

## YLEISEN TENTIN TENTTILOMAKE - GENERAL EXAM FORM

Opiskelija täyttää / Student fills in

<b>Opiskelijan nimi / Student name:</b> Click here to enter text.	<b>Opiskelijanumero / Student number:</b> Click here to enter text.
--	--

Opettaja täyttää / Lecturer fills in

<b>Opintojakson koodi / The code of the course:</b> 721066S	
<b>Opintojakson (tentin) nimi / The name of the course or exam:</b> Principles of Econometrics	
<b>Opintopistemäärä / Credit units:</b> 6  Mikäli kyseessä on välikoe, opintopistemääräksi täytetään 0 op. 0 ECTS Credits is used for mid-term exams.	
<b>Tiedekunta / Faculty:</b> OyKKK / OBS	
<b>Tentin pvm / Date of exam:</b> 30.1.2019	<b>Tentin kesto tunteina / Exam in hours:</b> 3 h
<b>Tentaattori(t) / Examiner(s):</b> Sanna Huikari	<b>Sisäinen postiosoite / Internal address:</b> 6 OyKKK
<b>Tentissä sallitut apuvälineet / The devices allowed in the exam:</b> <input checked="" type="checkbox"/> Funktiolaskin / Scientific calculator <input checked="" type="checkbox"/> Ohjelmoitava laskin / Programmable calculator <input checked="" type="checkbox"/> Muu tentissä sallittu materiaali tai apuvälineet. Tarkenna alla. / Other material or devices, allowed in the exam. Specify below. <b>Two-sided hand-written A4-sheet, which must be included into the answer sheet</b> <input type="checkbox"/> Tentissä ei ole sallittua käyttää apuvälineitä / The devices are not allowed in the exam	
<b>Muut tenttiä koskevat ohjeet opiskelijalle (esimerkiksi kuinka moneen kysymyksen opiskelijan tulee vastata) / Other instructions for students e.g. how many questions he/she should answer:</b> Click here to enter text.	

1. (25 p.) What are the root causes of terrorism? Poverty? Repressive political regimes? Religious or ethnic conflicts arising from heterogeneous populations? In this problem you will take a look at some empirical evidence on cross-country sources of terrorism. Number of countries included in analysis is 76 and the variables that are used are defined as follows:

Variable	Definiton
<i>lnftmpop</i>	Natural logarithm of the number of fatalities from terrorist incidents in the country, 1998-2004, per million population
<i>lngdppc</i>	Natural logarithm of the GDP per capita in the country
<i>higdppc</i>	Binary variable, which equals one if GDP per capita in the country is greater than or equal to the median in the data set, and which equals zero otherwise
<i>lackpf</i>	Index of the lack of political freedoms, 1-7 scale, 1 = extremely liberal political freedoms,....., 7 = extremely limited political freedoms
<i>ethnic</i>	Index of ethnic fractionalization in the country (0 to 1 scale, 0 = one ethnic group, 1 = more than one ethnic group)
<i>religion</i>	Index of religious fractionalization in the country (0 to 1 scale, 0 = one religion, 1 = more than one religion)
<i>mideast, latinam, east-europe, africa, eastasia</i>	= 1 if the country is in the indicated region, = 0 otherwise

Table 1 contains results from six estimated regressions. The dependent variable in all regressions is natural logarithm of the number of fatalities from terrorist incidents in the country (1998-2004) per million population. Use the results to answer the following questions. **Remember to show your work. No credit for answers if you do not show your work.**

- Using regression (1), test the hypothesis that the coefficient on *lngdppc* is zero, against the alternative that it is nonzero, at the 5% significance level. Explain in words what the result of the test and the value of the estimated coefficient mean. (3 p)
- Compare the results in regressions (1) and (2) regarding the effect of GDP on the number of fatalities from terrorism. Explain why the conclusions differ. (3 p)
- Using regression (3), is there evidence that the relationship between GDP and the number of fatalities from terrorism is nonlinear? Using regression (3), is there evidence that the relationship between lack of political freedom and the number of fatalities from terrorism is nonlinear? (3 p)
- Using regression (5), test the null hypothesis (at the 5% significance level) that the coefficients on the "regional dummies" all are zero, against the alternative hypothesis that at least one is nonzero. Explain in words what does the result mean. (3 p)
- Using regression (4), test the hypothesis that the population coefficients on *ethnic* and *religion* are both zero, against the alternative that at least one coefficient is nonzero. Explain in words what hypothesis you have tested, and what your conclusion is. (3 p)
- Using regression (4), estimate the effect on *lnftmpop* of changing from *lackpf* = 7 (extremely limited political freedoms) to *lackpf* = 5 (some political freedoms), holding constant the values of the other regressors in regression (4). Explain in words what this means. (3 p)
- According to the results of regression (6), is there a difference in how the number of fatalities from terrorism is affected by the lack of political freedom between countries with higher than median GDP and countries with lower than median GDP? (3 p)
- Interpret the signs of the slope coefficients *lackpf* and *higdppc*  $\times$  *lackpf* in regression (6). (4 p)

