Example of poverty and inequality analysis using PIP/Statistics Online

By Gildas Bopahbe Deudibe and Minh Cong Nguyen

This document describes a simple way to conduct poverty and inequality analysis using LSMS/HBS surveys available in the Poverty and Inequality Platform's (PIP) Statistics Online (SOL) tool. SOL relies on the Global Monitoring Database (GMD), which is a harmonized data collection managed by the World Bank based on household surveys data from national statistics offices. At the moment, registered users can have access to two global collections of survey data including Global Monitoring Database (GMD) and High Frequency Phone Survey (HFPS-COVID19) in the platform. Due to limited resources, users will be able to load and run maximum five (5) notebooks, and for each session user will have 60 minutes on the platform. The session will be auto disconnected if there is no activity in 15 minutes.

In our specific example, the analysis uses the two latest publicly available household budget surveys in Tunisia from 2010 and 2015. We will focus on estimating poverty at the international poverty lines of $3.20 and $5.50 (using 2011 purchasing power parities). This example is organized according to the following table of contents:

Table of contents

- Figure 7: Poverty rate in by region
- Figure 8: Poverty rate in by area of residence

1. Additional analysis

- Figure 9: Population pyramid
- Figure 10: Growth Incidence Curve

I. Loading of data

First, we load the survey data from the system using the default system syntax, datalibweb. Then, we will append the two years of survey data (2010 and 2015) together as this is useful to analyze patterns over time (or at least between 2 points in time). Datalibweb is a standard syntax to query survey data within SOL, and the syntax is automatically populated when we select the survey data in the data catalog tree. SOL will merge on the fly with the auxiliary data of Consumer Price Index (CPI) and Purchasing Power Parities (PPP) which are used for calculating international poverty rates.

In [6]:
```
datalibweb, country(TUN) year(2015) type(GMD) module(ALL) surveyid(TUN_2015_NSHBCSL_V01_M_V04_A_GMD) filename(TUN_2015_NS

 tempfile TUN2015
 save `TUN2015'
```

```
use  "data/TUN_2015_NSHBCSL_V01_M_V04_A_GMD_ALL.dta",  clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

. quietly {
.  cap gen code = "TUN"
. gen datalevel = 2
. gen survname = "NSHBCSL"
. merge m:1 code year survname datalevel using "data/Final_CPI_PPP_to_be_used.d
> ta"
. keep if _merge==3
. drop _merge
.  }

.
.  tempfile TUN2015

.  save `TUN2015'
file /tmp/St00012.000001 saved

.
```

In [7]:
```
datalibweb, country(TUN) year(2010) type(GMD) module(ALL) surveyid(TUN_2010_NSHBCSL_V01_M_V05_A_GMD) filename(TUN_2010_NS

 append using `TUN2015' , force
```

```
use  "data/TUN_2010_NSHBCSL_V01_M_V05_A_GMD_ALL.dta",  clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

. quietly {
.  cap gen code = "TUN"
. gen datalevel = 2
. gen survname = "NSHBCSL"
. merge m:1 code year survname datalevel using "data/Final_CPI_PPP_to_be_used.d
> ta"
. keep if _merge==3
. drop _merge
.  }

.
.  append using `TUN2015' , force
(note: variable agecat was byte in the using data, but will be str1 now)
(note: variable weight_h was int, now float to accommodate using data's
       values)
(note: variable age was byte, now int to accommodate using data's values)
(label ppp_domain already defined)
(label cdlab already defined)
(label leducat7 already defined)
(label leducat5 already defined)
(label leducat4 already defined)
(label lmarital already defined)
(label lrelationharm already defined)
(label lyn already defined)
(label imp_wat_rec already defined)
(label imp_san_rec already defined)
(label male already defined)
(label urban already defined)
```

II. Data transformation

In order to derive international poverty and inequality indicators, it is essential to convert the welfare measure (income or consumption, depending on the measure and currency of welfare in the country) into a daily income/consumption per capita measured using the 2011 Purchasing Power Parity (PPP). Using the converted welfare measure, poverty headcount ratio at $3.20 and $5.50 (% of population) are calculated given the status of Tunisia as a lower middle-income country.

In [9]:
```
gen welint =welfare/cpi2011/icp2011/365

 gen headcount3_2 = (welint<3.2)*100
 label def headcount3_2 0 "Non poor" 100 "Poor"
 lab val headcount3_2 headcount3_2

 gen headcount5_5 = (welint<5.5)*100
 label def  headcount5_5 0 "Non poor" 100 "Poor"
 lab val  headcount5_5  headcount5_5


 tempfile data_all
 save `data_all'
```

```
gen welint =welfare/cpi2011/icp2011/365

.
.  gen headcount3_2 = (welint<3.2)*100

.  label def headcount3_2 0 "Non poor" 100 "Poor"

.  lab val headcount3_2 headcount3_2

.
.  gen headcount5_5 = (welint<5.5)*100

.  label def  headcount5_5 0 "Non poor" 100 "Poor"

.  lab val  headcount5_5  headcount5_5

.
.
.  tempfile data_all

.  save `data_all'
file /tmp/St00012.000002 saved
```

