

Take Home Mid-Term Exam
Due Monday October 31st, beginning of class

Instructions:

- You may use any books or notes you like
- Show your work
- Be specific, concise and precise
- Be careful when you make calculations!
- Answer all the different parts of the questions
- Explain your answers
- For hypothesis tests: state the level of significance you are considering. Consider whether your answer differs depending on the significance level.
- Sign the below statement and turn it in together with your answers at the beginning of class, Monday October 31st. Late exams will not be accepted!

I worked on this exam by myself using my books and notes. I have not communicated with anybody (other than Prof. Macours) about any of the material of the econometrics course in the period between receiving the exam and handing it in. I have not shown the exam to anybody.

Print Name: _____

Sign Name: _____

A) Firms and Bribes in Transition Countries (9 points)

The European Bank for Reconstruction and Development (EBRD) and the World Bank conducted a Business Environment and Enterprise Performance Survey on a random sample of more than 4000 firms in Central and Eastern Europe in 1999. The purpose of the survey was to understand the conditions for doing business in transition countries. One of the questions in the survey reads “On average, what percent of revenues do firms like yours typically pay per annum in unofficial payments to public officials?”

Table: Determinants of informal payments (as share of revenue)

	(1)	(2)
age of the firm (years)	-0.028 (0.009)	-0.028 (0.009)
size firm (number of employees)	-0.005 (0.001)	
log(size firm)		-0.494 (0.086)
share of firm's capital owned by foreign company (%)	-0.005 (0.006)	-0.004 (0.006)
share of firm's capital owned by state organization (%)	0.001 (0.005)	0.003 (0.005)
Constant	4.728 (0.164)	5.980 (0.307)
Observations	1921	1915
R-squared	0.0270	0.0326
Adj. R-squared	0.0250	0.0306

Standard errors in parentheses

The above table presents regression results that attempt to shed light on the possible determinants of such informal payments.

- 1) Give a careful interpretation of the coefficient (sign, magnitude and significance) of age in the first column. (1.5 point)
- 2) Calculate the 99% confidence interval for the coefficient of firm's size in the first column. (1.5 point)
- 3) If you had to choose one of the 2 model specifications, which one would you pick? Explain. (1 point)
- 4) Are the regressions BLUE? Explain. (3 points)
- 5) Can you conclude from these results that young, small start-up firms are at a serious disadvantage in transition countries because they have to pay higher bribes? Explain. (1 point)
- 6) Specify a model that would allow to test whether the effect of foreign (co-) ownership depends on the size of the firm? (1 point)

B) Explaining conflict in Nepal (14 points)

A researcher is interested in understanding what determines civil war. He finds district-level data on the number of conflict-related incidents in Nepal in 2004 (*nrinc*). The civil conflict in Nepal started in 1996 when the Maoist party declared the “People’s War”. By the end of 2004, the conflict had spread to all districts of the country. In 2004, every district had at least one conflict-related incident. The researcher also found district level data from 1996 indicating the level of economic development: income per capita (*incpc*); % of adults that are literate (*adlit*); and life expectancy (*lifexp*) expressed in years. In addition he has variables indicating geography (i.e. the % of area with more than 30 degrees slope, *slope*), and natural resources, (i.e. % of area covered with forest, *forest*). He specifies the following model:

$$\log(nrinc) = \beta_0 + \beta_1 slope + \beta_2 slope^2 + \beta_3 forest + \beta_4 incpc + \beta_5 lif exp + \beta_6 adlit + u$$

and obtains the following stata output:

```
. sum nrinc slope forest incpc lifexp adlit
```

Variable	Obs	Mean	Std. Dev.	Min	Max
nrinc	74	10.33784	7.194513	1	35
slope	74	51.97081	25.57282	0	92.69
forest	74	.5378378	.5861887	.01	2.99
incpc	74	1.10223	.370501	.53	2.734
lifexp	74	55.6473	6.149134	36	66.5
adlit	74	34.43419	9.774366	15.13	60.37

```
. regress lnrrinc slope slopesq forest incpc lifexp adlit
```

