vrije Universiteit amsterdam

School of Business and Economics

| | |
|---|---|
| Exam: | Empirical Economics |
| Code: | E_EOR3_EEC |
| Coordinator: | Bas van der Klaauw |
| Co-Reader: | Casper Burik |
| Date: | December 18, 2018 |
| Time: | 15.15-18.00 |
| Duration: | 2 hours and 45 minutes |
| Calculator allowed: | yes |
| Graphical calculator allowed: | no |
| Answer in: | English |
| Credit score: | For each question you can earn a maximum number of points as indicated in the individual questions. In total you can earn 100 points. |
| Grade: | The grades will be made public asap. |
| Inspection: | January 11, 2018 between 10.00 and 11.00am |
| Number of pages: | 13 (including this page). |

**Name :**

**Student number :**

**Good Luck!**

**Question 1. [30 points in total]**

(*i*). **[6 points]** Why is tie-breaking required and what is the difference between single tie-breaking and multiple tie-breaking?

(*ii*). **[6 points]** In a randomized experiment there is full compliance in the control group, but partial compliance in the control group. Explain why this experiment estimates the Average Treatment Effect on the Treated.

(*iii*). **[6 points]** What does the assumption of monotonicity imply?

(*iv*). **[6 points]** What is the difference between multinomial logit and ranked-order logit?

(*v*). **[6 points]** Give three threats to the internal validity of randomized experiments. For each reason give a short on sentence description.

**Question 2. [16 points total]**

A researcher has data from the Dutch flower auction, which is located in Aalsmeer, just South of Amsterdam. The Dutch flower auction is the largest auction house in the world and contains many auction halls in which different types of flowers are auctioned simultaneously. The auctioning starts early in the morning at 6.30am. From that time bidders enter the different auction halls. To give all sellers an equal opportunity for selling their flowers at a high price, every day the order in which flowers are auctioned is randomized. The researcher is interested in testing the auction theory that more bidders increase competition and thus prices should be higher when more bidders are present in the auction. Therefore, the researcher specifies the model

$$\text{Price}_i = \beta_0 + \beta_1 \text{Bidders}_i + U_i$$

where $\text{Price}_i$ is the transaction price of flowers $i$ at the auction and $\text{Bidders}_i$ is the number of bidders present when flowers $i$ were auctioned. The researcher uses data on one type of flowers and finds the following Stata output.

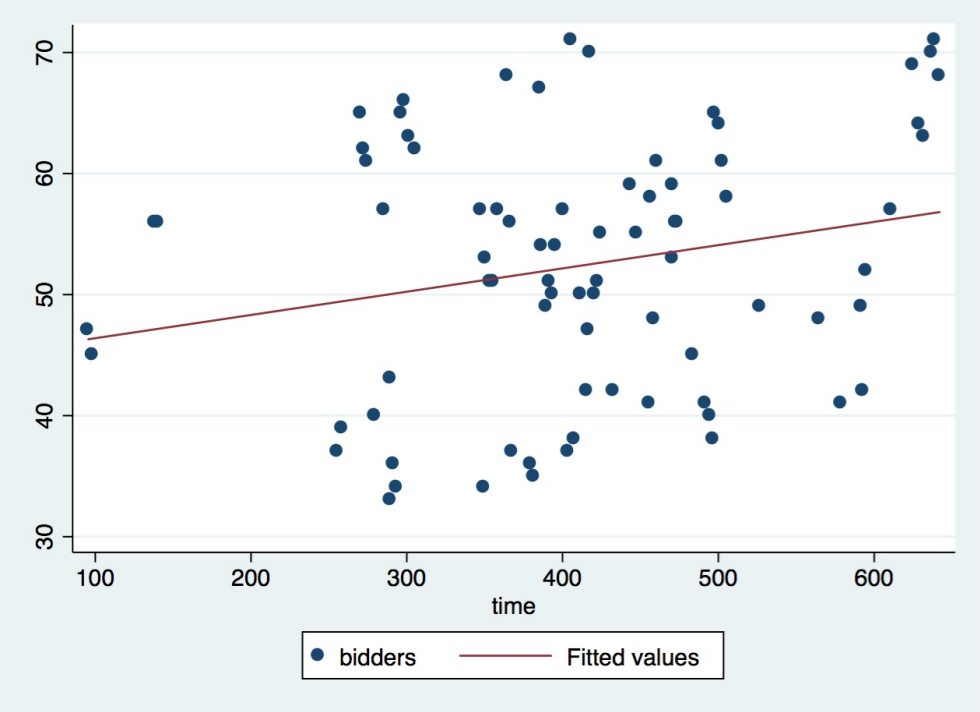
```
. reg price bidders

      Source |       SS       df       MS              Number of obs =      79
-------------+------------------------------           F(  1,    77) =   22.16
       Model |  154282.716     1  154282.716           Prob > F      =  0.0000
    Residual |  536085.158    77   6962.1449           R-squared     =  0.2235
-------------+------------------------------           Adj R-squared =  0.2134
       Total |  690367.873    78  8850.87017           Root MSE      =  83.439

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     bidders |   4.195584   .8912618     4.71   0.000     2.420855    5.970313
       _cons |  -49.48514   47.61859    -1.04   0.302    -144.3059     45.3356
------------------------------------------------------------------------------
```

($i$). [**5 points**] The researcher is happy to find a positive coefficient for the number of bidders. But do you believe that the estimated effect represents a causal effect? Briefly explain your answer.