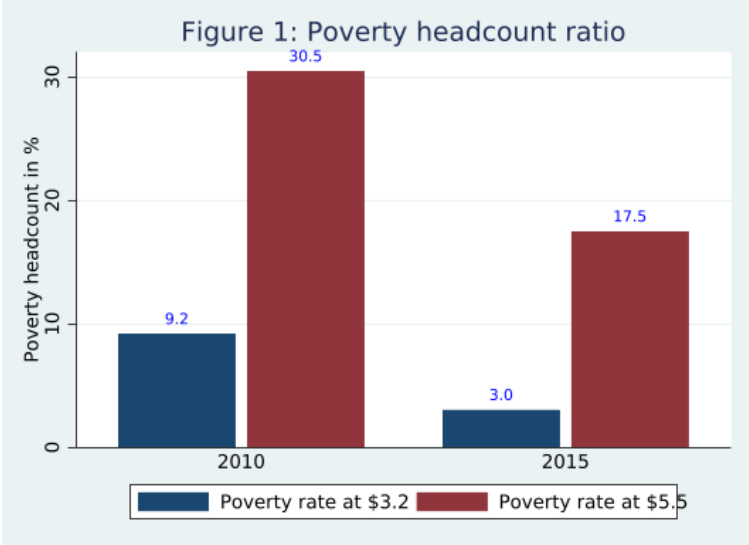
III. Overall analysis of poverty indicators

This section provides some examples to measure, describe, monitor, evaluate, and analyze poverty. It consists of graphs, tables, and explanatory text with numerous details as needed.

In [11]:
```
use `data_all' , clear
```

```
graph bar (mean) headcount3_2 headcount5_5 [aw=weight_p], over(year) bargap(10) legend(label(1 "Poverty rate at $3.2") l
graph export graph1.png, replace
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.
.  graph bar (mean) headcount3_2 headcount5_5 [aw=weight_p], over(year) bargap(
> 10) legend(label(1 "Poverty rate at $3.2") label(2 "Poverty rate at $5.5")) t
> itle("Figure 1: Poverty headcount ratio") ytitle("Poverty headcount in %") sa
> ving(graph1.gph, replace) blabel(bar, position(outside) format(%9.1f) color(b
> lue))
(note: file graph1.gph not found)
(file graph1.gph saved)

.  graph export graph1.png, replace
(note: file graph1.png not found)
(file graph1.png written in PNG format)
```



The graph above shows that the poverty rate decreases at both poverty lines namely $3.20 and $5.50. The next two graphs provide the cumulative and probability density distributions of welfare for each year to see how welfare distribution varies.

In [13]:
```
use `data_all'  , clear
cumul welint if year==2010 [aw=weight_p], gen(wel2010)
cumul welint if year==2015 [aw=weight_p], gen(wel2015)
stack wel2010 welint wel2015 welint, into(c temp) wide clear
line wel2010 wel2015 temp if temp<=50, sort saving(graph2.gph, replace) legend(label(1 "CDF 2010") label(2 "CDF 2015"))
graph export graph2.png, replace
```

```
use `data_all'  , clear
```

```
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  cumul welint if year==2010 [aw=weight_p], gen(wel2010)

.  cumul welint if year==2015 [aw=weight_p], gen(wel2015)

.  stack wel2010 welint wel2015 welint, into(c temp) wide clear

.  line wel2010 wel2015 temp if temp<=50, sort saving(graph2.gph, replace) lege
> nd(label(1 "CDF 2010") label(2 "CDF 2015")) title("Figure 2: Cumulative densi
> ty function") ytitle("CDF") xtitle("per capita welfare PPP$2011")
(note: file graph2.gph not found)
(file graph2.gph saved)

.  graph export graph2.png, replace
```

