IO_HW2

Mahdi Shahrabi (Collaborated with Anna Shchetkina)

9/22/2022

Question 1

```
library(haven)
library(plm)
library(stargazer)
```

```
##
## Please cite as:
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

```
library(AER)
```

```
## Loading required package: car
## Loading required package: carData
## 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
```

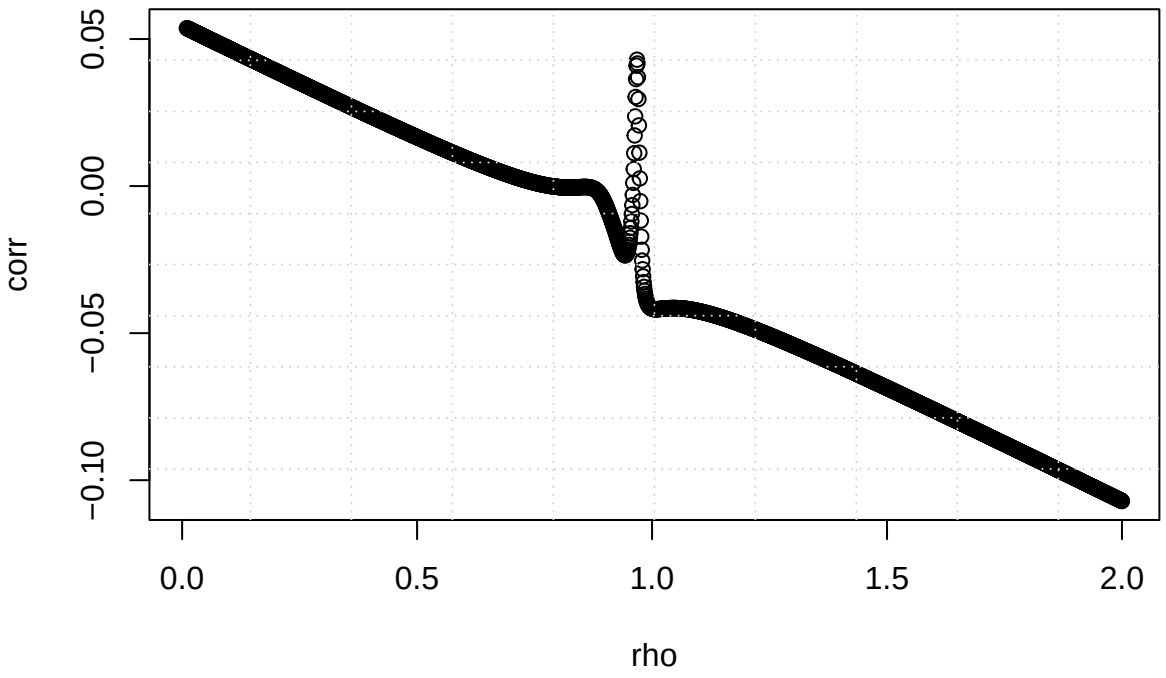
(B)

```
# Reading Data
df<-read_dta('GMdata.dta')
# Set it as panel
data <- pdata.frame(df, index=c("index","yr"))
# Make it balanced
# bdata <- make.pbalanced(data,("shared.individuals"))
bdata <- data

# Making Lagged Values
bdata$L1ldsal <- lag(bdata$ldsal, 5)
```

```
bdata$L2ldsal <- lag(bdata$ldsal, 10)
bdata$L3ldsal <- lag(bdata$ldsal, 15)

bdata$L1lemp <- lag(bdata$lemp, 5)
bdata$L2lemp <- lag(bdata$lemp, 10)
bdata$L3lemp <- lag(bdata$lemp, 15)

bdata$L1ldnpt <- lag(bdata$ldnpt, 5)
bdata$L2ldnpt <- lag(bdata$ldnpt, 10)
bdata$L3ldnpt <- lag(bdata$ldnpt, 15)


bdata$L1ldrst <- lag(bdata$ldrst, 5)
bdata$L2ldrst <- lag(bdata$ldrst, 10)
bdata$L3ldrst <- lag(bdata$ldrst, 15)

