

STATA 'tricks': Creating professional graphs to enhance communication of statistical results

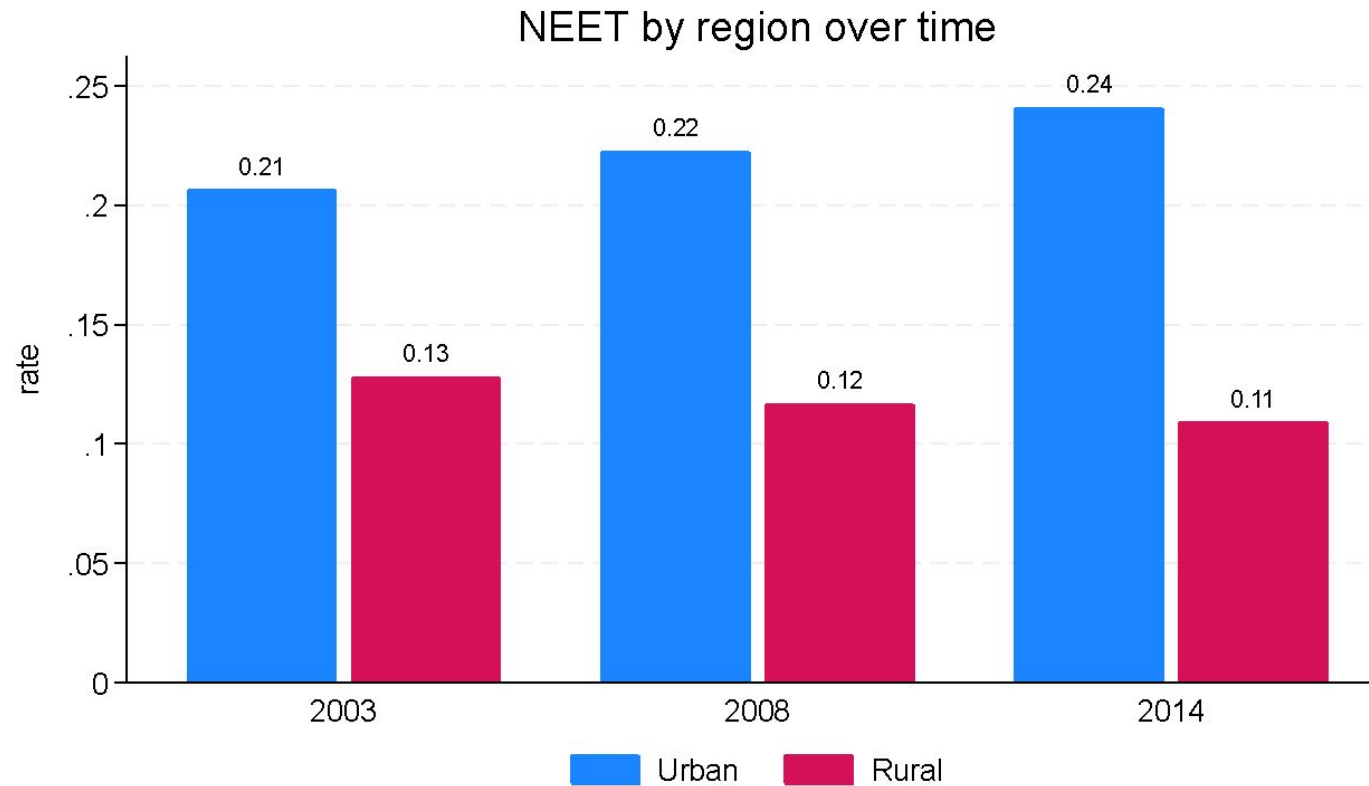
Dr. Paul Sirma – Economist, AIR

Prof. Ashu Handa, AIR Institute Fellow and Kenan Eminent Professor of Public Policy, University of North Carolina

Outline for session one

- Descriptive statistics of your data
 - Means table (required in a paper), other tables by group
- Graphs—much more user friendly
 - Bar graphs by group
 - Lowess graphs for two-way relationships (by group)
 - Walk you through the STATA commands for these graphs, focusing on the options to enhance the look and readability of the graphs
 - Remember, we want to make it as easy as possible for the reader/referee to get the main message

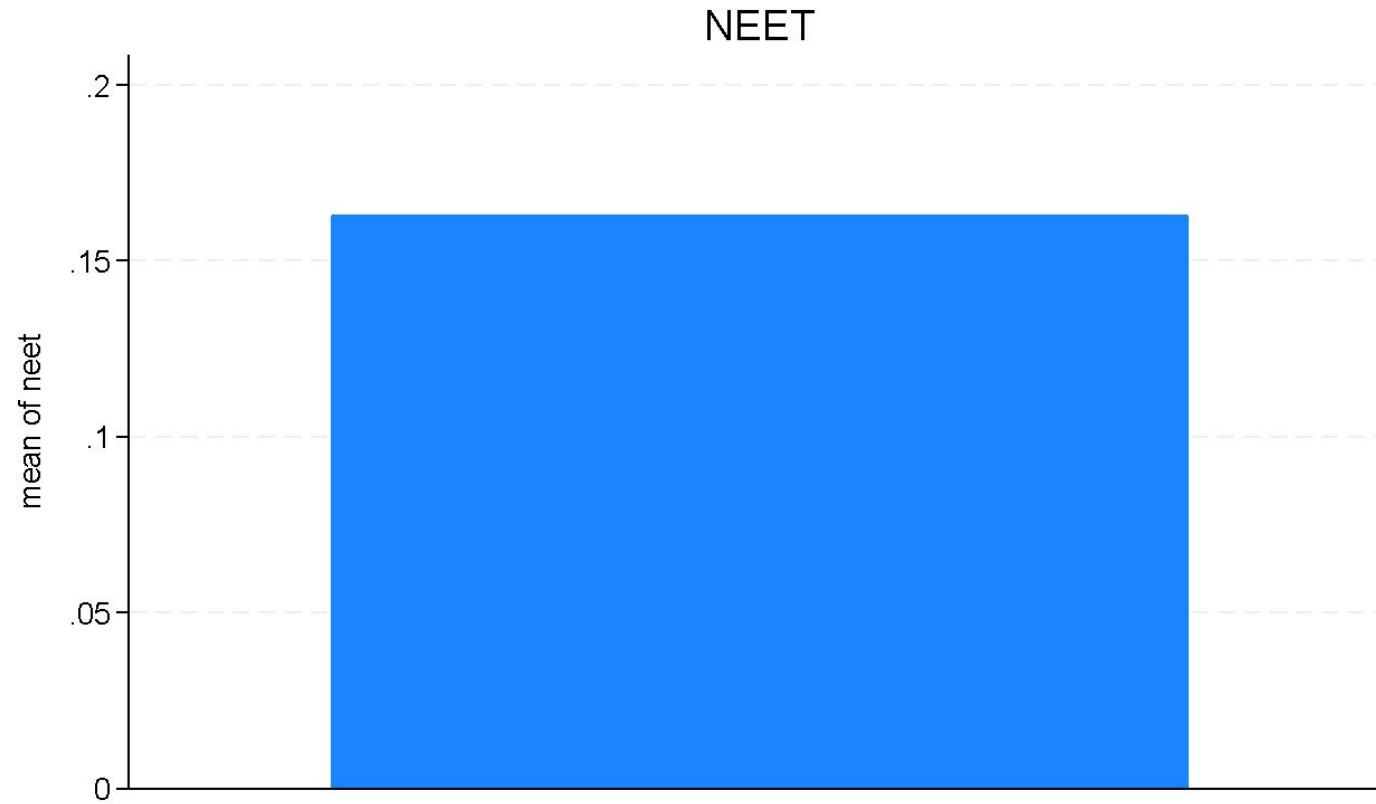
Goal: show the trend in NEET (Not in Education, Employment or Training) rates in rural and urban areas over time



