Assignment 4 - EDS 241: Environmental Policy Evaluation

Juliet Cohen

03/10/2022

Read in data:

```
data <- read.csv("/Users/juliet/Documents/MEDS/EDS_241_Env_Policy_Eval/Assignments/Assignment_4/EDS241_A
```

This question will ask you to estimate the price elasticity of demand for fresh sardines across **56 ports** located in **4 European countries** with **monthly** data from **2013 to 2019**. The data are contained in the file EU_sardines.csv, which is available on Gauchospace. Each row in the data file is a combination of port location (where the fish is landed and sold) in a given year and month. You can ignore the fact that the sample is not balanced (the number of monthly observations varies across ports).

For the assignment, you will need the following variables: **year, month, country, port (port where sardines are landed and sold), price_euro_kg (price per kg in €), and volume_sold_kg (quantity of sardines sold in kg).** In the questions below, I use `log()` to denote the natural logarithm.

1 (a) Estimate a bivariate regression of log(volume_sold_kg) on log(price euro_kg). What is the price elasticity of demand for sardines? Test the null hypothesis that the price elasticity is equal to -1.

```
data_log <- data %>%
  mutate(log_volume_sold_kg = log(volume_sold_kg),
         log_price_euro_kg = log(price_euro_kg))

model_price_vol <- lm(log_volume_sold_kg ~ log_price_euro_kg, data = data_log)
summary(model_price_vol)
```

```
##
## Call:
## lm(formula = log_volume_sold_kg ~ log_price_euro_kg, data = data_log)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8302 -1.9304 -0.2278  2.0553  6.8837
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         7.75906    0.04299  180.50   <2e-16 ***
## log_price_euro_kg -1.54534    0.07168  -21.56   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 2.714 on 3986 degrees of freedom
## Multiple R-squared:  0.1044, Adjusted R-squared:  0.1042
## F-statistic: 464.8 on 1 and 3986 DF,  p-value: < 2.2e-16
```

```r
model_price_vol$coefficients[[2]]
```

```
## [1] -1.545335
```

```r
# if price increases 1%, the volume sold decreases by 1.545%

#model_price_vol_table <- tidy(model_price_vol)
# model_price_vol_table %>%
#   select(term, estimate, std.error, p.value, conf.low, conf.high) %>%
#   kable()

# F-test for non-weak and relevant instruments (Lecture 9, slides 13-14)
linearHypothesis(model_price_vol, c("log_price_euro_kg=-1"), white.adjust = "hc2")
```

```
## Linear hypothesis test
##
## Hypothesis:
## log_price_euro_kg = - 1
##
## Model 1: restricted model
## Model 2: log_volume_sold_kg ~ log_price_euro_kg
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1   3987
## 2   3986  1 48.724 3.436e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# the f-stat is greater than 10
```

- The price elasticity of demand for sardines is -1.5453352, meaning that if the price of sardines increases by 1%, the volume sold decreases by -1.5453352 percentage points. We can reject the null hypothesis that the price elasticity of demand is equal to zero because the p-value is shown as significant at the 0.001 level. Our `linearHypothesis()` output shows that the F-statistic is 48.724, which greater than 10.

