**UCL DEPARTMENT OF POLITICAL SCIENCE**

**SCHOOL OF PUBLIC POLICY**

# **Essay Front Page**

Please complete this sheet electronically and attach it as the front page

for each essay submitted.

## Candidate Number: JYDW1

## Essay Title, Question Number/s or Text of Question/s you have answered:

## Final Essay Part A and B

## Essay Number: 1

## Module Code: POLS0012

## Module Title: Causal Analysis in Data Science

## Submission Date: 31/01/2024

## Word Count: 2995

**The essay cover sheet does not count towards the word count on your submission.**

Please tick this box if you **do not** agree to your essay being used anonymously in future teaching

**PART A**

**QUESTION 1:**

a)

i) Using first-stage regression, the causal estimate is -0.249. This indicates that a 1% increase in the distance (km) from the Turkish coast leads to a 0.0025 decrease in the probability of a municipality receiving immigrants. A negative coefficient suggests that municipalities closer to the coast are more likely to receive immigrants. A p-value less than 2.2e-16 confirms that the relationship between distance and the treatment is statistically significant at the 5% level.

ii) The F-statistic for this first stage result is 152.59, again clustered by municipality.

iii) The robust F-statistic (far exceeding the benchmark of 10) and significant negative coefficient validate the distance from the coast as a relevant instrument for assessing the impact of immigrant contact on local attitudes. The output suggests that islands closer to Turkey were more likely to receive immigrants, validating its use as an instrument for the probability of immigrant contact. Using this instrument helps in isolating the effect of receiving immigrants from other confounding factors, and affects the outcome only through the treatment.

b)

i) The randomization assumption implies that the placement of immigrants across the Greek islands was as good as random, influenced primarily by geographic proximity to Turkey. This assumption suggests that the variation in immigrant contact across different islands is not influenced by other confounding factors, such as pre-existing attitudes towards immigrants or socio-economic conditions, thus allowing for a causal interpretation of the study's findings.

ii) The exclusion restriction assumption means that the only way distance from the Turkish coast affects hostility towards immigrants is its impact on immigrant arrivals. This seems highly likely, as islands closer to Turkey did not have significantly different pre-crisis characteristics compared to those further away, as highlighted in the author’s placebo tests on voting and demographic indicators.

c)

The second-stage coefficient (LATE) for the asylum seeker component clustered by municipality is 0.218 with a standard error of 0.054. It is statistically significant at the 5% level with a p-value of 6.682e-05. For the immigrant component, the coefficient is 0.212 with a standard error of 0.07. It is statistically significant at the 5% level with a value of 0.003.

d)

i) The type of municipality that is a complier is either one that has received immigrants and is also less than 40 km from the coast or one that has not received immigrants and is greater than 40 km from the coast. An always-taker is a municipality that always receives immigrants, regardless of how far it is from the coast.

ii) The proportion of compliers is 0.931, and the intent-to-treat effect is 0.222

iii) Using the previous estimates, the Complier Average Causal Effect (CACE) of receiving immigrants on hostility towards asylum seekers is 0.239. This means that in compliers, receiving immigrants increases hostility towards asylum seekers by 0.239 units on the summary scale.

iv) We used a second-stage least squares model as it considers the two-stage estimation procedure by adjusting the standard errors of the estimated coefficients. This is crucial for hypothesis testing and for calculating accurate p-values. The coefficient is 0.239 and the output is statistically significant at the 5% level.

**QUESTION 2:**

a)

Using a regression discontinuity design (RDD) allows us to estimate causal effects with the LATE and interpret randomization where treatment isn’t randomly assigned by exploiting a naturally occurring cutoff (the threshold for winning an election). This is superior to using mean comparisons between Democratic and Republican sheriffs, as these could be biased due to confounding variables that influence both sheriff elections and local law enforcement policies. RDD design allows us to control for these confounders by comparing outcomes in jurisdictions where election results are close to the threshold, assuming that these counties are similar in all respects except for the party of the sheriff.

b)

The first possible violation of the RDD assumption is composite treatment, which can occur at the threshold when policies other than the treatment also change. In the case of this study, if an elected Democratic or Republican brought along a set of other correlated policy changes (for example, changes in funding that do not change at the vote share threshold), then the treatment is effectively composite and it would be difficult to attribute observed effects solely to the party of the elected sheriff. I argue that this is not very plausible, because Sheriffs effectively do not have much power to enforce those policies.

