Answers - Final Exam

Diego Díaz

01-06-2020

I.A - Describe the research question in this paper in words. Explain, in words and math, the ideal experiment one might want to use to answer this question.

The paper studies the default effect, that is, putting households in a default option in the context of time-varying residential pricing programs for electricity. More specifically, the research question is whether there is an impact of defaulting individuals into two different electricity pricing programs, called *Time-of-Use* (TOU) and *Critical Peak Pricing* (CPP), in two outcomes: whether there is an effect of assigning the default option in terms of the individuals staying in the pricing program or not, and second, on average electricity consumption in response to time-varying electricity prices.

An ideal experiment for this question requires a Randomized Control Trial (RCT) experiment since we require random assignment of the treatment. Since we are interested in studying the default option for two pricing programs and there are two programs, there are 4 default options (*opt-in* and *opt-out* for each program), therefore, there are 4 treatment groups ans 1 control group. Ideally we would have a group of N households and we would assign each treatment randomly among a portion of them.

Let's consider one of the 4 treatments. We are interested in estimating:

$$\tau_i = Y_{it}(TOU_{it}^{opt-in} = 1) - Y_{it}(TOU_{it}^{opt-in} = 0)$$

Where Y_{it} is the average electricity consumption for household i during period t and TOU_{it}^{opt-in} indicates the treatment status for the household in the TOU pricing program during period t. Although we can't observe τ_i since we can't observe a household in both states, having the treatment assigned randomly means that both groups, households in the opt-in group and those not being offered anything, are balanced. Given that both groups are equal in terms of both observable and unobservable characteristics on average, we can estimate the average treatment effect, τ_{ATE} (the impact in average electricity consumption), by calculating the difference in the outcome variable between both groups.

The dataset that the experiment would produce is the treatment binary indicator for households $TOU_{it}^{opt-in} \in \{0,1\}$, including those for TOU opt-out and both default options for the CPP pricing program as well. It would also include the measurement of average electricity consumption when the treatment is in effect (Y_{it}), measured in the terms of units of energy ($e.g. : [kWh]$). Using the potential outcomes framework, we estimate the τ_{ATE} for treated households as:

$$\tau^{ATE} = E[Y_{it}|TOU_{it}^{opt-in} = 1] - E[Y_{it}|TOU_{it}^{opt-in} = 0]$$

Differences between average consumption among households gives us an unbiased estimate of τ_{ATE}:

$$\hat{\tau}^{ATE} = \overline{Y_t}(TOU_{it}^{opt-in} = 1) - \overline{Y_t}(TOU_{it}^{opt-in} = 0)$$

I.B- Explain, in words and math, what treatment parameter the authors are recovering in Equation (1). Describe each term in the estimating equation in words, including a discussion of the unit of analysis. Explain how – if at all – this estimate differs from the population-wide average treatment effect of time-varying electricity pricing on electricity consumption. What assumptions are required in order for Equation (1) to recover the causal effect of interest? Do you think these assumptions are likely to be satisfied in this context? Why or why not? Include references to evidence presented in the paper to support your conclusion.

Equation (1) is the following:

$$y_{it} = \alpha + \beta_{ITT} Z_{it} + \gamma_i + \tau_t + \varepsilon_{it}$$

The treatment parameter the authors are interested in is β_{ITT}, which represents the average intent to treat effect because it does not capture the effect of being in the pricing program, but rather, that of being assigned a default option to the pricing program Z_{it}. The other terms in the equation are described below:

y_{it}: Hourly electricity consumption for household i in hour t. α: Base level of average electricity consumption for households. Zit: Indicator variable indicating if household i was encouraged to be in the treatment group (value of one) and zero otherwise. Starting on June 1, 2012. γ_i: Household fixed effect that captures systematic differences in consumption across households. τ_t: Fixed effect for the hour t of the measurment. Given the cyclical nature of electricity consumption, this is important. ε_{it}: Error term capturing unobservable noise.

The unit of analysis is household-hour as we are measuring household hourly electricity consumption in hour t. Given the research question the authors attempt to answer, the only possible design is at the household level. Studying default options for larger units such as blocks or counties would introduce different issues as it would be more likely to have unbalanced groups. As for the time unit, making it hourly makes sense given the cyclical component of electricity consumption. The pricing program is more likely to affect the distribution of consumption rather than the daily or weekly average.

The treatment parameter β_{ITT} differs from the poopulation-wide average treatment effect because of the presence of non-compliers. The treatment in this context is being in a particular pricing program, while the intent to treat is being encouraged to participate, as being placed in a default option and allowing to opt-in or out. Since not all households remain on the treatment assigned, we can capture the average intend to treat effect.

The specification used in equation (1) is a difference-in-differences (DID) approach using data from the pre-treatment and treatment periods. One of the assumptions required for the approach to work is **unconfoundedness**, that is, assuming that being assigned a default option (Z_i) is independent (orthogonal) to other determinants of energy consumption. Another assumption requiered for the estimator to capture the causal effect is that the different groups have **parallel trends** in the outcome variable before the treatment. This assumption is satisfied inmediatly in the authors' experiment because households are selected randomly and therefore both groups are balanced, therefore, both groups will have the same trend and the specification will be able to recover the intend to treat effect. The authors directly show evidence of this assumption in figure A1 (appendix page A9), where they show in a graph how electricity usage varies through the day for both treatments and control group in the pre-treatment period. The second asumption that we require is that the treatment and control groups are balanced. Although this is unavoidable when the treatment is assigned randomly, the authors show that it is the case in table 1 (page 35), where they compare the average of consumption of electricity per day and other variables between groups. One last assumption is the **stable unit treatment values**, that is, that electricity consumption of household i is not affected by the participation decisions of other households.

