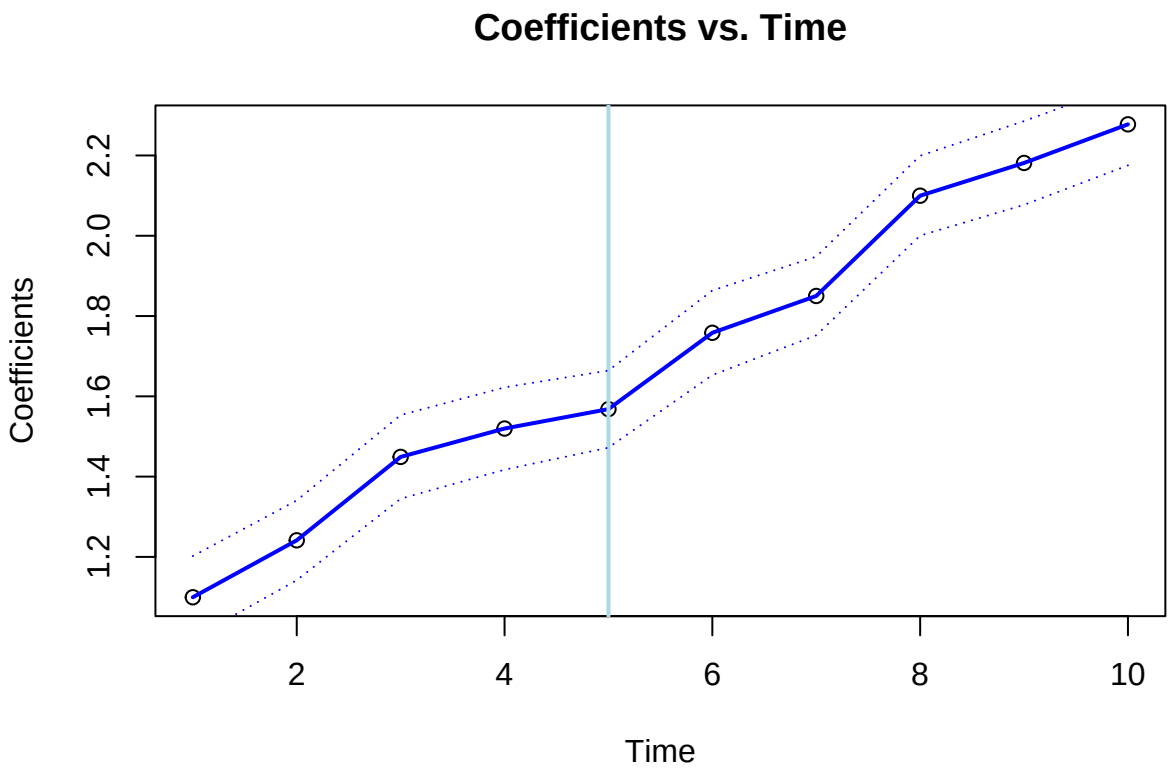
Table 4: FE regression with SE clustered by town

	Dependent variable:		
	y		
	(1)	(2)	(3)
after	0.0001 (0.022)	0.002 (0.017)	
treatment	1.376*** (0.023)	−0.289*** (0.037)	
x1		−0.562*** (0.021)	−0.562*** (0.021)
x2		0.332*** (0.006)	0.335*** (0.006)
after:treatment	0.658*** (0.033)	0.657*** (0.024)	0.657*** (0.024)
Constant	1.173*** (0.015)	0.456*** (0.024)	
Observations	5,000	5,000	5,000
R^2	0.698	0.816	0.453
Note:			*p<0.1; **p<0.05; ***p<0.01