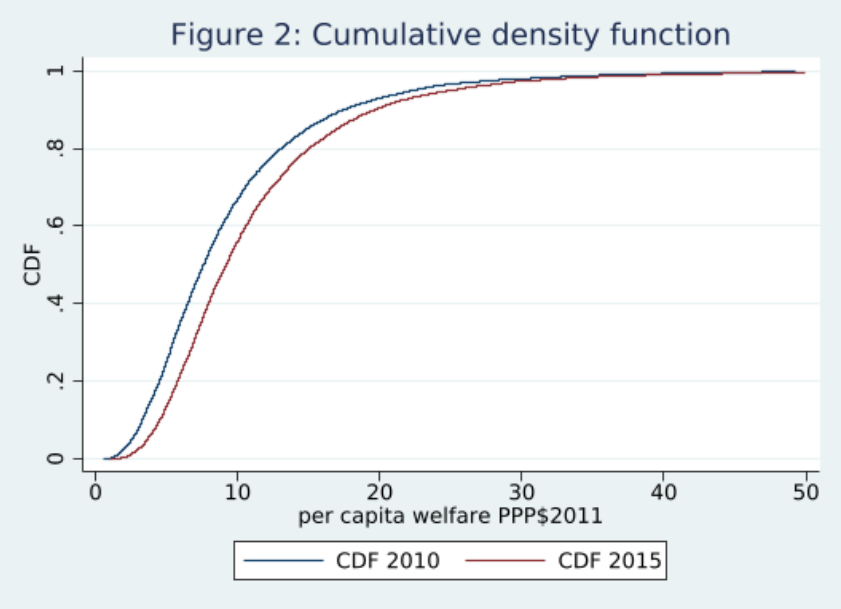
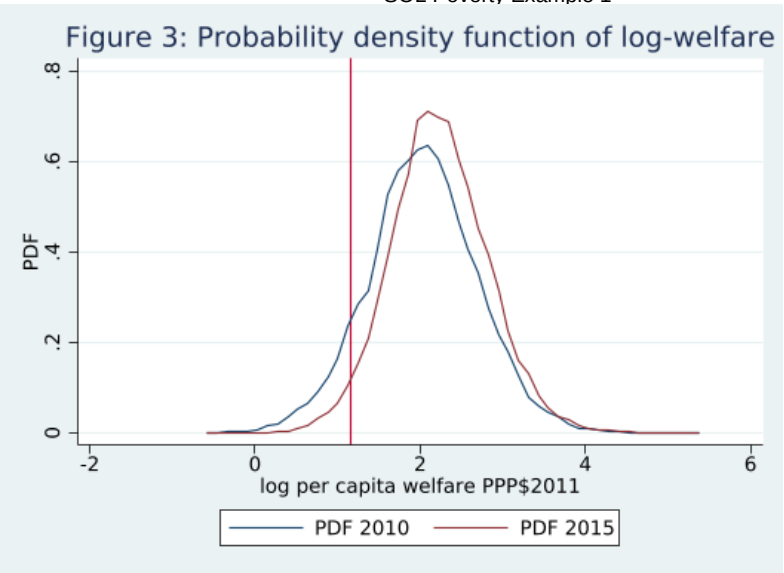


Figure 2 indicates that household per capita welfare distribution in 2015 stochastically dominates that of 2010, meaning that the poverty rate in 2015 is lower at any given poverty line.

In [15]:
```
use   `data_all'   ,   clear
 gen log_wel = log(welint)
 kdensity log_wel [aw=weight_p], nograph generate(x fx)
 kdensity log_wel if year==2010 [aw=weight_p], nograph generate(fx0) at(x)
 kdensity log_wel if year==2015 [aw=weight_p], nograph generate(fx1) at(x)
 label var fx0 "2010 density estimate"
 label var fx1 "2015 density estimate"
 line fx0 fx1 x, xline(1.1631508) sort saving(graph3.gph, replace) legend(label(1 "PDF 2010") label(2 "PDF 2015")) title(
 graph export graph3.png, replace
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  gen log_wel = log(welint)

.  kdensity log_wel [aw=weight_p], nograph generate(x fx)

.  kdensity log_wel if year==2010 [aw=weight_p], nograph generate(fx0) at(x)

.  kdensity log_wel if year==2015 [aw=weight_p], nograph generate(fx1) at(x)

.  label var fx0 "2010 density estimate"

.  label var fx1 "2015 density estimate"

.  line fx0 fx1 x, xline(1.1631508) sort saving(graph3.gph, replace) legend(lab
> el(1 "PDF 2010") label(2 "PDF 2015")) title("Figure 3: Probability density fu
> nction of log-welfare") ytitle("PDF") xtitle("log per capita welfare PPP$2011
> ")
(note: file graph3.gph not found)
(file graph3.gph saved)

.  graph export graph3.png, replace
(note: file graph3.png not found)
(file graph3.png written in PNG format)
```



Finally, the fat left tail of the welfare distribution in 2010, and the large mass of households that live on slightly less than the $3.20 (the line is represented by the red in log scale = 1.1631508) demonstrates the poverty reduction between 2010 and 2015. The analysis so far has only considered headcount poverty rates, which fail to consider how far poor households lie below the poverty line. One measure of the intensity of poverty is poverty gap which is defined as the average poverty gap in the population as a proportion of the poverty line. This ratio shows the average shortfall of the total population from the poverty line and is an important input to derive the minimum level of income required to secure the basic necessities or move out poor households above the poverty line.

There is another indicator that measures the severity of poverty in a population: the poverty severity index. It is calculated by averaging the square of the poverty gap ratio. This measure puts more weight the further a poor person's observed income falls below the poverty line.