Another potential violation relates to sorting around the threshold, whereby units can cheat their way into the treatment or control group by manipulating their value for the running variable. In the context of sheriff elections, this could look like electoral fraud (i.e. a Democratic or Republican sheriff), which is illegal, but the level of oversight in this level of election might be less than in a national election, so the threat is plausible. To evaluate this, I ran a McRary test to assess the density of the running variable at the cutoff.

After running a McRary sorting test, the reported p-value was 0.454, meaning that we cannot reject the null hypothesis of no discontinuous jump in the density of the running variable at the 5% level. The plot also confirms this violation visually (see Figure 1 below).

A graph of a graph of a running variable

Description automatically generated with medium confidence

Figure 1: McRary Density Plot of Running Variable

c)

The optimal bandwidth of 2.19 implies that within 2.19 units of the cutoff for RDD analysis, there is an optimal bias-variance tradeoff. This yields a LATE estimate suggesting a 14.9 percentage point decrease in ICE detention enforcement when transitioning from just below to just above the cutoff, with a marginal significance of 0.091 at the 10% level. suggesting a potential, though not definitive, causal effect of electing a Democratic sheriff.

d)

The treatment effect (electing a Democratic sheriff) on the enforcement of immigration detentions is generally negative across the bandwidths, suggesting that Democratic sheriffs are less likely to comply with ICE detention requests compared to Republicans. The plot (Figure 2) below shows that at very low bandwidths, very little data is used in estimation, making the variance very high. At low bandwidths, the confidence interval is very wide. However, the confidence intervals include zero in all cases, which suggests that the estimates are not statistically significant and therefore the results are not sensitive to bandwidth choice.

Figure 2: Sensitivity of RD Result to the Choice of Bandwidth

A graph of a graph with a line

Description automatically generated with medium confidence

e)

Our model has used two-way fixed effects for county and year, clustering standard error by municipality. This is useful in a panel data setting where the model can control unobserved heterogeneity across the counties and years in the data. The coefficient indicated that electing a Democratic sheriff is associated with a 1.72 percentage point decrease in the share of ICE detention requests enforced. The p-value, however, was statistically insignificant at 5%, with a value of 0.612.

f)

In this case, RDD provides a marginally significant estimate and leverages a natural cutoff to control for confounders. The fixed-effects model can’t control time-invariant confounders (i.e. historical attitudes towards law enforcement) which can introduce bias as it influences both sheriff elections and the enforcement of ICE detention requests.

**PART B**

**Causal Question We Want to Answer**

Stunted growth is a prevalent issue among children in India and is an important indicator of nutrition status, as well as a predictor of various health outcomes (World Health Organization, 2014). Traditional school nutrition programs often overlook the diverse dietary needs of undernourished children, especially in resource-poor settings (Cecil and Barton, 2020). There is a gap in the literature about the benefit of a tailored nutritional plan for students in resource-poor settings, likely due to funding limitations constraints (Zenebe et al., 2018, Jamaluddine et al., 2020). However, we argue this relationship can be researched effectively through available resources and local context. This understanding will allow for the creation and implementation of early intervention measures that have the potential to massively improve healthy development across the child’s life course.

This proposed study aims to address this gap and investigate whether the implementation of personalized nutrition plans during school lunches, customized to the unique metabolic profiles of students, can lead to significantly improved outcomes in preventing stunted growth as opposed to traditional one-size-fits-all approaches across primary schools in India. We have decided to use a randomized control trial (RCT) design and have taken measures to eliminate or minimize any threats to causality upfront.