I.C- Explain, in words and math, what treatment parameters the authors are recovering in Equation (2). Describe each term in the estimating equation in words, including a discussion of the unit of analysis. Explain how – if at all – this estimate differs from the population-wide average treatment effect of time-varying electricity pricing on electricity consumption. Describe two different procedures to estimate this treatment parameter. What assumptions are required in order for Equation (2) to recover the causal effect of interest? Do you think these assumptionsare likely to be satisfied in this context? Why or why not? Include references to evidence presented in the paper to support your conclusion.

Equation (2) is the following:

$$y_{it} = \alpha + \beta_{LATE}\text{Treat}_{it} + \gamma_i + \tau_t + \varepsilon_{it}$$

y_{it}, α, γ_i, τ_t, and ε_{it} are defined as in equation (1). Average consumption of electricity on hour t, base level of consumption, household fixed effect and hour fixed effect, and remaining error term, respectively. The main difference in this specification is in the treatment variable, which now it's expressed as $Treat_{it}$ and it indicates if household i was actually enrolled in treatment starting June 1st, 2012, and zero otherwise. The authors instrument for $Treat_{it}$ using the randomized encouragement to the treatment Zit from equation (1). This is, therefore, a DID instrumental variables (IV) approach, and since the treatment variable is actual treatment, the treatment parameter that the authors recover is β_{LATE} (localized average treatment effect), which measures the average impact on household electricity consumption of households enrolling in one of the pricing programs. The estimator will capture the population wide treatment effects if we assume we have constant treatment effects among households, which is a reasonable assumption when controlling for household fixed effects as we are, sin we can capture differences in other covariates such as income in this way. Given the specification of the authors and the fact that they can observe *compliers* and *always takers*, they can also estimate the effect on this groups separately. Assumptions required for the estimator to recover the actual $LATE$ is **unconfoundedness**, that is, assuming that being assigned a default option (Z_i) is independent (orthogonal) to other determinants of energy consumption. We also need an **exclusion restriction** for the treatment variable, that is, that the assignment of the default option in a pricing program (with the ability to opt-out) only affects electricity consumption indirectly by an effect on participation on the actual program. Thirdly, a **monotonicity assumption** which states that assignment to a default program weakly increases the probability of participating in the program, and the main assumption of instrumental variables, which is that the instrument is correlated with the treatment variable. In this case, the fulfillment of this last assumption is direct since the design of the experiment. Finally we are also assuming the **stable unit treatment values** assumption holds, which means that the treatment assignment of a household does not affect other untreated households. These assumptions are likely to be satisfied in this context. Some can be shown as the instrumental variables assumption of correlation between the instrument and treatment. This one is clear by the fact that complacent customers comprise more than 75% of the sample (page 3). The **exclusion restriction** is tested in table A6 (page A15), which estimates the effect of encouragement to the treatment on households who did not enroll in the treatment, wither by not opt-in or by opt-out of treatment. Since most of the estimates (7 out of 8) are not statistically significant, we can be mostly confident that the treatment assignment does not directly impact electricity consumption.

I.D- Describe the main results of the paper. Include a discussion of (at a minimum) Tables 3 and 4, in which you interpret the estimated coefficients and describe their magnitudes. What is the main policy take-away of the paper? Describe the results in Table 6.

The paper concludes about the power of the default effect in the context of household electricity pricing programs. Alocating residential customers into one of 5 groups: 4 different treatment groups and 1 control group, the authors document that only 20% of the customers opt into the new pricing programs and over 90% choose to stay in the program if placed in the opt-out option. This turned out true for both CPP and TOU programs. To study the impact of the encouragement of treatment and actual treatment, two approaches are used (as described in parts B and C).

Table 3 shows the result from applying the regression of equation (1) described in part B, which estimates the Intent to treat effects (ITT), or the effect of encouraging to participate. Eight coefficients are estimated in the table, two for each treatment group depending on the hour or event of measurement. The statistically significant value of -0.129 for the CPP opt-in group during a critical event indicates that the treatment reduces electricity consumption by $0.129[kWh]$ during a critical event (data includes period between 4-7 pm during simulated CPP events in 2011, and 4-7 pm during actual events in 2012-1013). The opt-out group experiences a larger effect on average, a reduction of $-0.305[kWh]$ on electricity consumption, which is expected given the mentioned significance of the default effect, which causes more households to be treated on the opt-out scheme. For the TOU pricing program, the opt-in group reduced it's consumption by $0.091[kWh]$ and the opt-out group by $0.130[kWh]$, also indicating a strong default effect since the magnitude is higher for the opt-out group. All coefficients are statistically significant.

