

# Juanjo Carin

## W241 (Field Experiments) – Problem Set #1 – MIDS Spring 2015

January 25, 2015

### Contents

<b>W241 – Problem Set #1</b>	<b>1</b>
1. On the notation of potential outcomes: . . . . .	1
2. FE, exercise 2.2. . . . .	3
3. FE, exercise 2.3. . . . .	3
4. More practice with potential outcomes. . . . .	6
5. FE, exercise 2.5. . . . .	10
6. FE, exercise 2.6. . . . .	11
7. FE, exercise 2.8. . . . .	12
8. FE, exercise 2.9. . . . .	12
9. FE, exercise 2.12(a). . . . .	13
10. Revised version of Essay 1. . . . .	14

---

### W241 – Problem Set #1

#### 1. On the notation of potential outcomes:

##### a. Explain the notation $Y_i(1)$ .

$Y_i(1)$  stands for the outcome for the  $i$ th subject—i.e., it is a function of  $i$ —in an experiment, if that subject is exposed to the treatment.

##### b. Explain the notation $E[Y_i(1) \mid d_i = 0]$ .

$E[Y_i(1) \mid d_i = 0]$  is the expectation of  $Y_i(1)$ —the *treated potential outcome* of the  $i$ th subject—when one subject (the  $i$ th one) is selected at random from those subjects that were not treated.

So it refers to a specific realization, and it is not to be confused with  $E[Y_i(1) \mid D_i = 0]$ , which is the conditional expectation of the treated potential outcome of the  $i$ th subject, given that he or she was actually **not** treated, and hence considers all possible realizations of  $D_i$  (while  $d_i$  is just one of them).

c. Explain the difference between the notation  $E[Y_i(1)]$  and the notation  $E[Y_i(1) | d_i = 1]$ .

$E[Y_i(1)]$  is the expectation of  $Y_i(1)$ —the *treated outcome* of the  $i$ th subject—when one subject is sampled at random, while  $E[Y_i(1) | d_i = 1]$  is the expectation of  $Y_i(1)$ —the *treated potential outcome* of the  $i$ th subject—when one subject is selected at random from those subjects that were treated. I.e., in the latter case we are only considering a subgroup, from a specific realization of the treatment allocation.

d. Explain the difference between the notation  $E[Y_i(1) | d_i = 1]$  and the notation  $E[Y_i(1) | D_i = 1]$ . Use exercise 2.7 from FE to give a concrete example of the difference.

$E[Y_i(1) | d_i = 0]$  refers to a specific realization of the treatment allocation, while  $E[Y_i(1) | D_i = 1]$  is the average across all possible ways that  $d_i$  could have been allocated.

Let's use exercise 2.7 from FE—where 2 villages are allocated to the treatment group and the other 5 villages to the control group—to give a concrete example of the difference.

```
villages <- data.frame(Yi0 = c(10, 15, 20, 20, 10, 15, 15),
                      Yi1 = c(15, 15, 30, 15, 20, 15, 30))

combinations <- combn(dim(villages)[1], 2)

# expected_Yi1_di1 <- as.numeric(
#   tapply(villages$Yi1[combinations],
#         list(2:(length(villages$Yi1[combinations])+1) %/% 2),
#         mean))
expected_Yi1_di1 <- numeric(dim(combinations)[2])
for (i in c(1:dim(combinations)[2]))
  expected_Yi1_di1[i] <- mean(villages$Yi1[combinations[,i]])
expected_Yi1_di1

## [1] 15.0 22.5 15.0 17.5 15.0 22.5 22.5 15.0 17.5 15.0 22.5 22.5 25.0 22.5
## [15] 30.0 17.5 15.0 22.5 17.5 25.0 22.5

(specific_value_of_expected_Yi1_di1 <-
  expected_Yi1_di1[which((combinations[1,]==3 & combinations[2,] == 7) ==
                        TRUE)])

## [1] 30

(expected_Yi1_Di1 <- mean(expected_Yi1_di1))

## [1] 20
```

As we can see, there are 21 possible ways of sampling 2 villages at random from a list of 7. The expected value of  $Y_i(1)$  for each one of them is shown above. Since the particular realization we are supposed to have selected (the 15th one) has unusually high potential treated outcomes,  $E[Y_i(1) | d_i = 1]$  is 30 for that particular case. The average value of all possible realizations is  $E[Y_i(1) | D_i = 1] = 20$ .

