Advanced Quantitative Methods
Lab 6: Causal inference

Johan A. Elkink
jos.elkink@ucd.ie

6 April 2018

1. Open the `uswages.dta` data set. Check the teaching data page for more information on the variables.
```
library(rio)
us <- import("http://www.joselkink.net/wp-content/uploads/2013/01/uswages.dta")
```
   (a) Interpret all test results in the standard output.
   (b) Perform a test evaluating whether education and experience jointly contribute.
   (c) Perform a Chow test to see if the regression is different for urbanised versus rural respondents (*smsa*).

2. Open the `demdev.dta` data set and select cases from 1988 only.
```
library(rio)
dd <- subset(import("demdev.dta"), year == 1988)
```
   (a) Run a regression using model specification

$$energy2_i = \beta_0 + \beta_1 democracy_i + \beta_2 ipyears_i + \beta_3 catho80_i. \quad (1)$$

   (b) We can execute a full matching procedure as follows:[1]
```
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.
   (c) Estimate the model using nearest neighbour matching instead (check the help file of `matchit` to see how).

3. Continuing with the model in Eq. (1):
   (a) Using two stage least squares (e.g. `ivreg` or `tsls` in R), estimate a model using `precipitation` as instrument for `energy2`.
   (b) Run a Hausman test comparing the two models. What do you conclude?
   (c) Open `crime.dta`, which contains data on number of police and crime intensity, for elections years and non-election years. Regress crime on police.
   (d) Why would election year be a suitable instrument for policy? Estimate 2SLS using election year as instrument.
   (e) Perform a Hausman test for exogeneity of police.

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