# Making time-industry Dummy
bdata$d357_73 <- ifelse(bdata$yr==73 & bdata$sic3==357,1,0)
bdata$L1d357_73<- lag(bdata$d357_73, 5)
bdata$L2d357_73 <- lag(bdata$d357_73, 10)

bdata$d357_78 <- ifelse(bdata$yr==78 & bdata$sic3==357,1,0)
bdata$L1d357_78<- lag(bdata$d357_78, 5)
bdata$L2d357_78 <- lag(bdata$d357_78, 10)

bdata$d357_83 <- ifelse(bdata$yr==83 & bdata$sic3==357,1,0)
bdata$L1d357_83<- lag(bdata$d357_83, 5)
bdata$L2d357_83 <- lag(bdata$d357_83, 10)

bdata$d357_88 <- ifelse(bdata$yr==88 & bdata$sic3==357,1,0)
bdata$L1d357_88 <- lag(bdata$d357_88, 5)
bdata$L2d357_88 <- lag(bdata$d357_88, 10)

# Making Time dummies
bdata$d73 <- ifelse(bdata$yr==73,1,0)
bdata$L1d73<- lag(bdata$d73, 5)
bdata$L2d73 <- lag(bdata$d73, 10)

bdata$d78 <- ifelse(bdata$yr==78,1,0)
bdata$L1d78<- lag(bdata$d78, 5)
bdata$L2d78 <- lag(bdata$d78, 10)

bdata$d83 <- ifelse(bdata$yr==83,1,0)
bdata$L1d83<- lag(bdata$d83, 5)
bdata$L2d83 <- lag(bdata$d83, 10)

bdata$d88 <- ifelse(bdata$yr==88,1,0)
bdata$L1d88<- lag(bdata$d88, 5)
bdata$L2d88 <- lag(bdata$d88, 10)
```

```
R<-seq(0.01,2,0.001)
out<-data.frame()
# bdata<-bdata[!is.na(bdata$L1ldnpt),]
for (rho in R) {
```

```
bdata$ldsal_rho <- (bdata$ldsal - rho*bdata$L1ldsal)
bdata$lemp_rho <- (bdata$lemp - rho*bdata$L1lemp)
bdata$ldnpt_rho <- (bdata$ldnpt - rho*bdata$L1ldnpt)
bdata$ldrst_rho <- (bdata$ldrst - rho*bdata$L1ldrst)

bdata$d73_rho <- (bdata$d73-rho*bdata$L1d73)
bdata$d78_rho <- (bdata$d78-rho*bdata$L1d78)
bdata$d83_rho <- (bdata$d83-rho*bdata$L1d83)
bdata$d88_rho <- (bdata$d88-rho*bdata$L1d88)

bdata$d357_73_rho <- (bdata$d357_73-rho*bdata$L1d357_73)
bdata$d357_78_rho <- (bdata$d357_78-rho*bdata$L1d357_78)
bdata$d357_83_rho <- (bdata$d357_83-rho*bdata$L1d357_83)
bdata$d357_88_rho <- (bdata$d357_88-rho*bdata$L1d357_88)

fit_model <- ivreg(ldsal_rho~lemp_rho+ldnpt_rho+ldrst_rho+d73_rho+d78_rho+d83_rho+d88_rho+d357_73_rho+d
bdata$epsilon <- resid(fit_model)


mc <- cov(na.omit(bdata[c("epsilon","L2ldsal")]))[1,2]
temp <- na.omit(bdata[c("epsilon","L2ldsal")])
mc2 <- abs(mean(temp$epsilon*temp$L2ldsal))

out <- rbind( out,data.frame(rho=rho,corr=mc,corr_abs=abs(mc),mc2=mc2))

