

HEVAL5140 - Exam - Part I - Guide to answers

Part 1: Copayment

Context: In this paper will will examine questions related to the effect of eliminating copayments for youths when going to the physician. In order to do this, you will use data from Norway which tells you how often females in different age groups went to the physician (in different years). The copayment was eliminated in 2010. Before 2010 you had to pay about 15 EURO for every visits when if you were 12 or older. In 2010 the age threshold was increased to 16 years. This means that, for instance, 14 year old females had to pay in 2009, but not in 2010. The general question is whether this led to an increase in the number of visits to the physician.

Note that because students select different age groups as control group, they may get different number answers for some of the questions. Note also that the importance of the interpretation of the results and the discussion (since much of the calculation was aided by what we did in the class).

a) Describe

- i) Report number of observations (also per year, per age)
- ii) Accurate? Possible points for discussion:
 - 1) Is the reporting process the same across all years (e..g what if change to electronic reporting)
 - 2) Substitution: Young people may go to clinics without doctors, and old people may get help in nursing homes
 - 3) False reporting
 - 4) Bonus. Test if valid: Graph age distribution to check if something strange shows up

b) Make a figure

- Remember to label both axis
- Use average not total number of visits (as some did)
- Main comment: Increasing trend over time, but mainly an increase between 2009 and 2010
- Bonus if comment on the steepness of the curve and make a quick calculation of the percent increase between 2010 and 2011

c) Run a standard regression and d: Comment

Yes, there seems to be a (positive) time trend (significant), but the increase could also be the result of the intervention, so we need to include both variables (time and intervention dummy).

- Should comment on how large the trend is (bonus if calculate percentage)
- Both copayment dummy and trend dummy (year) are significant
- List some standard assumptions: (No omitted variables, no autocorrelation, no heteroskedasticity, linearity)
- Comment on key causal assumptions (constant causal effects? Are all confounding variables included?)
- May also comment on the relatively few observations (but also that behind each annual aggregate observation there are many individuals)
- Also bonus for comparing and discussing size of coefficients (time trend vs. copayment coefficient size)
- In order to assume that the effect is causal:
 - No omitted variables (all confounders are included)
 - Constant causal effect
- May also comment on technical assumptions: linearity, no autocorrelation and so on

e and f) Difference in difference

- Intuition: Compare “like with like” where one group has experienced the intervention and one has not
- Key assumption: Parallel trend (control group and intervention group)
- Useful (bonus) if also present the equations and explain the method (not just run the regression)
- Bonus point: In this case, the difference in difference method seems to fit well. May discuss potential weaknesses (parallel trends?)
- Present equation and results of regression
- Interpret the coefficients

g) Synthetic control

- Create a synthetic control by running a regression predicting the average number of visits for 14 year old females using two other age groups (using only data BEFORE the intervention)
- Use the synthetic control to predict developments after the intervention
- Calculate the difference between the predicted and the actual development
- Discuss results:
 - Interpret the coefficients
 - Note the problem of lack of data (have many age group, but not very many years)
 - Bonus point: Note the problem of exact identification (perfect match between synthetic control and actual visits before reform because lack of data)

h) Plausible

- Different answers are possible, but difference-in-difference with good controls is both intuitive and robust (synthetic control is difficult due to lack of data and many possible control groups)

i) Gender

- Yes, there are gender differences. Females seem to react more to the change in copay. Around 20% vs. Around 12% (Depending on age group you use)

Part 2: RHC and propensity score matching

1. RHC group has HIGHER mortality than non-RHC. Should comment that this is probably due to more severe patients in RHC groups.
2. Exactly which variables are not that important, as long as it shows some independent thought and is justified. Age should obviously be included. Also bonus if not just include only the same variables we did in class. Some discussion of trying to capture variables that indicates severity/frailty.
3. Histograms of the propensity score distribution for two groups. The comments should discuss the degree to which there is overlap.
4. Note that exact matching may only be possible if we make variables less detailed (ex. age)
5. Propensity score matching will reduce the difference relative to no adjustment, but for many it may still be the case that RHC has higher mortality. Should comment that this is probably due to inadequate adjustment for severity.
6. The key assumption is no omitted variables. Bonus if discuss that we do not need to include all kinds of variables, only those that affect both the treatment and the outcome.

Part III: Discuss article. Regression Discontinuity

- a) Should discuss that the general topic of the article is whether a specific intervention for premature babies has an effect on mortality and academic achievement later on.
- b) Size of effect: Here are some options that give points: The stated effect is in standard deviations. This could be converted to percentage. It can also be compared to the effect of other kinds of interventions or factors that we know something about (IQ, gender). In general these effects are quite large (The numbers may seem relatively small, but compared to what one might expect and the effect of many other things, the effect seems large)
- c) Assumptions behind regression discontinuity
 - i) Participation cannot be "manipulated"
 - ii) Sharp threshold (here)
 - iii) Not correlated with other factors (say parents compensate, read more etc if baby is below 1500 gram)
 - iv) More technical: assumption about functional form (linear? non-linear?)