The researcher finds that just after the auctioning starts there are, on average, fewer bidders are present. And that afterwards the average number of bidders increases. This is expressed in the figure below.




Based on this figure someone suggests to the researcher to use time as instrumental variable for the number of bidders.

(*ii*). **[6 points]** Do you believe that time can be an instrumental variable for the number of bidders? Explain your answer.

To investigate the idea of an instrumental variables approach further, the researcher decides to run the following regression in Stata.

```
. reg bidders time

      Source |       SS       df       MS              Number of obs =      79
-------------+------------------------------           F(  1,    77) =    4.14
       Model |  447.462768     1  447.462768           Prob > F      =  0.0453
    Residual |  8317.14483    77  108.014868           R-squared     =  0.0511
-------------+------------------------------           Adj R-squared =  0.0387
       Total |  8764.60759    78  112.366764           Root MSE      =  10.393

------------------------------------------------------------------------------
     bidders |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   .0192295   .0094478     2.04   0.045     .0004165    .0380425
       _cons |   44.47341   4.056705    10.96   0.000     36.39548    52.55134
------------------------------------------------------------------------------
```

(*iii*). [**5 points**] Explain why the researcher performs this regression and what is the conclusion?.

**Question 3. [13 points total]**

A researcher is conducting a randomized experiment. In the original power calculation she assumed that there would be full compliance both in the treatment and in the control group. This power calculation shows that the targeted minimum detectable effect size requires 4800 individuals in the experiment. In the original setup 50% of the individuals in the experiment are assigned to the treatment group.

(*i*). **[6 points]** The person funding the experiment states that proving the treatment is costly and suggests to only assign 25% of the individuals to the treatment group. How many individuals are then required to maintain the same minimum detectable effect size?

(*ii*). **[7 points]** It is decided to keep 50% assigned to the treatment group, but it turn out that noncompliance in the treatment group can be expected. How many individuals are required in the experiment is treatment participation is 80% in the treatment group and absent in the control group?

**Question 4. [16 points total]**

In a population, each individual is uniquely characterized by its color. There are three types of individuals, purple, blue and green individuals. The researcher wants to know how a treatment affects outcomes. The table below shows for each color how many individuals are in the treatment and control group and what the average outcomes are for individuals in the treatment and control group.

| colour | number of individual | | average outcome | |
|---|---|---|---|---|
| | treated | control | treated | control |
| purple | 50 | 50 | 10 | 8 |
| blue | 75 | 25 | 12 | 7 |
| green | 25 | 75 | 11 | 10 |
| red | 50 | 50 | 13 | 11 |

(*i*). **[8 points]** Compute the average treatment effect (in the full population).

(*ii*). [**8 points**] Compute the average treatment effect on the treated.

**Question 5.** [**11 points total**]

There are approximately 40,000 highway traffic fatalities each year in the US. Approximately one-fourth of fatal crashes involve a driver who was drinking. Suppose that you have obtained annual data for 48 neighboring US states covering the period 2011-2017. The outcome variable is the logarithm of the number of traffic deaths in a given state in a given year. The panel data also contain for each state and each year the alcohol tax (in US dollars for each liter alcohol). Suppose that you study the effectiveness of a higher alcohol tax on traffic fatalities using a static fixed effects model.

(*i*). [**5 points**] Write down the fixed effects panel data model to investigate the effectiveness of a higher beer tax on traffic fatalities. effects.

Suppose that you estimate your model using the within estimator. You find that the coefficient estimate on alcohol tax is -0.037 and the standard error is 0.011.

(*ii*). **[6 points]** How should you interpret the coefficient? And is this a causal effect?

**Question 6. [14 points total]**

A tennis club has five male tennis players and five female tennis players. The players decide to participate in a tournament four mixed doubles. This means that only a couple consisting of one female and one male player can participate. The players have some preferences about who will be their partner in the tournament.

| | Preferences | | | | | | Preferences | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | | (1) | (2) | (3) | (4) | (5) |
| W1 | M1 | M2 | M3 | M4 | M5 | M1 | W4 | W3 | W1 | W2 | W5 |
| W2 | M1 | M2 | M5 | M4 | M3 | M2 | W3 | W1 | W2 | W5 | W4 |
| W3 | M5 | M1 | M3 | M2 | M4 | M3 | W3 | W1 | W5 | W4 | W2 |
| W4 | M4 | M5 | M2 | M3 | M1 | M4 | W2 | W4 | W5 | W1 | W3 |
| W5 | M2 | M1 | M5 | M3 | M4 | M5 | W1 | W5 | W3 | W4 | W2 |

Couples of players have to register themselves for the tournament, but the trainer suggests to make the following couples.

| | |
|---|---|
| W1 | M1 |
| W2 | M2 |
| W3 | M3 |
| W4 | M4 |
| W5 | M5 |

(*i*). [**4 points**] Explain why the players do not follow the suggestion of the trainer.

One female player suggests to use deferred acceptance to make couples. And her idea is that the female players make a first proposal to the male players.

(*ii*). [**3 points**] The idea of the female player is that the female players make a first proposal to the male players. Why does she advise this?

(*iii*). **[7 points]** Will the resulting allocation be Pareto efficient?