}

plot(out[c("rho","corr")])
grid(10,10)
```



```
final<-out[out$corr_abs==min(out$corr_abs),]
final
```

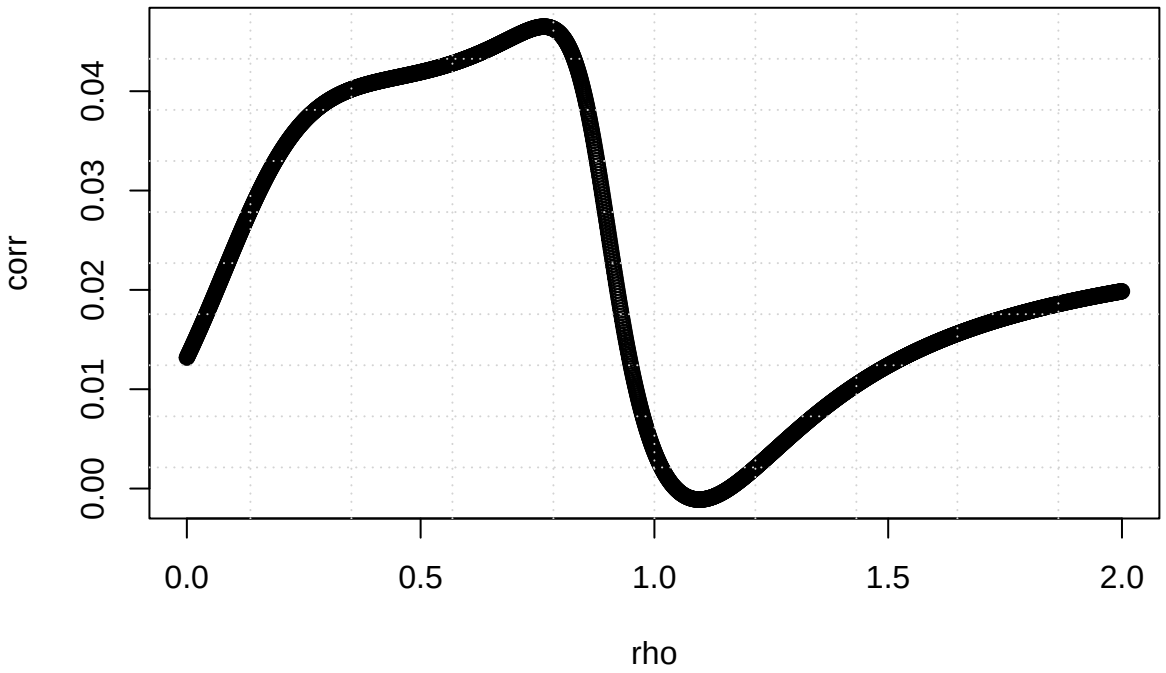
```
##       rho         corr    corr_abs          mc2
## 776 0.785 -7.05121e-07 7.05121e-07 7.040871e-07
```

```
rho<-0.785
bdata$ldsal_rho <- (bdata$ldsal - rho*bdata$L1ldsal)
bdata$lemp_rho <- (bdata$lemp - rho*bdata$L1lemp)
bdata$ldnpt_rho <- (bdata$ldnpt - rho*bdata$L1ldnpt)
bdata$ldrst_rho <- (bdata$ldrst - rho*bdata$L1ldrst)

bdata$d73_rho <- (bdata$d73-rho*bdata$L1d73)
bdata$d78_rho <- (bdata$d78-rho*bdata$L1d78)
bdata$d83_rho <- (bdata$d83-rho*bdata$L1d83)
bdata$d88_rho <- (bdata$d88-rho*bdata$L1d88)

bdata$d357_73_rho <- (bdata$d357_73-rho*bdata$L1d357_73)
bdata$d357_78_rho <- (bdata$d357_78-rho*bdata$L1d357_78)
bdata$d357_83_rho <- (bdata$d357_83-rho*bdata$L1d357_83)
bdata$d357_88_rho <- (bdata$d357_88-rho*bdata$L1d357_88)

