Problem Set2 Practical Questions

YULIA

3/20/2018

```
#installation_needed  <- TRUE
#loading_needed <- TRUE
#package_list <- c('foreign', 'xtable', 'plm','gmm', 'AER','stargazer','readstata13', 'boot', 'arm', 'l
#if(installation_needed){install.packages(package_list, repos='http://cran.us.r-project.org')}
#if(loading_needed){lapply(package_list, require, character.only = TRUE)}

# clear the global workspace

library(foreign)
library(ivpack)
```

```
## Loading required package: AER

## Loading required package: car

## Loading required package: lmtest

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: sandwich

## Loading required package: survival
```

```
setwd("/Users/yuliav/Downloads/")
my_data <- read.dta("fishdata.dta")
#summary(my_data)
```

1.Suppose you want to estimate the demand equation. You'd like to regress log of quantities in pounds of whiting per day on log of average daily price in US dollars per pound, dummies for day of the week, and control for weather on the shore. Do this via OLS.

```
form.ols <- qty ~ pricelevel + rainy + cold + day1 +day2 +day3 +day4
demand_fit <- lm(form.ols, data = my_data)
summary(demand_fit)
```

```
##
## Call:
## lm(formula = form.ols, data = my_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.23330 -0.34610  0.09256  0.41181  1.34425
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.286565   0.232468  39.948  < 2e-16 ***
## pricelevel  -0.613670   0.200616  -3.059  0.00283 **
## rainy        0.051075   0.177736   0.287  0.77441
## cold        -0.060147   0.134908  -0.446  0.65665
## day1         0.005496   0.207783   0.026  0.97895
## day2        -0.513751   0.204250  -2.515  0.01344 *
## day3        -0.567377   0.206943  -2.742  0.00721 **
## day4         0.069917   0.201330   0.347  0.72909
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6766 on 103 degrees of freedom
## Multiple R-squared:  0.2208, Adjusted R-squared:  0.1678
## F-statistic: 4.169 on 7 and 103 DF,  p-value: 0.0004448
```

2. Now do this demand analysis using stormy weather and mixed weather as instruments for price. Comment on the differences in the results.

```r
IV_weather_price_demand <- ivreg(qty ~ pricelevel + rainy + cold + day1 +day2 +day3 +day4  |  stormy + m
summary(IV_weather_price_demand)
```

```
##
## Call:
## ivreg(formula = qty ~ pricelevel + rainy + cold + day1 + day2 +
##     day3 + day4 | stormy + mixed + rainy + cold + day1 + day2 +
##     day3 + day4, data = my_data, x = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.17510 -0.42095  0.09309  0.48596  1.30057
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.67521    0.42054  23.007  < 2e-16 ***
## pricelevel  -1.06447    0.45155  -2.357  0.02029 *
## rainy        0.04289    0.18219   0.235  0.81433
## cold         0.01738    0.15451   0.113  0.91064
## day1        -0.05198    0.21890  -0.237  0.81278
## day2        -0.55182    0.21193  -2.604  0.01058 *
## day3        -0.60948    0.21526  -2.831  0.00557 **
## day4         0.06529    0.20625   0.317  0.75222
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.693 on 103 degrees of freedom
## Multiple R-Squared: 0.1826,  Adjusted R-squared: 0.127
## Wald test: 3.494 on 7 and 103 DF,  p-value: 0.002112
```

Now price level has a higher negative effect on quantity demanded.

3. Do you think the sailing weather variables (stormy, mixed) give us a good instrument for price in the demand equation? Explain and justify your answer.

Stromy weather –> price –> demand

No, I believe that price may not be very affected by the sailing weather variables and so we may have an irrelevant instrument.Also there can be channels other than price through which weather can influence demand. Eg: Stromy weather in the sea –> people are unwilling to go to the fish market –> less demand

But in case stormy/mixed weather indeed influences demand through price only, the instrument should be fine.

4. Do you think the sailing weather variables are strong instruments for price? Examine the first stage F statistic. Compute the Anderson Rubin confidence interval for the effect of price on demand. Is there a weak instruments problem?

```
first_stage_price_on_weather_demand <- lm(pricelevel ~ stormy + mixed + rainy + cold + day1 +day2 +day3
summary(first_stage_price_on_weather_demand)$fstatistic[1]
```

```
##    value
## 4.714005
```

F stat is 4.714 < 10 so we can worry about weak instrument problem.

```
ar_ci <- anderson.rubin.ci(IV_weather_price_demand, conflevel = 0.95)
ar_ci
```

```
## $confidence.interval
## [1] "[ -2.46695134664792 , -0.00483722346353988 ]"
```

The confidence interval looks nice, hence no weak instrument problem.