Table 4 shows the results of applying the approach from equation (2), which is a DID two stage least squares approach that uses the default assignment into a pricing program as an instrument of actual treatment to estimate the localized average treatment effect. By exploiting differences between the opt-in and opt-out group the authors estimate the average treatment effects on *always takers*, *complacents*, and *always takers together with complacents* (the group that was intended to treat). A coefficient is estimated for each of these groups, for each of the two pricing programs and for both *critical event hours* and *non-event day peak hours*, meaning that a total of 12 coefficients are estimated by this approach. Given that in this case, we are considering treatment to actually be enrolled in a pricing program, the treated in the opt-in group is going to be comprised of always takers only, while the treated in the outp-out group is going to be comprised of always takers and complacents. Separately, the effect for complacents can be estimated by removing the always takers from the opt-out group in the regression. We would expect in this analysis that the effect for always takers is larger than for complacents, meaning that always takers experience a larger decline in electricity consumption in comparison to complacents during critical event hours and non-event dat peak hours. This is true for both pricing programs. During critical event hours, always takers in the CPP program lower their consumption of electricity in $0.658[kWh]$, while the effect on the treated in the opt-out group (always takers and complacents) is $-0.33[kWh]$ and the effect of the treatment in the complacents is $-0.242[kWh]$. In the TOU program the effect is similar but lower for every group: always takers see a reduction in consumption of $0.48[kWh]$, always takers and complacents together have an impact of $-0.136[kWh]$ and complacents alonce experience a reduction of $0.051[kWh]$. All the coefficients mentioned are statistically significant. In the average treatment effects during non-event day peak hours we can also effect an important default effect as the complacents experience a reduction of consumption, however, the reductions during these hours are smaller than during critical event hours for both pricing programs, although higher in magnitude for the TOU group.

II.A- FUNACRONYM would like you to compare the average number of public goods (roads, schools, public buildings, etcetera) in towns with female-headed governments as compared with towns that have male-headed governments. Describe this comparison in math and words. Under what conditions would this comparison estimate the causal effect of female leaders on public goods provision? Provide two concrete examples of reasons why this comparison may be problematic.

We can compare the average number of public goods in towns with female-head governments to those with male-head goverments. This is what is called the naive estimator for the impact of female leaders in the provision of public goods and we can express it mathematically as:

$$\tau^{NAIVE} = \overline{Y}(female = 1) - \overline{Y}(female = 0)$$

Where female is the treatment indicator of whether the town has a female leader or not (1 it does, 0 it does not). Therefore, the first term on the right side is the average of Y (provision of public goods) of towns where there is a woman in power and the second is the average of public goods of towns with a man in power. The naive estimator can be equal to the average treatment effect if both groups were balanced before the treatment, that is, if both type of towns were similar in terms of observable and unobservable characteristics, which can be achieved by random assignment of the treatment. In this case, the comparison is likely to be a problem because there could be fundamental selection issues causing female-head towns to have higher provision of public goods. For instance, women could be more demanded in towns with higher amount of public goods because they care about solving problems unrelated to public goods. In this situation, if men are the ones that care about the provision of public goods, they might be more electable when the provision of public goods is low. This is the fundamental issue of selection and it happens when the outcome (provision of public goods) is correlated with the treatment assignment. Another example of this would be towns with higher income being more likely to elect women, which could happen if women care about issues like economic inequality and the poorest members of society instead of the average. A society that is richer might start caring more about these issues than a poorer society, and that causes a selection issue since richer places likely have more amount of public goods per capita.

II.B- FUNACRONYM gets it - this is not the best comparison. However, they have data on a bunch of other town characteristics: per-capita income, number of residents, year of incorporation, average population age, and share of gross city product devoted to manufacturing. Describe, using math and words, a comparison between female- and male-headed towns which leverages these administrative data. Under what conditions would this comparison estimate the causal effect of female leadership on public goods provision? Provide two concrete examples of reasons why this comparison may be problematic (different from what you described above).

The problem that we need to solve is the treatment and control groups being unbalanced in variables that affect the provision of public goods. If controlling by towns' characteristics, the treatment assignment is independent of the provision of public goods, then we could obtain an unbiased estimator for the average treatment effect by a regression approach, since both groups of towns would be balanced in other variables as well. This is called a selection on observables (SOO) approach and we can describe it with the following regression:

$$Y_i = \alpha + \tau female_i + \gamma X_i + \nu_i$$

Where α is a base level for the provision of public goods, τ is the average treatment effect, γ is a vector of coefficients of the town i's covariates that we are controlling for, X_i is a vector of those covariates, and ν_i is an error term uncorrelated with the provision of public goods. Given what we've explained so far, the assumption that we are making for this approach to work is that the treatment assignment is random after controlling for the covariates X, which include income, number of residents, year of incorporation, average population age, and share of city's product devoted to manufacturing.

This approach could be problematic because we are assuming that we are controlling for everything that impacts the provision of public goods that could be correlated with the treatment assignment. This is a strong assumption that we can't test and is something that is solved when we have random assignment of the treatment.

Examples of this approach being problematic are the following:

- We could have omitted variable bias if a variable such as a measurement of culture or prevalence of conservative ideas, which may affect treatment assignment as more conservative people might be less likely to vote for a woman to be the political leader of the town, affects the outcome, which may happen as conservatives might be more interested in public goods such as public parks.

- Another possible problem is omitting another variable as religion, which may be correlated with treatment assignment and preferences over public goods.

II.C- FUNACRONYM understands your concerns, but has some in-house machine learning experts. They tell you that they can use this same administrative data to solve your issues. Do you agree? Why or why not? Be specific.