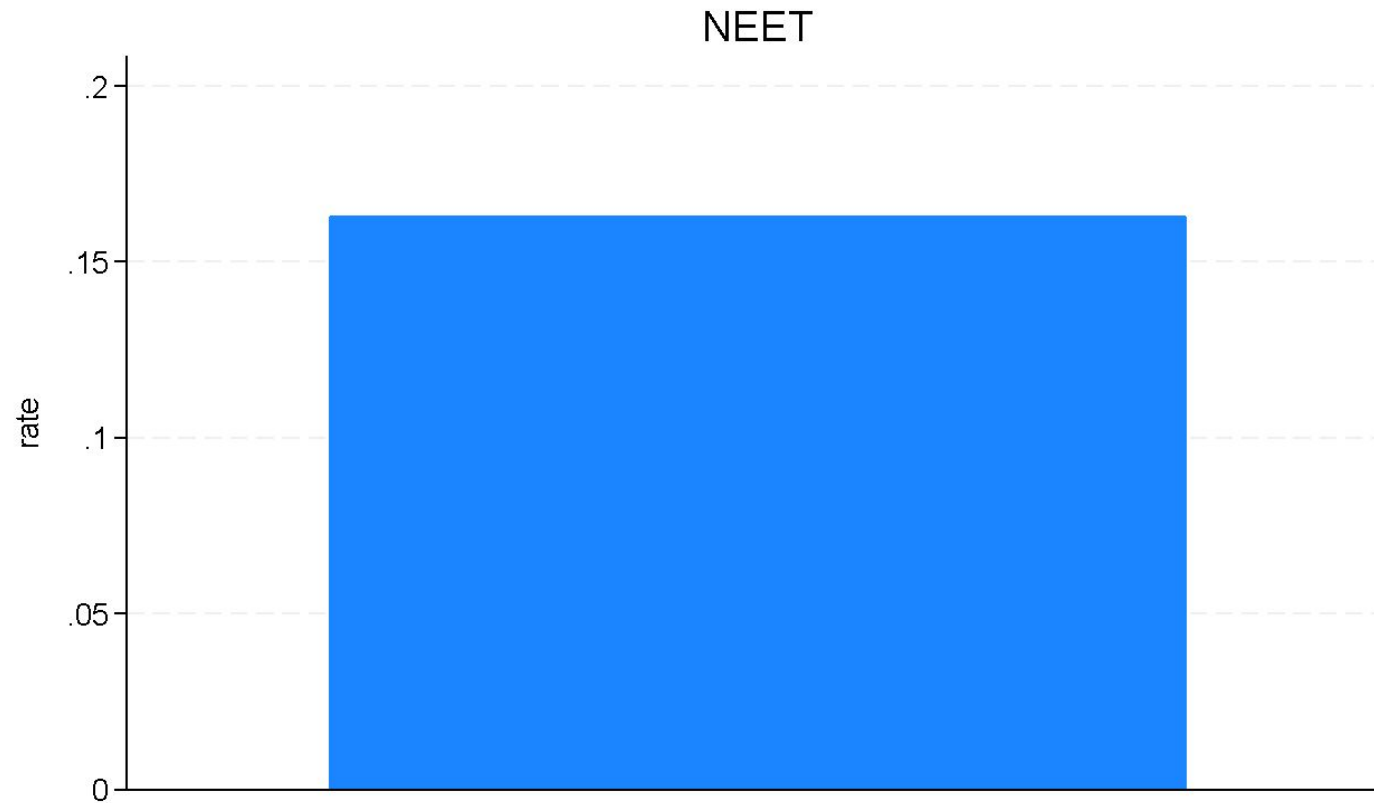
graph bar neet



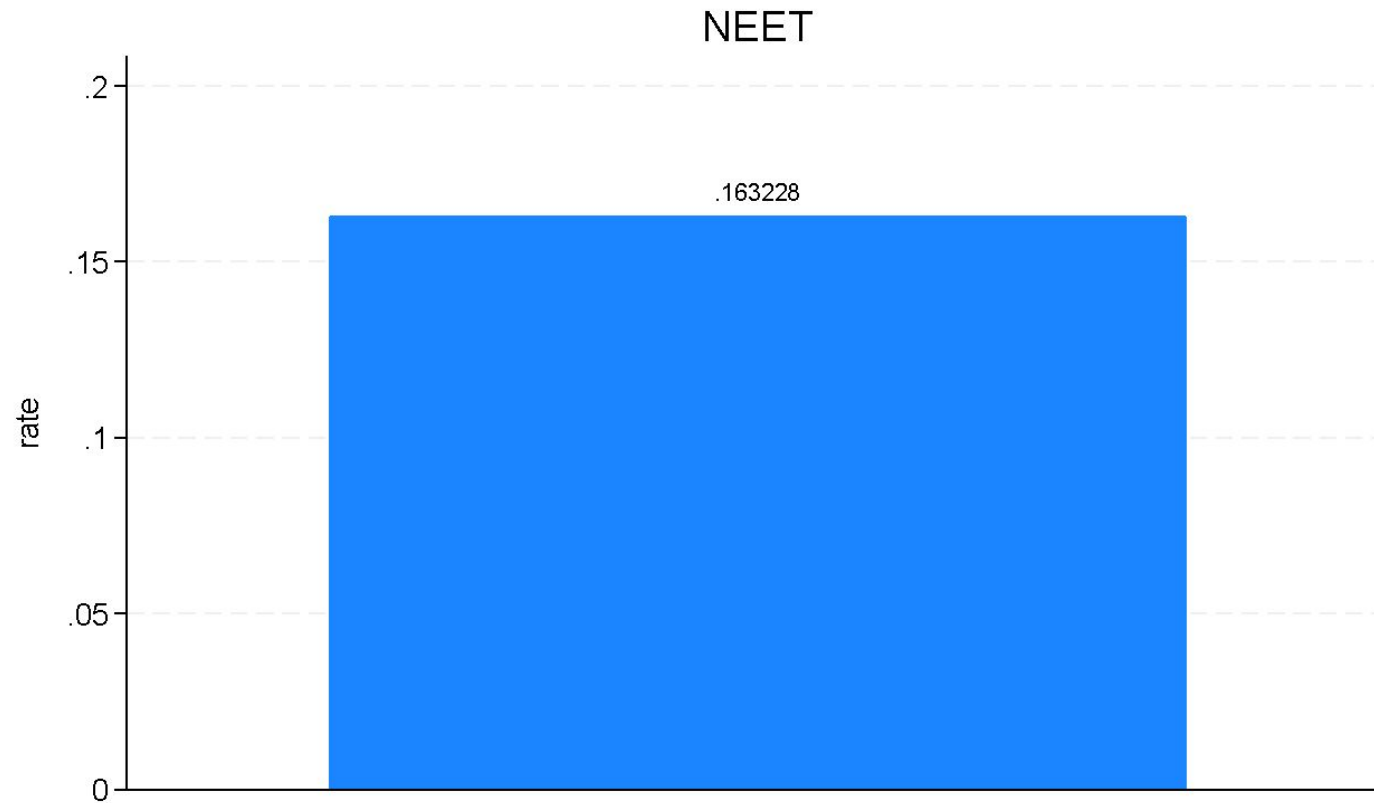
graph bar neet , title("NEET")



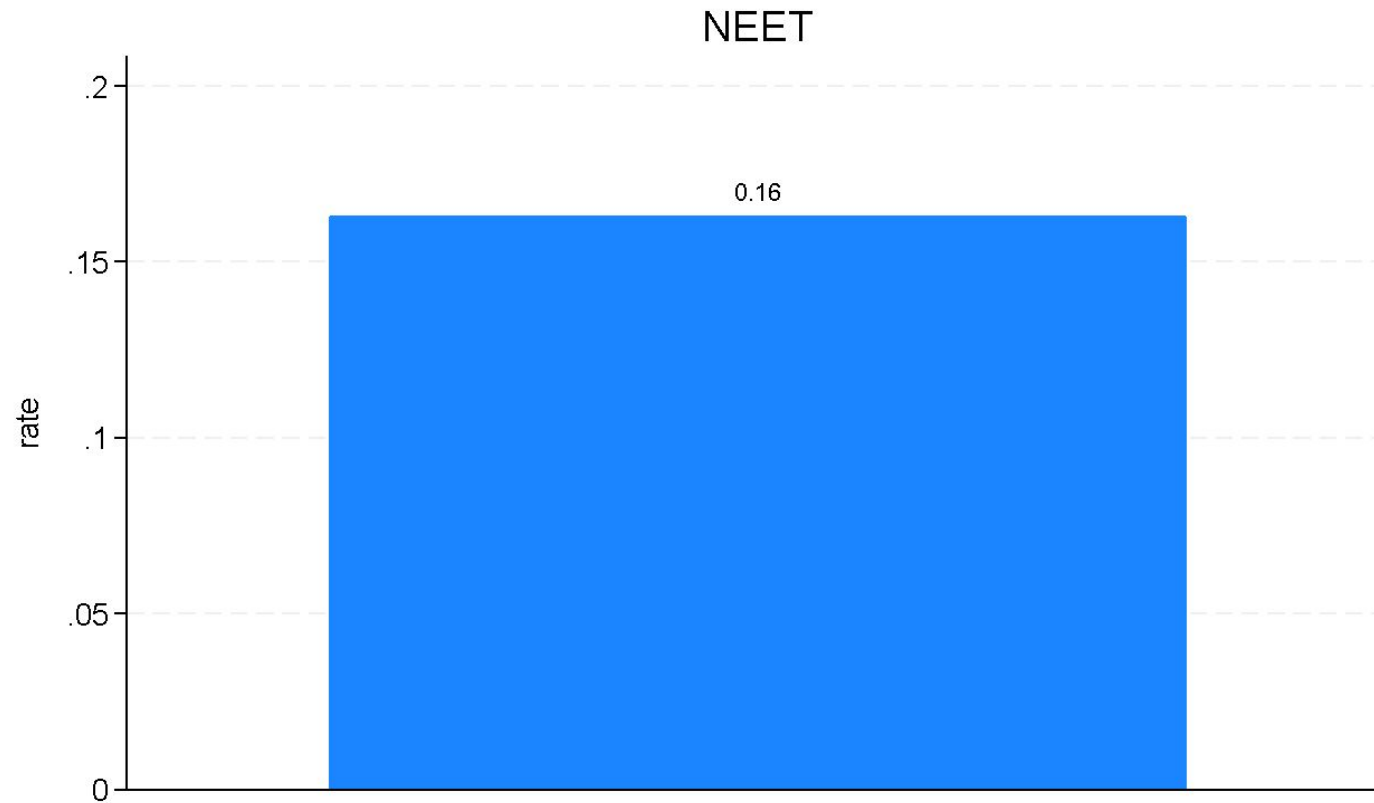
```
graph bar neet , title("NEET")  
ytitle("rate")
```



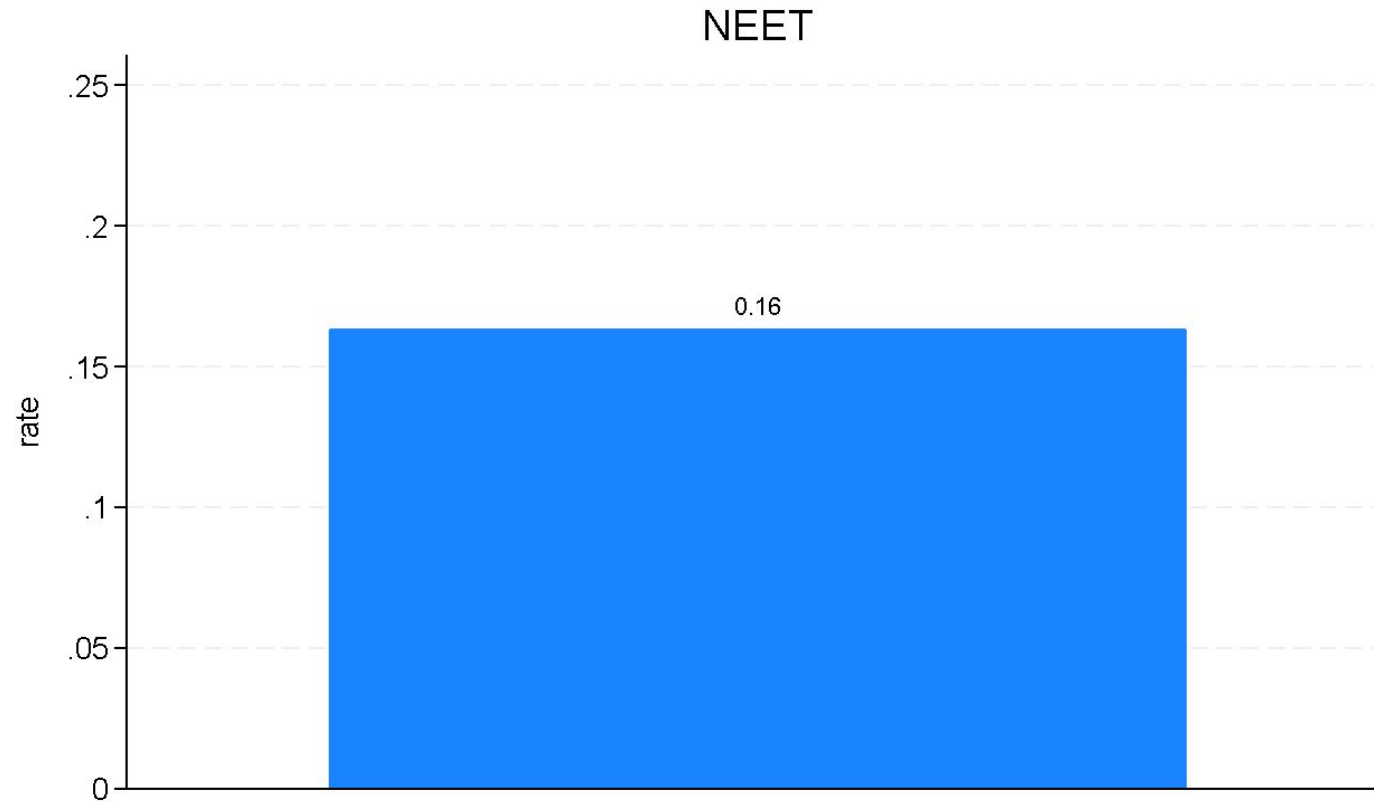
```
graph bar neet , title("NEET")  
ytitle("rate") blabel(bar)
```



```
graph bar neet , title("NEET")  
ytitle("rate") blabel(bar, format(%9.2f))
```

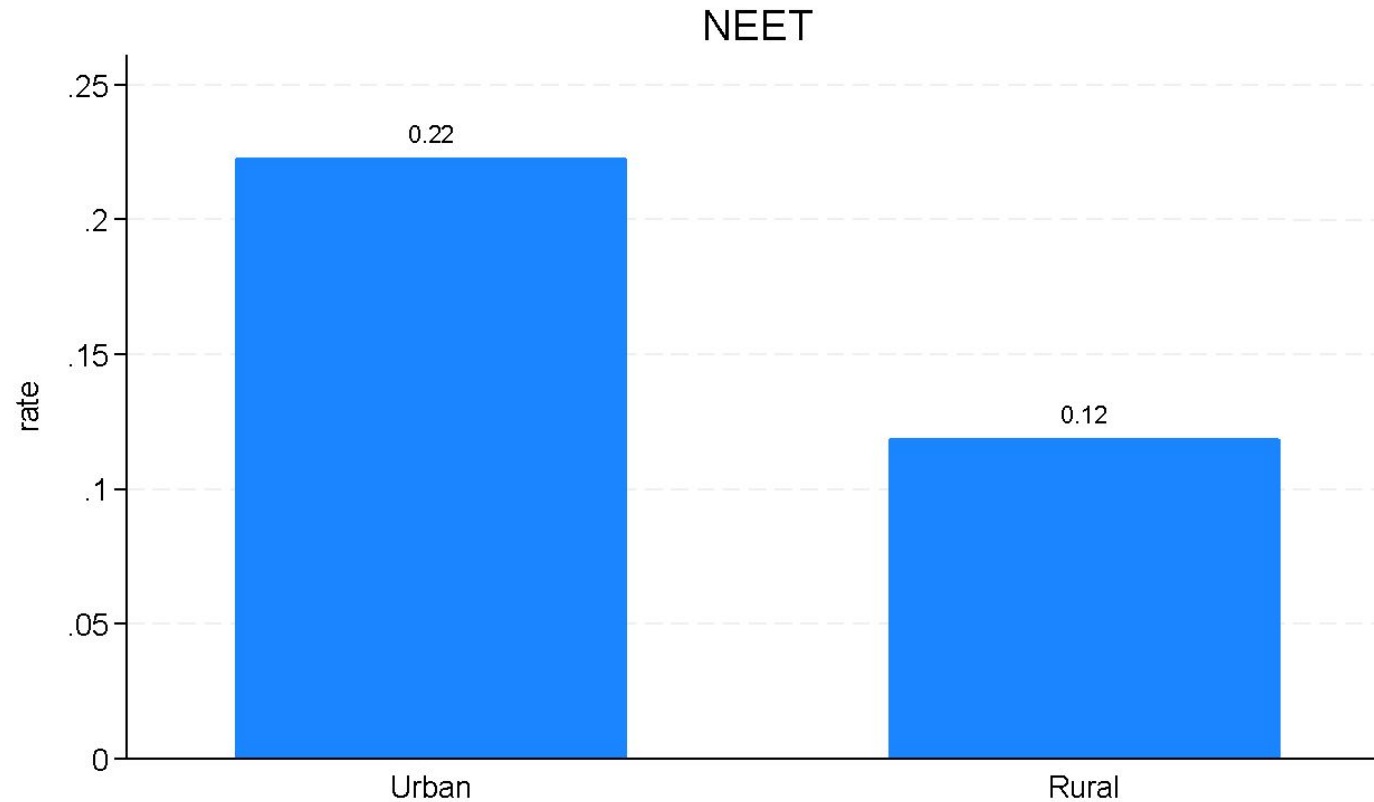


```
graph bar neet , title("NEET")
yttitle("rate") blabel(bar, format(%9.2f))
ylabel(0(.05).25)
```