2 (b) Like in Lecture 8 (see the IV.R script), we will use wind_m_s as an instrument for log(price_euro_kg). To begin, estimate the first-stage regression relating log(price_euro_kg) to wind_m_s. Interpret the estimated coefficient on wind speed. Does it have the expected sign? Also test for the relevance of the instrument and whether it is a "weak" instrument by reporting the proper F-statistic.

```r
# first stage regression
model_price_wind <- lm(log_price_euro_kg ~ wind_m_s, data = data_log)
```

```
summary(model_price_wind)
```

```
##
## Call:
## lm(formula = log_price_euro_kg ~ wind_m_s, data = data_log)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.4705 -0.3871 -0.0095  0.4060  3.7839
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.304888   0.026660  -11.44   <2e-16 ***
## wind_m_s     0.067346   0.005374   12.53   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5884 on 3986 degrees of freedom
## Multiple R-squared:  0.0379, Adjusted R-squared:  0.03766
## F-statistic:   157 on 1 and 3986 DF,  p-value: < 2.2e-16
```

```
# generate F-statistic
linearHypothesis(model_price_wind, c("wind_m_s=0"), white.adjust = "hc2")
```

```
## Linear hypothesis test
##
## Hypothesis:
## wind_m_s = 0
##
## Model 1: restricted model
## Model 2: log_price_euro_kg ~ wind_m_s
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1   3987
## 2   3986  1 144.65 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The estimated coefficient for wind speed is 0.0673459, which represents the percent increase in price per kg in € for every 1 percent increase in wind speed. Yes, this coefficient has the expected sign, because I presume that more windy conditions make fishing more difficult, which would result in the cost of sardines to increase. The F-statistic is 144.65, which is greater than 10, indicating that this is a non-weak instrument.

3 (c) Estimate the TSLS estimator of the price elasticity of demand for sardines using wind_m_s as an instrument for log(price_euro_kg). What is the estimated price elasticity of demand for sardines?

```r
tsls_vol_price_wind <- ivreg(log_volume_sold_kg ~ log_price_euro_kg | wind_m_s, data = data_log)
summary(tsls_vol_price_wind)
```

```
##
## Call:
## ivreg(formula = log_volume_sold_kg ~ log_price_euro_kg | wind_m_s,
##     data = data_log)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8626 -1.9790 -0.2333  2.0950  6.2354
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         7.75534    0.04331  179.08   <2e-16 ***
## log_price_euro_kg -1.08802    0.37003   -2.94   0.0033 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.728 on 3986 degrees of freedom
## Multiple R-Squared: 0.09529, Adjusted R-squared: 0.09506
## Wald test: 8.646 on 1 and 3986 DF,  p-value: 0.003297
```

- The estimated price elasticity of demand for sardines is -1.0880152, which represents the percent decrease in volume of sardines sold for every 1% increase in price, using wind as an instrument.

4 (d) Repeat the exercise in (c), but include fixed effects for each year, month, and country. [Hint: you can use the command "as.factor(country) + as.factor(year) + as.factor(month)" to the ivreg function in R]. Report the estimated price elasticity of demand and the F-statistic testing for relevant and non-weak instruments.

```r
fe_model <- lm(log_price_euro_kg ~ as.factor(country) + as.factor(year) + as.factor(month) + wind_m_s, 
summary(fe_model)
```

```
##
## Call:
## lm(formula = log_price_euro_kg ~ as.factor(country) + as.factor(year) +
##     as.factor(month) + wind_m_s, data = data_log)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6014 -0.3748 -0.0034  0.3912  3.7650
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    -0.337243   0.066518  -5.070 4.16e-07 ***
## as.factor(country)Italy        -0.041400   0.025745  -1.608 0.107899
## as.factor(country)Portugal      0.106499   0.077530   1.374 0.169630
```

```
## as.factor(country)United Kingdom -0.101376   0.069142  -1.466 0.142673
## as.factor(year)2014              -0.020595   0.033495  -0.615 0.538676
## as.factor(year)2015               0.012680   0.033432   0.379 0.704514
## as.factor(year)2016              -0.040329   0.033471  -1.205 0.228318
## as.factor(year)2017               0.003731   0.033471   0.111 0.911247
## as.factor(year)2018              -0.052021   0.034156  -1.523 0.127829
## as.factor(year)2019              -0.060297   0.042669  -1.413 0.157694
## as.factor(month)2                 0.035241   0.045957   0.767 0.443234
## as.factor(month)3                 0.065962   0.044724   1.475 0.140330
## as.factor(month)4                 0.116455   0.044961   2.590 0.009628 **
## as.factor(month)5                 0.098076   0.044355   2.211 0.027081 *
## as.factor(month)6                 0.113646   0.045905   2.476 0.013340 *
## as.factor(month)7                 0.143882   0.047931   3.002 0.002700 **
## as.factor(month)8                 0.187778   0.049108   3.824 0.000133 ***
## as.factor(month)9                 0.051515   0.047748   1.079 0.280705
## as.factor(month)10               -0.111702   0.047234  -2.365 0.018083 *
## as.factor(month)11               -0.124349   0.047209  -2.634 0.008471 **
## as.factor(month)12                0.055052   0.048557   1.134 0.256965
## wind_m_s                          0.072528   0.007397   9.804  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5821 on 3966 degrees of freedom
## Multiple R-squared:  0.06299,    Adjusted R-squared:  0.05802
## F-statistic: 12.69 on 21 and 3966 DF,  p-value: < 2.2e-16
```