It depends. Machine learning can be used to do a more sophisticated selection on observables design, but it will still be subject to the assumptions mentioned in part (B). The experts could be considering two different approaches. The first one is applying the following model to predict the outcome $\hat{Y}_i$:

$$Y_i = \alpha + \tau female_i + f(\mathbf{X}_i) + \nu_i$$

The idea is the same as in part (b), that $E[\nu|female, \mathbf{X}] = 0$, however, the model will choose the covariates that are important for predicting the number of public goods, which are not necessarily the ones that matter for wether the town has a woman as leader. In this case the assumption will likely fail and the estimator will be biased.

Another approach by the ML experts might be to predict (with a LASSO approach) the number of public goods in each town only as a function of the covariates in the administrative data, that is, predict $\hat{Y}_i$ as a function of X. Next, predict the treatment ($female$) as a function of X, also with LASSO, to finally estimate the impact of having a woman in power by using the residuals from both regressions to estimate τ. This is done by regressing:

$$Y_i^{residual} = \alpha + \tau female_i^{residual} + \varepsilon_i$$

This approach would work in terms of producing an unbiased estimator of τ only if the assumptions from part (B) hold. Mainly that $E[\nu|female, \mathbf{X}] = 0$.

II.D- FUNACRONYM forgot to tell you that, in India, certain local government positions are "reserved" for women – meaning only women can run for office to fill these seats (this is, again, a Real Thing!). They inform you that towns are selected to have reserved seats based on their political party. Towns where the party that rules the state also won the last town election are required to have female leaders, whereas other towns can elect either women or men to office. They suggest that you use an instrumental variables approach leveraging this new piece of information. Describe, in math and words, what this IV approach would look like. Under what conditions would this approach estimate the causal effect of female leadership on local public goods provision? Provide two reasons why these conditions may not be satisfied in this setting.

The approach would use the new variable (let's call it $reserved_i$) to isolate the variation in the number of public goods that only on having a woman in power. In math, this approach consists in a two-stage regression. On the first stage, the instrument ($reserved$) is used in the following regression:

$$female_i = \alpha + \gamma\ reserved_i + \beta X_i + \eta_i$$

Next, we store the predicted values of $female_i$, called $\hat{female}_i$, and regress the number of public goods in this new treatment variable:

$$Y_i = \alpha + \tau_{IV}\ \hat{female}_i + \delta X_i + \nu_i$$

Where τ_{IV} is our instrumental variables estimator and it is an unbiased estimator of the local average treatment effect of female leadership as long as the following assumptions hold:

- The instrument and the the treatment are correlated, or $Cov(reserved_i, female_i) \neq 0$, which we can test by looking at an F statistic of the first stage and see if it is high (usually > 20).
- The instrument does not impact the provision of public goods directly, that is, the fact that there are reserved positions for women in leadership positions in towns does not influence the provision of public goods directly, but only indirectly by making it more likely that a woman gets to power. This is what's called the exclusion restriction and it can be expressed mathematically as $Cov(reserved_i, \nu_i) = 0$. This assumption cannot be tested as we don't observe the errors.

Next we give two reasons of why these assumption might fail in this case:

- If women are always preferred instead of men for leadership positions in the towns under study, it will be the case that having reserve positions for them does not change anything. This will mean that $Cov(reserved_i, female_i) = 0$, breaking our first assumption.
- Towns having higher standard of living for the average citizen probably makes them more likely to worry about gender equality and to implement reserved positions for women. At the same time, this makes them more likely to have more public goods, breaking our exclusion assumption.

II.E- In Uttar Pradesh, an Indian state, all towns are put on a list, ordered by the share of women in the population. Each election cycle, the top 500 towns on the list are required to reserve the leadership positions for women (though, be warned – official rules aren't always perfectly followed). FUNACRONYM asks you whether you can use Uttar Pradesh as a test case. Describe, in math and words, the research design you would use to leverage this new information. Be sure to include a regression equation. Under what conditions would this approach estimate the causal effect of female leadership on public goods provision? For whom is this causal effect identified?

We can use this data to draw conclusions for towns that satisfy certain conditions in Uttar Pradesh, FUNACRONYM will have to decide afterwards whether the conditions in other states are similar enough for extrapolation. Given the way Uttar Pradesh sets *reserved* positions for women, we can implement a regression discontinuity design (RD) with a cutoff set around the town ranked 500 for the share of women in the population. We must consider than given that official rules are not always perfectly followed, the probability of having a woman in power given the reserve seating is not 1, but it should certainly be higher than when there is no reserve seating. Fortunately this can be tested in the data with a simple F-test on a regression of rank below and after 500 to having a woman in power. This is the same regression as $female_i = \alpha + \gamma\ rank_i^{\leq 500} + \eta_i$. Given this consideration, we can implement a fuzzy RD approach just as we would implement an instrumental variable approach around the cutoff, by a two stage least squares regression.

