# Income Relative

*Henri Makika*

*7/3/2019*

## Income relative

```
library(readxl)
library(stargazer)
library(AER)
library(MatchIt)

Dados = read_excel("~/Videos/Inverno 2019/Aula 3/Data_RelativeIncome.xls")

summary(Dados)
```

```
##      rural            mulher         anos_est         rndpcap
##  Min.   :0.0000   Min.   :0.000   Min.   : 0.000   Min.   :   28.88
##  1st Qu.:0.0000   1st Qu.:0.000   1st Qu.: 4.000   1st Qu.:  459.51
##  Median :0.0000   Median :0.000   Median : 8.000   Median :  781.91
##  Mean   :0.2875   Mean   :0.266   Mean   : 7.404   Mean   : 1281.20
##  3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:11.000   3rd Qu.: 1352.27
##  Max.   :1.0000   Max.   :1.000   Max.   :15.000   Max.   :36446.79
##                                   NA's   :9
##   meanrndpcap         sufrnd
##  Min.   : 267.5   Min.   :0.0000
##  1st Qu.: 603.7   1st Qu.:0.0000
##  Median : 807.2   Median :0.0000
##  Mean   :1113.4   Mean   :0.2649
##  3rd Qu.:1333.8   3rd Qu.:1.0000
##  Max.   :6579.3   Max.   :1.0000
##
```

### Regression simples

Il est question ici de l'article du professeur Gori (voir Gori 2013).

```
linear = lm(sufrnd ~ log(meanrndpcap), data = Dados)
```

### Regression multiple

```
linmultiple = lm(sufrnd ~ log(meanrndpcap) + log(rndpcap), data = Dados) #+ #(anos_est) +
                  #mulher + rural, data = Dados)
#summary(linmultiple)
stargazer(linear, linmultiple, type = "text", digits = 5,
          column.labels = c("Lin-model", "Multi-model"),
          keep.stat = c('n','rsq','adj.rsq','f'), out = "mrd.txt")
```

```
##
## ===============================================================================
```

```
##                                 Dependent variable:
##                     ---------------------------------------------------
##                                         sufrnd
##                               Lin-model                 Multi-model
##                                  (1)                       (2)
## ------------------------------------------------------------------------
## log(meanrndpcap)              0.14598***                -0.03054**
##                               (0.01230)                 (0.01456)
##
## log(rndpcap)                                            0.19582***
##                                                         (0.00968)
##
## Constant                      -0.73109***               -0.84067***
##                               (0.08420)                 (0.08001)
##
## ------------------------------------------------------------------------
## Observations                    3,628                     3,628
## R2                              0.03742                   0.13512
## Adjusted R2                     0.03715                   0.13464
## F Statistic           140.94990*** (df = 1; 3626) 283.16600*** (df = 2; 3625)
## ========================================================================
## Note:                                        *p<0.1; **p<0.05; ***p<0.01
```

Interpretation se dit en terme de probabilité, puisque la variable dependente est binaire. variation de sufficence de revenu du voisin sur soi est de 0.14. (modèle 1). C'est 0.14% de probabilité de satisfaction. De (1) le modèle est super-estimé. Pourquoi ? Parce qu'on a pas considéré d'autres variables importantes pour conclure le modèle. La satisfaction à diminuer du fait que nous avons considéré d'autres variables de controles. Il ne plus possible de dire qu'on a sur-estimé la regression.

## Exercice 2

Voir article Gori et al (). Titre : Saude e mercado do trabalho.

```
Donne = read_excel("~/Videos/Inverno 2019/Aula 3/Data_HealthIncome.xls")

summary(Donne)
```

```
##      rural          declarante        boasaude          lnrnd
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :-10.628
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.0000   1st Qu.:  1.088
##  Median :0.0000   Median :1.0000   Median :1.0000   Median :  1.463
##  Mean   :0.0561   Mean   :0.5133   Mean   :0.8358   Mean   :  1.338
##  3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:  2.001
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :  6.500
##     escolar            fem             idade            cp
##  Min.   : 0.000   Min.   :0.0000   Min.   :10.00   Min.   :0.0000
##  1st Qu.: 6.000   1st Qu.:0.0000   1st Qu.:26.00   1st Qu.:0.0000
##  Median :11.000   Median :0.0000   Median :36.00   Median :0.0000
##  Mean   : 9.181   Mean   :0.4321   Mean   :37.22   Mean   :0.1586
##  3rd Qu.:11.000   3rd Qu.:1.0000   3rd Qu.:47.00   3rd Qu.:0.0000
##  Max.   :15.000   Max.   :1.0000   Max.   :89.00   Max.   :1.0000
##       empr
##  Min.   :0.00000
##  1st Qu.:0.00000
```

```
##  Median :0.00000
##  Mean   :0.04988
##  3rd Qu.:0.00000
##  Max.   :1.00000
```

### Regression simple

```
linsimple = lm(boasaude ~ (lnrnd) + fem + idade + escolar + declarante + rural,
               data = Donne)
```

### Variables instrumentales (IV)

```
intrument = ivreg(boasaude ~ lnrnd + fem + idade + escolar + declarante + rural | .-lnrnd +
                    cp + empr, data = Donne)

stargazer(linsimple, intrument, type = "text", digits = 5,
          column.labels = c("", ""),
          keep.stat = c('n','rsq','adj.rsq','f'), out = "mrd.txt")
```

```
##
## ===================================================
##                          Dependent variable:
##              ---------------------------------------
##                              boasaude
##                            OLS              instrumental
##                                                variable
##
##                            (1)                  (2)
## ---------------------------------------------------
## lnrnd                   0.01088***            0.00224
##                         (0.00137)            (0.01094)
##
## fem                    -0.03393***          -0.03829***
##                         (0.00504)            (0.00745)
##
## idade                  -0.00546***          -0.00532***
##                         (0.00019)            (0.00027)
##
## escolar                 0.01565***           0.01683***
##                         (0.00067)            (0.00163)
##
## declarante               0.00201              0.00181
##                         (0.00498)            (0.00499)
##
## rural                   -0.02369**           -0.03136**
##                         (0.01069)            (0.01440)
##
## Constant                0.89598***           0.89361***
##                         (0.01094)            (0.01134)
##
## ---------------------------------------------------
```

```
## Observations            21,212               21,212
## R2                     0.09487              0.09318
## Adjusted R2            0.09462              0.09292
## F Statistic  370.44270*** (df = 6; 21205)
## ===================================================
## Note:                      *p<0.1; **p<0.05; ***p<0.01
```