_____

Empirical Exercise 10.2

Calculations for this exercise are carried out in the STATA file **EE_10_2.do.**

The solutions will reference the following regression results.

	(1)	(2)	(3)	(4)	(5)
log_GDPPC	0.236	0.235	0.083	0.054	0.025
	(0.012)	(0.012)	(0.031)	(0.042)	(0.054)
	[0.212, 0.259]	[0.211, 0.259]	[0.021, 0.146]	[-0.030, 0.137]	[-0.057, 0.120]
Controls	No	No	No	No	Yes
State Effects	No	No	Yes	Yes	Yes
Time Effects	No	Yes	No	Yes	Yes
F-Statistics and p-values testing exclusion of groups of variables					
Time Effects		9.31		5.73	4.61
		(0.000)		(0.00)	(0.000)
Age Variables					2.12
					(0.08)
Age, educ., & pop. variables					1.44
					(0.21)

Controls: Age, education and population variables

a. The dataset is unbalanced because data are available over different years for different countries. As an example, data on *Dem_ind* for Andorra are available for 1970, 1995, and 2000, but for Afghanistan data are available for 1960, 1965, … , 2000.

b. (i) Values of *Dem_ind* range from 0.0 to 1.0. The mean in 0.50 and the standard deviation is 0.37. The 10th, 25th, 50th, 75th, and 90th percentiles are 0.0, 0.17, 0.5, 0.83, and 1.0.

(ii) The value for the U.S. in 2000 is *Dem_ind* = 1.0. The average for the nine years in the sample is 0.99.

(iii) The value for Libya in 2000 is *Dem_ind* = 0.0. The average for the nine years in the sample is 0.11.

(iv) This involves looking through the data set. Do it. It's fun and interesting!

c. (i) The coefficient is 0.24 with a standard error of 0.01. The 95% confidence interval is 0.21 to 0.26. The coefficient is large, as described below.

(ii) A 20% increase in GDP per capita implies that *log_gdp* increases by approximately 0.20, so that *Dem_ind* is predicted to increase by approximately

_____

$0.20 \times 0.24 = 0.048$, or about 1/10 of the standard deviation in the dataset. The 95% confidence for the effect is (approximately) 0.20×0.21 to 0.20×0.26 or 0.042 to 0.052.

(iii) Clustered standard errors are needed because of country-specific omitted factors in the regressions. The clustered standard error for *log_gdp* is 0.012; the unclustered standard error is smaller (0.007) because it ignores the positive within-country correlation of the errors.

d. (i) Countries have different histories, institutions, social structures, and religions. All affect their preference for demography and may also be correlated with economic development, and therefore per-capital income.

(ii) The estimated coefficient falls by a factor of 3, to 0.083 with a standard error of 0.032. The estimated effect, while significantly smaller is still statistically significant at the 1% significance level. The estimated effect in (c.ii) also falls by a factor of 3.

(iii) Data for Azerbaijan is available only for 2000. These data are completely absorbed the country-specific fixed effect and therefore have no effect on the estimated coefficient on *log_gdppc*.

(iv) The demand for democracy is contagious and sweeps across countries (remember the "Arab Spring" of 2012). To the extent that these events coincide with global changes in economic conditions they are correlated with *log_gdppc* and will therefore lead to omitted variable bias.

(v) The estimated coefficient falls further to 0.05, approximately 1/5 of the value that omits time and country fixed effects. The estimate is not statistically significant at the 10% level.

(vi) Regression (5) in the table includes population, age, and education. Jointly, these variables are not statistically significant in the regression, although the age variables are significant at the 10% level. When these variables are included, the estimated coefficient on *log_gdppc* falls further to 0.03 with a standard error of 0.05.

e. After controlling for omitted variables – particularly country-fixed effects – there is little evidence of an income effect on the demand for democracy.