First stage:

$$female_i = \alpha + \gamma\ rank_i^{\leq 500} + \beta X_i + \eta_i$$

Since the instrument of being ranked in the first 500 positions is the same as having reserved seating for women, we can also write the first stage as:

$$female_i = \alpha + \gamma\ reserved_i + \beta X_i + \eta_i$$

We use the predicted values of $female$, $\widehat{female}$ to estimate the second stage:

$$Y_i = \alpha + \tau_{IV}\ \widehat{female}_i + \delta X_i + \nu_i$$

The main consideration in the RD approach is that we have to perform the regression on a selected bandwidth about the cutoff point. Selecting the size of the bandwidth is tricky since we have to consider a bias-variance tradeoff, the higher the bandwidth the more we reduce the variance but we increase the bias. The important thing is that we can leverage the new information obtained (depending on whether we have enough) to obtain, at least theoretically, an unbiased estimator of the local average treatment effect (for compliers only) of having a woman in power in the provision of public goods. A consideration to keep in mind is that this estimator is theoretically valid around the cutoff point. The conditions required are the same as the ones we need for the instrumental variable approach:

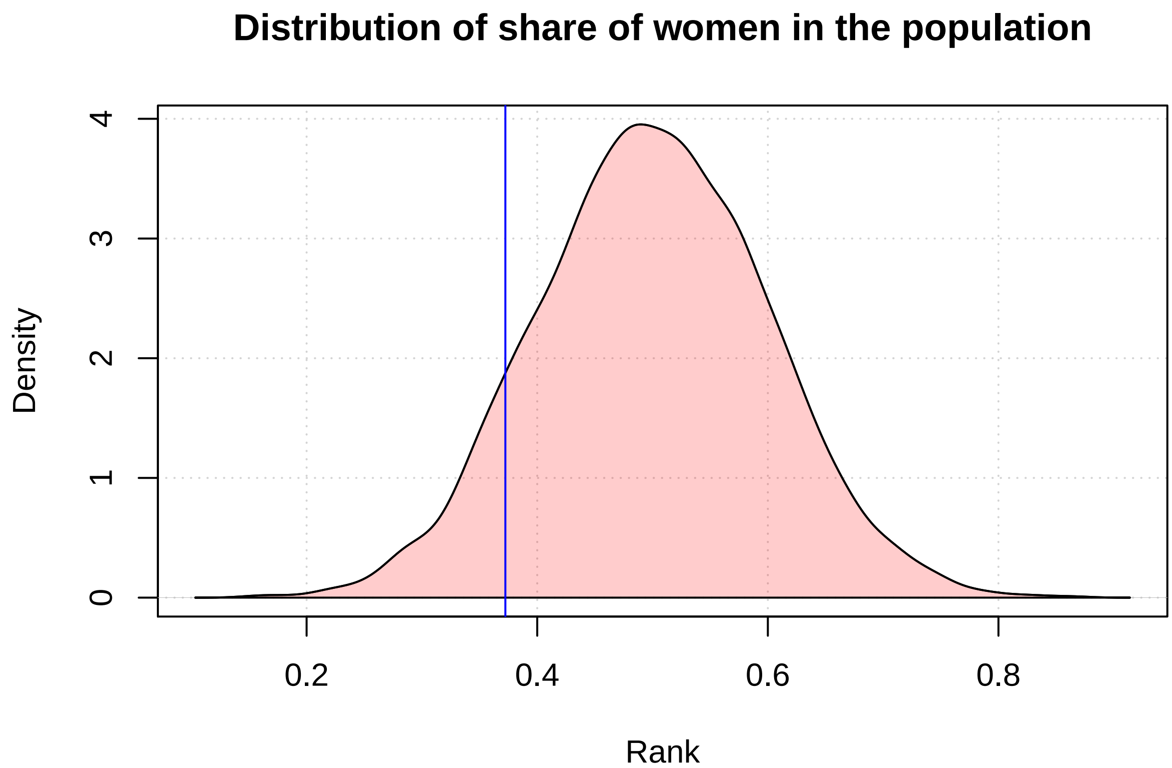
- The cutoff point in share of women in the population (or rank) produces a discontinouity in the expectation of having a woman in power. In math: $E[female_i|rank_i \leq 500] \neq E[female_i|rank_i > 500]$.
- Independence of the treatment (having a woman in power) and the share of women in power (or rank) around the cutoff only affects the provision of public goods through the treatment variable ($female$).
- Exclusion restriction: Being on each side of the cutoff only affects the provision of public goods indirectly through having a woman in power.
- Monotonicity: The probability of having a woman in power is different at different sides of the cutoff point.
- Covariate smoothness: The covariates of the number of public goods are continuous around the cutoff.

II.F: FUNACRONYM likes this idea, and is willing to share data with you to try this out. Use the dataset contained in final_exam_2020.csv. What empirical tests would you like to perform, prior to attempting to estimate the effect of female leadership on public goods provision, to provide evidence in support of the identifying assumption(s)? Perform at least two tests (hint: these should be simple graphical exercises). What do they tell you about the validity of the identifying assumption(s) in this case?

Although we can't prove the identifying assumption, we can provide evidence in favor of it.