These indicators complement the headcount ratio and are known as FGT indicators, which are a family of poverty metrics. The formula for these measures is provided as follow:

$$FGT_\alpha = \frac{1}{N}[\sum_{i=1}^{H}(\frac{z - y_i}{z})^\alpha]$$

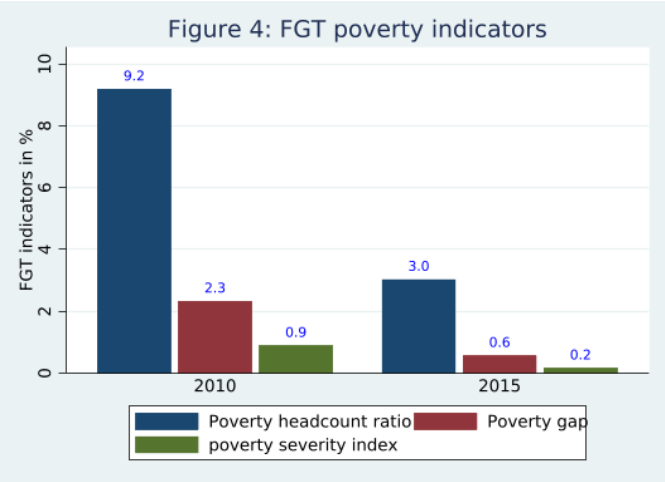
where z is the poverty threshold, N is the number of individuals in the survey (economy), H is the number of poor (those with incomes at or below z), y_i is the income of each individual i. The higher the value of α, the greater the weight place on the poorest individuals.

The graph below provides estimates of FGT indicators including poverty headcount ratio that we already showed in Figure 1, using the $3.2 poverty line.

In [17]:
```
use `data_all' , clear
 gen zref =3.2
 gen p0=100*(welint<zref) if welint!=.
 gen  p1=((zref-welint)/zref)*p0  if  welint!=.
 gen p2=(((zref-welint)/zref)^2)*p0 if welint!=.

 graph bar (mean) p0 p1 p2 [aw=weight_p], over(year) bargap(10) legend(label(1 "Poverty headcount ratio") label(2 "Povert
 graph export graph4.png, replace
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  gen zref =3.2

.  gen p0=100*(welint<zref) if welint!=.

.  gen p1=((zref-welint)/zref)*p0 if welint!=.

.  gen p2=(((zref-welint)/zref)^2)*p0 if welint!=.
```

```
.
.  graph bar (mean) p0 p1 p2 [aw=weight_p], over(year) bargap(10) legend(label(
> 1 "Poverty headcount ratio") label(2 "Poverty gap") label(3 "poverty severity
>  index")) title("Figure 4: FGT poverty indicators") ytitle("FGT indicators in
>  %") saving(graph4.gph, replace) blabel(bar, position(outside) format(%9.1f)
> color(blue))
(note: file graph4.gph not found)
(file graph4.gph saved)

.  graph export graph4.png, replace
(note: file graph4.png not found)
(file graph4.png written in PNG format)
```



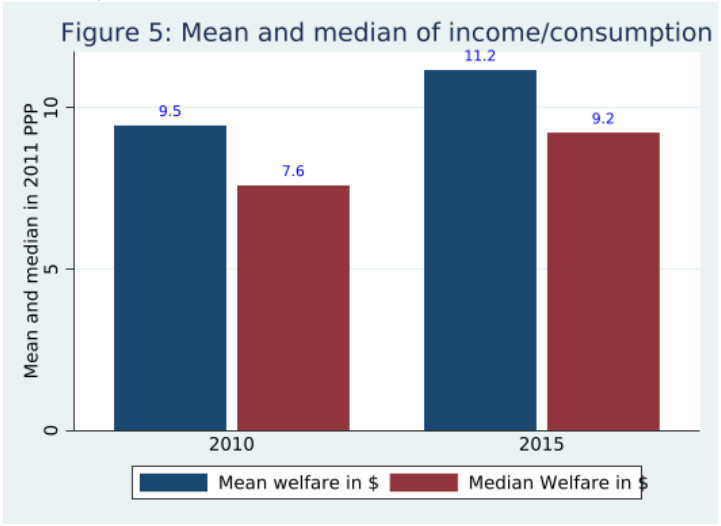
The above graph not only show a reduction in the poverty rate, but also in its intensity. This means that poverty alleviation between 2010 and 2015 have benefited the poorest households as well. The next graphs will demonstrate by how much poorer households have benefited from such a reduction using various inequality measures. One basic inequality measure is to compare mean and median income (or welfare in general).

4. Overall analysis of Inequality

In [19]:
```
graph bar (mean) welint (median) welint [aw=weight_p], over(year) bargap(10) legend(label(1 "Mean welfare in $") label(2
 graph export graph6.png, replace
```

```
graph bar (mean) welint (median) welint [aw=weight_p], over(year) bargap(10)
> legend(label(1 "Mean welfare in $") label(2 "Median Welfare in $")) title("Fi
> gure 5: Mean and median of income/consumption") ytitle("Mean and median in 20
> 11 PPP") saving(graph6.gph, replace) blabel(bar, position(outside) format(%9.
> 1f) color(blue))
(note: file graph6.gph not found)
(file graph6.gph saved)
```

```
.  graph export graph6.png, replace
(note: file graph6.png not found)
(file graph6.png written in PNG format)
```



Generally, when a mean income tends to be much greater than median income, inequality tend to be higher (we will confirm how inequality varies in next graph/tables with more rigorous inequality measures). In this case, median income grew much more than mean income from 2010 to 2015, suggesting inequality reduction. Let's check using the Gini index in the next table.