**Data We Will Use to Answer Question**

In our RCT design, our treatment group is primary schools across the Indian state of Kerala, who receive school meals tailored to their nutritional needs. The control group consists of schools that receive standard nutritional plans, not tailored to each student. This standard meal plan aligns with the mid-day meal scheme, a program sponsored by the Indian government. Control groups will receive the recommended nutritional standard of 450–700 kCal and 12–20 g of protein (Vastrad, 2023). Our primary outcome measure is the incidence of stunted growth after our predefined period of 2 academic school years.

We will be collecting baseline height-for-age data, measured by a doctor or nurse to ensure the highest validity for both treatment and control groups. For the treatment group, we will need to collect further data on metabolic profiles, including age, gender, weight, body composition, and any nutritional deficiencies via blood samples to create tailored nutritional plans. The use of a mobile health clinic to collect nutritional and height data will promote accessibility and accuracy of results. We will also be conducting follow-up assessments, and regularly measuring height-for-age every 4 months for both groups. To eliminate the threat of confounding variables relating to infectious disease, we will be providing deworming tablets to both groups.

We choose to use height-for-age as a measure, because it is a standardized and common measure of healthy growth in different health systems across the globe, meaning the results of our experiment can be easily understood in health policy outside this setting. We will be using the World Health Organization’s (WHO) definition of stunted growth, which is a height-for-age score more than two standard deviations below the WHO Child Growth Standard median (see Figure 1 below) (World Health Organization, 2014). As mentioned, stunted growth is an important measure for chronic undernutrition as it indicates that a child has received inadequate nutrients to support normal growth. Data on the socioeconomic status of each school will also be measured based on mean household income in the school’s region and accessed from publicly available census data. This will be used to stratify randomization, which we will discuss further in the next section.

|  |  |
| --- | --- |
| **Z-Score** | **Height-for-age** |
| Above 3 | Very tall |
| Above 2 | Normal |
| Above 1 | Normal |
| 0 (median) | Normal |
| Below -1 | Normal |
| Below -2 | Stunted |
| Below -3 | Severely Stunted |

Figure 1: WHO Defined Height-for-Age Z-Scores

**Method of Causal Inference, and Why is Our Study Design an Appropriate Method to Answer Our Question**

We have deliberately chosen RCT design for many reasons. Firstly, randomization into treatment and control groups ensures that any observed differences in outcomes can be confidently attributed to the treatment effect alone, offering a robust foundation for causal inferences. In the intricate landscape of nutritional research where a myriad of variables can influence growth outcomes, such as socio-economic status, baseline nutritional status, physical activity levels, and infection, the RTC provides a methodologically sound framework for minimizing confounding variables and promoting high internal validity. Moreover, the replicability of the RTC design enhances the generalizability and external validity of our findings.

**Specific Tests and Statistical Techniques Used to Apply Our Method to the Data**

In our study design, we emphasize the importance of randomization to ensure the elimination of selection into treatment bias. It is also crucial because it guards against potential confounding variables and systemic differences between treatment and control groups, making them counterfactuals and comparable at baseline. Ultimately, we can make assumptions based on what the treatment group would have done if put in the control group, and vice versa. Our method of choice is block randomization. We have chosen to pre-stratify 12 schools based on socioeconomic status (SES) as a continuous variable divided into three stratum (low, medium, and high). Stratifying by SES ensures that both groups have a balanced distribution of schools with similar socioeconomic characteristics. This reduces the risk of confounding, as SES is often associated with various health and educational outcomes independently of the intervention, through access to resources, parental education, and general health status. We will follow this by running balance tests, using T-tests specifically to compare means of continuous variables between treatment and control groups within each SES stratum. By running this test, we can confirm whether our randomization process was effective.

Figure 2: School Stratified Block Randomization Flowchart

A diagram of a school

Description automatically generated

These measures have been put in place so that we can be confident in an unbiased ATE estimate. The ATE represents the average difference in the incidence of stunted growth between the treatment and control groups and can determine the direction and strength of our relationship. A t-test will return a p-value to determine the statistical significance of this relationship.

