Written Exam for the B.Sc. or M.Sc. in Economics Summer 2019

Applied Econometric Policy Evaluation

Take-home exam

June 18, 2019, from 10:00am to 10:00pm

This exam consists of 8 pages in total.

Answers only in English.

Be careful not to cheat at exams!
Exam cheating is for example if you:

- Copy other people's texts without making use of quotation marks and source referencing, so that it may appear to be your own text
- Use the ideas or thoughts of others without making use of source referencing, so it may appear to be your own idea or your thoughts
- Reuse parts of a written paper that you have previously submitted and for which you have received a pass grade without making use of quotation marks or source references (self-plagiarism)
- Receive help from others in contrary to the rules laid down in part 4.12 of the Faculty of Social Science's common part of the curriculum on cooperation/sparring

You can read more about the rules on exam cheating on your Study Site and in part 4.12 of the Faculty of Social Science's common part of the curriculum.

Exam cheating is always sanctioned by a written warning and expulsion from the exam in question. In most cases, the student will also be expelled from the University for one semester.

Practical instructions for the take-home exam

Read the entire exam before you respond. Answer every question in each problem. The exam consists of five problems in total.

The exam can be answered in groups of a **maximum of 2 students**. Hand-in a single report for the entire group **and specify each group member's contribution to the report**.

You must submit a comprehensive report with relevant tables and figures. The front page of the report must use the template available at `https://eksamen.ku.dk/`. Fill in the exam numbers of all group members on the front page. The second page of the template must specify which paragraphs and/or sections of the report is answered by which group member. This page may not contain other information.

Prepare one STATA do-file generating all tables and figures that appear in your report. The program must produce tables and figures in the same order as they appear in the report. Comments should clearly indicate which table or figure appearing in the report is being produced. Make sure that the do-file can be executed without any errors. The do-file must include the exam numbers of all group members.

The report must not exceed 12 (normal) pages. This includes the main text, tables and figures in the report, but not the front page and the list summarizing each group member's contribution to the report.

For the exam in Applied Econometric Policy Evaluation, a normal page is defined as a text document with the following attributes[1]:

- A4 format
- Font size set to 12
- Line spacing set to 1.5
- Margins (left/right/top/bottom) of at least 2.5 cm

The exam ends **June 18 at 22.00 (10:00pm)**. The report and the STATA do-file must be uploaded electronically no later than 22.00 (10:00pm).

Uploading your report

Each group must hand-in only one report in total. One student hands in the report by uploading it to University of Copenhagen's Digital Exam system and then adding the rest of the group members to the hand-in. Go to the website `https://eksamen.ku.dk/` and click on 'Log in as student'. Use your regular KU login and password to enter Digital Exam. Click on 'Applied Econometric Policy Evaluation' in your assignments. On the page 'Information about the hand-in', you must add the other group member to the handed-in answer (if you are in a group). Click on 'Add member' and follow the instructions on Digital Exam to invite your fellow group members. Group members will be added to the handed-in answer as soon as they **accept** your invitation.

[1]The Study Handbook for the Economics program defines a normal page as 2,400 characters, but for this exam, a normal page is instead defined in terms of format, font size, line spacing and margins.

Next, go to 'Upload hand-in' to upload your files. Each group must upload two files:

1. The report itself must be uploaded as a PDF file. The filename must start with the letter R followed by the exam numbers of all members of the group separated by _ ("underscore").
2. The STATA do-file must be uploaded as a file in plain text format (.txt). The filename must start with the letter P followed by the exam numbers of all members of the group separated by _ ("underscore").

Use the same combination of exam numbers for both files.

> Example: A group of two members with exam numbers 72 and 174 will submit the following files:
>
> 1. R_72_174.pdf
> 2. P_72_174.txt

If needed, a free PDF converter is available at `www.pdf995.com`.

If you have problems accessing the Digital Exam system at the deadline of the take-home exam or if you have difficulties with the upload function you must e-mail your answer to `samf-fak@samf.ku.dk` within 22:30 (10:30pm). Handing in your exam answer by e-mail requires that you describe the problems and provide screen dumps that document this.

Access to data

For the take-home exam, there are several data sets available on the Digital Exam website (`https://eksamen.ku.dk/`). Follow the instructions below to pick the correct data set for your group:

1. Determine the **lowest** number among the exam numbers of the group members. Use the **last** digit of the **lowest** exam number as your "group number".

   > Example: A group of two members with exam numbers 72 and 174 will have "2" as the last digit of the lowest exam number.

2. Download the STATA file groupdataX.dta from the Digital Exam website, where X is equal to the group number.

   > Example: The group from before downloads groupdata2.dta from the Digital Exam website.