First we provide what is called a manipulation test to show evidence that the towns can't sort themselves into different sides of the cutoff point. If towns can manipulate their share of population by gender, they could rank higher/lower in purpose, producing a selection problem in our approach. We show evidence to show that this is not the case by plotting the distribution of share of women in the population:

```
library(ggplot2)

final_exam_2020 <- read.csv("C:/Google Drive/Program Eval/data/final_exam_2020.csv")

dens = density(final_exam_2020[,1])
plot(dens, xlab = "Rank",
         main = 'Distribution of share of women in the population',
         panel.first = grid())
polygon(dens, density = -1, col = rgb(1,0,0,0.2))
abline(v=0.3723761, col='blue')
```

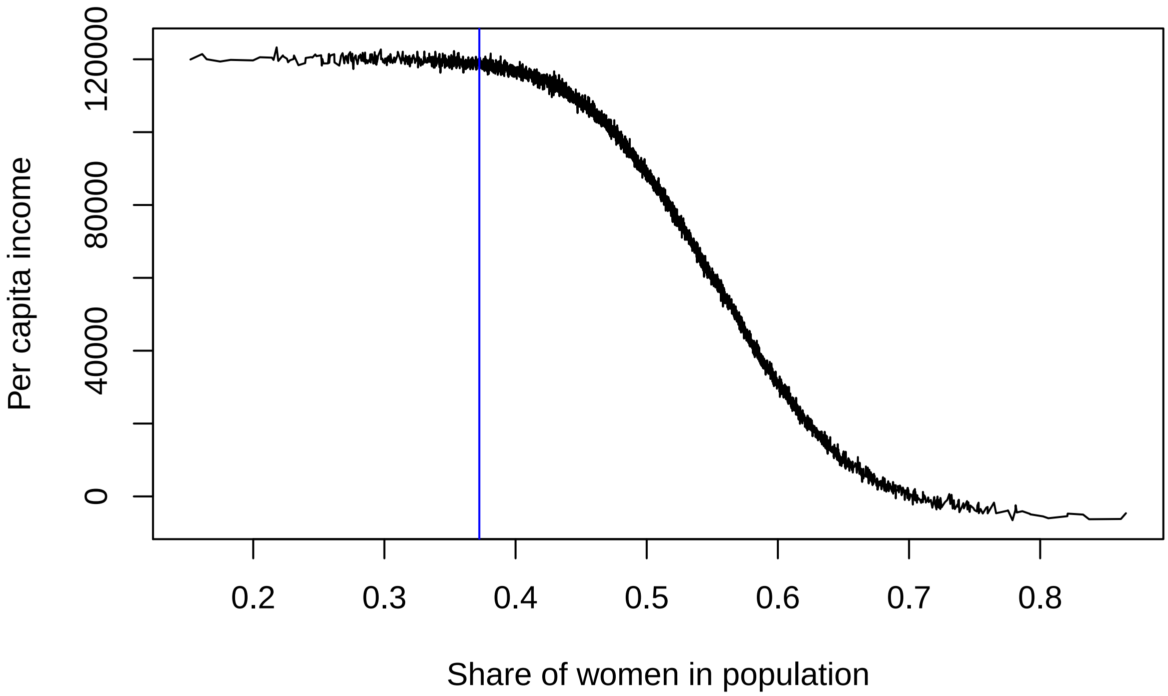


As we can see, we have what seems to be a normal distribution in the share of women in the population without any discontinouity around the cutoff.

The next test we can provide is of covariate smoothness, which is that other covariates don't have a discontinouity around the cutoff. We show this with Income per capita by plotting this variable in the y-axis with share of women in the x-axis.

As we can see in the data, the cutoff is $shareofwomen = 0.3724$, which we plot to see evidence for a discontinouity around that line.

```
plot(final_exam_2020$share_women, final_exam_2020$per_capita_income_rupees, type='l', xlab="Share of wom
abline(v=0.3723761, col='blue')
```



As we can see, there is no discontinouity around the cutoff point, which provides evidence in favor of our approach. Both tests provide evidence in support of the identifying assumptions for the reasons mentioned.

II.G- Plot the relationship between a town's position on the list and its likelihood of having a female leader. Describe what you're plotting, using a definition from the course. Plot the relationship between the probability of having a female leader and public goods provision. Describe what you're plotting, using a definition from the course. Informed by these plots, write down your preferred regression equation(s) for estimating the causal effect of female leadership on public goods provision. Defend your choice of bandwidth and any functional form choices you make.

Since our approach is a fuzzy regression discontinouity, the probability of treatment does not change from 0 to 100% around the cutoff. Given this we have towns with both female leaders and male leaders at every section of the running variable (share of women in the population or rank). To show that the probability of having a woman in power changes with being ranked in the first 500 places, we need to divide the data into bins and show an aggregated measure.

```
library(dplyr)
```