**Potential Limitations of Our Approach**

There is the potential issue of two-sided non-compliance because children may not adhere to their respective nutrition plans. Ethically, we are unable to force children to eat their meals. However, this risk alone is relatively small. There is also the risk of children in both treatment and control groups eating foods that their parents may have packed for them. To mitigate this issue, we will send out pamphlets to parents to strongly advise against sending their children to school with packed snacks or lunches.

The Hawthorne Effect is another potential limitation of this study design if participants know they are in an experiment and change their behavior outside the experiment. This is a threat particularly in the treatment group, because families might become more reliant on the school’s meal for nutritional value, and therefore eat less nutritionally high foods outside the experiment (threatening the study’s validity). We propose utilizing deception techniques so that students and parents are unaware of the specific goal of the study beyond tracking healthy growth, which would explain regular visits with the doctor. Debrief sessions will be conducted following the experiment as per ethical guidelines.

There is also the issue of attrition if children drop out or take extended periods of leave from school. This issue is not uncommon in many schools in low- and middle-income countries, where financial, transportation, or health concerns can act as a barrier to school attendance. One thing our study design can implement is tracking and documenting reasons for dropout, to guide future study improvements. Because our study design collects interim data at multiple time points, we can collect partial data analysis even if some participants do not complete the entire study. We do, however, intentionally refrain from collecting data too longitudinally, which also alleviates this risk. For this reason, we will not implement instrumental variables.

We also justify in this section further ethical considerations relating to our study design. By providing schools in the control group with a well-rounded nutritional plan that aligns with guidelines set by local government, while it is not personalized, we argue that ethically, it is better to receive such a plan than nothing at all. In terms of deception, we support the decision to enhance the methodological quality of the study. We also believe that by ensuring more valid results, in the long run, the contribution of the knowledge will benefit many others. As part of our study, we will seek approval from an appropriate ethical review board, especially considering the study involves minors.

**What Assumptions are Needed to Yield a Causal Effect**

The main assumption needed to yield a causal effect in any RCT design is randomization. This ensures that both groups are statistically similar in all respects before the treatment, thus any differences observed after the treatment can be attributed to the treatment itself. This also means that both groups can serve as valid counterfactuals for each other. Because of block randomization and balance tests in our study design, we can be confident that this assumption is met.

Another assumption is the Stable Unit Treatment Value Assumption (SUTVA). This means that the treatment of one individual does not affect the outcomes of another, preventing a 'spillover' effect. Because our study design has treatment and control groups physically separated as different schools, this assumption is also met.

Third, missingness at random is an important assumption because it avoids a bias in the kinds of participants who drop out and the missing data. In the design of our study, we anticipate the potential issues of missing data and have implemented strategies to minimize them. This includes careful follow-up, and robust data collection methods mentioned previously.

**How Valid Will Our Results Be? Discuss and Justify any Robustness Tests Used to Mitigate These Concerns**

By ensuring proper randomization, we can be confident in high internal validity, but this could be at the expense of external validity (as is common with most RCT designs). This study was conducted across only 2 school years, so we cannot establish prolonged exposure to personalized nutrition plans. We have gone to multiple schools across Kerala and calculated the ATE across these geographically stratified schools making the sample as representative as possible. We also have chosen a field experiment design to make the treatment and research setting as close to the real world as possible. To improve external validity, we propose that this experiment be run again across different states in India, to assess the larger generalizability of the results.

**Conclusion**

In conclusion, the methodology of this study will establish the causal effect of personalized treatment plans on students in India, with threats to causal inference taken care of in the design upfront. The findings of this experiment can be used to drive policy to target stunted growth and undernutrition in other LMIC countries and help reach the WHO target of a 40% reduction in stunted growth by 2025 (World Health Organization, 2014).

References

CECIL, J. E. & BARTON, K. L. 2020. Inter-individual differences in the nutrition response: from research to recommendations. *Proc Nutr Soc,* 79**,** 171-173.

