Statistics/Data analysis

```
*legend
putexcel set "${xls_tool}", sheet(legend) modify

global row = 3
foreach aggregate in market_income $comp_list {
        qui putexcel E${row} = "`aggregate'"
        global row = ${row} + 1
        foreach var in $`aggregate' {
                qui putexcel E${row} = "`var'"
                global row = ${row} + 1
        }
}

*net cash position
global row = 1
foreach group_var in all decile decile_final hh_type county strata2  {

        use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_
> concepts}.dta", clear

        if "`group_var'" == "all" {
                gen all = 1
        }

        if "`group_var'" == "decile" {
                _ebin ${rank_var} [aw = $weight], gen(decile) nq(10)
        }

        if "`group_var'" == "decile_final" {
                _ebin final_income [aw = $weight], gen(decile_final) nq(10)
        }

        groupfunction [pw = $weight], sum(${income_list} ${comp_list} ${program_list
> }) by(`group_var') norestore

        foreach var in $income_list $comp_list $program_list {
                qui replace `var' = `var' / 10^6
        }

        sort `group_var'
        order `group_var' ${income_list} ${comp_list} ${program_list}

        export excel "${xls_tool}", sheet(totals) sheetmodify cell(A${row}) firstrow
> (variables) keepcellfm

        global row = ${row} + _N + 2

}

* recipients
putexcel set "${xls_tool}", sheet(recipients, replace) modify
putexcel A1 = "househols" A5 = "individuals"

use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_concepts
> }.dta", clear
groupfunction, sum(hh_weight ${program_list}) by(hhID) norestore
foreach var in $program_list {
        replace `var' = (`var' != 0)
}
```

```
groupfunction [pw = hh_weight], sum(${program_list}) norestore
foreach var in $program_list {
        replace `var' = `var' / 10^3 // thsd households
}
export excel "${xls_tool}", sheet(recipients) sheetmodify cell(A2) firstrow(variable
> s) keepcellfm

use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_concepts
> }.dta", clear
foreach var in $program_list {
        replace `var' = (`var' != 0)
}
groupfunction [pw = $weight], sum(${program_list}) norestore
foreach var in $program_list {
        replace `var' = `var' / 10^3 // thsd individuals
}
export excel "${xls_tool}", sheet(recipients) sheetmodify cell(A6) firstrow(variable
> s) keepcellfm

* Gini and Poverty
use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_concepts
> }.dta", clear
gen all = 1
sp_groupfunction [aw = $weight], gini(${income_list}) theil(${income_list}) poverty(
> ${income_list}) povertyline(${povline})  by(all)
gen concat = variable +"_"+ measure+"_"
order concat, first

export excel "${xls_tool}", sheet(gini_poverty) sheetreplace first(variable)

* Gini and Poverty by groups
global row = 1
foreach group_var in hh_type county strata2 {

        use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_
> concepts}.dta", clear

        sp_groupfunction [aw = $weight], poverty(${income_list}) povertyline(${povli
> ne}) by(`group_var')

        keep if measure == "fgt0"

        decode `group_var', generate(`group_var'_str)
        gen concat = variable +"_" + `group_var'_str
        order concat, first
        drop `group_var'_str

        export excel "${xls_tool}", sheet(gini_poverty) sheetmodify cell(I${row}) fi
> rstrow(variables) keepcellfm

        global row = ${row} + 100


}
```

```
*================================================================================
*by decile
use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_concepts
> }.dta", clear

        _ebin ${rank_var} [aw = $weight], gen(decile) nq(10)

        foreach var in $SSC $direct_taxes $indirect_taxes SSC direct_taxes indirect_
> taxes  {
                replace `var' = -`var'
        }

sp_groupfunction [aw = $weight], mean(${program_list} ${comp_list} ${income_list})
> by(decile)
        gen concat = variable +"_"+ measure+"_" +string(decile)
        order concat, first
export excel "${xls_tool}", sheet(all) sheetreplace first(variable)


* Marginal contributions
use "${postSimulationData}\12_${countryName}_${simulationName}_${output_inc_concepts
> }.dta", clear

gen all = 1

global MC_list


foreach inc in market market_pens net_market gross disposable consumable final {
        foreach var in $program_list $comp_list {
            gen double `inc'_`var' = `inc'_income - `var'
                global MC_list $MC_list `inc'_`var'
        }
}

sp_groupfunction [aw = $weight], gini(${MC_list}) poverty(${MC_list}) povertyline(${
> povline})  by(all)

keep if measure == "fgt0" | measure == "gini"
gen concat = variable +"_"+ measure

keep concat variable measure value
order concat concat variable measure value
sort concat
export excel "${xls_tool}", sheet(MC) sheetreplace first(variable)


shellout using "${xls_tool}"
```