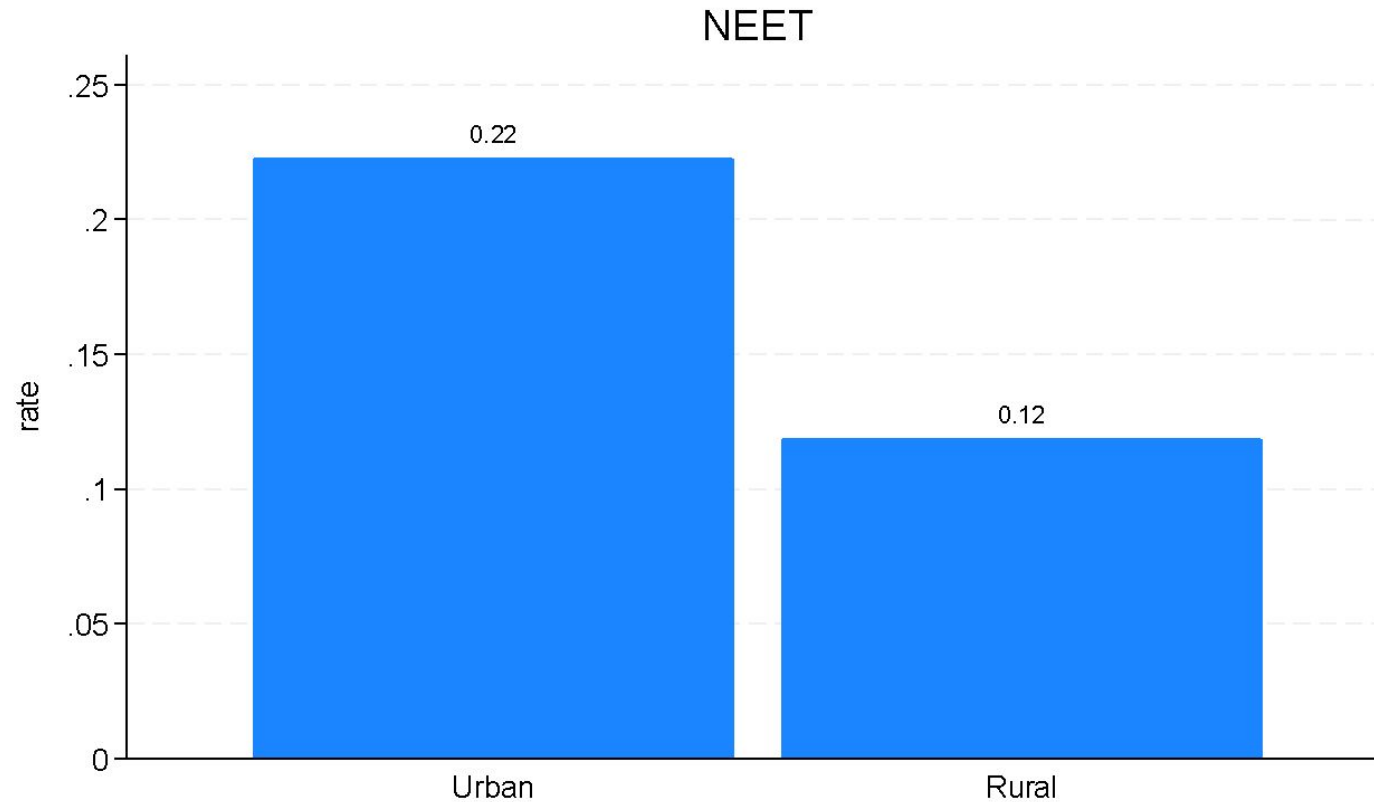


NEET for urban and rural

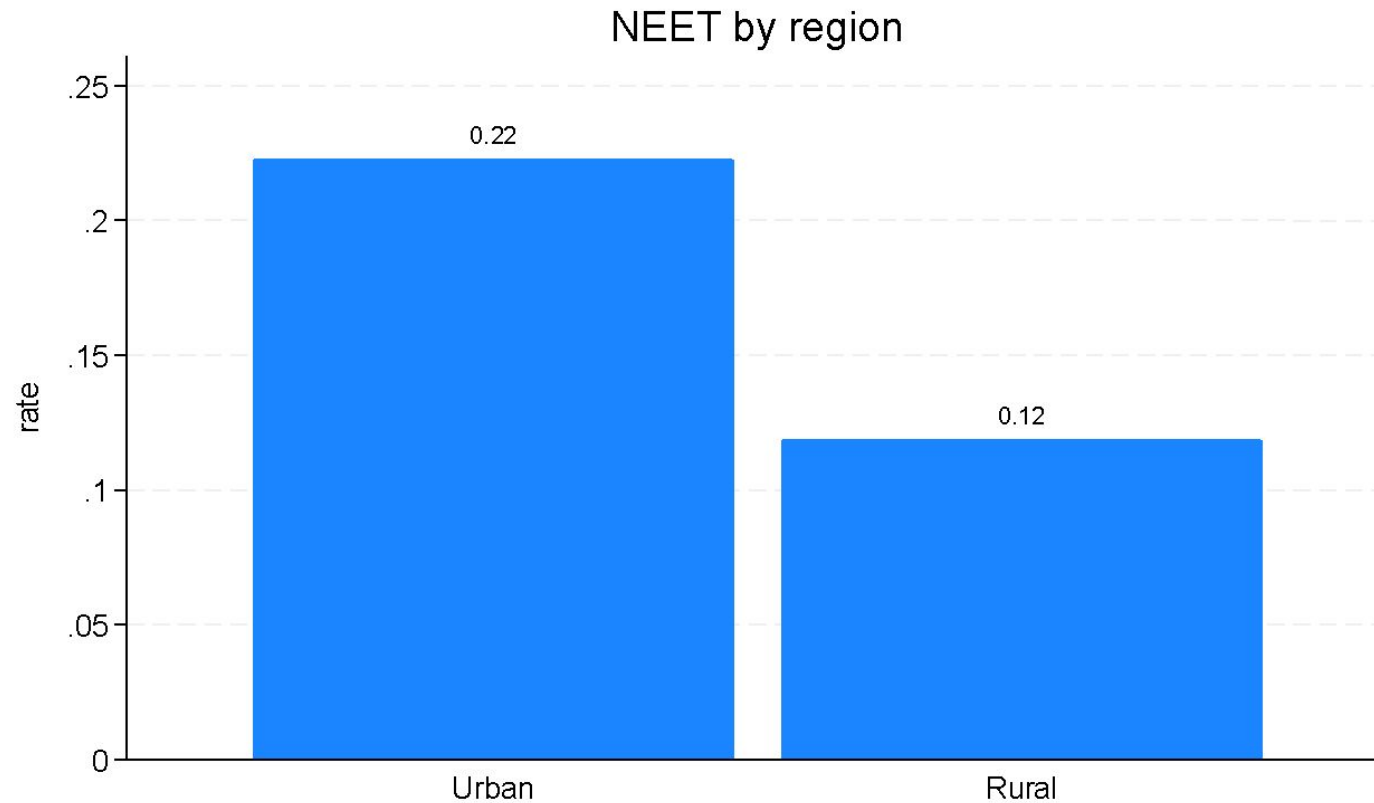
```
graph bar neet , title("NEET") ytitle("rate")
blabel(bar, format(%9.2f)) ylabel(0(.05).25)
over(rural)
```



```
graph bar neet , title("NEET") ytitle("rate")  
blabel(bar, format(%9.2f)) ylabel(0(.05).25)  
over(rural, gap(10))
```

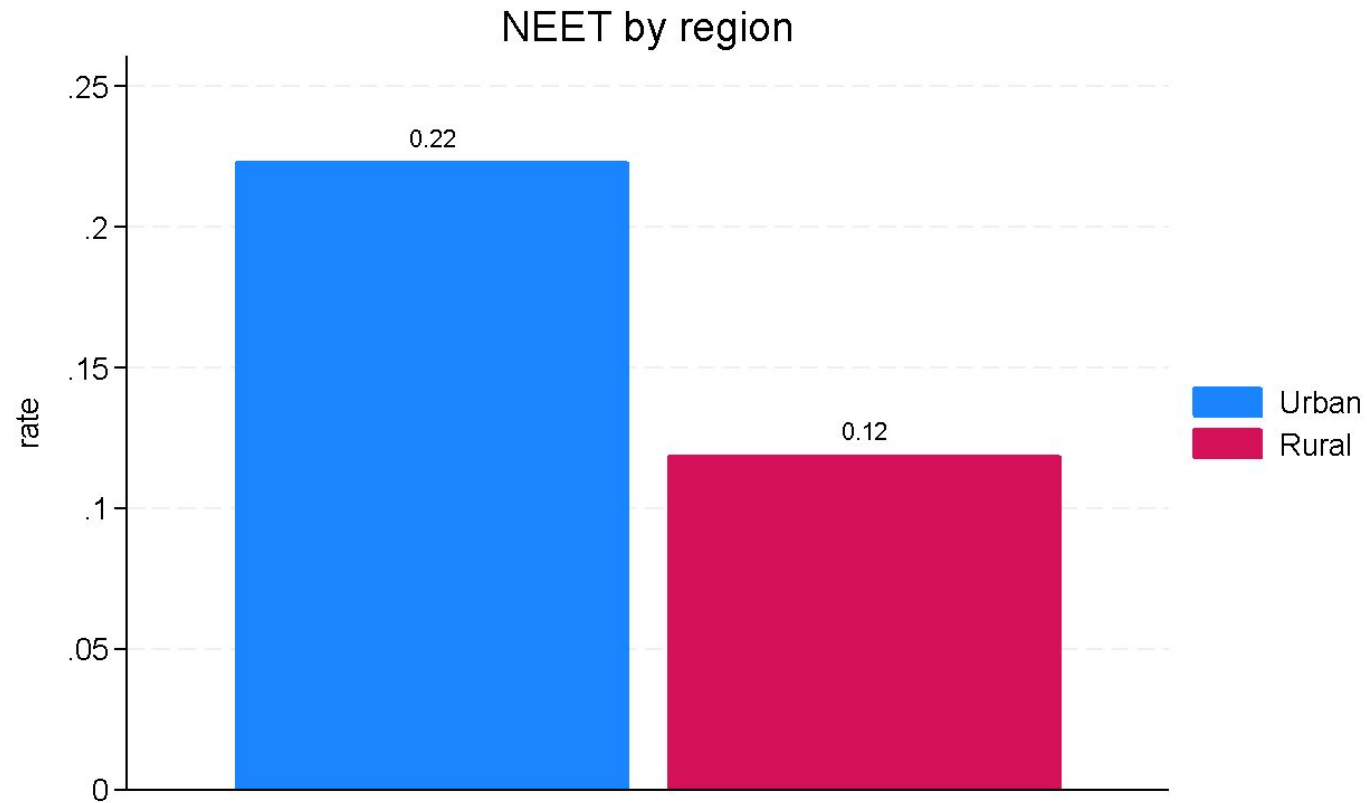


```
graph bar neet , title("NEET by region")  
ytitle("rate") blabel(bar, format(%9.2f))  
ylabel(0(.05).25) over(rural, gap(10))
```