3. Download the data to your computer.
4. Open the data in STATA and execute the **describe** command to ensure the data appears operative.

If you have trouble selecting or opening the data, you can contact Søren Leth-Petersen on telephone 3532 3084 or Daniel le Maire on telephone 3532 3063 during the period 10.00am to 12.00pm (noon) on June 18.

After this, no additional help will be provided for the exam.

Introduction to the assignment:

"Measuring the effect of housing subsidies on rents"

Many countries have housing subsidy schemes. The aim of housing subsidy schemes is to level out-of-pocket housing expenses of recipients. Housing subsidies are a supplement to public social insurance schemes and are typically targeted low-income households living in rented housing. A typical feature of housing subsidy programs is that they grant subsidies depending on the characteristics of the household, for example according to income level and household size, and the level of subsidy tends to (weakly) increase with the size of the dwelling, albeit at a decreasing rate, so that the average per square metre subsidy is smaller in large dwellings than in small dwellings.

Specifically, the housing subsidy is designed so that there is a subsidy depending on the size of the dwelling according to the following schedule:

Table 1: m2 subsidy rate (m2rate)

Floor area (m^2)	rate per m^2
base subsidy	150
reduction in subsidy for $m^2 \geq 30$	-100
additional reduction in subsidy for $m^2 \geq 55$	-50

The full housing subsidy is given to households with a total household income below a level corresponding to 1.5 times the full-time earnings at the minimum wage, i.e. when income of the household before taxes is below 24,000 DKK per month. When income exceeds this level, the subsidy is reduced with 30 percent of each additional DKK earned above the threshold. The subsidy increases with 250 DKK per month for each child for the first four children. Finally, the housing subsidy can never exceed 85 percent of the rent, and it can never be negative.

The housing subsidy is granted according to the following formula:

$$\begin{aligned} subsidy_i = & min[0.85 \times rent_i, [150 \times 1(m_i^2 < 30) \times m_i^2 + 50 \times 1(30 \leq m_i^2 < 55) \times m_i^2 \\ & - 0.3 \times max\left[0, income_i - 24,000\right] + min\left[\#children_i \times 250, 1000\right]]] \end{aligned}$$

An important issue with housing subsidies is, however, that they arguably affect housing subsidy recipients' willingness to pay for rental housing, and this could affect the rents set by landlords. The objective here is to quantify the effect of housing subsidies on rents. You have access to a data set including observations for individuals who receive housing subsidies. The data set includes information about the characteristics of the recipient, such as family structure, age, income, and the characteristics of the dwelling such as square footage, the rent, and the subsidy received. The data are recorded for 2016.
The variables available are summarized in Table 1 below.[2]

[2]The data used for this exam are simulated.

Table 2: List of variables

Variable name	*Description*
pnr	Social security number (anonymized)
age	Age by end of 2016
adults	Number of adults
children	Number of children
income	Monthly income, DKK
m2	size of the dwelling in square metres
rent	monthly rent, DKK
subsidy	monthly housing subsidy, DKK

Problem 1 (15%):

1. Provide a descriptive analysis of the variables in your data set using relevant summary statistics. Examples of relevant aspects to include are number of observations, the distribution of rents, subsidies, income, square footage and age. Note that these may not be the only aspects of the data that are relevant to describe. The descriptive analysis may include both table(s) and relevant graphical illustration(s) of the data

2. Graphical analysis. Construct binned scatter plots showing how the rent and the subsidy are related to the size of the dwelling, i.e. two panels. Insert vertical lines at $m2 = (30, 55)$. Comment on your findings.

Problem 2 (25%):

1. Answer the three key questions for an empirical analysis:

    1. What is the causal relationship of interest?
    2. What is the ideal experiment that would capture the causal relationship?
    3. Based on the housing subsidy formula and the descriptive evidence from Problem 1, what is a relevant identification strategy?

2. Consider the following regressions:

$$rent_i = \beta_0^{30} + \beta_1^{30} D_i^{30-54} + \beta_2^{30}(m2_i - 30) + \beta_3^{30} D_i^{30-54} \times (m2_i - 30) + \beta_4^{30} X_i + u_i \quad (1)$$
$$rent_i = \beta_0^{55} + \beta_1^{55} D_i^{55+} + \beta_2^{55}(m2_i - 55) + \beta_3^{55} D_i^{55+} \times (m2_i - 55) + \beta_4^{55} X_i + u_i \quad (2)$$

    $rent_i$ is monthly rent paid by household i. D^{30-54} is a dummy variable taking the value one if the dwelling is sized 30-54 square metres and D^{55+} is a dummy variabel taking the value one if the dwelling is sized 55 square metres and above. X is a vector of control variables including income, age, number of children, and number of adults. Explain how each parameter in the equation is to be interpreted.