Table 1: Gini Index

In [21]:
```
ginidesc welfare , by(year) m(a1) gk(a2)
```

```
ginidesc welfare , by(year) m(a1) gk(a2)

Pyatt's Inequality decomposition

---------------------------------
Desig and |
Grupos    |      Indice        Part.
----------+----------------------
DESCOMP.  |
Between   |       0.071       21.341
Overlap   |       0.076       22.870
Within    |       0.185       55.789
----------+----------------------
TOTAL     |
Gini      |       0.332      100.000
-----------------------------------------
Stored in matrix a1

Gini Coefficient by subgroups
of year

----------------------
        K |     Gini_k
 ---------+-----------
     2010 |      0.350
     2015 |      0.310
----------------------
Stored in matrix a2

.
```

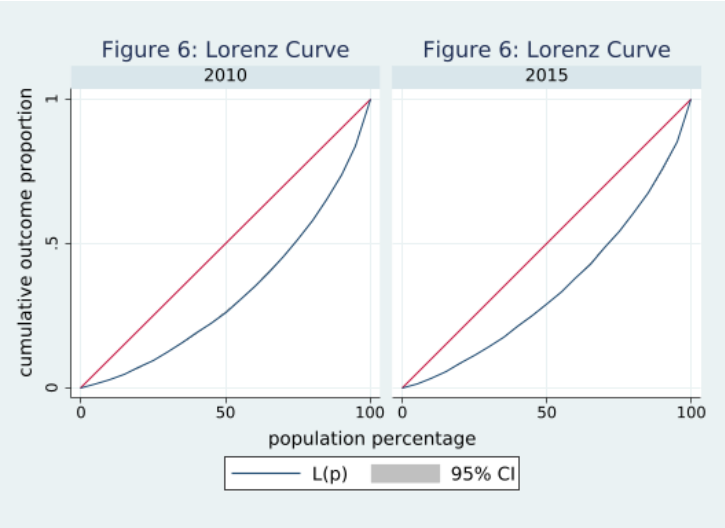
From the table above, the Gini index decreases from 0.35 in 2010 to 0.31 in 2015, meaning reduction in inequality. Although the difference is small and couldn't probably be identified from a graph, let's plot a Lorenz curve for each year in the next figures.

In [23]:
```
use `data_all'  , clear
 lorenz estimate welint , over(year)
 lorenz graph, aspectratio(1) xlabel(, grid) title("Figure 6: Lorenz Curve") saving(graph5.gph, replace)
 graph export graph5.png, replace
```

```
use `data_all'  , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  lorenz estimate welint , over(year)
```

```
L(p)                            Number of obs    =    155,428

         2010: year = 2010
         2015: year = 2015

------------------------------------------------------------------------------
      welint |      Coef.   Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------------------------
2010         |
           0 |          0  (omitted)
           5 |   .0108397   .0000783       .0106863    .0109931
          10 |   .0270237   .0001531       .0267237    .0273238
          15 |   .0465667   .0002225       .0461307    .0470028
          20 |   .0692854    .000305       .0686877    .0698831
          25 |   .0951009   .0003862        .094344    .0958578
          30 |   .1235025    .000459       .1226028    .1244022
          35 |   .1543544   .0005338       .1533083    .1554006
          40 |   .1877577   .0006097       .1865627    .1889527
          45 |   .2238978   .0006895       .2225465    .2252492
          50 |    .262906   .0007676       .2614014    .2644105
          55 |   .3051071    .000847       .3034469    .3067673
          60 |   .3508059   .0009255        .348992    .3526198
          65 |    .400255   .0010017       .3982917    .4022183
          70 |      .4539   .0010743       .4517943    .4560057
          75 |   .5125124   .0011409       .5102762    .5147486
          80 |   .5773638   .0011918       .5750279    .5796996
          85 |   .6506229   .0012267       .6482186    .6530273
          90 |   .7351357   .0012322       .7327205    .7375508
          95 |   .8372479   .0011579       .8349784    .8395173
         100 |          1           .              .           .
-------------+----------------------------------------------------------------
2015         |
           0 |          0  (omitted)
           5 |   .0142353   .0000587       .0141202    .0143504
          10 |   .0338775     .00011       .0336619    .0340931
          15 |    .056966   .0001596       .0566532    .0572788
          20 |   .0828212   .0002089       .0824118    .0832306
          25 |   .1112339   .0002583       .1107275    .1117402
          30 |   .1420172   .0003076       .1414143    .1426201
          35 |   .1752825   .0003596       .1745776    .1759875
          40 |   .2109916   .0004105       .2101871    .2117962
          45 |   .2491723   .0004615       .2482679    .2500768
          50 |   .2898329   .0005121       .2888291    .2908366
          55 |   .3332458   .0005626       .3321431    .3343485
          60 |   .3797344   .0006127       .3785336    .3809352
          65 |   .4294792   .0006609        .428184    .4307745
          70 |   .4829872   .0007069       .4816017    .4843728
          75 |   .5408506   .0007491       .5393825    .5423188
```

```
          80 |    .6042865    .0007826       .6027527     .6058204
          85 |    .6748033    .0008096       .6732166       .67639
          90 |    .7551522    .0008142       .7535565      .756748
          95 |    .8508476     .000778       .8493227     .8523725
         100 |           1           .              .            .
-------------+------------------------------------------------------------

.  lorenz graph, aspectratio(1) xlabel(, grid) title("Figure 6: Lorenz Curve")
> saving(graph5.gph, replace)
(note: file graph5.gph not found)
(file graph5.gph saved)

.  graph export graph5.png, replace
(note: file graph5.png not found)
(file graph5.png written in PNG format)
```