Of course,  $E[Y_i(1) | D_i = 1] = E[Y_i(1) | D_i = 0] = E[Y_i(1)]$ .

```
(expected_Yi1 <- mean(villages$Yi1))

## [1] 20

expected_Yi1_di0 <- numeric(dim(combinations)[2])
for (i in c(1:dim(combinations)[2]))
  expected_Yi1_di0[i] <- mean(villages$Yi1[-combinations[,i]])
(expected_Yi1_Di0 <- mean(expected_Yi1_di0))

## [1] 20
```

---

## 2. FE, exercise 2.2.

We want to illustrate that  $E[Y_i(0)] - E[Y_i(1)] = E[Y_i(0) - Y_i(1)]$ :

```
villages$minus_AT <- villages$Yi0 - villages$Yi1
(expected_Yi0_minus_Yi1 <- mean(villages$minus_AT))

## [1] -5

expected_Yi0 <- mean(villages$Yi0)
expected_Yi0 - expected_Yi1

## [1] -5
```

---

## 3. FE, exercise 2.3.

a. Fill in the number of observations in each of the nine cells.

```
library(knitr)
Prob_table <- as.matrix(table(villages[, c("Yi0", "Yi1")]))
rownames(Prob_table) <- paste("Yi(0)=", rownames(Prob_table), sep = "")
colnames(Prob_table) <- paste("Yi(1)=", colnames(Prob_table), sep = "")
kable(Prob_table, format = "markdown")
```

	Yi(1)=15	Yi(1)=20	Yi(1)=30
Yi(0)=10	1	1	0
Yi(0)=15	2	0	1
Yi(0)=20	1	0	1

b. Indicate the percentage of all subjects that fall into each of the nine cells. (These cells represent what is known as the joint frequency distribution of  $Y_i(0)$  and  $Y_i(1)$ .)

```
Prob_table <- round(prop.table(Prob_table), 3)
kable(Prob_table, format = "markdown")
```

	Yi(1)=15	Yi(1)=20	Yi(1)=30
Yi(0)=10	0.143	0.143	0.000
Yi(0)=15	0.286	0.000	0.143
Yi(0)=20	0.143	0.000	0.143

c. At the bottom of the table, indicate the proportion of subjects falling into each category of  $Y_i(1)$ . (These cells represent what is known as the marginal distribution of  $Y_i(1)$ .)

```
Prob_table <- rbind(Prob_table,
  as.numeric(round(prop.table(margin.table(
    table(villages[, c("Yi0", "Yi1")]), 2)), 3))
rownames(Prob_table)[4] <- "Marginal dist. of Yi(1)"
kable(Prob_table, format = "markdown")
```

	Yi(1)=15	Yi(1)=20	Yi(1)=30
Yi(0)=10	0.143	0.143	0.000
Yi(0)=15	0.286	0.000	0.143
Yi(0)=20	0.143	0.000	0.143
Marginal dist. of Yi(1)	0.571	0.143	0.286

d. At the right of the table, indicate the proportion of subjects falling into each category of  $Y_i(0)$  (i.e., the marginal distribution of  $Y_i(0)$ ).

```
Prob_table <- cbind(Prob_table,
  round(c(as.numeric(prop.table(margin.table(
    table(villages[, c("Yi0", "Yi1")]), 1))), 1), 3))
colnames(Prob_table)[4] <- "Marginal dist. Yi(0)"
kable(Prob_table, format = "markdown")
```

	Yi(1)=15	Yi(1)=20	Yi(1)=30	Marginal dist. Yi(0)
Yi(0)=10	0.143	0.143	0.000	0.286
Yi(0)=15	0.286	0.000	0.143	0.429
Yi(0)=20	0.143	0.000	0.143	0.286
Marginal dist. of Yi(1)	0.571	0.143	0.286	1.000

e. Use the table to calculate the conditional expectation that  $E[Y_i(0) | Y_i(1) > 15]$ .

All the values we need to know in order to calculate

$$\begin{aligned} E[Y_i(0) | Y_i(1) > 15] &= \sum_{Y_i(0)} Y_i(0) Pr[Y_i(0) | Y_i(1) > 15] = \\ &= \sum_{Y_i(0)} Y_i(0) \frac{Pr[Y_i(0), Y_i(1) > 15]}{Pr[Y_i(1) > 15]} \end{aligned}$$

already appear in the latest table (for the sake of clarity, fractions rather than decimals are used below).

$$\begin{aligned} E[Y_i(0) | Y_i(1) > 15] &= (10) \left( \frac{\frac{1}{7} + 0}{\frac{1}{7} + \frac{2}{7}} \right) + (15) \left( \frac{0 + \frac{1}{7}}{\frac{1}{7} + \frac{2}{7}} \right) + (20) \left( \frac{0 + \frac{1}{7}}{\frac{1}{7} + \frac{2}{7}} \right) = \\ &= (10 + 15 + 20) \left( \frac{1/7}{3/7} \right) = \frac{45}{3} = 15 \end{aligned}$$