```
linearHypothesis(fe_model, c("wind_m_s=0"), white.adjust = "hc2")
```

```
## Linear hypothesis test
##
## Hypothesis:
## wind_m_s = 0
##
## Model 1: restricted model
## Model 2: log_price_euro_kg ~ as.factor(country) + as.factor(year) + as.factor(month) +
##     wind_m_s
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1   3967
## 2   3966  1 77.658 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# same regression with fixed effects for year, month, and country
fe_tsls_vol_price_wind <- ivreg(log_volume_sold_kg ~ log_price_euro_kg +
                                + as.factor(country) + as.factor(year)
                                + as.factor(month) | as.factor(country)
                                + as.factor(year) + as.factor(month)
                                + wind_m_s, data = data_log)
# ivreg() is running 2 first stage least regressions, first with the coefficients on the right side
# of | and then with the coefficients on the left
summary(fe_tsls_vol_price_wind)
```

```
##
## Call:
## ivreg(formula = log_volume_sold_kg ~ log_price_euro_kg + +as.factor(country) +
##     as.factor(year) + as.factor(month) | as.factor(country) +
##     as.factor(year) + as.factor(month) + wind_m_s, data = data_log)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.2940 -1.8317 -0.1353  1.9969  6.2894
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       7.33742    0.20781  35.309  < 2e-16 ***
## log_price_euro_kg                -1.25004    0.46393  -2.694  0.00708 **
## as.factor(country)Italy          -0.68925    0.12970  -5.314 1.13e-07 ***
## as.factor(country)Portugal        1.71563    0.34614   4.957 7.48e-07 ***
## as.factor(country)United Kingdom -0.07422    0.31428  -0.236  0.81332
## as.factor(year)2014               0.14610    0.15281   0.956  0.33909
## as.factor(year)2015               0.18487    0.15221   1.215  0.22461
## as.factor(year)2016               0.21335    0.15320   1.393  0.16382
## as.factor(year)2017               0.07400    0.15224   0.486  0.62692
## as.factor(year)2018              -0.09137    0.15508  -0.589  0.55580
## as.factor(year)2019               0.03602    0.19688   0.183  0.85483
## as.factor(month)2                 0.06866    0.20972   0.327  0.74339
## as.factor(month)3                 0.51583    0.20489   2.518  0.01185 *
## as.factor(month)4                 0.91433    0.20297   4.505 6.84e-06 ***
## as.factor(month)5                 1.14887    0.20370   5.640 1.82e-08 ***
## as.factor(month)6                 1.14474    0.20164   5.677 1.47e-08 ***
## as.factor(month)7                 1.40047    0.21047   6.654 3.24e-11 ***
## as.factor(month)8                 1.26382    0.21692   5.826 6.12e-09 ***
## as.factor(month)9                 1.31072    0.21298   6.154 8.29e-10 ***
## as.factor(month)10                0.72059    0.22958   3.139  0.00171 **
## as.factor(month)11                0.48128    0.22575   2.132  0.03308 *
## as.factor(month)12                0.06683    0.21920   0.305  0.76049
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.648 on 3966 degrees of freedom
## Multiple R-Squared: 0.1522,  Adjusted R-squared: 0.1477
## Wald test: 9.963 on 21 and 3966 DF,  p-value: < 2.2e-16
```

- **The estimated price elasticity of demand is -1.250. The F-statistic is 77.658, which is greater than 10, so these instruments are non-weak and relevant.**