**Table 1. Determinants of Terrorism**

	(1)	(2)	(3)	(4)	(5)	(6)
<b>Dependent variable:</b>	<i>lnftmpop</i>	<i>lnftmpop</i>	<i>lnftmpop</i>	<i>lnftmpop</i>	<i>lnftmpop</i>	<i>lnftmpop</i>
<b>Regressors:</b>						
<i>lngdppc</i>	-0.312 (0.150)	-0.279 (0.204)	0.057 (0.215)	0.047 (0.245)	-0.065 (0.283)	—
<i>lngdppc</i> <sup>2</sup>	—	—	-0.034 (0.081)	—	—	—
<i>higdppc</i>	—	—	—	—	—	-2.84 (1.00)
<i>lackpf</i>	—	0.036 (0.143)	1.65 (0.539)	1.466 (0.546)	1.691 (0.597)	-0.248 (0.165)
<i>lackpf</i> <sup>2</sup>	—	—	-0.194 (0.062)	-0.170 (0.062)	-0.199 (0.067)	—
<i>higdppc</i> × <i>lackpf</i>	—	—	—	—	—	0.645 (0.238)
<i>ethnic</i>	—	—	—	0.656 (0.958)	0.878 (0.972)	—
<i>religion</i>	—	—	—	-1.243 (0.970)	-1.573 (1.225)	—
<i>Regional dummies (Mideast, latinam, easteurope, africa, eastasia)?</i>	No	No	No	No	Yes	No
<i>Intercept</i>	-2.05 (0.213)	-2.20 (0.623)	-4.81 (1.08)	-4.43 (1.29)	-3.58 (1.60)	-0.72 (0.84)
<b>F-statistics testing the hypothesis that the population coefficients on the indicated regressors are all zero:</b>						
<i>lngdppc, lngdppc</i> <sup>2</sup>	—	—	0.11 (0.900)	—	—	—
<i>lackpf, lackpf</i> <sup>2</sup>	—	—	4.95 (0.010)	3.81 (0.027)	4.47 (0.015)	—
<i>ethnic, religion</i>	—	—	—	1.01 (0.371)	1.25 (0.294)	—
<i>Regional dummies</i>	—	—	—	—	0.81 (0.524)	—
<b>Regression summary statistics</b>						
$\bar{R}^2$	0.049	0.037	0.123	0.119	0.087	0.085
<i>SER</i>	1.84	1.86	1.77	1.767	1.799	1.808

Notes: Heteroskedasticity-robust standard errors are given in parentheses under estimated coefficients, and *p*-values are given in parentheses under *F*-statistics. The “regional dummies” included in regression (5) are *Mideast*, *latinam*, *easteurope*, *africa*, and *eastasia* (the omitted case is Western Europe combined with North America).

2. (6 p.) Consider the following two possible sources of bias in a regression study:

- Omitted variable bias;
- Functional form misspecification.

i) Give briefly a general definition and the consequences of these biases. (3 p)

ii) In your judgment are these two sources of bias plausibly important in the application presented in question 1? Why or why not? Be brief (a few sentences for each) but specific. (3 p)

3. (9 p.) Internal and external validity. Question 3 consists of 6 multiple choice questions. Each question has only one correct (or clearly best) answer. You get +1.5 points for a right answer, zero points for no answer, and -0.75 points for a wrong answer. Maximum points from this question are therefore 9 points, minimum is 0. Write the letters that indicate your answers in answering sheet.

1) By including another variable in the regression, you will

- A) decrease the regression R<sup>2</sup> if that variable is important.
- B) eliminate the possibility of omitted variable bias from excluding that variable.
- C) look at the t-statistic of the coefficient of that variable and include the variable only if the coefficient is statistically significant at the 1% level.
- D) decrease the variance of the estimator of the coefficients of interest.

2) Errors-in-variables bias

- A) is present when the probability limit of the OLS estimator is given by  $\hat{\beta}_1 \xrightarrow{p} \beta_1 + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_w^2}$
- B) arises when an independent variable is measured imprecisely.
- C) arises when the dependent variable is measured imprecisely.
- D) always occurs in economics since economic data is never precisely measured.

3) A statistical analysis is internally valid if

- A) its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- B) statistical inference is conducted inside the sample period.
- C) the hypothesized parameter value is inside the confidence interval.
- D) the statistical inferences about causal effects are valid for the population being studied.

4) Misspecification of functional form of the regression function

- A) is overcome by adding the squares of all explanatory variables.
- B) is more serious in the case of homoskedasticity-only standard error.
- C) results in a type of omitted variable bias.
- D) requires alternative estimation methods such as maximum likelihood.

5) Sample selection bias occurs when

- A) the choice between two samples is made by the researcher.
- B) data are collected from a population by simple random sampling.
- C) samples are chosen to be small rather than large.
- D) the availability of the data is influenced by a selection process that is related to the value of the dependent variable.

6) The true causal effect might not be the same in the population studied and the population of interest because

- A) of differences in characteristics of the population
- B) of geographical differences
- C) the study is out of date
- D) all of the above

**Large-Sample Critical Values for the  $t$ -statistic from the Standard Normal Distribution**

	Significance Level		
	10%	5%	1%
<b>2-Sided Test (<math>\neq</math>)</b>			
Reject if $ t $ is greater than	1.64	1.96	2.58
<b>1-Sided Test (<math>&gt;</math>)</b>			
Reject if $t$ is greater than	1.28	1.64	2.33
<b>1-Sided Test (<math>&lt;</math>)</b>			
Reject if $t$ is less than	-1.28	-1.64	-2.33

**TABLE 3 Critical Values for the  $\chi^2$  Distribution**

Degrees of Freedom	Significance Level		
	10%	5%	1%
1	2.71	3.84	6.63
2	4.61	5.99	9.21
3	6.25	7.81	11.34
4	7.78	9.49	13.28
5	9.24	11.07	15.09
6	10.64	12.59	16.81
7	12.02	14.07	18.48

**Large-Sample Critical Values for the  $F$ -statistic from the  $F_{m,\infty}$  Distribution**

Reject if $F >$ Critical Value			
Degrees of Freedom ( $m$ )	Significance Level		
	10%	5%	1%
1	2.71	3.84	6.63
2	2.30	3.00	4.61
3	2.08	2.60	3.78
4	1.94	2.37	3.32
5	1.85	2.21	3.02
6	1.77	2.10	2.80
7	1.72	2.01	2.64