fit_model <- ivreg(ldsal_rho~lemp_rho+ldnpt_rho+ldrst_rho+d73_rho+d78_rho+d83_rho+d88_rho+d357_73_rho+d

summary(fit_model)
```

```
##
## Call:
## ivreg(formula = ldsal_rho ~ lemp_rho + ldnpt_rho + ldrst_rho +
##     d73_rho + d78_rho + d83_rho + d88_rho + d357_73_rho + d357_78_rho +
##     d357_83_rho + d357_88_rho | d73_rho + d78_rho + d83_rho +
##     d88_rho + d357_73_rho + d357_78_rho + d357_83_rho + d357_88_rho +
##     bdata$L2ldsal + bdata$L2lemp + bdata$L2ldnpt + bdata$L2ldrst,
##     data = bdata, na.action = na.exclude)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.696640 -0.130528 -0.006865  0.113970  1.152163
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.82049    0.07577  10.829  < 2e-16 ***
## lemp_rho     0.51703    0.13694   3.776 0.000174 ***
## ldnpt_rho    0.43938    0.08724   5.037 6.09e-07 ***
## ldrst_rho    0.09630    0.06944   1.387 0.165999
## d78_rho      0.49663    0.04306  11.533  < 2e-16 ***
## d357_78_rho -2.40128    0.12369 -19.414  < 2e-16 ***
## d357_83_rho -1.03204    0.07461 -13.832  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2161 on 675 degrees of freedom
## Multiple R-Squared: 0.8589,  Adjusted R-squared: 0.8576
## Wald test: 424.9 on 6 and 675 DF,  p-value: < 2.2e-16
```

(C)

```
R<-seq(0,2,0.001)
out<-data.frame()

for (rho in R) {


bdata$ldsal_rho2 <- (bdata$ldsal - rho*bdata$L1ldsal) - (bdata$L1ldsal - rho*bdata$L2ldsal)
bdata$lemp_rho2 <- (bdata$lemp - rho*bdata$L1lemp) - (bdata$L1lemp - rho*bdata$L2lemp)
bdata$ldnpt_rho2 <- (bdata$ldnpt - rho*bdata$L1ldnpt) - (bdata$L1ldnpt - rho*bdata$L2ldnpt)
bdata$ldrst_rho2 <- (bdata$ldrst - rho*bdata$L1ldrst) - (bdata$L1ldrst - rho*bdata$L2ldrst)


bdata$d73_rho2 <- (bdata$d73-rho*bdata$L1d73)-(bdata$L1d73-rho*bdata$L2d73)
bdata$d78_rho2 <- (bdata$d78-rho*bdata$L1d78)-(bdata$L1d78-rho*bdata$L2d78)
bdata$d83_rho2 <- (bdata$d83-rho*bdata$L1d83)-(bdata$L1d83-rho*bdata$L2d83)
bdata$d88_rho2 <- (bdata$d88-rho*bdata$L1d88)-(bdata$L1d88-rho*bdata$L2d88)

bdata$d357_73_rho2 <- (bdata$d357_73-rho*bdata$L1d357_73)-(bdata$L1d357_73-rho*bdata$L2d357_73)
bdata$d357_78_rho2 <- (bdata$d357_78-rho*bdata$L1d357_78)-(bdata$L1d357_78-rho*bdata$L2d357_78)
bdata$d357_83_rho2 <- (bdata$d357_83-rho*bdata$L1d357_83)-(bdata$L1d357_83-rho*bdata$L2d357_83)
bdata$d357_88_rho2 <- (bdata$d357_88-rho*bdata$L1d357_88)-(bdata$L1d357_88-rho*bdata$L2d357_88)


fit_model <- ivreg(ldsal_rho2~lemp_rho2+ldnpt_rho2+ldrst_rho2+d73_rho2+d78_rho2+d83_rho2+d88_rho2+d357_
bdata$epsilon2 <- resid(fit_model)

mc <- cor(na.omit(bdata[c("epsilon2","L3ldsal")]))[1,2]
temp <- na.omit(bdata[c("epsilon2","L3ldsal")])

out <- rbind( out,data.frame(rho=rho,corr=mc,corr_abs=abs(mc)))
}

plot(out[c("rho","corr")])
grid(10,10)
```




```
final<-out[out$corr_abs==min(out$corr_abs),]
final
```

```
##        rho          corr     corr_abs
## 1160 1.159 -4.061012e-06 4.061012e-06
```

