Universidad Carlos III de Madrid

Applied Economics

May 2019

Exam duration: 2 hours and 30 minutes

Type of Exam: 1

Identification number

Name

Group

DO NOT DETACH ANY SHEET FROM THE EXAM.

DO NOT OPEN THE EXAM BEFORE YOU ARE TOLD TO DO SO.

Read carefully the following instructions:

(A) The exam contains **4 multiple choice questions** (4 points each) and 17 **short-answer questions** (points given in each question).

(B) Space is limited. Only use the space allocated specifically for the answer of each question. You can use the back of the exam sheets as scrap paper. We will not grade answers that lay outside the allocated space. Answer in a clear and concise manner.

Questions	**Points**	**Grade**
Section I - 5 short answer questions	$5 \times 5 = 25$	
Section II - 5 short answer questions	$5 \times 5 = 25$	
Section III - 2 short answer questions	$1 \times 6 + 1 \times 3 = 9$	
- 4 multiple-choice questions	$4 \times 4 = 16$	
Section IV - 5 short answer questions	$5 \times 5 = 25$	
Total	**100**	

Answer Sheet

Type: 1

Multiple Choice Questions

	a	b	c	d
12	◯	◯	◯	◯
13	◯	◯	◯	◯
14	◯	◯	◯	◯
16	◯	◯	◯	◯

SECTION I. POLITICAL REPRESENTATION OF WOMEN - LPM, PROBIT

We have a representative sample of 1052 electoral candidates contesting political office.
We have the following variables:
- *elected* is equal to 1 if candidate was elected; 0 otherwise.
- *woman* is equal to 1 if candidate is a woman; 0 otherwise.
- *age* indicates the age of candidate.
- *university* is equal to 1 if candidate has university education; 0 otherwise.
- *white_collar* is equal to 1 if candidate has white-collar job; 0 otherwise.

We are interested in the probability of getting elected given candidate's observable characteristics.
$Pr\left(elected = 1 | woman, age, university, white\_collar\right) =$
$= F\left(\beta_0 + \beta_1 woman + \beta_2 age + \beta_3 university + \beta_4 white\_collar\right) \qquad (*)$

Hint: after a probit estimation, the value for '$f(beta'x)$ at mean of independent vars' is the corresponding pdf evaluated at the average values of all the regressors.

1. [5 points] What is the average probability of getting elected? What is the average probability of getting elected by gender? Ignoring other factors, does the probability of getting elected significantly differ by gender? Report the p-value of the test.

2. [5 points] Considering the Probit model (∗), what is the probability of getting elected for a 35 years old woman with university education and white-collar job? What would be the probability of getting elected for a similar woman without university education? Can university degree be considered a significant factor for the probability of getting elected?

3. [5 points] Considering the Probit model (∗) and *age* as a continuous variable, what is the contribution of an increase in one year of age to the probability of being elected for an average individual? Does age impact the probability of candidate to get elected?

4. [5 points] Use Probit estimates of model $(*)$ to test that the electoral benefits of university education are exactly offset by the penalty for being a woman. Explain your answer using an appropriate test, stating the null hypothesis, the alternative hypothesis, the test statistic, and your conclusion.

5. [5 points] Propose an enlarged Probit model for $Pr(elected = 1|.)$ that allows for a differential effect of university education by gender. Using your new model, how would you test that the effect of university education does not differ by gender? State the null hypothesis in terms of the parameters of your model, state the test statistic that you would use and its distribution under the null hypothesis(reporting if necessary the number of degrees of freedom).

SECTION II. WAGES PAID TO POLITICIANS - IV METHODS

This section aims to investigate if wages paid to politicians can be an effective tool to attract educated individuals into politics. We have access to a cross-sectional survey of municipal mayors with the following information:

- *leduc* is the logarithm of mayor's years of education (the outcome variable)
- *lwage* is logarithm of monthly wage paid to mayor in a given municipality.
- *woman* is equal to 1 if mayor is female; 0 otherwise.
- *lage* represents logarithm of mayor's age.
- *lkmsq* represents logarithm of municipal area measured in square kilometers.
- *lpop* represents logarithm of municipal population.

The considered model looks like as follows:

$$leduc_i = \beta_0 + \beta_1 lwage_i + \beta_2 woman_i + \beta_3 lage_i + \epsilon_i$$

You are however concerned that better-qualified politicians might have set higher wages for themselves, which could make wages endogenous. Your colleagues advise you that municipal area and population could be good instruments for wage, as mayors in larger municipalities tend to receive higher wages. On the other hand, you suspect that some politicians may try to persuade citizens to move to their municipality, so that wages could increase.

6. [5 points] Interpret the coefficients on *lwage* in columns (1) and (3) in the table reported in the statistical output for Section II. Compare both results and comment if they suggest something about the exogeneity of *lwage*. Do you reach the same conclusion if you compare coefficients on *lwage* in columns (1) and (5)?

7. [5 points] Is it possible to test if *lwage* is exogenous in the models estimated in columns (3) and (5)? And in column (7)? If exogeneity can be tested with the available information, perform the test(s), and explain your conclusion.

8. [5 points] Can you test that *lpop* and *lkmsq* are relevant instruments for *lwage*? Can the instruments be considered strong? If possible, support your answers with test(s), and explain your conclusion(s).

9. [5 points] Is it possible to test the exogeneity of instruments in columns (3) and (5)? And in column (7)? Explain your answer. In addition, if exogeneity can be tested with the available information, perform the test(s), and explain your conclusion(s).