3. Estimate equation (1) by OLS using only observations for $m2 \in (0, 45)$. Subsequently, estimate equation (2) by OLS using only observations for $m2 \in (35, 80)$. Insert the two sets of parameter estimates for $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ in a table and comment on the results. Is there any evidence that housing subsidies affect rents?

4. Add squared terms of all the variables that include $m2_i$ in (1) and (2). Estimate these equations by OLS and report in a table.

    (a) What is (are) the parameter(s) of interest?

    (b) Does this change the conclusion about the effect of housing benefits on the housing rents?

5. Estimate the model while, first, calculating robust standard errors and, next, calculating errors clustered by m^2. Explain why this might be relevant and describe whether using robust and clustered standard errors might change you conclusions. Which type of error calculation is to be preferred.

Problem 3 (15%):

1. Construct a simple histogram of the square metres to check whether there is bunching at the thresholds. [Hint: No formal test is asked for.] Depict the histogram with vertical lines at the thresholds. Would you intuitively expect bunching at the thresholds?

2. Is there sorting that leads to discontinuities or kinks in background characteristics?

3. Are results sensitive to observations positioned exactly in the cut-off points?

Problem 4 (20%):

Irrespective of the conclusions reached in Problem 2 and 3, consider the following regression equation only for the $m^2 = 30$ cutoff.

$$rent_i = \alpha_0^{30} + \alpha_1^{30} D_i^{30-54} + \alpha_2^{30}(m2_i - 30) + \alpha_3^{30} subsidy_i + \alpha_4^{30} X_i + u_i \qquad (3)$$

This equation can be estimated by 2SLS, where $D_i^{30-54} \times (m2_i - 30)$ is used as an instrumental variable for $subsidy_i$.

1. What can be learned from estimating equation (3) by 2SLS compared to estimating equation (1) by OLS?

2. Estimate and insert the parameter estimates for the first stage equation and the equations of interest in a table and comment on the results. Estimation of equation (3) and the associated first stage equation should be based only on observations where $m2 \in (0, 45)$.

3. Is the estimated effect of the subsidy on the monthly rent similar whether a sharp or fuzzy design is used?

Problem 5 (25%):

The Working Environment Authority (WEA) carries out inspections in order to prevent occupational accidents and diseases and to ensure safe and healthy work places. If a WEA inspection uncovers health and safety problems it can lead to various responses. These include guidance or a ruling that the company must carry out specific actions. The WEA can also issue fines or notify the police in the event of serious violations of the Working Environment Act.

The WEA works under fairly explicit rules stipulating how to choose firms for inspection. The probability of a visit by WEA depends on the following six variables: i) Number of employees, ii) Whether the firm has been established within the last two years, iii) Whether the WEA has visited the firm within the last three years iv) Previous assessment of the WEA in case of a WEA inspection at the firm within the last three years, v) Industry factors such as the number of reported on-the-job injuries last year, noise, ergonomic work environment etc. aggregated into a single industry risk factor using the 19 industry standard grouping, vi) The number of reported on-the-job injuries last year.

How the WEA selects firms to visit is unknown to you as a researcher and may change from year to year. However, WEA is only allowed to select firms for inspection based on the variables mentioned above. This means that if two firms have the exact same values of these variables, it may still be the case that only one of the firms actually gets inspected by the WEA, even though both firms have the same probability of being inspected.

You have access to a panel data set for 2008-2015 with the following variables:

- Firm id
- Year
- Number of employees in the firm
- A dummy for whether the firm has been established within the last two years.
- Industry codes, the 19 industry standard grouping
- An indicator variable for visit by the WEA within the last three years
- Previous assessment if assessment was made within the last three years
- Date of inspection if inspection occurred in the current year
- Assessment of WEA if visit made in the given year
- The number of on-the-job injuries in the given year
- Firm profits in the given year

Ideally, the WEA's inspections should reduce on-the-job injuries without imposing too large costs related to improving the working conditions and safety. You have been asked to examine the effect of WEA's inspections on a) the number of on-the-job injuries and the b) the profits of the firms.

1. What identification strategy would you propose?
2. How will you implement the proposed identification strategy?
3. Can you provide a check for the validity of the identification strategy?
4. How can one examine the short-term effects of inspection (the effect on e.g. firm profits in the same year as the inspection) and medium-term effects (the effects on e.g. firm profits 1-2 years after the WEA visit)?