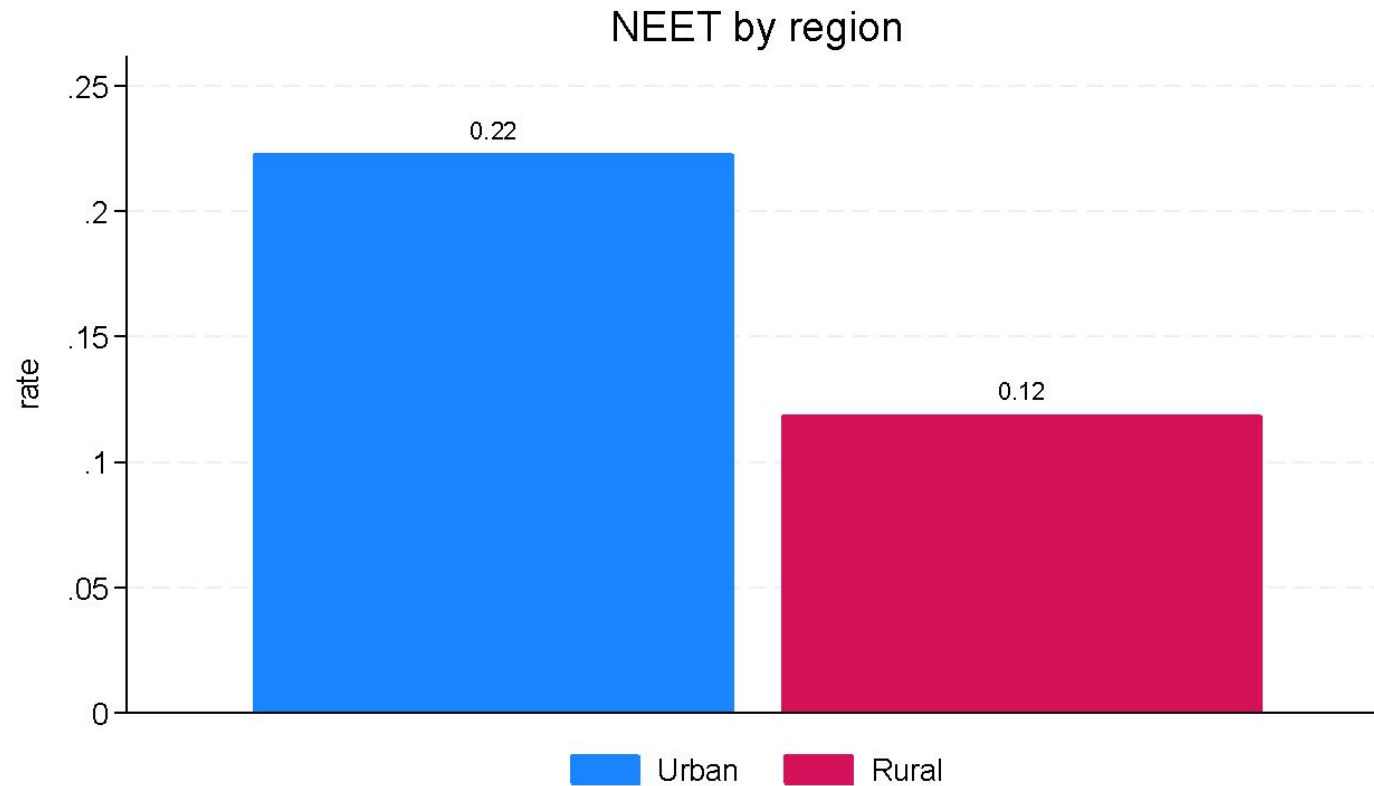


Different colors for urban and
rural

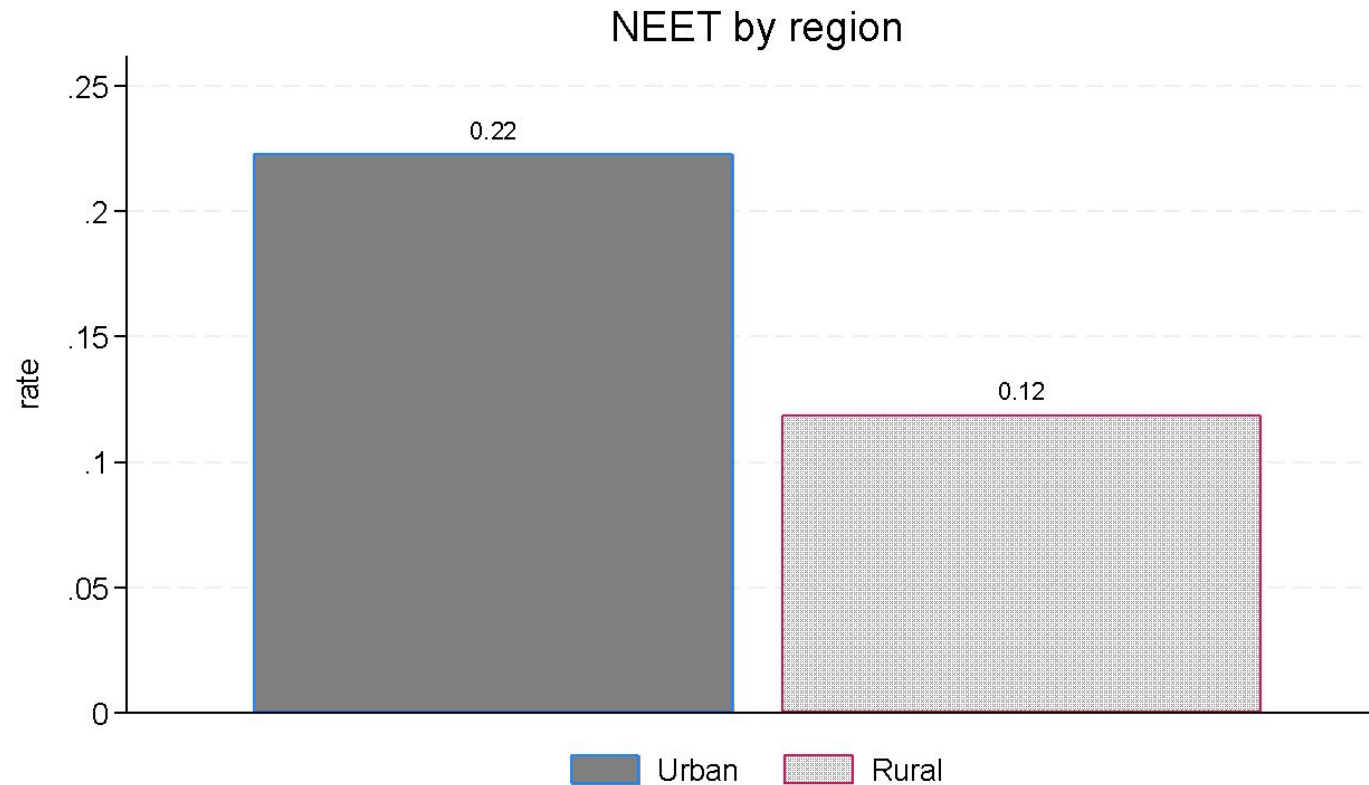
```
graph bar neet , title("NEET by region")
yttitle("rate") blabel(bar, format(%9.2f))
ylabel(0(.05).25) over(rural, gap(10)) asyvars
```



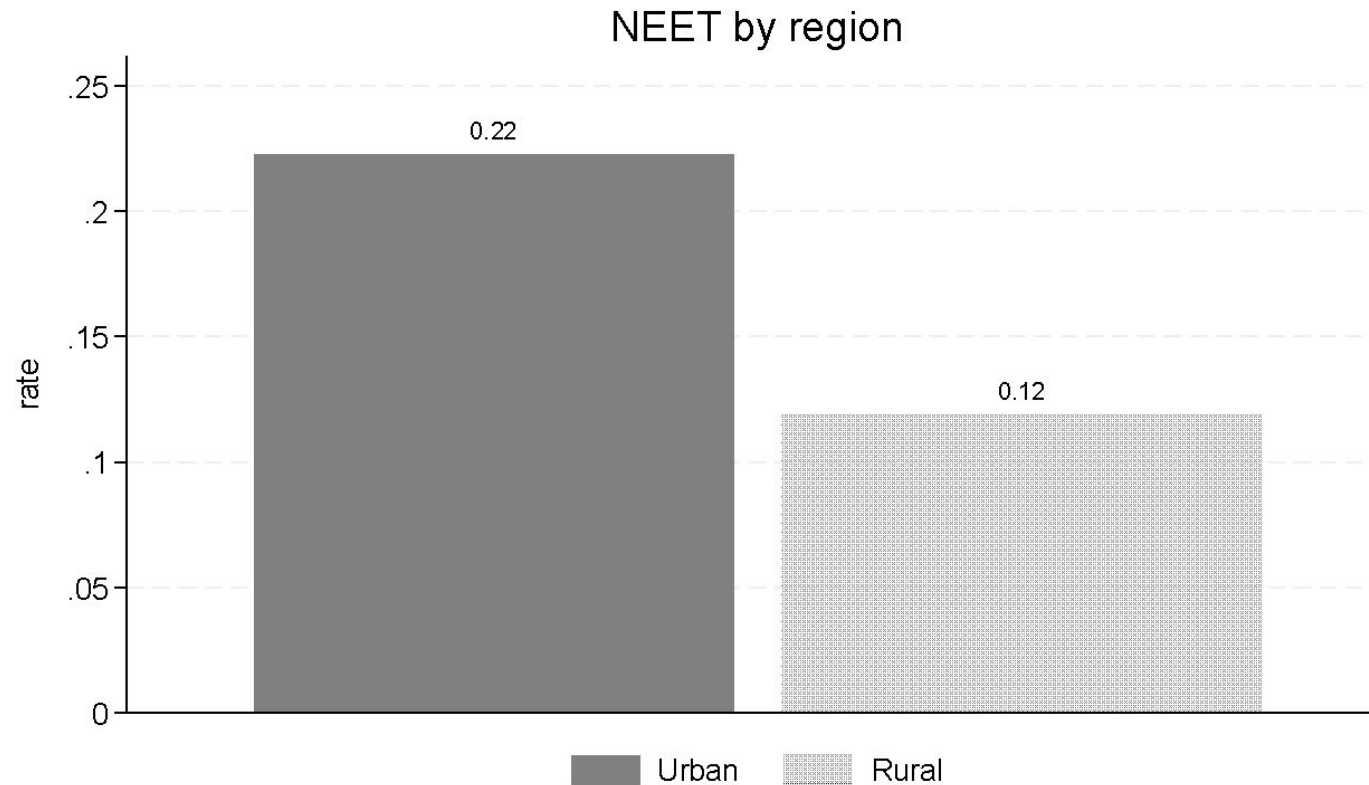
```
graph bar neet , title("NEET by region") ytitle("rate")  
blabel(bar, format(%9.2f)) ylabel(0(.05).25) over(rural,  
gap(10)) asyvars legend(pos(6) row(1) )
```



```
graph bar neet , title("NEET by region") ytitle("rate")  
blabel(bar, format(%9.2f)) ylabel(0(.05).25) over(rural,  
gap(10)) asyvars legend(pos(6) row(1) ) bar(1, fcolor(gray))  
bar(2, fcolor(gray%30))
```

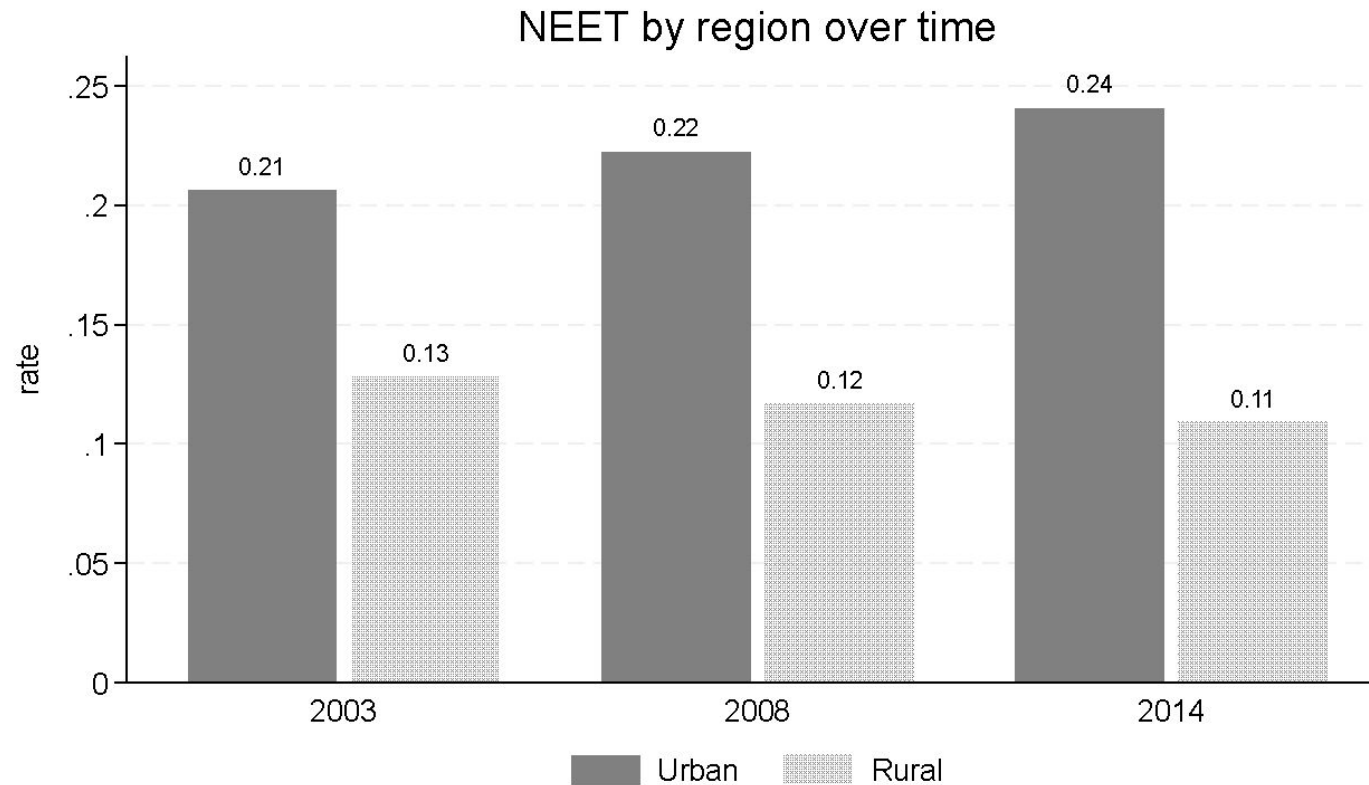


```
graph bar neet , title("NEET by region") ytitle("rate")
blabel(bar, format(%9.2f)) ylabel(0(.05).25) over(rural,
gap(10)) asyvars legend(pos(6) row(1)) bar(1, fcolor(gray)
lc(none)) bar(2, fcolor(gray%30) lc(none))
```