JAMALUDDINE, Z., CHOUFANI, J., MASTERSON, A. R., HOTEIT, R., SAHYOUN, N. R. & GHATTAS, H. 2020. A Community-Based School Nutrition Intervention Improves Diet Diversity and School Attendance in Palestinian Refugee Schoolchildren in Lebanon. *Curr Dev Nutr,* 4**,** nzaa164.

WORLD HEALTH ORGANIZATION 2014. Global nutrition targets 2025: stunting policy brief.

ZENEBE, M., GEBREMEDHIN, S., HENRY, C. J. & REGASSA, N. 2018. School feeding program has resulted in improved dietary diversity, nutritional status and class attendance of school children. *Ital J Pediatr,* 44**,** 16.

Appendix

# QUESTION 1

#a)

# i

load("2023essay\_q1.Rda")

library(lmtest)

library(AER)

library(multiwayvcov)

first\_stage\_model <- lm(treatment ~ logdistance, data = g)

clustered\_se<- coeftest(first\_stage\_model, cluster.vcov(first\_stage\_model, g$munid))

print(clustered\_se)

# ii

wald\_test\_result <- waldtest(first\_stage\_model, vcov=cluster.vcov(first\_stage\_model,g$munid))

print(wald\_test\_result)

#c)

second\_stage\_asylum <- ivreg(score\_asylum ~ treatment | logdistance, data = g)

second\_stage\_immig <- ivreg(score\_immig ~ treatment | logdistance, data = g)

coeftest(second\_stage\_asylum, vcov=cluster.vcov(second\_stage\_asylum, g$munid))

coeftest(second\_stage\_immig, vcov=cluster.vcov(second\_stage\_immig, g$munid))

#d)

#ii

itt <- mean(g$score\_asylum[g$low\_distance==1]) - mean(g$score\_asylum[g$low\_distance==0])

print(itt)

prop.c <- sum(g$treatment[g$low\_distance==1])/length(g$treatment[g$low\_distance==1])-sum(g$treatment[g$low\_distance==0])/length(g$treatment[g$low\_distance==0])

print(prop.c)

summary(lm(g$treatment ~ g$low\_distance))

#iii

CACE <- itt/prop.c

print(CACE)

#iv (p-value as output from 2nd stage least squares)

CACE\_pvalue <- ivreg(score\_asylum ~ treatment | low\_distance, data = g)

summary(CACE\_pvalue)

#QUESTION 2

load("2023essay\_q2.Rda")

library(rdd)

#b)

mccrary\_result <- DCdensity(sheriff$running\_var, verbose = TRUE)

#c)

band <- IKbandwidth(sheriff$running\_var, sheriff$share\_detained\_sheriff)

print(band)

rdtest <- RDestimate(share\_detained\_sheriff~ running\_var, data = sheriff)

summary(rdtest)

#d)

rdests=rdci.up=rdci.down <- c()

thresholds <- seq(from=2,to=10,by=1)

for(i in 1:length(thresholds)){

rdest <- RDestimate(share\_detained\_sheriff~running\_var,bw=thresholds[i], data=sheriff)

rdests[i] <- rdest$est[1]

rdci.up[i] <- rdests[i] + 1.96\*rdest$se[1]

rdci.down[i] <- rdests[i] -1.96\*rdest$se[1]

}

pdf("thresholds.pdf")

plot.new()

plot(rdests,

type="l",

lwd=2,

ylim=c(min(rdci.down), max(rdci.up)),

xaxt="n",

xlab="Threshold",

ylab="Estimate")

axis(1,

at=1:length(thresholds),

labels= thresholds)

abline(h=0)

lines(rdci.up,

lty=3)

lines(rdci.down,

lty=3)

legend("topright",

c("RD Estimate","95% Confidence Interval"),

lty=c(1,3))

dev.off()

#e)

install.packages("fixest")

library(fixest)

fe\_model <- feols(share\_detained\_sheriff ~ treat | county\_id + year, data = sheriff)

print(fe\_model)