We could have done it directly from R:

```
Prob_table2 <- cbind(Prob_table[, 1], Prob_table[, 2] + Prob_table[, 3])
colnames(Prob_table2) <- c("Yi(1)=15", "Yi(1)>15")
kable(Prob_table2, format = "markdown")
```

	Yi(1)=15	Yi(1)>15
Yi(0)=10	0.143	0.143
Yi(0)=15	0.286	0.143
Yi(0)=20	0.143	0.143
Marginal dist. of Yi(1)	0.571	0.429

```
sort(unique(villages$Yi0)) %*% Prob_table2[1:3, 2] / Prob_table2[4, 2]
```

```
##      [,1]
## [1,] 15
```

Or, even simpler:

```
mean(villages[villages$Yi1 > 15, "Yi0"])
```

```
## [1] 15
```

f. Use the table to calculate the conditional expectation that  $E[Y_i(1) | Y_i(0) > 15]$ .

$$\begin{aligned} E[Y_i(1) | Y_i(0) > 15] &= \sum_{Y_i(1)} Y_i(1) Pr[Y_i(1) | Y_i(0) > 15] = \\ &= \sum_{Y_i(1)} Y_i(1) \frac{Pr[Y_i(1), Y_i(0) > 15]}{Pr[Y_i(0) > 15]} = (15) \left( \frac{1/7}{2/7} \right) + (20) \left( \frac{0}{2/7} \right) + (30) \left( \frac{1/7}{2/7} \right) = \\ &= (15 + 30) \left( \frac{1}{2} \right) = \frac{45}{2} = 22.5 \end{aligned}$$

We could have done it directly from R:

```
Prob_table3 <- rbind(Prob_table[1, ] + Prob_table[2, ], Prob_table[3, ])
rownames(Prob_table3) <- c("Yi(0)<=15", "Yi(0)>15")
kable(Prob_table2, format = "markdown")
```

	Yi(1)=15	Yi(1)>15
Yi(0)=10	0.143	0.143
Yi(0)=15	0.286	0.143
Yi(0)=20	0.143	0.143
Marginal dist. of Yi(1)	0.571	0.429

```
sort(unique(villages$Yi1)) %*% Prob_table3[2, 1:3] / Prob_table3[2, 4]
```

```
##      [,1]
## [1,] 22.5
```

Or, even simpler:

```
mean(villages[villages$Yi0 > 15, "Yi1"])
```

```
## [1] 22.5
```

#### 4. More practice with potential outcomes.

We are interested in the hypothesis that children playing outside leads them to have better eyesight.

a. Compute the individual treatment effect for each of the ten children. Note that this is only possible because we are working with hypothetical potential outcomes; we could never have this much information with real-world data. (We encourage the use of computing tools on all problems, but please describe your work so that we can determine whether you are using the correct values.)

The individual Treatment Effect ( $TE_i$ ) is simply  $Y_i(1) - Y_i(0)$  for each one of the children.

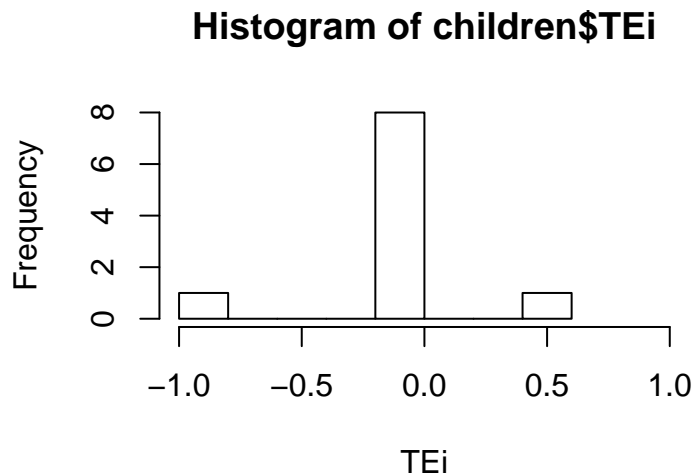
```
children <- data.frame(Yi0 = c(1.1, 0.1, 0.5, 0.9, 1.6,
                             2.0, 1.2, 0.7, 1.0, 1.1),
                      Yi1 = c(1.1, 0.6, 0.5, 0.9, 0.7,
                             2.0, 1.2, 0.7, 1.0, 1.1))
rownames(children) <- paste("Child", rownames(children))
children$TEi <- children$Yi1 - children$Yi0
kable(children, format = "markdown")
```