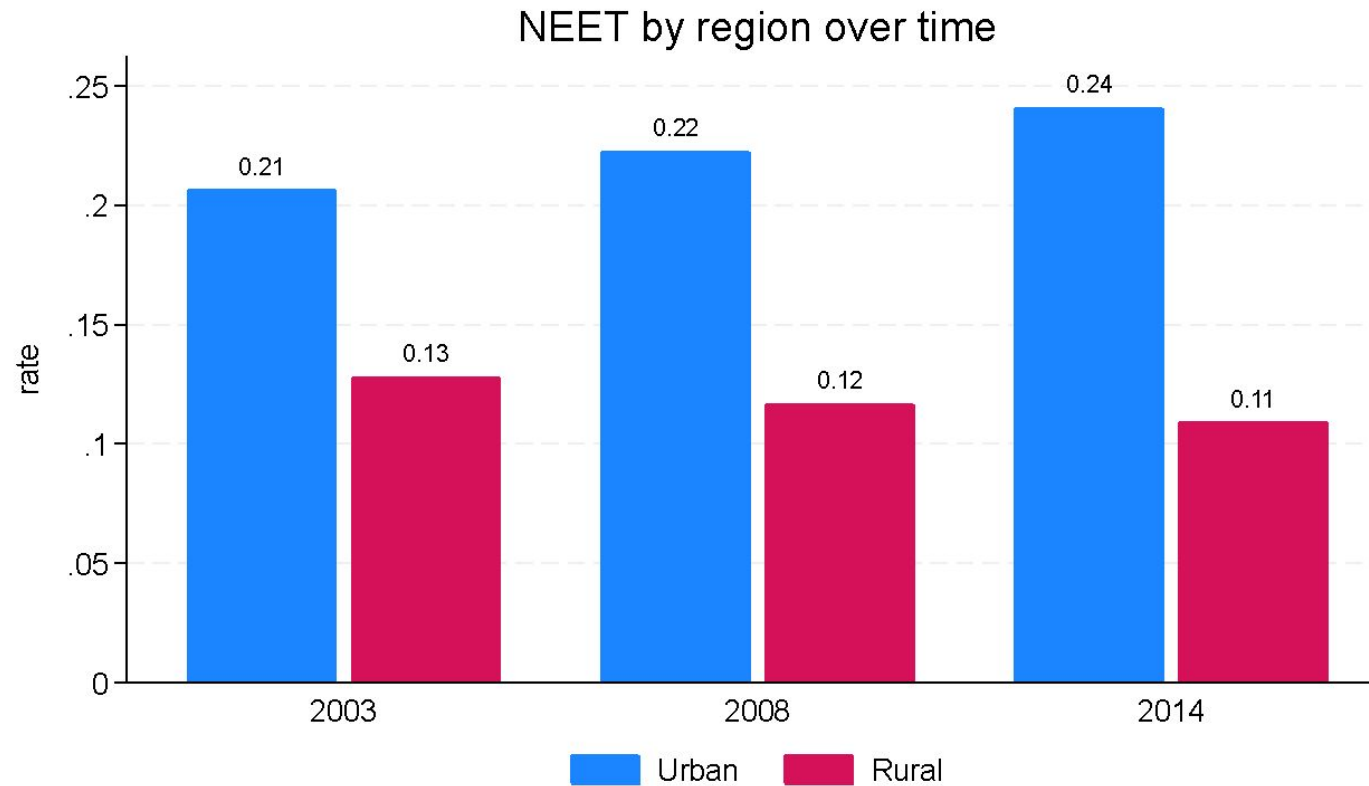


NEET rates in urban and rural
areas over time

```
graph bar neet , title("NEET by region over time") ytitle("rate")
blabel(bar, format(%9.2f)) ylabel(0(.05).25) over(rural, gap(10)) asyvars
legend(pos(6) row(1) ) bar(1, fcolor(gray) lc(none)) bar(2,
fcolor(gray%30) lc(none)) over(year)
```

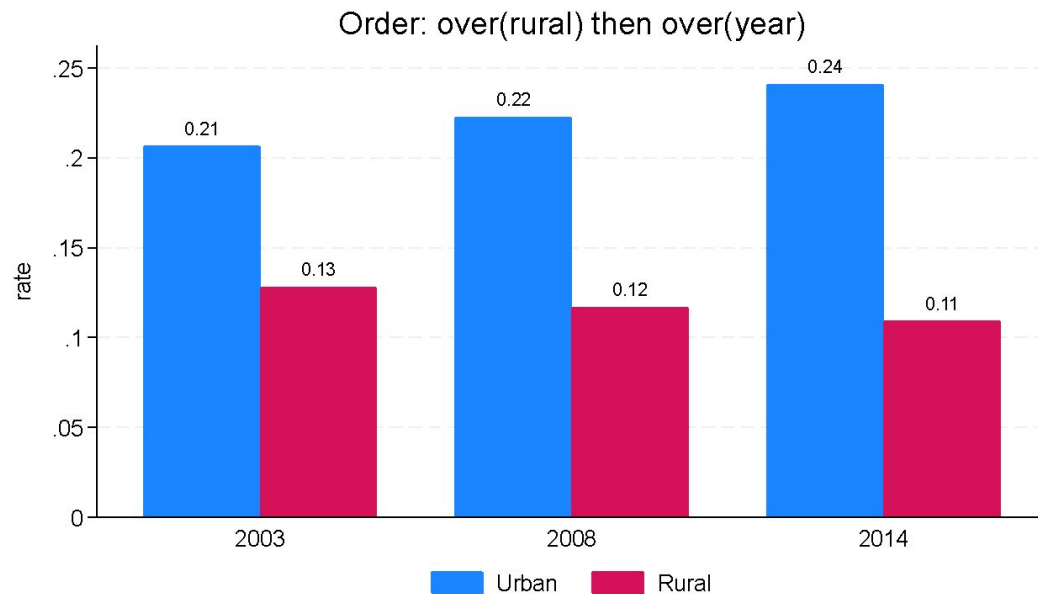


```
graph bar neet , title("NEET by region over time") ytitle("net  
enrollment") blabel(bar, format(%9.2f)) ylabel(0(.05).25)  
over(rural, gap(10)) asyvars legend(pos(6) row(1) )  
over(year)
```

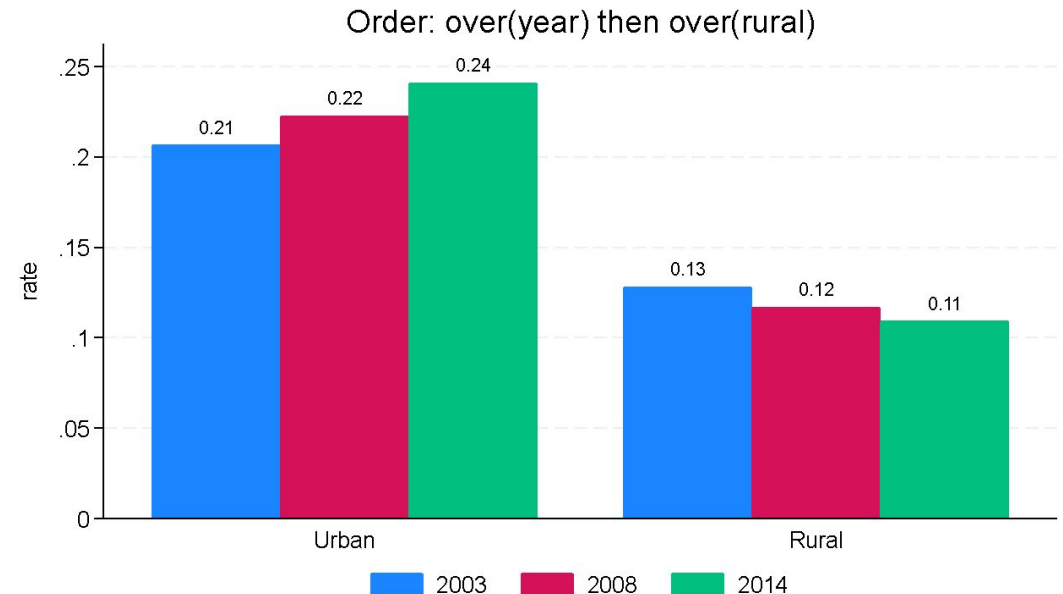


Order the 'over()' options according to what you want to present

```
graph bar neet , title("NEET enrollment by region over  
time") ytitle("rate") blabel(bar, format(%9.2f))  
ylabel(0(.05).25) asyvars legend(pos(6) row(1) )  
over(rural) over(year)
```

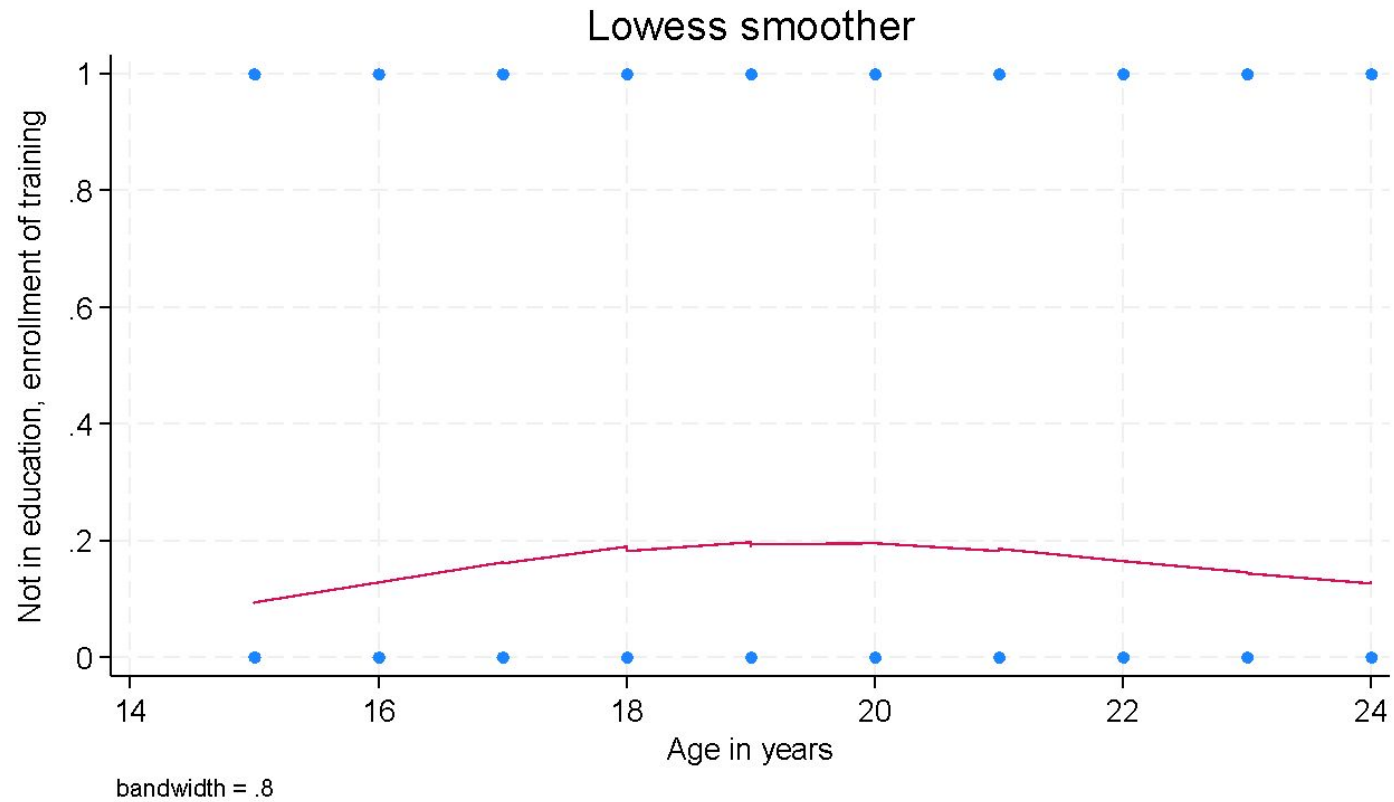


```
graph bar neet , title("NEET enrollment by region over  
time") ytitle("rate") blabel(bar, format(%9.2f))  
ylabel(0(.05).25) asyvars legend(pos(6) row(1)) over(year)  
over(rural)
```