From the Lorenz curve above, it is noticeable that the area between the Lorenz curve and the diagonal line shrunk, suggesting reduction in income inequality.

5. Analysis of poverty measure by geographical attributes

5.1. Poverty by region (first level geographical desegregation)

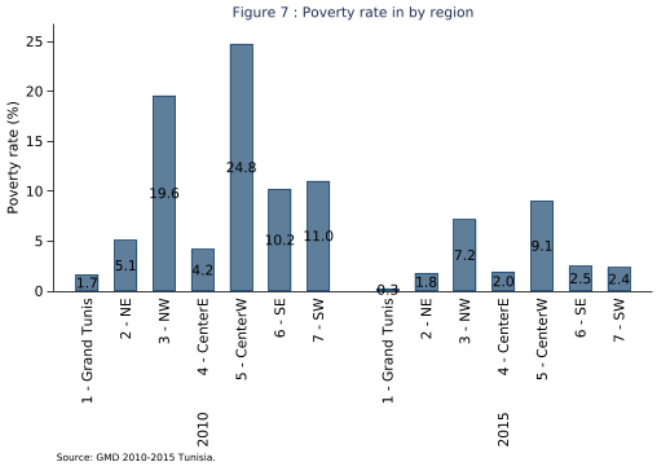
This paragraph provides poverty estimates, particularly headcount ratio, by geographical attributes such as province/region, areas of

residence, etc.

In [27]:
```
use `data_all' , clear
 graph bar headcount3_2 [aw=weight_p],  inten(*0.7) o(subnatid, lab(labsi(small) angle(vertical))) o (year, lab(labsi(sma
 graph export graph10.png, replace
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  graph bar headcount3_2 [aw=weight_p],  inten(*0.7) o(subnatid, lab(labsi(sma
> ll) angle(vertical))) o (year, lab(labsi(small) angle(vertical))) blab(bar, p
> os(center) format(%3.1f) size(8pt)) ti(Figure 7 : Poverty rate in by region,
> si(small)) note("Source: GMD 2010-2015 Tunisia. ", si(*.7)) graphregion(c(white))
> ysize(6) xsize(6.5) legend(si(vsmall) r(3)) yti("Poverty rate (%)", si(small))
>  ylab(,labs(small) nogrid angle(0)) scheme(s2color) saving(graph10.gph,
>  replace)
```

```
.  graph export graph10.png, replace
(note: file graph10.png not found)
(file graph10.png written in PNG format)
```

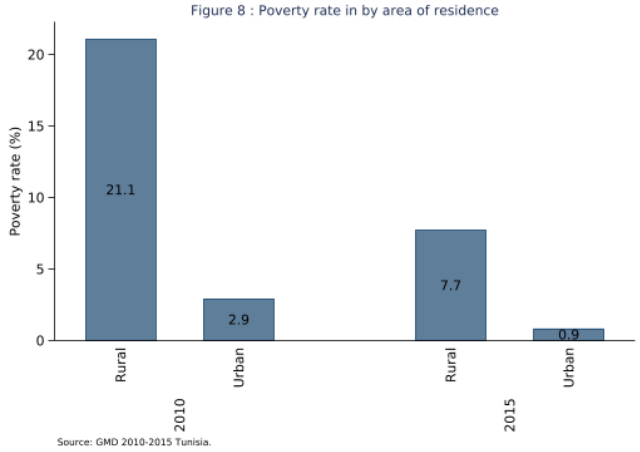


Considerable disparities in poverty rates persist across regions in both 2010 and 2015.

In 2010, poverty rates ranged from a low rate of 1.7 and 5.1 percent in Grand Tunis and Centre Est to a high of 25 and 20 percent in Center West and Northwest regions, respectively. The same pattern is observed in 2015 with a drastic reduction in each region. The next graph displays poverty headcount ratios by urban-rural.