	Yi0	Yi1	TEi
Child 1	1.1	1.1	0.0
Child 2	0.1	0.6	0.5
Child 3	0.5	0.5	0.0
Child 4	0.9	0.9	0.0
Child 5	1.6	0.7	-0.9
Child 6	2.0	2.0	0.0
Child 7	1.2	1.2	0.0
Child 8	0.7	0.7	0.0
Child 9	1.0	1.0	0.0
Child 10	1.1	1.1	0.0

So for each of the ten children it is 0, 0.5, 0, 0, -0.9, 0, 0, 0, 0, 0.

b. In a single paragraph, tell a story that could explain this distribution of treatment effects. What might cause some children to have different treatment effects than others?

First, let's check the distribution of treatment effects.



Most of the children (80% of them) have a treatment effect equal to zero—i.e., their treated outcome equals their untreated outcome—, one of them (Child 2) has a treated outcome higher (0.5) than his/her untreated outcome, and another one (Child 5) has a treated outcome much lower (-0.9) than his/her untreated outcome. These two cases correspond to a child with very low visual acuity if untreated (0.1)—i.e., after not playing outside—, and another child whose visual acuity if untreated is quite high (1.6; only one other child has a higher visual acuity).

One possible explanation is that playing outside only have an effect when visual acuity is far from normal—a negative effect if visual acuity is much better than normal (though Child 6 contradicts that), and a positive effect if visual acuity is much worse than normal. Another plausible explanation (which does not necessarily contradicts the previous one) is that the treatment is very small—at least for 8 of the children—, and so is its effect. I.e., the time spent playing outside was not very different in the 8 children whose treatment effect equals zero. E.g., for those 8 children the outcome was measured playing an average of 9.5 and 10.5 hours per week, in each case. Why would the treatment effect be not equal to zero for Children 2 and 6 in that case? Maybe Child 2 was considered untreated after playing outside a very short time (say 1 hour per week), so 9 hours more per week were beneficial for him/her), while Child 5 was considered treated after playing outside very long periods (say 20 hours per week), so he/she was too much exposed to sunlight.

c. For this population, what is the treatment effect (ATE) of playing outside.

```
ATE <- mean(children$TEi)
```

**ATE = -0.04**

d. Suppose we are able to do an experiment in which we can control the amount of time that these children play outside for three years. We assign the odd-numbered children to treatment and the even-numbered children to control. What is the estimate of the ATE?

An estimate of the ATE (using experimental data) can be obtained by means of the following formula:

$$\widehat{ATE} = E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 0]$$

So we just need to implement this fomrula in R

```
children2 <- children[, 1:2]
children2[c(1:dim(children2)[1]) %% 2 == 1, "Yi0"] <- NA
children2[c(1:dim(children2)[1]) %% 2 == 0, "Yi1"] <- NA
(estimate_ATE <- mean(children2$Yi1, na.rm = TRUE) - mean(children2$Yi0, na.rm = TRUE))
```

```
## [1] -0.06
```

to obtain  **$\widehat{ATE} = -0.06$** .

e. How different is the estimate from the truth? Intuitively, why is there a difference?

**$\widehat{ATE} - ATE = -0.02$**

The difference is easily explained: we have not used  $D_i$  but  $d_i$  (a single realization of the experiment). Had we tested all possible combinations of treatment allocation, we would have obtained, on average, the real value of the ATE.



f. We just considered one way (odd-even) to split the children into an experiment. How many different ways (every possible ways) are there to split the children into a treatment versus a control group (assuming at least one person is always in the treatment group and at least one person is always in the control group)?

$$\sum_{i=1}^{N-1} \binom{N}{i} = \sum_{i=1}^9 \binom{10}{i} = \binom{10}{1} + \dots + \binom{10}{9} = \frac{10!}{1!9!} + \frac{10!}{2!8!} + \dots + \frac{10!}{8!2!} + \frac{10!}{9!1!} = (10!) \left( \frac{1}{5!5!} + (2) \left( \frac{1}{1!9!} + \frac{1}{2!8!} + \frac{1}{3!7!} + \frac{1}{4!6!} \right) \right)$$

We could code this in R

```
sum(as.numeric(lapply(c(1:9), function(x) choose(10,x))))
```

```
## [1] 1022
```

but an even easier way to calculate this is  $2^{10} - 2 = 1022$ .