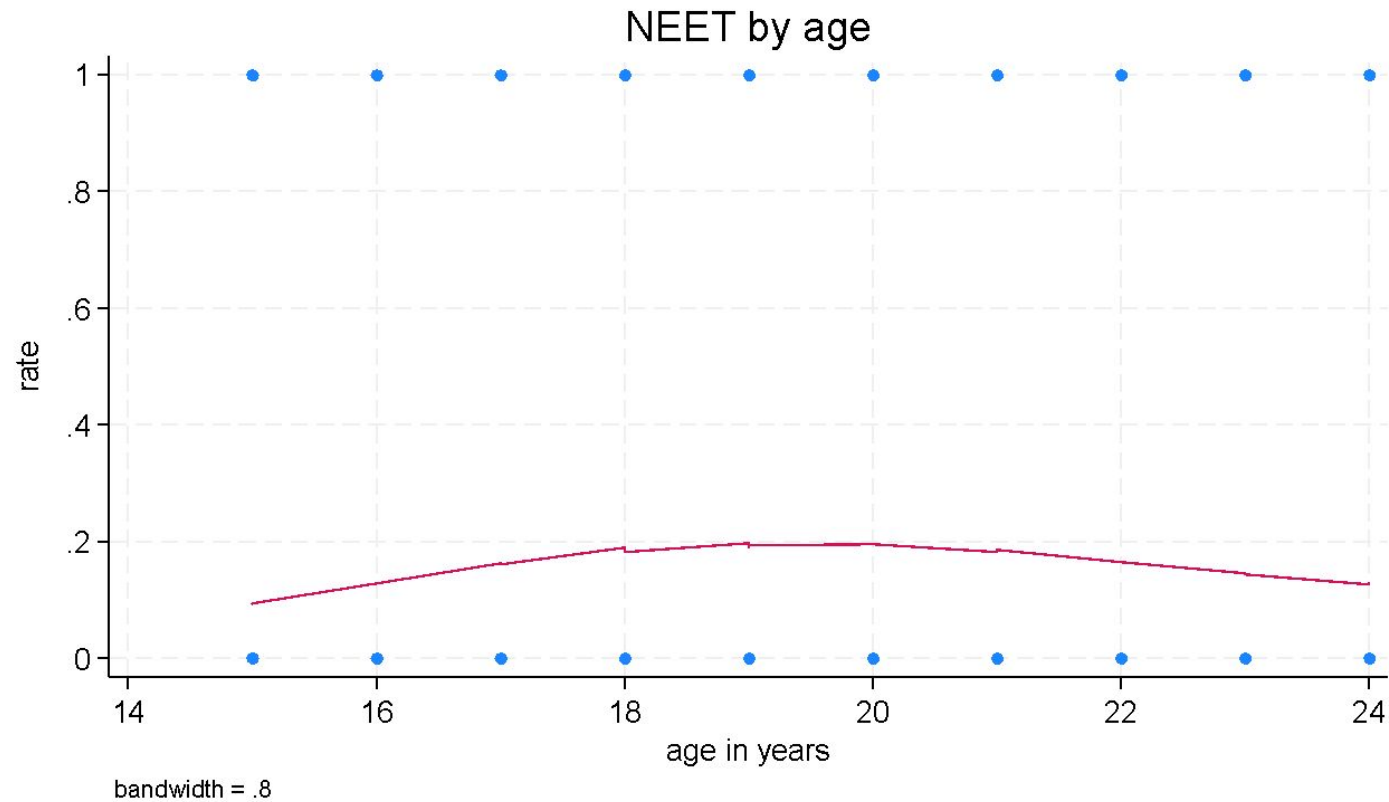


Locally Weighted Scatterplot Smoothing (lowess)

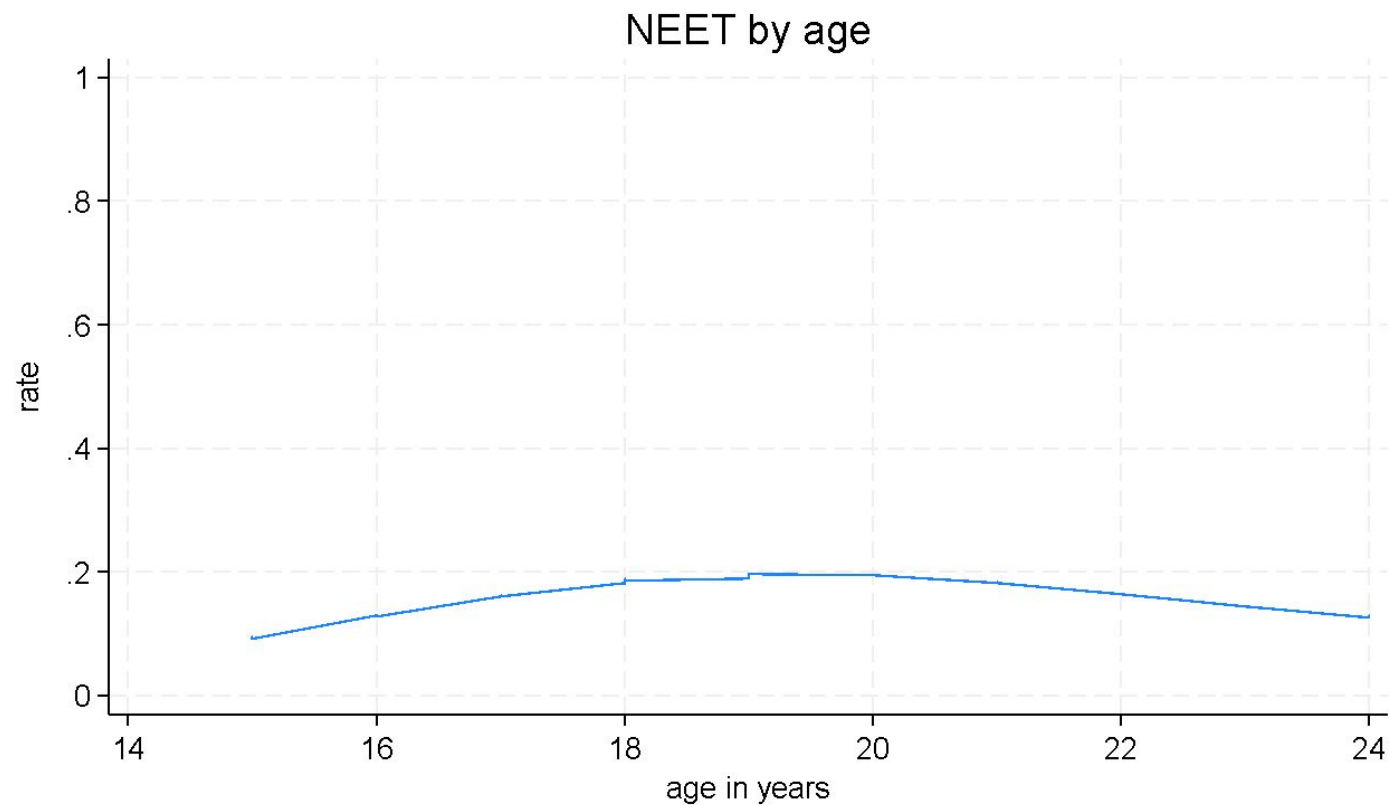
lowess neet age



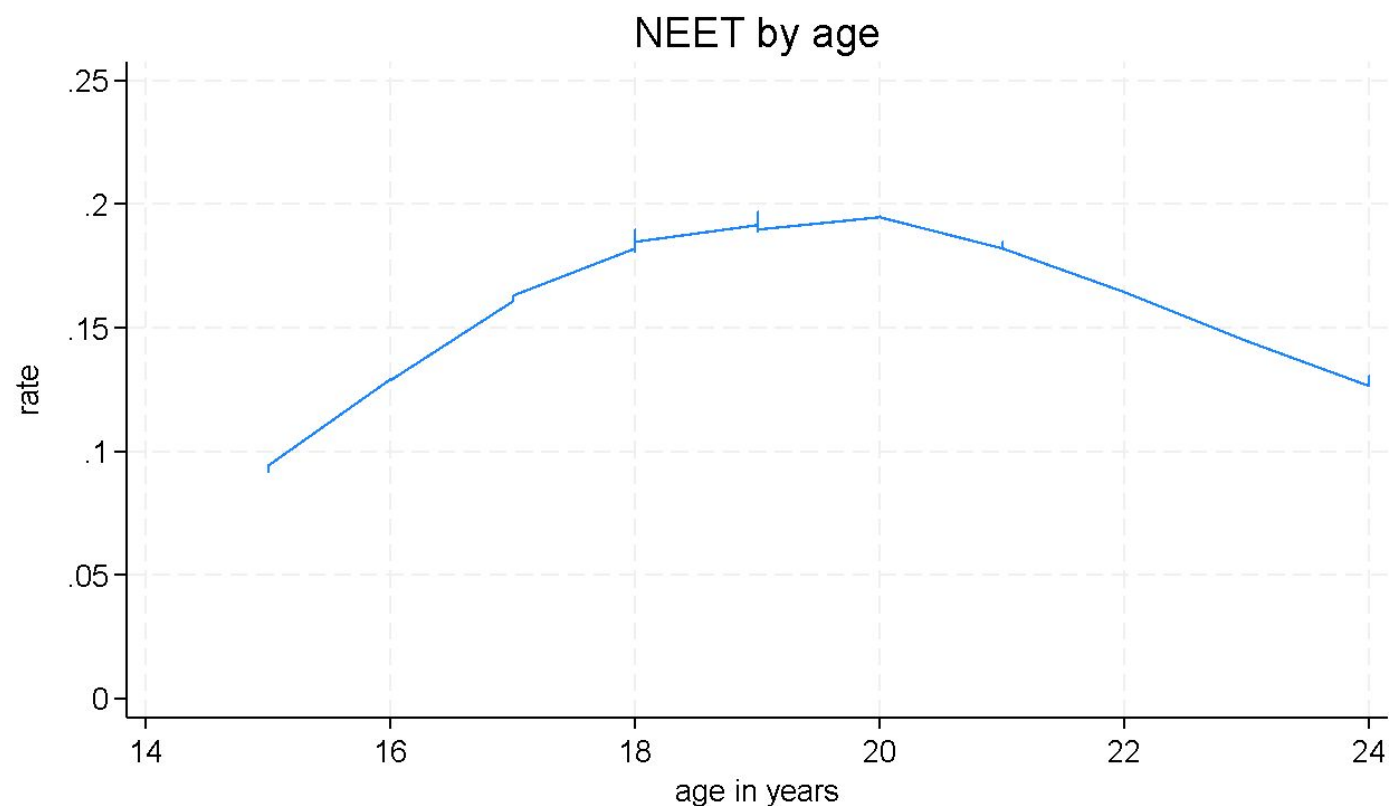
```
lowess neet age , title("NEET by age")  
ytitle("rate") xtitle("age in years")
```