10. [5 points] Taking into account all the analysis above, can you conclude that wages paid to politicians attract better educated candidates into politics? Which column in the table corresponds to the preferred model?

SECTION III. WAGES PAID TO POLITICIANS - DIFF-IN-DIFF

This section investigates the impact of wages paid to politicians on electoral candidacy of white-collar workers using difference-in-differences approach. We have information about Czech municipalities in 2014 and 2018 elections. The wages of Czech municipal politicians are set centrally by the government. Specifically, before the year 2018, there was a substantive wage premium for elected officials in large municipalities. In 2018 the central government decided to unify wages leveling up salaries for elected officials in small municipalities. This effectively meant that wages rose faster in "small" municipalities compared to "large" municipalities.

Our dataset contains the following variables:
- $white\_collar$ is the share of white-collar candidates in a given municipality
- $small$ is a binary variable equal to 1 for "small" municipalities; and 0 otherwise
- $after$ is a binary variable that takes value 1 for 2018 elections; and 0 for 2014 elections
- $interact$ is an interaction term between variables $small$ and $after$

We use these variables to estimate the following model using OLS:

$$white\_collar_{it} = \beta_0 + \beta_1 small_i + \beta_2 after_t + \beta_3 interact_{it} + \epsilon_{it}$$

11. [6 points] Compute the percentage of white-collar candidates a) in small municipalities in 2014, b) in small municipalities in 2018, and c) in large municipalities in 2014.

12. [4 points] Was the share of white-collar candidates in 2018 significantly higher in small compared to large municipalities?

    (a) No, as β_2 is not statistically different from zero at conventional significance levels.

    (b) Yes, as the sum of β_2 and β_3 is statistically greater than zero at the 5% level.

    (c) No, as the sum of β_1 and β_3 is not statistically greater than zero at the 5% level.

    (d) Yes, as β_3 is statistically larger than zero at the 5% level.

13. [4 points] Is the share of white-collar candidates in large municipalities higher in 2018 compared to 2014?

    (a) Yes, as β_1 is negative and statistically larger than zero at the 5% level.

    (b) Yes, as the sum of β_2 and β_3 is greater than zero at the 1% level.

    (c) No, as the sum of β_1 and β_3 is not greater than zero at the 10% level.

    (d) No, as β_2 is not statistically larger than zero at the 10% level.

14. [4 points] Is the share of white-collar candidates in small municipalities higher in 2018 compared to 2014?

    (a) Yes, as β_1 is statistically larger than zero at the 5% level.

    (b) Yes, as the sum of β_2 and β_3 is statistically greater than zero at the 5% level.

    (c) Yes, as β_3 is statistically larger than zero at the 5% level.

    (d) No, as β_2 is not statistically larger than zero at standard significance levels.

15. [3 points] Under the parallel paths assumption, what would be the percentage of white-collar candidates in 2018 small municipalities if the wage reform had not been implemented?

16. [4 points] Under the parallel paths assumption, are wages paid to politicians an effective tool to attract individuals from white-collar occupations to electoral candidacy?

    (a) Yes, as the sum of β_2 and β_3 is statistically greater than zero at the 5% level.

    (b) Yes, as β_3 is statistically greater than zero at the 5% level.

    (c) No, because the share of white-collar workers among candidates in small municipalities is smaller compared to large municipalities.

    (d) No, as the share of white-collar workers among candidates did not increase in 2018 compared to 2014.

SECTION IV. ECONOMIC PERFORMANCE AND RE-ELECTION CHANCES - PANEL METHODS

We examine the hypothesis that good economic conditions and expansionary fiscal policies help politicians get re-elected. We use data about a panel of 55 countries over the 1986-2003 period. Specifically, we use the following variables:

- *country* identifies the country.
- *year* identifies the year in which elections occur. Elections are typically held every 4 years.
- *reelect* is the binary outcome equal to 1 if country leader was re-elected in a given year.

The regressors are the following:

- *ddef* is the change in government surplus ratio to GDP in the election year compared to the previous year. (For example, if we observe 1 percentage point growth in the ratio of government surplus to GDP across years, the variable takes value 0.01)
- *gdp_gr* is a per-capita GDP growth in the last year of leader's term. (For example, if we observe 1% per-capita GDP growth, the variable takes value 0.01)
- *maj* is a dummy equal to 1 if a country has a majoritarian electoral system, 0 otherwise.
- *nd* stands for "new democracy". It is a dummy variable equal to 1 for the first 4 election periods after a country turned into a democracy, 0 otherwise.

17. [5 points] According to Model 1, how much would leader's re-election chances change if government surplus growth in the election year increased by additional 0.5 percentage points and GDP growth increased by additional 2 percentage points? Are the two variables individually significant?

18. [5 points] Name two possible sources of unobserved heterogeneity in Model 1. What would large differences between the estimates of Model 2 and Model 1 suggest about such heterogeneity? Explain.

19. [5 points] Write down the `gretl` command, which produces estimates of Model 2.

20. [5 points] Name a test that we can use to choose between Model 2 and Model 3. State the null and the alternative hypotheses of this test. Which model is more efficient under the null hypothesis? If the `gretl` output contains the results of this test, use it to choose between the two models. Justify your answer.

21. [5 points] Suppose you have a balanced panel of 10 countries and 5 electoral periods for each country. Suppose you want to estimate a fixed effects model by manually including country fixed effects and time (electoral periods) effects into the regression model. How many dummy variables would be included in the model next to the constant and the 4 main explanatory variables ($ddef$, gdp_gr, maj, and nd)? How many additional variables are included in a random effects model next to the constant and the aforementioned 4 main variables? Explain.