5. Suppose you want to estimate the supply equation. You'd like to regress log of quantities in pounds of whiting per day on log of average daily price in US dollars per pound and control for sailing weather (stormy, mixed).Do this via OLS.

```
form.ols <- qty ~ pricelevel + stormy + mixed
supply_fit <- lm(form.ols, data = my_data)
summary(supply_fit)
```

```
##
## Call:
## lm(formula = form.ols, data = my_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.48170 -0.36218  0.09401  0.49591  1.26881
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.0207     0.1975  45.672   <2e-16 ***
## pricelevel   -0.4308     0.2337  -1.843   0.0681 .
## stormy       -0.2791     0.1898  -1.471   0.1444
## mixed        -0.1169     0.1692  -0.691   0.4910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7168 on 107 degrees of freedom
## Multiple R-squared:  0.09135,    Adjusted R-squared:  0.06588
## F-statistic: 3.586 on 3 and 107 DF,  p-value: 0.0162
```

6. Now do this supply analysis using dummies for day of the week and weather on the shore as instruments for price. Comment on the differences in the results.

```
IV_weather_price_supply <- ivreg(qty ~ pricelevel + stormy + mixed + day1 + day2 + day3 + day4  |  storm
summary(IV_weather_price_supply)
```

```
##
## Call:
## ivreg(formula = qty ~ pricelevel + stormy + mixed + day1 + day2 +
##     day3 + day4 | stormy + mixed + rainy + cold + day1 + day2 +
##     day3 + day4, data = my_data, x = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.15265 -0.40586  0.06747  0.49265  1.21782
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.47957    1.91823   4.942 3.01e-06 ***
## pricelevel  -0.82144    2.49221  -0.330   0.7424
## stormy      -0.09921    1.02352  -0.097   0.9230
## mixed        0.05667    0.51802   0.109   0.9131
## day1        -0.01391    0.39910  -0.035   0.9723
## day2        -0.54438    0.29400  -1.852   0.0669 .
## day3        -0.60627    0.29575  -2.050   0.0429 *
## day4         0.05703    0.19995   0.285   0.7761
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6775 on 103 degrees of freedom
## Multiple R-Squared: 0.2186,  Adjusted R-squared: 0.1655
## Wald test: 3.723 on 7 and 103 DF,  p-value: 0.001244
```

7. Do you think these are good instruments for the supply equation? Explain and justify your answer.

Weather on shore may cause more/less people to come to fish market and hence influence the price for fish. If the price is higher sellers may be wiliing to supply more. However, in case of fish selling I do not think that there is a possibility to increase supply with higher price. The fish is already there and sailors can't just go out to the sea and get some more after observing a high price. Thus, I think the channel between weather on shore on some particular day and price for fish and supply of fish is weak. The instrument can be irrelevant.

8. Do you think these are strong instruments for price? Examine the first stage F statistic. Compute the Anderson Rubin confidence interval for the effect of price on demand. Is there a weak instruments problem?

```
first_stage_price_on_weather_supply <- lm(pricelevel ~ stormy + mixed + rainy + cold + day1 + day2 + da
summary(first_stage_price_on_weather_supply)$fstatistic[1]
```

```
##    value
## 4.714005
```

F stat is 4.714 < 10 so we can worry about weak instrument problem.

```
ar_ci <- anderson.rubin.ci(IV_weather_price_supply, conflevel = 0.95)
ar_ci
```

```
## $confidence.interval
## [1] "Whole Real Line"
```

We get "Whole Real Line" as confidence interval which means that confidence set is empty. Infinite confidence sets appear mainly when instruments are weak. In these cases, we have little or no information about the parameter of interest, which is correctly pointed out by these confidence sets.

9. Perform a simple nonparametric bootstrap to estimate the standard errors for the coefficient of price on quantity supplied, ignoring that this is time series data for now.

```
library(boot)
```

```
##
## Attaching package: 'boot'

## The following object is masked from 'package:survival':
##
##     aml

## The following object is masked from 'package:car':
##
##     logit
```

```
linear_model <- qty ~ pricelevel + stormy + mixed
coef_picker <- function(formula, data, indices) {
  d <- data[indices,] # allows boot to select sample
  fit <- lm(formula, data=d)
  return(summary(fit)$coef[1,1])
}
results <- boot(data= my_data, statistic=coef_picker, R=1000, formula=linear_model)
results
```

```
##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
##
```

```
## Call:
## boot(data = my_data, statistic = coef_picker, R = 1000, formula = linear_model)
##
##
## Bootstrap Statistics :
##     original        bias    std. error
## t1* 9.020742 -0.0003971439   0.1862587
```

10. Now perform a nonparametric block time-series bootstrap of the standard errors for the effect of price on quantity supplied, with a fixed block length of 5 periods. Check for sensitivity to the block length choice.

```
linear_model <- qty ~ pricelevel + stormy + mixed
coef_picker <- function(formula, data, indices) {
  d <- data[indices,] # allows boot to select sample
  fit <- lm(formula, data=d)
  return(summary(fit)$coef[1,1])
}

#bootstrap_price_quantiry <- tsboot(pricelevel, coef_picker, 1000, "fixed", l = 5)
#options(digits=2)
#mean.ts <- round((mean(bootstrap_lynx_mean$t)),digits=4)
#ses.ts<- round(sd(bootstrap_lynx_mean$t),digits=4)
#mean.ts
#ses.ts
```