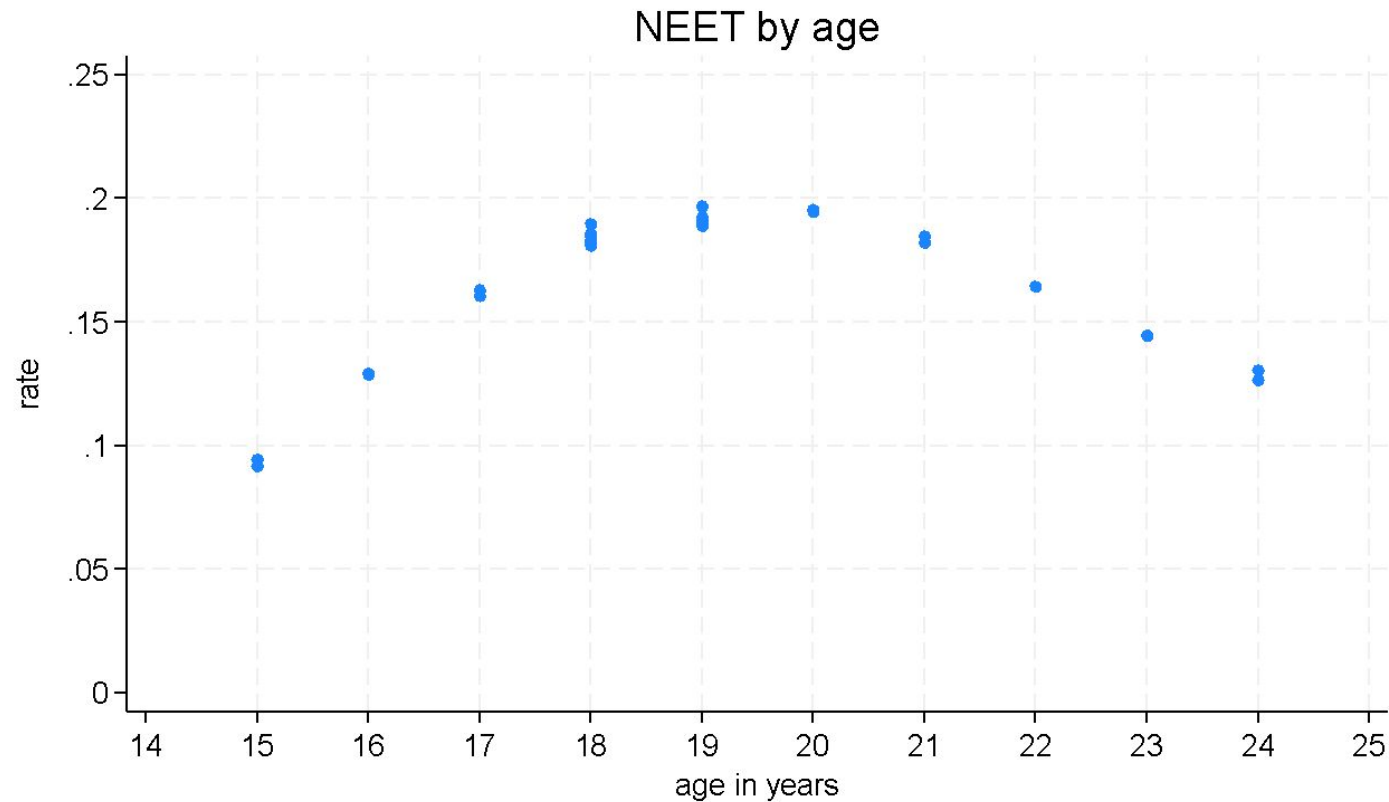
twoway lowess neet age, title("NEET by age") ytitle("rate")
xtitle("age in years") ylabel(0(0.2)1)



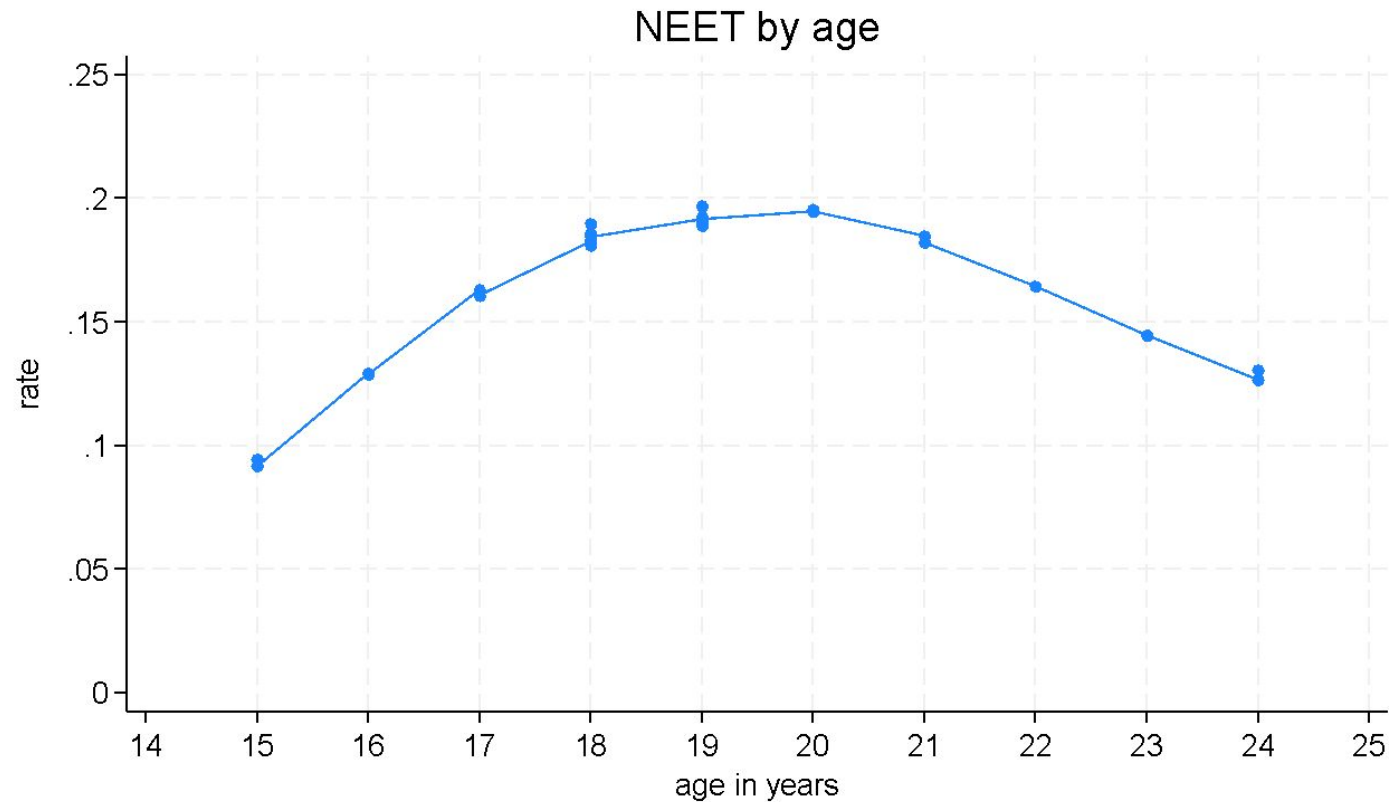
```
twoway lowess neet age , title("NEET by age") ytitle("rate")  
xtitle("age in years") ylabels(0(0.05)0.25)
```



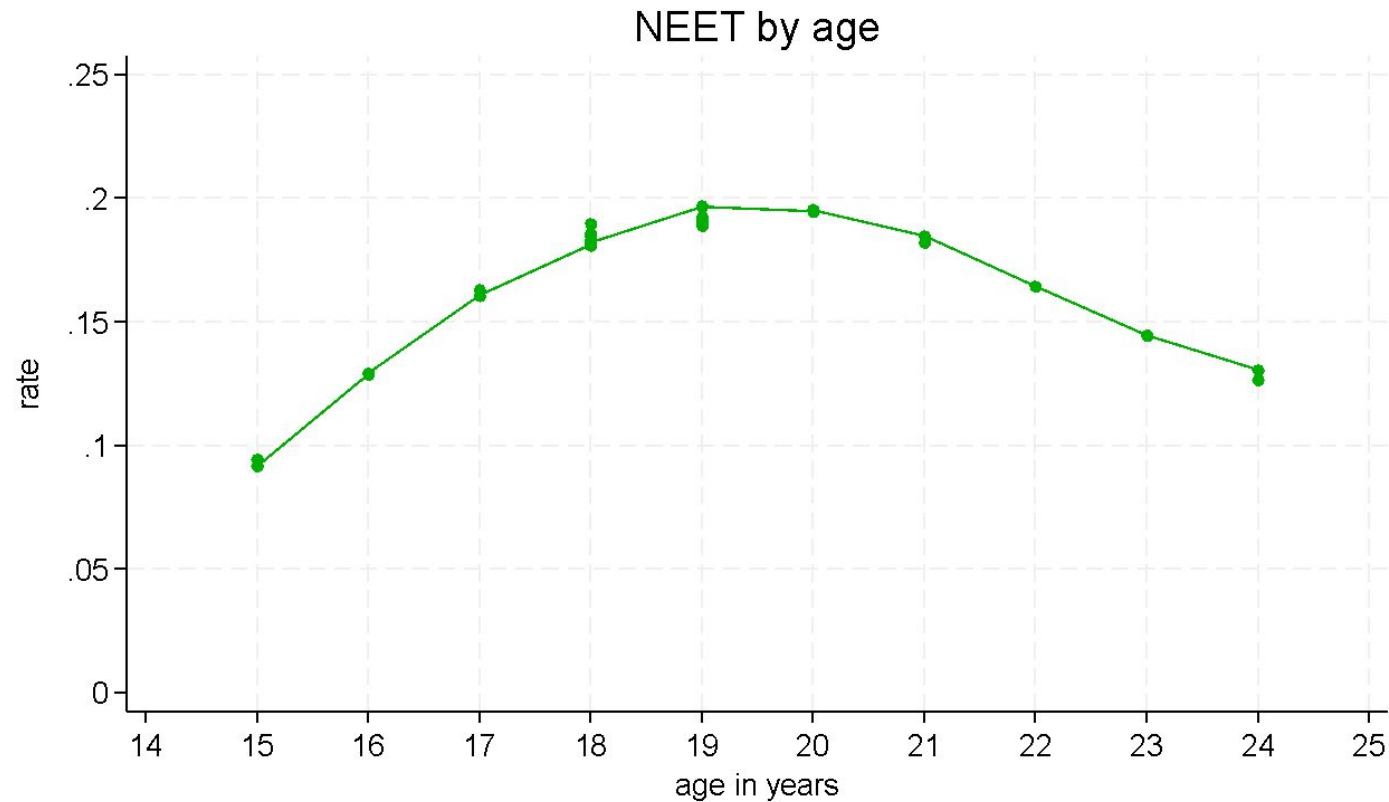
```
twoway lowess neet age , title("NEET by age") ytitle("rate")  
xtitle("age in years") xlabel(14(1)25) ylabel(0(0.05)0.25)  
recast(scatter)
```



```
twoway lowess neet age , title("NEET by age") ytitle("rate")  
xtitle("age in years") xlabels(14(1)25) ylabels(0(0.05)0.25)  
recast(scatter) connect(line)
```



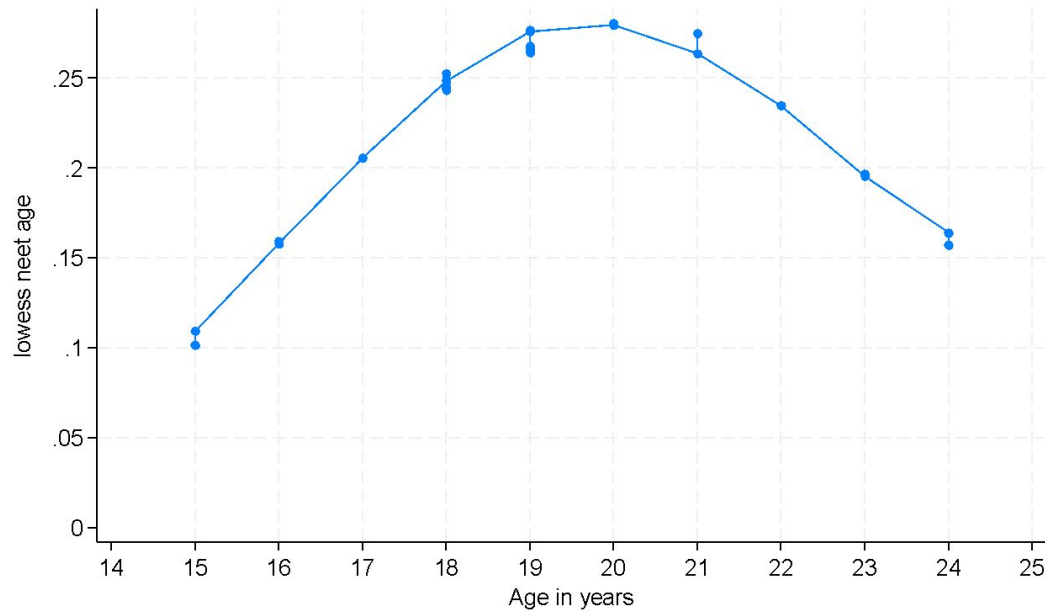
```
twoway lowess neet age , title("NEET by age") ytitle("rate")  
xtitle("age in years") xlabels(14(1)25) ylabels(0(0.05)0.25)  
recast(scatter) connect(l) color(midgreen)
```



