Advanced Quantitative Methods
Lab 6: Causal inference

Johan A. Elkink
jos.elkink@ucd.ie

2 March 2017

1. Open the `demdev.dta` data set and select cases from 1988 only.
```
library(rio)
dd <- subset(import("demdev.dta"), year == 1988)
```

2. Run a simple regression using model specification

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

3. 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.

4. Estimate the model using nearest neighbour matching instead (check the help file of `matchit` to see how).

5. Run an OLS regression estimating:

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

6. Using two stage least squares (e.g. `ivreg` or `tsls` in R), estimate a model using `precipitation` as instrument for `energy2`.

7. Run a Hausman test comparing the two models. What do you conclude?

8. Open `crime.dta`, which contains data on number of police and crime intensity, for elections years and non-election years. Regress crime on police.

9. Why would election year be a suitable instrument for policy? Estimate 2SLS using election year as instrument.

10. 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.