

Advanced Quantitative Methods

Homework 3

Johan A. Elkink
jos.elkink@ucd.ie

Due March 28, 2014, 5pm

Please submit by email in PDF format. Add R code in a separate .R file, or SPSS code in a separate .sps file, or Stata code in a separate .do file, or the code for any other package you use separately. Note that if you do not use Latex, there is an "Equation" entry in Microsoft Word under "Insert" that will allow you to include matrices and other mathematical equations.

(5%) of the grade is used for an overall evaluation of the clarity and presentation of your code.

Data

For all questions, we will use the demdev.dta data set, using only data from 1988:

```
dd <- read.dta("demdev.dta")  
dd <- dd[dd$year == 1988, ]
```

You will also need the demdev-W-1988.dta data set.

Questions

1. In this question, we will look at the impact of democracy on energy consumption, the latter being a proxy for economic development.

- (a) Run a simple regression using model specification

$$energy2_i = \beta_0 + \beta_1 democracy_i.$$

- (b) Add further controls, estimating the model

$$energy2_i = \beta_0 + \beta_1 democracy_i + \beta_2 ipyears_i + \beta_3 catho80_i.$$

- (c) We can execute a full matching procedure as follows:¹

¹matchit() cannot properly deal with missing data, so we use the filtered data output from the linear regression.

```

library(MatchIt)

m1 <- lm(energy2 ~ democracy + ipyears + catho80, data = dd)

msample <- matchit(democracy ~ ipyears + catho80,
                  data = m1$model, method = "full")

summary(m2 <- lm(energy2 ~ democracy, data = match.data(msample)))

```

Estimate this model.

- (d) (5%) Estimate the model using nearest neighbour matching instead (check the help file of `matchit` to see how).
 - (e) (5%) Produce a regression table of publishable standard including all four regressions.
 - (f) (10%) Write a short report (at least 300 words) interpreting these results, discussing the impact of democracy on development. Discuss explicitly the use of the matching estimator and differences with OLS.
2. In this question we will attempt to estimate the impact of economic development (proxied by energy consumption) on the level of democracy:

$$polity2_i = \beta_0 + \beta_1 energy2_i + \beta_2 ioscore_i + \beta_3 catho80_i.$$

- (a) Estimate the baseline model using OLS.
 - (b) (10%) Using two stage least squares (e.g. `ivreg` or `tsls` in R), estimate a model using precipitation as instrument for energy2.
 - (c) (5%) Run a Hausman test comparing the two models. What do you conclude?
 - (d) (10%) Write a short report (at least 400 words) interpreting these results, discussing what does it tells you about the relationship between development and democracy. Discuss explicitly the use of the IV estimator and differences with OLS.
3. In this question, we will work with the same baseline regression model:

$$polity2_i = \beta_0 + \beta_1 energy2_i + \beta_2 ioscore_i + \beta_3 catho80_i.$$

We will extend this model by adding a spatially lagged dependent variable:

$$polity2_i = \beta_0 + \rho W \cdot polity2 + \beta_1 energy2_i + \beta_2 ioscore_i + \beta_3 catho80_i,$$

whereby W is an $n \times n$ matrix identifying neighbouring states and ρ a coefficient to estimate the level of spatial autocorrelation. To estimate a model of this kind, we need the following loglikelihood function (Ward and Gleditsch, 2008, 35):

$$\ell = \ln|I - \rho W| - \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(2\sigma^2) - \frac{e'e}{2\sigma^2},$$

where

$$e = (y - \rho W y - X\beta).$$

We will use the following approximation:

$$\ln|I - \rho W| = \prod_i (1 - \rho \omega_i),$$

whereby ω_i is the i th eigenvalue of W .²

- (a) (8%) Finish the following loglikelihood function in R for the spatial autoregressive model described above:

```
ll <- function(theta, y, X, W) {  
  
  s2 <- exp(theta[1])  
  rho <- 1 / (1 + exp(-theta[2]))  
  beta <- theta[-c(1,2)]  
  
  n <- length(y)  
  I <- diag(n)  
  
}
```

- (b) (8%) Using

```
X <- model.matrix(polity2 ~ energy2 + ioscore + catho80, data = dd)  
y <- dd$polity2  
W <- as.matrix(read.dta("demdev-W-1988.dta"))
```

to generate the data matrices, estimate the spatial regression model using `optim()`.

- (c) (8%) Using the output from `optim()`, calculate the standard errors for the β -coefficients.
(d) (5%) Perform t -tests for the β -coefficients, reporting the p -values.
(e) (5%) Calculate the estimated values of ρ and σ^2 .
(f) (6%) Produce a regression table of publishable standard including all three regressions from Questions 2 and 3.
(g) (10%) Write a short report (at least 300 words) interpreting these results, discussing what does it tells you about the relationship between development and democracy and the level of spatial autocorrelation in democracy.

²In R this will be `prod(1 - rho * eigen(W)$value)`.

Grade conversion scheme

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

Ward, Michael D and Kristian Skrede Gleditsch. 2008. *Spatial regression models*. Vol. 155 Sage.