In [29]:
```
use `data_all' , clear
 graph bar headcount3_2 [aw=weight_p],  inten(*0.7) o(urban, lab(labsi(small) angle(vertical))) o (year, lab(labsi(small)
 graph export graph10.png, replace
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  graph bar headcount3_2 [aw=weight_p],  inten(*0.7) o(urban, lab(labsi(small)
>  angle(vertical))) o (year, lab(labsi(small) angle(vertical))) blab(bar, pos(
> center) format(%3.1f) size(8pt)) ti(Figure 8 : Poverty rate in by area of res
> idence, si(small)) note("Source: GMD 2010-2015 Tunisia. ", si(*.7)) graphregi
> on(c(white)) ysize(6) xsize(6.5) legend(si(vsmall) r(3)) yti("Poverty rate (%
> )", si(small)) ylab(,labs(small) nogrid angle(0)) scheme(s2color) saving(grap
> h10.gph, replace)

.  graph export graph10.png, replace
(file graph10.png written in PNG format)
```



6. Additional analysis

Additional analysis can be performed depending on the project's goal. For example, in this example it could be relevant to analyze poverty by other demographics variables such as sex of household head, education of the head, or even poverty across age-groups. Such analysis is easy to perform and is similar to the one displayed in Figures 7 and 8.

In this last section, we provide additional graphs that could be added to a poverty analysis. This includes a growth incidence curve and an age pyramid.

6.1. Pyramid of age

In [31]:
```
use `data_all', clear
 gen agecat2 = 5 * floor(age/5)
 tab agecat2 , m
 lab var agecat2 "Age in categories"


 gen cmale = 1 if male ==1 & year ==2015
 gen cfemale=1 if male==0 & year ==2015

 preserve
 collapse (sum) cmale cfemale [pw=weight_h*hsize], by(agecat2)
 sum
 tempfile pyramide
 save `pyramide'

 restore
```

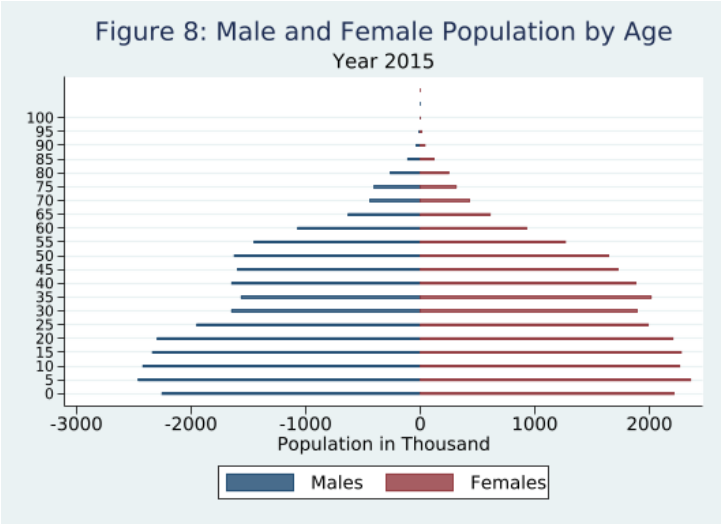
```
use `data_all', clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  gen agecat2 = 5 * floor(age/5)
(85 missing values generated)

.  tab agecat2 , m

    agecat2 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |     12,760        8.21        8.21
          5 |     12,567        8.09       16.30
         10 |     12,641        8.13       24.43
         15 |     13,323        8.57       33.00
         20 |     12,888        8.29       41.29
         25 |     11,792        7.59       48.88
         30 |     10,926        7.03       55.91
         35 |     10,611        6.83       62.74
         40 |     10,194        6.56       69.29
         45 |      9,614        6.19       75.48
         50 |      9,396        6.05       81.52
         55 |      7,999        5.15       86.67
         60 |      6,418        4.13       90.80
         65 |      4,283        2.76       93.56
         70 |      3,544        2.28       95.84
```

```
          75 |       2,861        1.84       97.68
          80 |       2,068        1.33       99.01
          85 |         956        0.62       99.62
          90 |         356        0.23       99.85
          95 |         126        0.08       99.93
         100 |          16        0.01       99.94
         105 |           2        0.00       99.94
         110 |           2        0.00       99.95
           . |          85        0.05      100.00
-------------+-----------------------------------
       Total |     155,428      100.00

.  lab var agecat2 "Age in categories"

.
.
.  gen cmale = 1 if male ==1 & year ==2015
(103,901 missing values generated)

.  gen cfemale=1 if male==0 & year ==2015
(101,898 missing values generated)

.
.  preserve

.  collapse (sum) cmale cfemale [pw=weight_h*hsize], by(agecat2)

.  sum

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     agecat2 |         23          55    33.91165           0        110
       cmale |         24     1094729    932820.2           0    2460800
     cfemale |         24     1106958    944809.1           0    2364348

.  tempfile pyramide

.  save `pyramide'
file /tmp/St00012.000004 saved

.
.  restore

.
.
```