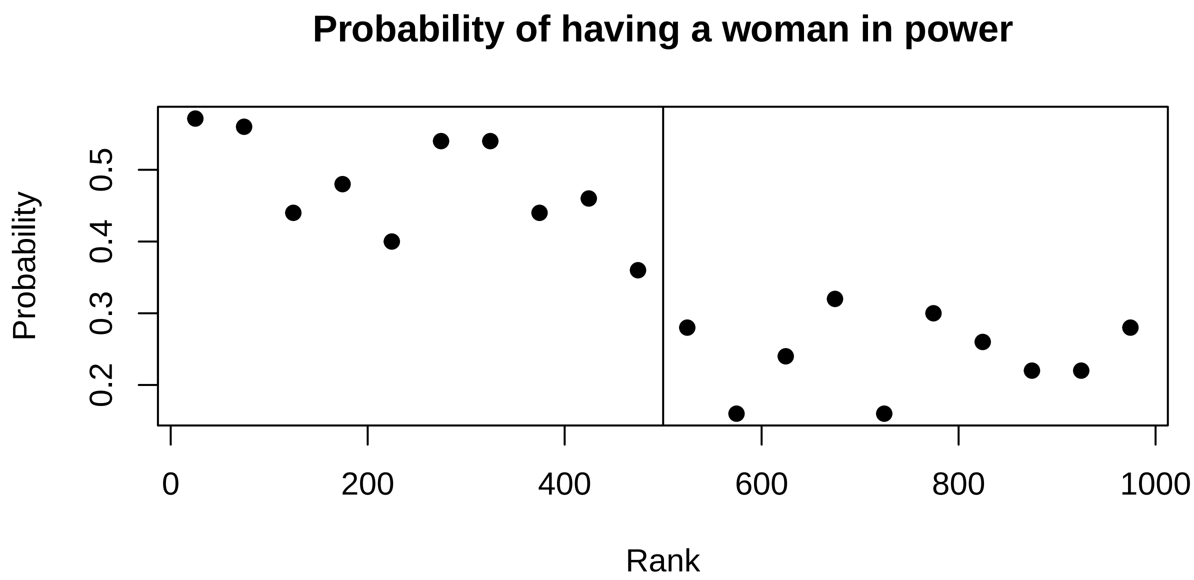
```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
final_exam_2020$bin = floor((final_exam_2020$list_rank - 500)/50)
mort = final_exam_2020 %>% group_by(bin) %>% summarise(f_leader =
mean(female_leader), med_share = median(list_rank))
```

Here, we're binning our data by every 10 positions in the ranking in Uttar Pradesh, taking the median share within each bin.

Plotting this data results in the following graph:

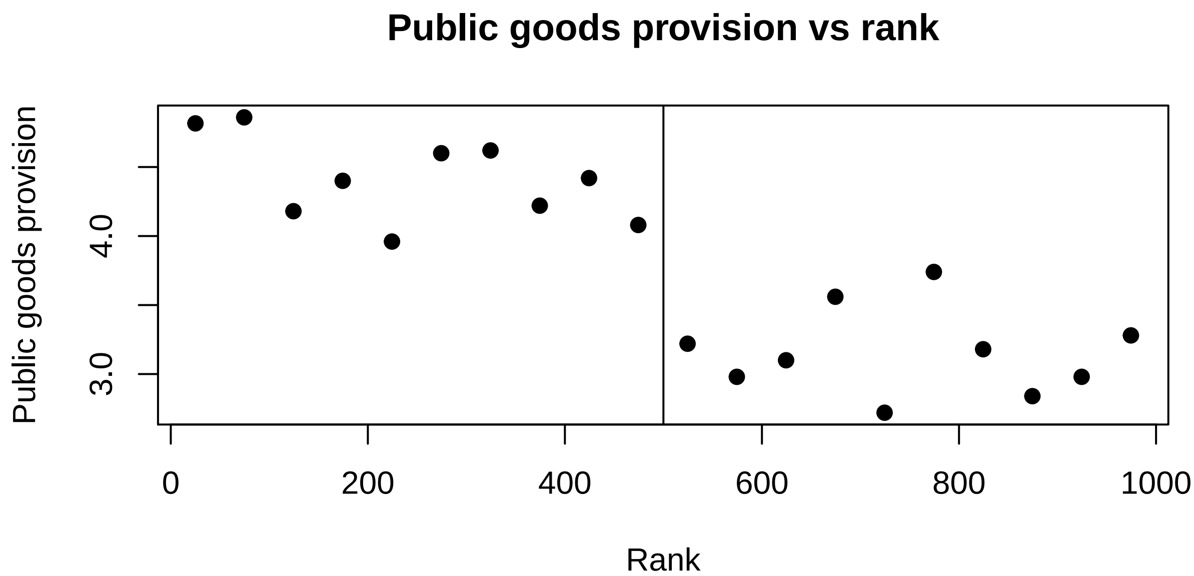
```
plot(mort$med_share[1:20], mort$f_leader[1:20], main="Probability of having a woman in power", xlab="Ra
     ylab = "Probability", pch=19)
abline(v=500)
```



As we can see, the probability of having a woman in power changes sharply around the cutoff of the running variable ($rank = 500$). Using a definition from the course, we are plotting the probability of treatment assignment, as in lecture 14 (slide 8/23).

For the next plot, as there is a typo in the question, we are actually interested in how the provision of public goods (outcome) changes with the running variable (rank). We proceed as in the last plot, binning the data and plotting the outcome:

```
final_exam_2020$bin = floor((final_exam_2020$list_rank - 500)/50)
mort = final_exam_2020 %>% group_by(bin) %>% summarise(outcome =
mean(public_goods_number), med_share = median(list_rank))
plot(mort$med_share[1:20], mort$outcome[1:20], main="Public goods provision vs rank", xlab="Rank",
     ylab = "Public goods provision", pch=19)
abline(v=500)
```



We can observe that the public goods privision has a sharp decrease at the rank's cutoff point of 500. Using a definition from the course, we can call this as the outcome versus the running variable (class 14, slide 9/23).

Given the discontinouity in treatment assignment, outcome, and the evidence in favor of the RD approach, we can implement our RD approach by a two stage least squares regression as explained in (E) using the observations around the cutoff, for which we need a selection of bandwidth. As we will consider only those towns ranked in the bandwidth around the cutoff.

