

# Lab 7: TabLego and LPM

*Julieth Santamaria*

*March 15, 2019*

## TabLego

To install TabLego on your computers, follow the instructions in file ‘tabLego Introduction.pdf’, in the ‘functions’ folder. You can also read this file to check how the command works.

1. Unzip the tabLego zip file and put it in the ‘functions’ folder

```
# RUN THIS ONLY ONCE!
install.packages('lmtest')
install.packages('sandwich')
setwd("G:/My Drive/U of M/TA/TA APEC3003/APEC 3003 - 2019") # Your 'functions' folder
install.packages("tabLego", repos=NULL, type="source")
```

## Application (measurement error)

1. Set the working directory and load nscg17.

```
setwd("G:/My Drive/U of M/TA/TA APEC3003/APEC 3003 - 2019/APEC 3003 R work/labs/")
load("../data/nscg17.rdata")
```

2. Run the following code that creates variables we have already used in past labs

```
nscg17 <- within(nscg17, {
  # Salary
  salary[salary >= 9999998 | salary==0] <- NA

  # Potential experience
  exper <- 2017-dgryr

  # Gender
  female <- NA
  female[gender=="F"]<-1
  female[gender=="M"]<-0

  #Education
  bachelors<-NA
  bachelors<-as.numeric(dgrdg==1)
  masters<-NA
  masters<-as.numeric(dgrdg==2)
  phd<-NA
  phd<-as.numeric(dgrdg==3)
  professional<-NA
  professional<-as.numeric(dgrdg==4)
})
```

3. Assume that experience is observed with error. Let’s check two options:

- Measurement error 1: You observe experience + two years

- Measurement error 2: You observe experience with a random error of mean two and S.D.=5 Create these variables and run three regressions. One of salary on observed potential experience; a second regression of salary on experience measured with the first type of error; and a third regression of salary on experience measured with the second type of error.

```
nscg17 <- within(nscg17, {
  N=nrow(nscg17)
  exper.error1 <- exper+2
  exper.error2 <- exper+ rnorm(N,2,5)
})
reg1<-lm(salary~exper, data = nscg17)
reg2<-lm(salary~exper.error1, data = nscg17)
reg3<-lm(salary~exper.error2, data = nscg17)
```

#### 4. Let's create a nice table

```
library(openxlsx)
library(tabLego)
library(sandwich)
library(lmtest)
library(zoo)
```

```
# 1. Import labels of the independent variables (You can also label some statistics).
varTags <- read.xlsx("varTags_example1.xlsx")
print(varTags)
```

```
##           xvars           var.tags
## 1 (Intercept)           Intercept
## 2      exper      Potential experience
## 3 exper.error1 Experience with error 1
## 4 exper.error2 Experience with error 2
```

```
# 2. Create the list of regressions you will include in the table
regsdof <- addRegs(list(reg1, reg2, reg3))
```

```
# 3. Define how characteristics of your table
tab1 <- textTable(reg.frame=regsdof,
  regnums=c(1,2,3), # To set the order of the regressions
  decs=2, # Sets the number of decimals to 2
  var.tags=varTags # To use the labels
)
```

```
# 4. Display your table
textTablePrint(tab1,
  before="Table 1: Regressions of salary on experience",
  after=c("Significance: 0.01=***, 0.05=**, 0.1=*",
  "Note: This table displays the summarized results of running a",
  "regression of salary on experience [a], and experienced measured",
  "with two types of error.")
```

```
## Table 1: Regressions of salary on experience
## -----
##           [a]           [b]           [c]
##           Intercept 73322.17*** 71397.05*** 74203.73***
##           (473.22)   (522.46)   (488.57)
## Potential experience 962.56***
```

```
##                               (30.73)
## Experience with error 1          962.56***
##                               (30.73)
## Experience with error 2          810.09***
##                               (27.58)
##                               N      69344      69344      69344
##                               r.squared  0.018      0.018      0.015
## -----
## Significance: 0.01=***, 0.05=**, 0.1=*
## Note: This table displays the summarized results of running a
## regression of salary on experience [a], and experienced measured
## with two types of error.
```

The first type of measurement error does not affect the estimate because the variance of the error is zero. However, the second type of error makes the coefficient be biased towards zero. This is called **Attenuation Bias**

## Your turn! (Application using a linear probability model)

1. Create a dummy for whether the person worked for pay or profit during the week for February 1, 2017/

```
nscg17<-within(nscg17,{
  work<-as.numeric(wrkg=="Y")
})
with(nscg17,table(wrkg,work,useNA='always'))
```

```
##      work
## wrkg   0   1 <NA>
##  N   14041  0   0
##  Y     0 69631  0
## <NA>   0   0   0
```

2. Run a linear probability model of employment on experience, education and sex, adding one variable at a time.

```
reg4<-lm(work~exper,data=nscg17)
reg5<-lm(work~exper+masters+phd+professional,data=nscg17)
reg6<-lm(work~exper+masters+phd+professional+female,data=nscg17)
```

3. Create a file of variable tags and save it in the labs folder. Name the file 'varTags\_example2.xlsx'
4. Create a nice table!

```
varTags2 <- read.xlsx("varTags_example2.xlsx")

regsdF <- addRegs(list(reg4, reg5, reg6))
tab2 <- textTable(reg.frame=regsdF, regnums=c(1,2,3), decs=2,var.tags=varTags2)

textTablePrint(tab2,
  before="Table 2: Regressions of employment",
  after=c("Significance: 0.01=***, 0.05=**, 0.1=*"))
```

```
## Table 2: Regressions of employment
## -----
##                               [a]      [b]      [c]
##      Intercept  1.00***  1.00***  1.04***
##                               (0.00)  (0.00)  (0.00)
## Potential experience -0.01*** -0.01*** -0.01***
```

```

##          (0.00)  (0.00)  (0.00)
##      Master's degree  -0.01*** -0.01**
##          (0.00)  (0.00)
##      Professional degree  0.06*** 0.06***
##          (0.01)  (0.01)
##          PhD degree  0.03*** 0.03***
##          (0.00)  (0.00)
##          Female  -0.08***
##          (0.00)
##          N 83672  83672  83672
##          r.squared 0.106  0.108  0.119
## -----
## Significance: 0.01=***, 0.05=**, 0.1=*

```

5. Interpret the coefficients. How are LPMs different to linear regressions?