This is because

$$\sum_{i=0}^N \binom{N}{i} = 2^N$$

and hence

$$\sum_{i=1}^{N-1} \binom{N}{i} = 2^N - 2$$

g. Suppose that we decide it is too hard to control the behavior of the children, so we do an observational study instead. Children 1-5 choose to play an average of more than 10 hours per week from age 3 to age 6, while Children 6-10 play less than 10 hours per week. Compute the difference in means from the resulting observational data.

```
children3 <- children[, 1:2]
children3[1:5, 1] <- NA
children3[6:10, 2] <- NA
t.test(children3$Yi0, children3$Yi1, alternative = "two.sided")
```

```
##
## Welch Two Sample t-test
##
## data: children3$Yi0 and children3$Yi1
## t = 1.8176, df = 5.861, p-value = 0.1202
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.1557541 1.0357541
## sample estimates:
## mean of x mean of y
## 1.20 0.76
```

The difference in means is  $-0.44$ . That difference is not statistically significant at the 0.05 level (though it is with a significance level of 0.2, two-sided).

**h. Compare your answer in (g) to the true ATE. Intuitively, what causes the difference?**

The difference in means of the previous observational study is a 1000% (i.e., 11 times) bigger than the true ATE. This is because the first five children have, on average, low visual acuity (4 of them have a visual acuity below 1), while the remaining five children have, on average, normal to high visual acuity (three of them have a visual acuity about 1, while one child has a visual acuity equal to 2, and only the other has low visual acuity, 0.7).

---

## 5. FE, exercise 2.5.

**a. Discuss the strengths and weaknesses of each approach.**

The three approaches are easy to implement, for such a small number of participants (each one is just a little bit more difficult than the previous one).

The 1st and 2nd methods are equal, from a random point of view. In both cases—provided that the coin is fair and that the assignment of each card is independent to each other—, each subject has the same probability ( $p = \frac{1}{2}$ ) of being asked to donate 30 or 60 minutes. But, since the number of subjects is small ( $N = 6$ ), there could be much more people being assigned to one group than to the other.

$$Pr[x \text{ subjects assigned to one of the two groups}] = \binom{N}{i} \left(\frac{1}{2}\right)^N$$

$$\begin{aligned} Pr[\text{no subjects in one group}] &= Pr[\text{six subjects in one group}] = \\ &= \binom{6}{0} \left(\frac{1}{2}\right)^6 = \frac{1}{2^6} = \frac{1}{64} = 0.015625 \end{aligned}$$

$$\begin{aligned} Pr[\text{one subject in one group}] &= Pr[\text{five subjects in one group}] = \\ &= \binom{6}{1} \left(\frac{1}{2}\right)^6 = (6) \left(\frac{1}{2^6}\right) = \frac{3}{32} = 0.09375 \end{aligned}$$

$$\begin{aligned} Pr[\text{two subjects in one group}] &= Pr[\text{four subjects in one group}] = \\ &= \binom{6}{2} \left(\frac{1}{2}\right)^6 = \left(\frac{6 \times 5}{2}\right) \left(\frac{1}{2^6}\right) = \frac{15}{64} = 0.234375 \end{aligned}$$

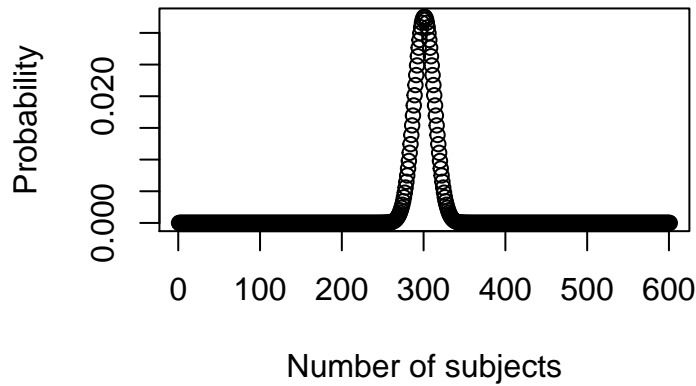
$$Pr[\text{three subjects in one group}] = \binom{6}{3} \left(\frac{1}{2}\right)^6 = \left(\frac{6 \times 5 \times 4}{3 \times 2}\right) \left(\frac{1}{2^6}\right) = \frac{20}{64} = 0.3125$$

The probability of having a group with no subjects or only one subject would be  $(2)(0.015625 + 0.09375) = 0.21875$ .

The 3rd method ensures *complete random assignment*, with exactly 3 subjects in each one of the two groups.