First stage:

$$female_i = \alpha + \gamma \ rank_i^{\leq 500} + \beta X_i + \eta_i$$

We use the predicted values of $female$, $f\hat{emale}$ to estimate the second stage:

$$Y_i = \alpha + \tau_{RD} \ f\hat{emale}_i + \delta X_i + \nu_i$$

The selection of bandwidth allows us to estimate an average treatment effect for the compliers, that is, for towns that actually have a woman in power. The choice of bandwidth creates a trade-off between bias and variance. The LATE is estimated just in the cutoff, and moving away from it produces bias in the estimator. However, decreasing the size of the bin increases variance while reducing bias.

We use the rdbwselect() function from the R package rdrobust to select the bandwidth.

```
library(rdrobust)


summary(rdbwselect(final_exam_2020$public_goods_number,final_exam_2020$list_rank, c=500, fuzzy=final_exa
```

```
## Call: rdbwselect
##
## Number of Obs.                  5000
## BW type                        mserd
## Kernel                    Triangular
## VCE method                        NN
##
## Number of Obs.                  499        4501
## Order est. (p)                    1           1
## Order bias  (q)                   2           2
## Unique Obs.                     499        4501
##
## =======================================================
##                   BW est. (h)    BW bias (b)
##             Left of c Right of c  Left of c Right of c
## =======================================================
##      mserd   246.744    246.744    380.568    380.568
## =======================================================
```

We find that it's best to select a bandwidth around the cutoff of the rank of 247 units.

II.H- Finally, estimate the causal effect of female leadership on public goods provision. What do you find? Interpret your results. Advise FUNACRONYM: should they expand female leadership to all towns?

With the bandwidth selected we estimate a two stage instrumental variables regression in R. We use ranking in the first 500 places (or reserved seating) as an instrument for having a woman in power, and then use the predicted values of the first stage as our treatment variable in the second stage, of which the coefficient will give us the **LATE for the compliers in the cutoff**. We obtain:

```
bw <- final_exam_2020[final_exam_2020$list_rank >= (500 - 247),]
bw <- bw[bw$list_rank < (500 + 247),]

first_stage <- lm(female_leader~reservation, bw)
d.hat <- fitted.values(first_stage)
tsls <- lm(public_goods_number~d.hat+per_capita_income_rupees+manufacturing_product_share+incorp_year+nu
summary(tsls)
```

```
##
## Call:
## lm(formula = public_goods_number ~ d.hat + per_capita_income_rupees +
##     manufacturing_product_share + incorp_year + number_of_residents,
##     data = bw)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8482 -2.0046 -0.8771  2.2835  6.4956
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  4.822e+01  2.144e+01   2.249   0.0250 *
## d.hat                        3.335e+00  1.966e+00   1.696   0.0905 .
## per_capita_income_rupees    -2.719e-04  1.131e-04  -2.405   0.0166 *
## manufacturing_product_share -1.007e+01  5.492e+00  -1.834   0.0672 .
## incorp_year                 -6.211e-03  9.145e-03  -0.679   0.4974
```

```
## number_of_residents         -1.748e-07  2.271e-07  -0.770   0.4418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.52 on 488 degrees of freedom
## Multiple R-squared:  0.07183,    Adjusted R-squared:  0.06232
## F-statistic: 7.553 on 5 and 488 DF,  p-value: 7.59e-07
```

As we can see, we find an estimate for our treatment variable (d.hat) of 3.335, indicating that having a woman in power increases the provision of public goods in 3.335 units. The coefficient is statistically significant. However, since our estimator is valid for compliers around the cutoff we cannot recommend FUNACRONYM to implement the measure for all towns, but we can recommend that they expand it to a small amount of towns to the right of the cutoff.

BONUS: Find an example of a popular press article describing a study which would not pass muster for this class. Describe, in a few sentences, the study, and the main problem with the study, through the lens of this course. Attach the article, in PDF form, to your exam when you turn it in.

We attach the study referenced in the following news article about the *impact of the response of Angela Merkel in COVID-19 anxiety* from Psychology Today.(click here to go to article)

The study described was published as a correspondence in the Journal of Public Health called *Not all world leaders use Twitter in response to the COVID-19 pandemic: impact of the way of Angela Merkel on psychological distress, behaviour and risk perception* (click here to go to study). The study (and article) state that the speeches given by Angela Merkel reduced measures of anxiety and depression en Germany. The authors measured self-reported anxiety, depression, and risk of catching, suffering consequences, and dying from COVID-19 by surveying 12.244 people daily from 10 March to 24 March 2020. They followed the individuals during that period of time and report the trajectory of the average scores for the variables mentioned, attributing differences to the speeches given by the german chanciller Angela Merkel and other policy responses.

The problem with the approach is that it is comparing the average of a group of individuals that were treated (by hearing the speeches) with themselves before they were treated. Not having a control group of similar individuals who did not hear the speeches makes it impossible to isolate the trend that would have been experienced by the treated individuals regardless of the treatment. Since the authors do not perform this type of differences-in-differences approach, they are estimating a naive time series estimator τ^{TS}_{NAIVE}, which is biased when we don't sustract the counterfactual trend and have no evidence in favor of the **parallel trends assumption** holding.