

EDS241: Assignment 4

Charles Hendrickson

03/10/2022

1 Data

```
# Load the data
EU_sardines_data <- data.frame(read.csv("EU_sardines.csv"))

# Take the log of volume_sold and price_euro_kg and create new columns for them
EU_sardines_data <- EU_sardines_data %>%
  mutate(log_volume_sold_kg = log(volume_sold_kg)) %>%
  mutate(log_price_euro_kg = log(price_euro_kg))
```

2 Question (a)

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

The price elasticity of demand for sardines is -1.545

We reject the null hypothesis that the price elasticity is equal to -1 because our linear hypothesis test tells us that the p-value is less than 0.001 and is statistically significant.

```
# Bivariate regression of log(volume_sold_kg) on log(price euro_kg)
model_1 <- lm_robust(log_volume_sold_kg ~ log_price_euro_kg, data = EU_sardines_data)

# Get summary of the model
summary(model_1)

##
## Call:
## lm_robust(formula = log_volume_sold_kg ~ log_price_euro_kg, data = EU_sardines_data)
##
## Standard error type: HC2
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
## (Intercept)      7.759    0.04302  180.34 0.000e+00   7.675   7.843 3986
## log_price_euro_kg -1.545    0.07813  -19.78 3.805e-83  -1.699  -1.392 3986
##
## Multiple R-squared:  0.1044 ,    Adjusted R-squared:  0.1042
## F-statistic: 391.3 on 1 and 3986 DF,  p-value: < 2.2e-16

# Linear hypothesis test
linearHypothesis(model_1, 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
##
##      Res.Df Df    Chisq      Pr(>Chisq)
## 1      3987
## 2      3986   1 48.724 0.000000000002946 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3 Question (b)

(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.

The estimated coefficient on wind speed is **0.06735**. This means that for every additional meter per second increase in wind speed, there is a **0.06735** percent increase in the log price of sardines. It does have the expected sign (positive) because increased wind speeds would make fishing effort more difficult, thus there should be a positive increase in the log price of sardines.

The F-statistic is **144.65**, which means the `wind_m_s` instrument is relevant and non-weak (Lecture 9, slide 14).

```
# Regress log(price_euro_kg) on wind_m_s
model_2 <- lm_robust(log_price_euro_kg ~ wind_m_s, data = EU_sardines_data)

# Get summary table of the regression coefficients
summary(model_2)

##
## Call:
## lm_robust(formula = log_price_euro_kg ~ wind_m_s, data = EU_sardines_data)
##
## Standard error type:  HC2
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## (Intercept) -0.30489    0.027309  -11.16 1.609e-28 -0.35843 -0.25135 3986
## wind_m_s      0.06735    0.005599   12.03 9.379e-33  0.05637  0.07832 3986
##
## Multiple R-squared:  0.0379 ,    Adjusted R-squared:  0.03766
## F-statistic: 144.7 on 1 and 3986 DF,  p-value: < 2.2e-16

# F-test for non-weak and relevant instruments
linearHypothesis(model_2, 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
##
##   Res.Df Df    Chisq Pr(>Chisq)
## 1     3987
## 2     3986   1 144.65  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4 Question (c)

- (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?

The two stage least squares (TSLS) regression of the price elasticity of demand for sardines using `wind_m_s` as an instrument for `log(price_euro_kg)` shows us that `-1.08802` is the estimated price elasticity of demand for sardines.

```
# Two stage least squares regression
tsls1 <- ivreg(log_volume_sold_kg ~ log_price_euro_kg | wind_m_s, data = EU_sardines_data)

# Summary of TSLS
summary(tsls1)

##
## Call:
## ivreg(formula = log_volume_sold_kg ~ log_price_euro_kg | wind_m_s,
##       data = EU_sardines_data)
##
## 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
```

5 Question (d)

- (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.

The estimated price elasticity of demand for sardines with fixed effects for each year, month, and country is `-1.25004`

```
# Completes two stage least regression with fixed effects for each year, month, and country
tsls2 <- ivreg(log_volume_sold_kg ~ log_price_euro_kg +
  as.factor(country) + as.factor(year) + as.factor(month) |
  wind_m_s + as.factor(country) + as.factor(year) + as.factor(month),
  data = EU_sardines_data)
```

```
# Get summary table of TSLS
summary(tsls2)
```

```
##
## Call:
## ivreg(formula = log_volume_sold_kg ~ log_price_euro_kg + as.factor(country) +
##       as.factor(year) + as.factor(month) | wind_m_s + as.factor(country) +
##       as.factor(year) + as.factor(month), data = EU_sardines_data)
##
## 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 0.0000001129314
## as.factor(country)Portugal    1.71563    0.34614   4.957 0.0000007476568
## 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 0.0000068372477
## as.factor(month)5       1.14887    0.20370   5.640 0.0000000181902
## as.factor(month)6       1.14474    0.20164   5.677 0.0000000146801
## as.factor(month)7       1.40047    0.21047   6.654 0.0000000000324
## as.factor(month)8       1.26382    0.21692   5.826 0.0000000061221
## as.factor(month)9       1.31072    0.21298   6.154 0.0000000008287
## 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
##
## (Intercept)          ***
## log_price_euro_kg      **
## as.factor(country)Italy    ***
## as.factor(country)Portugal ***
## as.factor(country)United Kingdom
## as.factor(year)2014
## as.factor(year)2015
## as.factor(year)2016
## as.factor(year)2017
## as.factor(year)2018
```

```
## as.factor(year)2019
## as.factor(month)2
## as.factor(month)3          *
## as.factor(month)4          ***
## as.factor(month)5          ***
## as.factor(month)6          ***
## as.factor(month)7          ***
## as.factor(month)8          ***
## as.factor(month)9          ***
## as.factor(month)10         **
## as.factor(month)11         *
## as.factor(month)12
## ---
## 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 F-statistic is 77.658, which means the instruments are relevant and non-weak (Lecture 9, slide 14).

```
# Can't get the f-statistic from ivreg
model_3 <- lm_robust(log_price_euro_kg ~ wind_m_s + as.factor(country) + as.factor(year) + as.factor(month))

# F-test for non-weak and relevant instruments (Lecture 9, slides 13-14)
linear_hypoth_model_3 <- linearHypothesis(model_3, c("wind_m_s = 0"), white.adjust = "hc2")

linear_hypoth_model_3
```

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