3.5

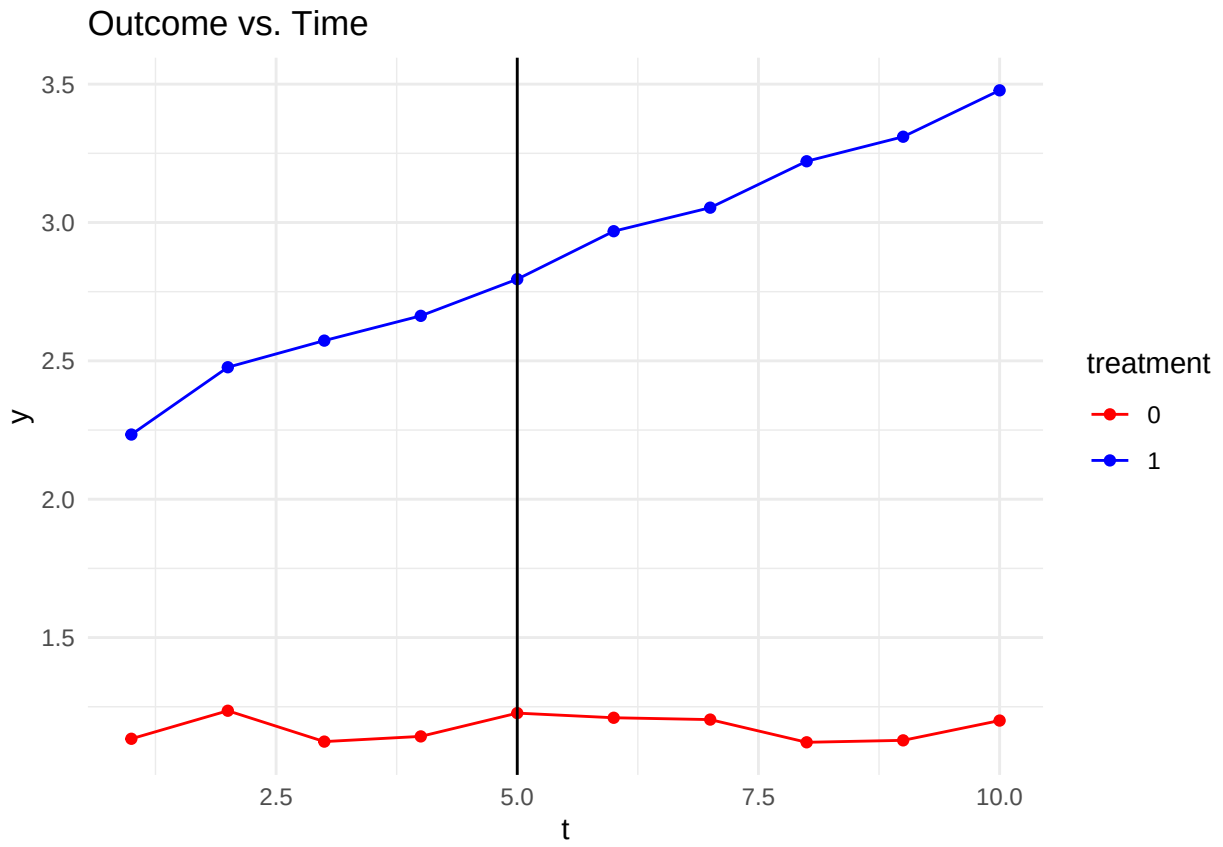
```
ate <- rep(0, 10)
se <- rep(0, 10)
for (i in 1:10) {
  indicator <- d.did$t == i
  lm.fit <- lm(y ~ treatment, data=d.did, subset=indicator)
  ate[i] <- coef(summary(lm.fit))[2, 1]
  se[i] <- coef(summary(lm.fit))[2, 2]
}
se.bands <- cbind(ate+2*se, ate-2*se) # 95% confidence interval
```

```
plot(1:10, ate, xlab="Time", ylab="Coefficients")
lines(ate, lwd=2, col="blue")
abline(v=5, col="lightblue", lwd=2)
matlines(1:10, se.bands, lwd=1, col="blue", lty=3)
title(main="Coefficients vs. Time")
```




```
d.plot <- d.did %>%
  group_by(treatment, t) %>%
  summarise(y=mean(y), .groups="drop")
d.plot$treatment <- factor(d.plot$treatment)
```

```
ggplot(d.plot, aes(x=t, y=y), color=treatment) +
  geom_line(aes(color=treatment)) +
  geom_point(aes(color=treatment)) +
  geom_vline(xintercept=5) +
  scale_color_manual(values=c("red", "blue")) +
  theme_minimal() + ggtitle("Outcome vs. Time")
```



The parallel pre-trend assumption does not seem to hold.

3.6

```
sub.did <- d.did[d.did$t %in% 1:5, ]
sub.did$placebo.after <- ifelse(sub.did$t %in% 4:5, 1, 0)
```

```
did.placebo <- lm(y ~ placebo.after*treatment, data=sub.did)
did.placebo1 <- plm(y ~ placebo.after*treatment, data=sub.did,
                    effect="individual", model="pooling", index=c("town", "t"))
stargazer(did.placebo, single.row=T, header=F, keep.stat=c("rsq", "n"),
          title="Placebo Test")
```

The placebo test is performed using a fake treatment group. The fake treatment group is constructed before the program. We should expect insignificant impact in order to support parallel pre-trend assumption. In table 5, however, we see that the coefficient of interest (interaction term) is statistically different from zero. The placebo test reveals significant impact. Therefore, the parallel pre-trend assumption is not supported.

3.7

This suggests that the method used in the first two subsections gave us biased coefficient as we do not have the parallel trends assumption holding for all time periods before the treatment hits.

Table 5: Placebo Test

	Dependent variable:
	y
placebo.after	0.020 (0.033)
treatment	1.263*** (0.029)
placebo.after:treatment	0.281*** (0.047)
Constant	1.165*** (0.021)
Observations	2,500
R^2	0.598
Note:	*p<0.1; **p<0.05; ***p<0.01