Lecture: Analyzing the household: smoothing and adjusting
Classroom assignment
Iván Mejía-Guevara

1. Use the earnings (yle) and unincorporated (yls) income variables from your survey data (the ones you have estimated using NTA methods):

a. Use the "mean" command in Stata to construct the age profiles and confidence intervals (assume you have a simple random sample design survey)
b. Identify whether you have high variability problems for some age groups and characteristics of the profile you migh want to keep
c. Identify the "t" statistic used by stata to construct the confidence interval
c. Compute the coefficient of variation for the unsmooted age profile

2. Please use the "supsmu" R code to smooth the age profiles:
a. Think first in the strategy you want to follow: remember the guidelines provided in the lecture.
b. Smooth the profiles using 4 span options: 0.05, 0.1, 0.3 and "cv".
c Compute the coefficient of variation for the smoothed age profile. Do the same for the unsmoothed profile and compare.

3. Construct the final labor income age profile:
a. Use your macro controls to adjust the earnings and unincorporated income profiles.
b. Sum yle, yls up to obtain the yl.
c. Graph your results.

Hint: You may use the following R code example

******Friedman's SuperSmoother and confidence intervals*******

```
#input data
nta<-read.table("data.txt", header=T, na.strings=c(".","NA"))
age<-nta$age
yle_u<-nta$yl                   #unsmoothed age profile
se<-nta$se                      #standard error
cil<-nta$cil                    #confidence interval (lower value)
pop<-nta$sample                         #population by age
t<-(yle_u[20]-cil[20])/se[20] #t-statistic
zeros<-rep(0,91)
sp<-0.05                        #span selection
J<-91*sp                        #neighborhood

# smoothing
test<-supsmu(age[12:91], yle_u[12:91], pop[12:91], span="0.05")
yle_s<-append(zeros[1:11],replace(test$y, test$y<=0,0))
```

```
# confidence interval for smoothed profile
cils<-yle_s-t*se/J       #lower
cius<-yle_s+t*se/J      #upper

# plot confidence intervals
png("yle.png")
plot(age,yle_s,"l", col="dark red",ylab="mexican pesos",main="yle: sna 1993", xlim=c(0,90))
points(age,cils,"l", col="dark blue")
points(age,cius,"l", col="dark blue")
axis(1, at=5*0:range(age)[2], xlab="age")
dev.off()

# macro-adjustment
mcyle<-2251576293524.85              #macro-control for yle
beta_yleu<-mcyle/(pop%*%yle_u)
yleu<-beta_yleu*yle_u
beta_yles<-mcyle/(pop%*%yle_s)
yle<-beta_yles*yle_s

# plot final profile
postscript(file="yle.eps",
          paper="special",
          width=8,
          height=8,
          horizontal=FALSE)
             yvalues = runif(50)
plot(age,yleu,"l", col="dark red",ylab="mexican pesos",main="yle: sna 1993", xlim=c(0,90))
points(age,yle,"l", col="dark blue")
axis(1, at=5*0:range(age)[2], xlab="age")
dev.off()

#save profile
write.csv(yle,"yle.csv")
```