**b. In what ways would your answer to (a) change if the number of subjects were 600 instead of 6?**

The 1st and 2nd methods would still be identical in terms of randomness. But now the probability of dissimilar groups (in terms of number of subjects) would be extremely low. E.g., the probability of a group with more than 325 or less than 275 subjects is 0.041%.



So that disadvantage that we had with only 6 subjects now disappears. The 3rd methods remains the same, in terms of probabilities.

But now the 2nd and 3rd methods are difficult to implement (or at least time-consuming). The 1st methods, on the other hand, would be much more feasible.

c. What is the expected value of  $D_i$  (the assigned number of minutes) if the coin toss method is used? What is the expected value of  $D_i$  if the sealed envelope method is used?

The expected value of  $D_i$  will be 45 minutes, whatever the method and the number of subjects is. For the sealed envelope method, it is quite trivial to demonstrate it:

$$E[D_i] = \sum D_i Pr[\text{being assigned to } D_i] = (30 + 60) \left(\frac{1}{2}\right) = \frac{90}{2} = 45$$

For the coin toss methods, it can also be demonstrated that:

$$\begin{aligned} E[D_i] &= \sum D_i Pr[\text{being assigned to } D_i] = \\ &= \sum_{x=0}^N \frac{60x+30(N-x)}{N} Pr[x \text{ subjects assigned to } D_i = 60] = \\ &= \sum_{x=0}^N \frac{30(N+x)}{N} \binom{N}{x} p^x (1-p)^{N-x} = \\ &= (30) \sum_{x=0}^N \binom{N}{x} p^x (1-p)^{N-x} + \left(\frac{30}{N}\right) \sum_{x=0}^N x \binom{N}{x} p^x (1-p)^{N-x} = \\ &= (30)(1) + \left(\frac{30}{N}\right) (Np) = (30)(1+p) = (30) \left(\frac{3}{2}\right) = 45 \end{aligned}$$

## 6. FE, exercise 2.6.

This is an observational study because the sample has been randomly selected, but the treatment (taking a preparatory class) has not been randomly assigned. Therefore, we cannot ensure that the only difference between both groups is having been exposed to the treatment or not.

**7. FE, exercise 2.8.**

- a. Interpret the apparent effects of the treatments on the proportion of applicants who have their residence verified and the speed with which verification occurred.**

Each of the 3 treatments has a slight apparent effect on the proportion of applicants who have their residence verified, compared to the control group: that proportion increased from 95.2% to 100%. Only paying the bribe had an effect on the speed with which verification occurred, compared to the control group: the median number of days was reduced by about 54%.

- b. Interpret the apparent effects of the treatments on the proportion of applicants who actually received a ration card.**

The NGO treatment had an apparent negative effect on the proportion of applicants who actually received a ration card, compared to the control group: that proportion decreased from 23.8% to 16.7% (a 30% decrease). The RTIA treatment had an apparent positive effect (a 265% increase, from 23.8% to 87%), as well as the Bribe treatment (a 320% increase, from 23.8% to 100%).

- c. What do these results seem to suggest about the effectiveness of the Right to Information Act as a way of helping slum dwellers obtain ration cards?**

The RTIA seems to be very effective on obtaining a ration card.

**8. FE, exercise 2.9.**

- a. Critically evaluate this assumption.**

The Selection Bias  $E[Y_i(0) | D_i = 1] - E[Y_i(0) | D_i = 0]$  may not equal zero, and hence the estimate of the ATE might be biased. Winning the lottery has a lot to do with playing it, and those who play more have more chances to win a bigger prize or more prizes. Hence, the two groups may be different in other ways rather than the treatment (having won more than \$10,000 in the lottery). For example, people with lower income tend to play more, and their views about the state tax, had not they won the lottery, might be quite different to those who didn't win so much or anything at all (maybe because they didn't play so much or anything at all, maybe because they have higher incomes).

- b. Suppose the researcher were to restrict the sample to people who had played the lottery at least once during the past year. Is it now safe to assume that the potential outcomes of those who report winning more than \$10,000 are identical, in expectation, to those who report winning little or nothing?**

Following the previous reasoning, the expectations may be closer in this case, but still not identical. Reducing the sample to people who play makes the two groups more similar, but there may still be some differences between both groups, apart from having received the treatment. Maybe those who won more than \$10,000 play, on average, many times a year—hence increasing their chances—, and this might be related to other facts also related to views about the state tax (like their income, as mentioned before).

