# Term/Exam paper, HEVAL5140

## Part I: Health insurance and health spending

In this part we will investigate the relationship between insurance and health expenditure. The key question is: Do people with insurance have higher health spending than those without insurance?

To answer this, we will use data based on the Medical Expenditure Panel Survey from 2011 (USA). I have made the file ("meps2011small") available:

In xls (excel) format: https://www.dropbox.com/s/so70a3t7wa5gqp8/studmepsinsurance.xls?dl=1

You can also read more about the data and an explanation of all the variables here:

http://meps.ahrq.gov/mepsweb/data_stats/download_data_files_detail.jsp?cboPufNumber=HC-147

There are lots of variables – more than 2000 – but you should not worry about all. The important issue is not to use all kinds of variables, but to show that you know what you are doing! The key variable is unins11 (which is 1 if the person does not have insurance) and totexp11 (which is the total sum of health expenditure). Remember: You may have to recode some of the variables before you use them – for instance coding variables to 0-1 instead of 1 -2 (and 3, 4) and so on. This is done using the option "Transform" and "recode." (You may also want to save your results and the commands in a syntax or do file as you go along – and use it as an appendix if you want to).

### Questions

1. Describe the data (How many observations do you have, what are the means of the most relevant variables, how many have insurance and so on).
2. Make a figure that shows health spending in different age groups and health spending for those with and without insurance.
3. Run a simple regression to estimate the causal effect of insurance on health spending. Choose the variables and observations and type of analysis that you think is best suited (explain your reasoning and the variables you include) to estimate the effect (but do not include hundreds of variables!).
4. Estimate the same effect using the method of propensity scores. Explain and examine the results.
5. If you get different results, why do you think they differ? If the results are similar, explain when one would expect regression and matching to give similar results.
6. Draw a graph of mean health spending in the different age categories for those who have insurance and those who do not have insurance. Recall that that in the US most people get government insurance when they turn 65 years old. Does this lead to a jump or decrease in health spending? Does the jump differ depending on whether the person has insurance? How would you explain the results (briefly)?

## Part 2: The effect of copayment on visits to the physician

In 2010 copayment for visits to the physician was eliminated for individuals under the age of 16 in Norway. Before 2010 the age limit was 12 years old. Your task is to evaluate to what extent this reduction led to a change in the number of visits to the physician. You may use the data provided.

Data access (csv format):

Female: https://www.dropbox.com/s/5w4l7fvus79wvew/femaleVisitsToPhysician.csv?dl=1

Male: https://www.dropbox.com/s/74vc6yqguz4z0rx/maleVisitsToPhysician.csv?dl=1

a) Create a figure of the average number of visits for 12-15 year olds over time
b) Use a simple regression ("Interrupted Time Series") to examine if the intervention in 2010 influenced the number of visits.
c) Report and interpret the effect from the regression.
d) Discuss potential problems with using the regression results as a measure of the causal effect of the intervention
e) Describe the intuition behind the difference-in-difference method
f) Use the difference-in-difference method to estimate the effect of the intervention
g) Do you believe the difference-in-difference method captures the causal effect?
h) Examine if the effect may affect genders differently
i) Challenge: Discuss what other methods you might use to estimate the effect of the intervention (in words, no new calculations are necessary)

## Part 3: The effect of intervention for premature babies, low birth-weight

Read the article (You do not have to memorize or understand all the statistical details):

*Early Life Health Interventions and Academic Achievement*

Prashant Bharadwaj, Katrine Vellesen Løken and Christopher Neilson

American Economic Review, 2013, vol. 103, issue 5, 1862-91

a) In your own words, what is the main message of the article?
b) Discuss the size of the measured effect. Would you say it is small or large?
c) Present the key assumptions behind the regression discontinuity design and discuss whether you think the assumptions apply in this case?