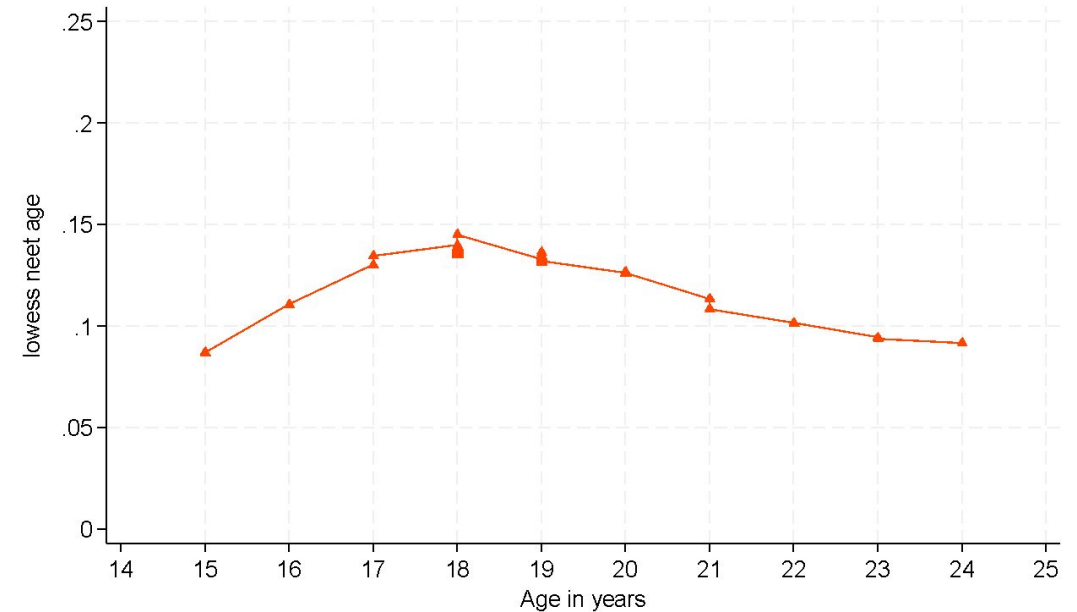
NEET by age in rural and
urban areas

Enrollment for urban on the left and rural on the right

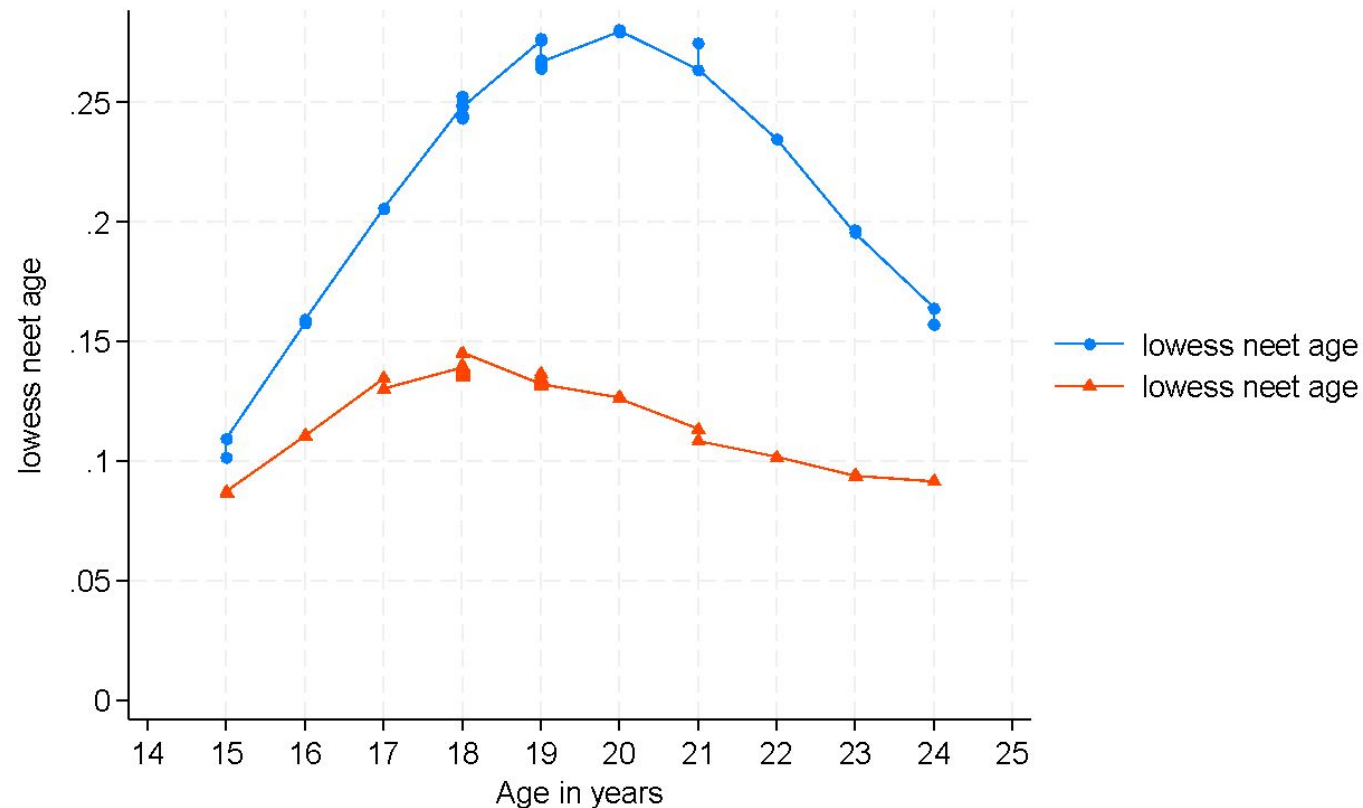
```
tw lowess neet age if rural==0,  
xlabels(14(1)25) ylabels(0(0.05)0.25)  
recast(scatter) connect(l) color(midblue)
```



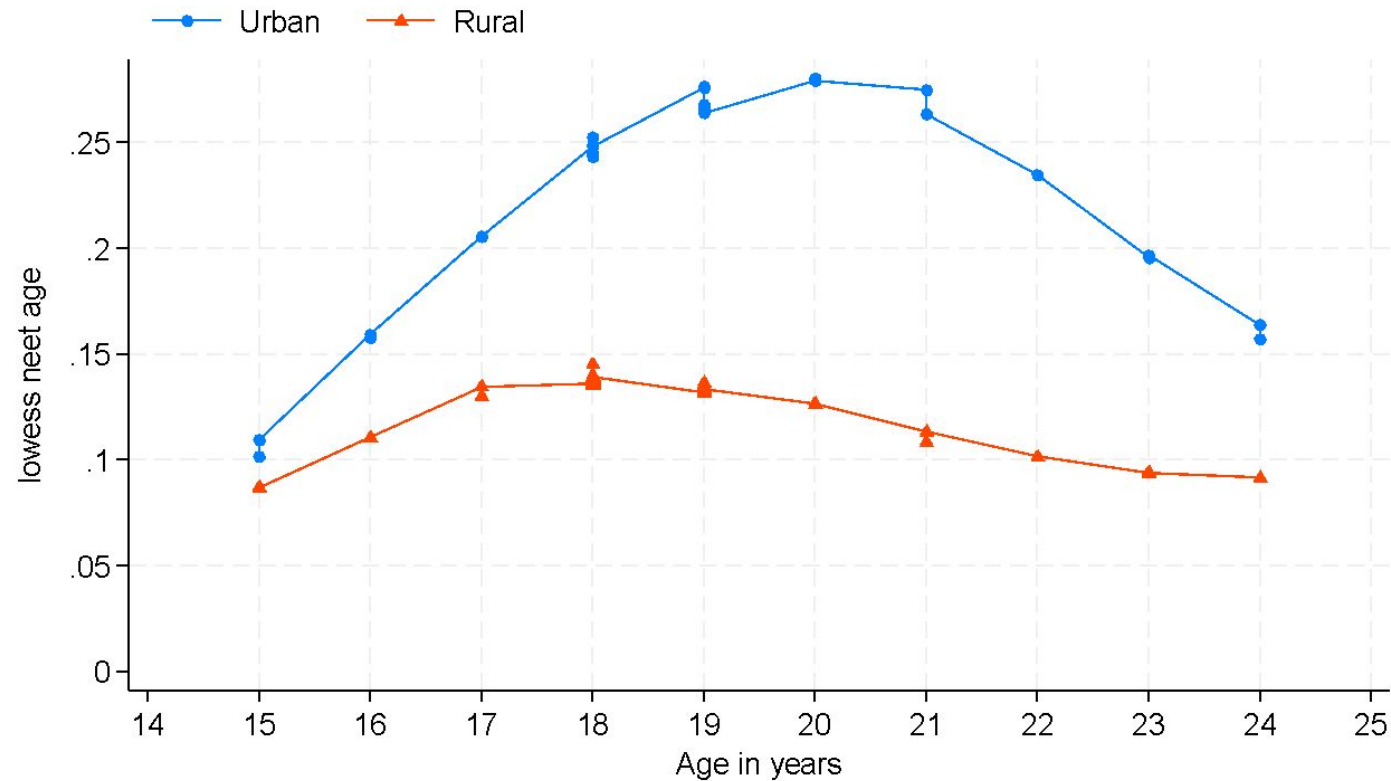
```
tw lowess neet age if rural==1,  
xlabels(14(1)25) ylabels(0(0.05)0.25)  
recast(scatter) connect(l) color(orange_red)  
msymbol(triangle)
```



```
tw lowess neet age if rural==0 , recast(scatter) connect(l)
color(midblue) || lowess neet age if rural==1 , xlabel(14(1)25)
ylabel(0(0.05)0.25) recast(scatter) connect(l) color(orange_red)
msymbol(triangle)
```



```
tw lowess neet age if rural==0 , recast(scatter) connect(l) color(midblue) ||  
lowess neet age if rural==1 , xlabel(14(1)25) ylabel(0(0.05)0.25)  
recast(scatter) connect(l) color(orange_red) ms(triangle) leg(pos(11) row(1)  
order(1 "Urban" 2 "Rural") ) name(neet_age, replace)
```



Extra slides not covered

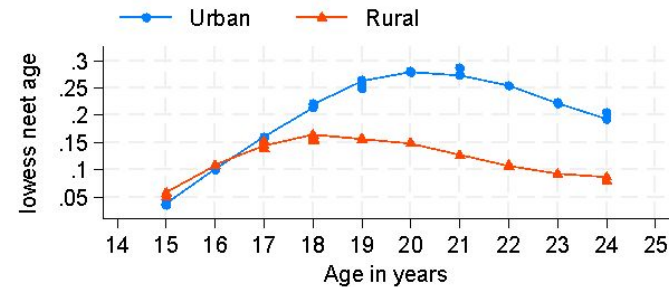
Simple loop for repeating tasks

```
levelsof year, local(years) // store values of variable year to use later
foreach y in `years' { // use the stored values of years in local year
  preserve
    keep if year == `y'
    tw lowess neet age if rural==0 , recast(scatter) connect(l)
    color(midblue) || lowess neet age if rural==1, xlabel(14(1)25)
    recast(scatter) connect(l) color(orange_red) ms(triangle) leg(pos(11)
    row(1) order(1 "Urban" 2 "Rural") ) title("`y'", span pos(11) )
    name(neet_age_`y', replace)
  restore
}
```

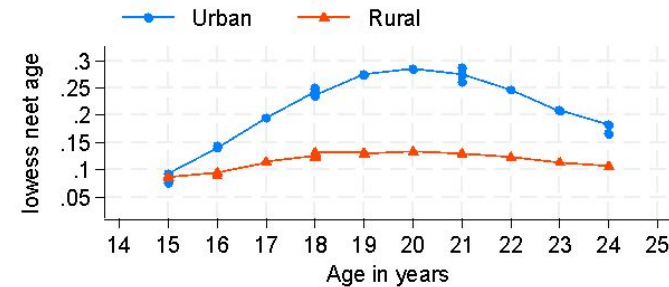
graph combine neet_age_2003 neet_age_2008
neet_age_2014, xcomm ycomm title("NEET by age in urban
and rural areas over time")

NEET by age in urban and rural areas over time

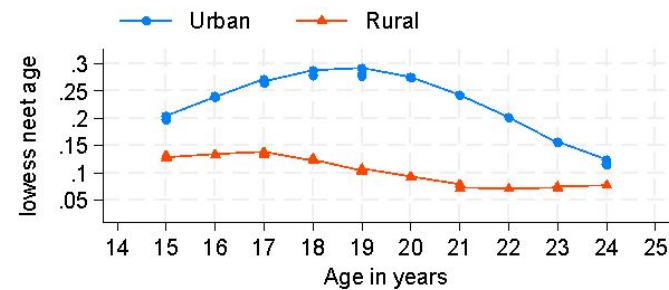
2003



2008



2014



```
grc1leg neet_age_2003 neet_age_2008 neet_age_2014,  
legendfrom(neet_age_2008) xcomm ycomm title("NEET by age in urban  
and rural areas over time") name(neet_byage_combined_1legend,  
replace)
```

NEET by age in urban and rural areas over time