In [32]:
```
use   `pyramide', clear
 replace cmale = - cmale/1000
 replace cfemale= cfemale/1000
```

```
twoway bar cmale agecat2, horizontal xvarlab(Males) || bar cfemale agecat2, horizontal xvarlab(Females) || , ylabel(0(5)
graph export graph3a.png, replace
```

```
use  `pyramide', clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  replace cmale = - cmale/1000
(23 real changes made)

.  replace cfemale= cfemale/1000
(23 real changes made)

.
.  twoway bar cmale agecat2, horizontal xvarlab(Males) || bar cfemale agecat2,
> horizontal xvarlab(Females) || , ylabel(0(5)100, angle(horizontal) valuelabel
>  labsize(*.8)) xtitle("Population in Thousand") ytitle("")  legend(label(1 Ma
> les) label(2 Females)) title("Figure 8: Male and Female Population by Age") s
> ubtitle("Year 2015") saving(graph3a.gph, replace)
(note: file graph3a.gph not found)
(file graph3a.gph saved)

.  graph export graph3a.png, replace
(note: file graph3a.png not found)
(file graph3a.png written in PNG format)
```



6.2. Growth Incidence Curve (GIC)

The Growth Incidence Curve (GIC) captures graphically the annualized growth rate of per capita income/consumption for every percentile

of the income/consumption distribution between two points in time (in this case 2010 and 2015). Over the last half-decade, a period of solid decline in inequality, income growth of the households at the bottom of the income distribution in most countries was significantly higher than those at the top.

The Growth Incidence Curve is a conceptually useful tool to analyze the impact of aggregate economic growth over a wide range of the distribution.

In [34]:

```
use `data_all' , clear
 local nq=100
        _ebin welfare if year==2010, gen(q_2010) nq(`nq')
        _ebin welfare if year==2015, gen(q_2015) nq(`nq')
        cap mat drop a
        foreach ii of num 1/`nq'                {
              su welfare if year==2010 & q_2010==`ii', mean
               local bb1=r(mean)
              su welfare if year==2015 & q_2015==`ii', mean
               local bb2=r(mean)
              local _`ii'=(((`bb2'/`bb1')^(1/(2015-2010)))-1)*100
        mat a=nullmat(a) \ `ii', `_`ii'' , `bb1' , `bb2'
        }
        tempvar centiles changecons mean1 mean2
        cap mat coln a=`centiles' `changecons' `mean1' `mean2'
        mat li a

        preserve
        svmat a , names (col)
        keep if _n<=100
        tempfile GIC_data
        save `GIC_data'
        restore
```

```
use `data_all' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

.  local nq=100

.  _ebin welfare if year==2010, gen(q_2010) nq(`nq')

.  _ebin welfare if year==2015, gen(q_2015) nq(`nq')

.  cap mat drop a

.  foreach ii of num 1/`nq'               {
  2.        su welfare if year==2010 & q_2010==`ii', mean
  3.                local bb1=r(mean)
  4.        su welfare if year==2015 & q_2015==`ii', mean
  5.                local bb2=r(mean)
  6.        local  _`ii'=(((`bb2'/`bb1')^(1/(2015-2010)))-1)*100
  7.  mat a=nullmat(a) \ `ii', `_`ii'' , `bb1' , `bb2'
  8. }

.  tempvar centiles changecons mean1 mean2

.  cap mat coln a=`centiles' `changecons' `mean1' `mean2'

.  mat li a
```

```
a[100,4]
        __000000   __000001   __000002   __000003
  r1           1  14.621369   312.36492  618.00247
  r2           2  13.693708    415.6442  789.59408
  r3           3  13.286411   480.58945  896.73333
  r4           4  12.719105   534.52353  972.64529
  r5           5  12.341363   581.02636   1039.667
                             (omitted for printing)
```

```
. preserve

. svmat a , names (col)

. keep if _n<=100
(155,328 observations deleted)

. tempfile GIC_data

. save `GIC_data'
file /tmp/St00012.000006 saved

. restore
```

In [35]:
```
use `GIC_data' , clear
 replace `changecons' =. if _n==1 | _n==_N
 twoway (line `changecons' `centiles', lcolor(red)), title("Figure 9: Growth Incidence Curve")  yti("Annual growth in per
 graph export graph13a.png, replace
```

```
use `GIC_data' , clear
(By MNAPOV/World Bank & Institut National de la Statistique - Ministère de la P
> la)

. replace `changecons' =. if _n==1 | _n==_N
(2 real changes made, 2 to missing)

. twoway (line `changecons' `centiles', lcolor(red)), title("Figure 9: Growth
> Incidence Curve")  yti("Annual growth in percentile average consumption/incom
> e", size(small)) xlabel(0(10)100,labs(small)) xtitle("Percentile group", size
> (small))saving(graph13a.gph, replace)
(note: file graph13a.gph not found)
(file graph13a.gph saved)
```

```
.  graph export graph13a.png, replace
(note: file graph13a.png not found)
(file graph13a.png written in PNG format)
```