

Introduction to Statistics

homework 4

Johan A. Elkink
jos.elkink@ucd.ie

Due 12 December 2018, 5 pm

Note revised submission date—still 5 pm deadline.

Submission instructions

You will submit one **PDF** file^a including all interpretations. Please ensure that the **file name** contains your **UCD** student number and the module code (POL40950), in the following format: Homework4_POL40950_18123456.pdf (please note absence of spaces and be precise in your naming!).

Your name should not appear in either the file name or the file contents.

Please send both files to jos.elkink@ucd.ie

^aWord files will be sent back—note that newer versions of Word can easily save to PDF format.

Data

The data set we will use for this homework is based on the European Social Survey, using a random subset of Irish respondents in the 2010–11 wave of the survey. You can find the data on the teaching data page¹, including a description of each variable. While we use a subset of the data, with fewer variables and all missing data removed, you will be able to find the original data on the ESS website.² That said, while the description of the data on the website will be important, there is no requirement to access data for this homework.

Table 1: Explaining trust in politicians by interest in politics using linear regression

	Trust in politicians		
	1	2	3
Political interest	-.495*** (.107)	-.527*** (.113)	-.268** (.099)
Education		-.015 (.031)	-.039 (.027)
Age		-.005 (.006)	-.007 (.005)
Female		.120 (.201)	.174 (.172)
Satisfaction (economy)			.142** (.050)
Satisfaction (government)			.474*** (.046)
Constant	4.604*** (.320)	5.077*** (.732)	2.547*** (.655)
Observations	476	473	464
R ²	.043	.045	.311
Adjusted R ²	.041	.037	.301
Residual Std. Error	2.157 (df = 474)	2.167 (df = 468)	1.829 (df = 457)
F Statistic	21.388*** (df = 1; 474)	5.559*** (df = 4; 468)	34.300*** (df = 6; 457)
<i>Notes:</i>	*P < .05 **P < .01 ***P < .001		

Table 2: Explaining interest in politics by gender using linear regression

	Political interest		
	1	2	3
Female	.165* (.083)	.208** (.080)	.210** (.079)
Education		-.070*** (.012)	-.062*** (.012)
Age		-.013*** (.002)	-.011*** (.002)
Fine Gael			-.246 (.141)
Labour			-.317 (.184)
Other			.090 (.138)
Sinn Fein			-.162 (.221)
Constant	2.793*** (.060)	4.306*** (.219)	4.145*** (.258)
Observations	499	494	494
R ²	.008	.105	.134
Adjusted R ²	.006	.099	.121
Residual Std. Error	.928 (df = 497)	.884 (df = 490)	.873 (df = 486)
F Statistic	3.930* (df = 1; 497)	19.096*** (df = 3; 490)	10.698*** (df = 7; 486)

Notes:

*P < .05

**P < .01

***P < .001

Table 3: Explaining attitude towards immigrants by satisfaction with the economy and union membership using linear regression

	Attitude towards immigrants	
	1	2
Satisfaction (economy), S	.214*** (.052)	.272*** (.056)
Union member, U		1.102* (.504)
$S \times U$		-.425** (.152)
Constant	4.532*** (.172)	4.380*** (.184)
Observations	491	491
R ²	.033	.049
Adjusted R ²	.031	.043
Residual Std. Error	2.266 (df = 489)	2.253 (df = 487)
F Statistic	16.866*** (df = 1; 489)	8.297*** (df = 3; 487)

Notes:

*P < .05

**P < .01

***P < .001

Questions

1. Table 1 provides a set of multiple regressions explaining the level of trust in politicians by the level of interest in politics, including a set of control variables.³
 - (a) (10%) Why would education be an appropriate control variable?
 - (b) (20*%)⁴ Write a 300–400 word report in which you fully interpret the regression output (all three models)—what does it tell you about the relationship between interest in politics and trust in politicians?
Make sure you include at least:
 - Interpretation of regression coefficients and R^2 .
 - Interpretation of the size of the main effect.
 - Interpretation of the t -tests and F -test.
 - What control variables might be important to add?
2. Table 2 provides a set of multiple regressions explaining the level of political interest by gender⁵, with a set of control variables.
 - (a) (10%) How do we read the $-.246$ coefficient on the Fine Gael variable?
 - (b) (30*%) Write a 300–400 word report in which you fully interpret the regression output (all three models)—what does it tell you about gender and political interest?
Make sure you include at least:
 - Interpretation of regression coefficients and R^2 .
 - Interpretation of the size of the main effect.
 - Interpretation of the t -tests and F -test.
3. Table 3 provides a set of multiple regressions explaining the attitude towards immigrants (where immigrants are considered to make Ireland 0 ‘a worse place to live’ to 10 ‘a better place to live’) by union membership and satisfaction with the economy.
 - (a) (30*%) Write a 300–400 word report in which you fully interpret the regression output (both models)—what does it tell you about satisfaction with the economy and attitudes towards immigrants?
Make sure you include at least:
 - Interpretation of regression coefficients and R^2 .

¹<http://www.joselkink.net/data.php>

²<https://www.europeansocialsurvey.org/data/>

³Note that arguably the political interest variable should be treated as ordinal, not a scale variable, but you can proceed with this homework interpreting it as a scale variable.

⁴Where percentages are marked with an asterisk, positive marking will be applied—in all other cases negative marking.

⁵Note that the ESS survey takes gender as a binary variable, although there are important arguments to allow for a wider range of gender identities, not forcing respondents in this binary classification. The survey question itself refers to sex and is coded by the interviewer, possibly without asking the respondent.

- Interpretation of the size of the main effect.
- Interpretation of the interaction effect.
- Interpretation of the t -tests and F -test.

References

Grade conversion scheme

Score	Grade		Score	Grade		Score	Grade		Score	Grade	
	UCD	TCD		UCD	TCD		UCD	TCD		UCD	TCD
97-100%	A+	A+	85-87%	B	B	74-76%	C-	C	54-64%	E+	D
94-96%	A	A	83-84%	B-	B	71-73%	D+	C	44-53%	E	D
91-93%	A-	A	80-82%	C+	C+	68-70%	D	C	33-43%	E-	D
88-90%	B+	B+	77-79%	C	C	65-67%	D-	C	0-32%	F	F

Note that all results are entered in grades, not percentages, and the average score at the end of the module is therefore also based on the grades, not the percentages.