**9. FE, exercise 2.12(a).**

**a.** In this study, nature has assigned a particular realization of  $d_i$  to each subject. When assessing this study, why might one be hesitant to assume that  $E[Y_i(0) | D_i = 0] = E[Y_i(0) | D_i = 1]$  and  $E[Y_i(1) | D_i = 0] = E[Y_i(1) | D_i = 1]$ .

The excludability assumption may not be met. Because both groups (those who read more than three hours per day, and those who read less) may differ in other factors apart from the treatment, so receipt of the treatment would not be the (only) relevant causal agent—maybe those who read do so because they are peaceful, and prefer dialogue and reasoning to violence.

Those who do not read (*non-natural* readers:  $D_i = 0$ ) are violent ( $Y_i(0)$ )...and they probably still would be more violent than those who read (*natural* readers:  $D_i = 1$ ) even if the latter were forced to not read ( $Y_i(0)$ ). So it is likely that  $E[Y_i(0) | D_i = 0] > E[Y_i(0) | D_i = 1]$ .

And *natural* reader ( $D_i = 1$ ) may still be less violent ( $Y_i(1)$ ) than *non-natural* readers ( $D_i = 0$ ) even if the latter were obliged to read ( $Y_i(1)$ ). So it is likely that  $E[Y_i(1) | D_i = 0] > E[Y_i(1) | D_i = 1]$ .

---

## 10. Revised version of Essay 1.

The study I would like to comment is mentioned in this press release:

“New study shows clearly that **highly emotional and graphic anti-smoking advertisements increase quit attempts**”, statement of Matthew L. Myers, President, Campaign for Tobacco-Free Kids, October 9, 2012. [http://www.tobaccofreekids.org/press\\_releases/post/2012\\_10\\_09\\_graphic](http://www.tobaccofreekids.org/press_releases/post/2012_10_09_graphic)

More information about the study can be found at: [http://www.ajpmonline.org/article/S0749-3797\(12\)00527-2/fulltext](http://www.ajpmonline.org/article/S0749-3797(12)00527-2/fulltext)

The **causal claim** being made is the one mentioned in the title of the press release: “**highly emotional and graphic anti-smoking advertisements**” (as opposed to advertisements of other kind, or to no advertisements at all) “**increase quit attempts.**” According to the lead author of the study, “*The implications of this study are fairly straightforward: if the goal is to motivate smokers to try and quit smoking, antismoking advertisements should be designed to elicit a strongly negative emotional reaction.*” It is also mentioned in the press release that “*States that have conducted hard-hitting media campaigns as part of their tobacco prevention and cessation programs have reduced smoking rates faster and to lower levels than the nation as a whole.*”

The author also mentions that this kind of advertisements—“*especially those that use personal stories to evoke strong emotions and that realistically depict the devastating health consequences of tobacco use*”—, i.e., those which “*elicit a strongly negative emotional reaction*” or “*depict serious harm done by smoking or secondhand smoke in an authentic way*”, “*perform better than those that do not.*” The **mechanism**—which is out of the scope of this essay, in any case—is not explicitly mentioned, but it may be linked to psychological reasons, e.g., fears of pain and lost.

To arrive to the former conclusion, “*Researchers at RTI International and the New York State Department of Health analyzed data from more than 8,700 smokers from 2003 to 2010 to gauge the impact of the media component of the state’s tobacco prevention and cessation program. Whether measured by potential exposure or ad recall, smokers exposed to the ads were significantly more likely to have tried to quit in the past year. Smokers who recalled seeing emotional or graphic antismoking advertisements such as personal testimonials about smoking health consequences or images of diseased lungs, had increased odds of quitting in the previous 12 months of 29% compared to those who did not.*”

But the **purported causal effect could be overestimated, or it might not really exist at all.** I.e., there is a correlation between ad recall and quit attempts, but the former does not necessarily cause the latter. Those smokers who recalled seeing those adverts<sup>1</sup> might already have in mind—subconsciously or not—quitting. That might be a possible reason why they paid especial attention to them. And smokers who had no intention of quitting might not have paid attention to the advertisements, or even forced themselves to forget them.

Other possible explanation for that correlation—that would not imply causation either—is that smokers are not always sincere about their quit attempts, and tend to underestimate those attempts because (unless they succeed and actually quit) they feel like they have failed, and do not like to admit it. It could have happened that the smokers who did not recall the ads acted that way, while those who recalled the ads did not (and maybe even overestimated their own attempts) because—due to that effect of shame—admitting that they recalled the ads but they had not led them to attempt to quit would make them feel embarrassed.