Source	SS	df	MS			
Model	6.92065046	6	1.15344174	Number of obs =	74	
Residual	37.6812923	67	.562407348	F(6, 67) =	2.05	
				Prob > F =	0.0708	
				R-squared =	0.1552	
				Adj R-squared =	0.0795	
Total	44.6019428	73	.610985518	Root MSE =	.74994	

lnrrinc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
slope	.0235525	.0151442	1.56	0.125	-.0066755	.0537805
slopesq	-.0003024	.0001807	-1.67	0.099	-.0006631	.0000582
forest	-.09494	.2805439	-0.34	0.736	-.654908	.465028
incpc	-.2736008	.2877927	-0.95	0.345	-.8480375	.3008359
lifexp	-.0402146	.0188037	-2.14	0.036	-.077747	-.0026822
adlit	.0012046	.0107034	0.11	0.911	-.0201595	.0225687
_cons	4.41632	1.124552	3.93	0.000	2.171704	6.660935

- 1) Give a careful interpretation of the coefficient (sign, magnitude and significance) of *lifexp*. (1.5 point)
- 2) Give a careful interpretation of β_2 (1.5 point)

The researchers also estimates the following model:

```
. regress lnrrinc slope slopesq forest
```

Source	SS	df	MS			
Model	2.71245367	3	.904151224	Number of obs =	74	
Residual	41.8894891	70	.598421273	F(3, 70) =	1.51	
				Prob > F =	0.2193	
				R-squared =	0.0608	
				Adj R-squared =	0.0206	
Total	44.6019428	73	.610985518	Root MSE =	.77358	

lnrrinc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
slope	.0310612	.0152465	2.04	0.045	.000653	.0614695
slopesq	-.000376	.0001826	-2.06	0.043	-.0007402	-.0000119
forest	.2184414	.2616745	0.83	0.407	-.3034519	.7403347
_cons	1.605886	.3106698	5.17	0.000	.9862749	2.225498

- 3) Test whether the variables capturing the level of development are jointly significant. Write out the null-hypothesis and the alternative hypothesis, calculate the test and derive your conclusions. (3 points)

- 4) **True or False: based on the above estimates, we can accept the hypothesis that districts with higher life expectancy have a lower incidence of conflict in Nepal. T/F. Explain. (1 point)**
- 5) **Do you think it was a good idea to use the log of the dependent variable in this case? Explain. (1 point)**
- 6) **In the first model, the researcher was surprised to not to find a significant coefficient for the variable measuring income per capita. Imagine the researcher would estimate the first regression without including the life expectancy variable.**
 - a. **Do you expect the coefficient of *incpc* to become more negative, less negative or positive? Explain why. (2 points)**
 - b. **Do you expect the precision of the estimate of *incpc* to be higher or lower than in the regression shown above? Explain why. (1 points)**
- 7) **The researcher does some additional efforts and also finds data on district-level inequality in 2004. Do you think it would be a good idea to incorporate this information as an additional variable in the model? Explain. (1 point)**
- 8) **At what level of significance can you reject the hypothesis that geography and forest cover together don't have an effect on conflict-related incidents. (1 point)**
- 9) **Based on the regression results presented here, do you think the researcher should conclude that countries' natural resource endowments do not affect the likelihood of civil war? Explain. (1 point)**

C) Your own hypothesis (7 points)

- 1) **Formulate an interesting hypothesis that you would like to test. You can pick any type of hypothesis related to any subject you like. However, it must be your own personal and original hypothesis. Hence it cannot be a hypothesis related to a model we've discussed in class, nor a hypothesis you found in this midterm, in the book, the problem sets, or in material of last years' econometrics courses. (1 point)**
- 2) **Specify a model that would allow you to test your hypothesis. Make sure to specify a model that would give you unbiased estimates and motivate shortly why you think it would be unbiased. (3 points)**
- 3) **Specify the null-hypothesis and the alternative hypothesis. (1 point)**
- 4) **Indicate which test you would use to test your hypothesis and describe briefly how you would conduct this test. (2 points)**

D) Double check

Go back to the instructions at the beginning of the exam and double check that you've followed them for the answers in part A, B and C. Did you answer all questions, and all parts of each question (6 questions in part A, 9 questions in part B, 4 questions in part C)?