```
rho<-1.159
bdata$ldsal_rho2 <- (bdata$ldsal - rho*bdata$L1ldsal) - (bdata$L1ldsal - rho*bdata$L2ldsal)
bdata$lemp_rho2 <- (bdata$lemp - rho*bdata$L1lemp) - (bdata$L1lemp - rho*bdata$L2lemp)
bdata$ldnpt_rho2 <- (bdata$ldnpt - rho*bdata$L1ldnpt) - (bdata$L1ldnpt - rho*bdata$L2ldnpt)
bdata$ldrst_rho2 <- (bdata$ldrst - rho*bdata$L1ldrst) - (bdata$L1ldrst - rho*bdata$L2ldrst)


bdata$d73_rho2 <- (bdata$d73-rho*bdata$L1d73)-(bdata$L1d73-rho*bdata$L2d73)
bdata$d78_rho2 <- (bdata$d78-rho*bdata$L1d78)-(bdata$L1d78-rho*bdata$L2d78)
bdata$d83_rho2 <- (bdata$d83-rho*bdata$L1d83)-(bdata$L1d83-rho*bdata$L2d83)
bdata$d88_rho2 <- (bdata$d88-rho*bdata$L1d88)-(bdata$L1d88-rho*bdata$L2d88)

bdata$d357_73_rho2 <- (bdata$d357_73-rho*bdata$L1d357_73)-(bdata$L1d357_73-rho*bdata$L2d357_73)
bdata$d357_78_rho2 <- (bdata$d357_78-rho*bdata$L1d357_78)-(bdata$L1d357_78-rho*bdata$L2d357_78)
bdata$d357_83_rho2 <- (bdata$d357_83-rho*bdata$L1d357_83)-(bdata$L1d357_83-rho*bdata$L2d357_83)
bdata$d357_88_rho2 <- (bdata$d357_88-rho*bdata$L1d357_88)-(bdata$L1d357_88-rho*bdata$L2d357_88)


fit_model <- ivreg(ldsal_rho2~lemp_rho2+ldnpt_rho2+ldrst_rho2+d73_rho2+d78_rho2+d83_rho2+d88_rho2+d357_
```

```
summary(fit_model)
```

```
##
## Call:
## ivreg(formula = ldsal_rho2 ~ lemp_rho2 + ldnpt_rho2 + ldrst_rho2 +
##     d73_rho2 + d78_rho2 + d83_rho2 + d88_rho2 + d357_73_rho2 +
```

```
##     d357_78_rho2 + d357_83_rho2 + d357_88_rho2 | d73_rho2 + d78_rho2 +
##     d83_rho2 + d88_rho2 + d357_73_rho2 + d357_78_rho2 + d357_83_rho2 +
##     d357_88_rho2 + bdata$L3ldsal + bdata$L3lemp + bdata$L3ldnpt +
##     bdata$L3ldrst, data = bdata, na.action = na.exclude)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -11.98337  -0.94682  -0.02479   1.06170   6.32130
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.8973     2.7739   0.684    0.495
## lemp_rho2     -0.7363     3.7953  -0.194    0.846
## ldnpt_rho2     3.8581     6.7006   0.576    0.565
## ldrst_rho2    -0.5971     6.1459  -0.097    0.923
## d357_78_rho2  -0.6458     0.9895  -0.653    0.515
##
## Residual standard error: 2.021 on 209 degrees of freedom
## Multiple R-Squared: -9.679,  Adjusted R-squared: -9.883
## Wald test: 1.109 on 4 and 209 DF,  p-value: 0.3534
```