The **underlying assumption that would make the correlation between ad recall and quit attempts a causal effect** is that both groups of smokers were, on average, identical in every way except for the fact that one of them—the one with a higher rate of quit attempts in the 12 months prior to the end of

<sup>1</sup> The press release mentions that the exposure to the ads was “*measured by potential exposure or ad recall,*” and the study (see the second link in the first paragraph) states, “*Exposure was measured by self-reported confirmed recall and market-level gross rating points.*” Those references to “potential exposure” / “market-level GRPs” surely mean that the level of exposure to the TV ads was estimated by assigning the GRPs of the campaign to those smokers who self-reported having watched the TV programs during which the ads were shown. That would certainly be a very rough estimate, not very reliable, so I have preferred to talk only of “ad recall” (also because the press release mentions this far more than “potential exposure”). The experiment proposed in the next page uses a more reliable metric, which does not have the disadvantages of recall (mentioned later).

the study—was formed by people who were exposed to the ads (which is not exactly the same as recalling seeing them).

To be able to infer a causal effect, an **experiment** would be the only option. The **ideal** one would use (1) a sample of smokers as large as possible, as well as (2) **random assignment** to allocate subject to treatment or control group, ensuring:

- (3) **excludability** (i.e., that the treatment assignment has no effect on outcomes apart from the treatment itself—it does not set in motion other potential causes),
- (4) there is **no selection bias** (e.g., self-selection, which—if the treatment is the exposure to advertising—would be typical in any media: people decide which TV programs they watch, which journals they read, which websites they visit, or which content they search), and
- (5) **non-interference** (i.e., that untreated subjects are actually untreated, not even affected by the treatment received by the smokers in the other group<sup>2</sup>).

Unfortunately, some of these features—as already mentioned—are not feasible. E.g., conducting an experiment in TV would be challenging<sup>3</sup>, because we would need to apply the treatment—to show the TV ads—only to a certain group and not to the other<sup>4</sup>.

That’s why I propose to run an **online** experiment. For example, we could randomly assign the smokers within an **access panel**<sup>5</sup> to two groups, and show highly emotional and graphic anti-smoking online ads—regardless of the online platform, be it a desktop, mobile or video search engine, or a social network—only to those in the first group. After the anti-smoking advertising campaign, we would **compare the mean difference of quit attempts between both groups**.

This online experiment would allow us to **infer causality** if it exists, being able—due to the nature of the panel—to **generalize the results** to the whole population from which the panel was taken. We would just have to confirm that the smokers in our **sample**—a subset of the panel—are still **representative** of all the smokers in the whole population (it might also be argued that an access panel is not as representative of a population as a TV audience panel, especially in regions where Internet penetration is still low).

Using online advertising in a small target group would make the **cost** of the study to be quite **low** (as compared to a TV ad campaign). There might be some **arguments** against showing the ads only to such a reduced subset of the population (those smokers who are members of the access panel and have been randomly assigned to the treatment group), hence denying other smokers the benefits—be them big or small—of being exposed to a treatment that might lead them to try to quit. But the counterpart would be **knowing for sure whether these ads are effective**—i.e., the cause of smokers trying to quit—or not, and hence the way policy rulers could decide whether they should **invest in this kind of campaigns** or **take other measures against smoking**.

Of course, we should be aware of—and try to correct as far as possible—the **limitations of this non-perfect experiment**, which (for the reasons previously mentioned when explaining the ideal experiment) may have selection bias, does not ensure non-interference, only uses a relatively small sample which may not be fully representative, etc.

<sup>2</sup> In our context it would mean that people do not talk each other about “highly emotional” ads (while still being unaware of the treatment they are receiving).

<sup>3</sup> At least with the current broadcast TV technology.

<sup>4</sup> Some options (far from being optimal) would be:

Using a **TV audience panel** (see the next footnote): we could run the anti-smoking TV campaign only in certain TV programs. Hence, the treatment group would consist of those smokers within the panel that watched those programs, and those smokers who did not would belong of the control group... but the treatment would not be randomly allocated, so this would not be a “true” experiment—we could not ensure that people in the treatment group only differ from those in the control group in the fact that the former watch certain TV programs.

Using a **cable TV** (or an on-demand streaming media, although these usually do not push advertisements) **audience panel**: it would be possible—at least from a technical point of view—to show anti-smoking advertisements only to the treatment group, and not to the control group... but in this case we might **not be able to generalize the findings** to the whole population (i.e., to people who cannot afford or are not interested in these media).

<sup>5</sup> A panel is a representative sample of a population that agrees that their behavior is tracked. Depending on the kind of panel (access, consumer, audience), different behaviors (online, purchasing, TV) are tracked.