**Assignment weeks 3, 4 and 5**

# *To answer all the questions below, you must use Stata (and, specifically, DASP, if requested). Be concise and clear in your answers.*

# *The assignment is divided into three exercises (the points assigned to each exercise are indicated next to each exercise). Please answer (A) directly in this file after each question (Q) and please attach the \*.do file (do-file) that you generated. Rename both files as: “Assignment weeks 3-4-5 - Name, Surname”. Please submit this completed file and the \*.do through the virtual drop box (boîte de dépôt) in the course portal, no later than Tuesday, February 23 11:59 p.m. (*[*Québec time*](https://www.timeanddate.com/worldclock/converter.html?iso=20190227T045900&p1=189)*).*

***Please, organize your dofile by exercise. Feel free to make your comments/discussions in the do-file.***

# Exercise 1 (4%)

Assume that the population is composed of six individuals belonging to two population groups, 1 and 2. The following table shows the distribution of incomes for three different periods.

|  |  |  |  |
| --- | --- | --- | --- |
| *Group* | *inc1* | *inc2* | *inc3* |
| 1 | 1 | 2 | 2 |
| 1 | 2 | 2 | 4 |
| 1 | 9 | 2 | 18 |
| 2 | 3 | 6 | 2 |
| 2 | 6 | 6 | 4 |
| 2 | 27 | 6 | 18 |

* 1. Discuss if the following affirmations are true or false and why, and this, for the distribution *inc1*.

1. Based on the *Scale invariance principle* the income inequality of group1 is equal to that of group 2. Input the data and confirm your justifications by estimating the Gini index by the population group.

**A: (TRUE)**

**/\*inputing data\*/**

**input Group inc1 inc2 inc3**

**1 1 2 2**

**1 2 2 4**

**1 9 2 18**

**2 3 6 2**

**2 6 6 4**

**2 27 6 18**

**end**

**/\*Explaination: It is true that income inequality in group 1 and 2 are the same. First, we should know that using relative measure of inequality, multiplying all income by the same scaler would not change the relative differences. In the same vein, the scale invariance principle says that an inequality index should not change if all incomes are scaled by a common factor. In our case incomes in group 1 are multplied by 3 to get corresponding incomes for group 2.\*/**

**igini inc1, hgroup(Group)**

1. By considering the *Scale invariance principle* and the *Population principle,* the income inequality of the group1 is equal to that of the total population.

**A: Answer (FALSE)**

**/\*The population principle that states that the inequality should remain the same to replication of the population and in our case, scaling does not necessarily mean replication as such we are bound to find differences in inequality index.\*/**

**igini inc1, hgroup(Group) /\*Indeed the results confirm that the population is different from the group 1 inequality\*/**

1. The between group inequality of *inc1* is equal to that of *inc2.* Also, check this using the ***dentropyg*** DASP command (for instance, for theta=0).

**A: (TRUE)**

**/\***

**Explanation: The ratio between the average income of the two groups in the periods 1 was 4/12=1/3. And similarry the ration of average income of the two groups in period 2 was 2/6=1/3 that shows that the between group inequality are the same.**

**\*/**

**dentropyg inc1, hgroup(Group) theta(0)**

**dentropyg inc2, hgroup(Group) theta(0)**

**/\*We can see from these estimates that indeed the between group inequality for both inc1 and inc2 are the same= 0.143841**

1.2 Using the DASP command ***dentropyg***, decompose the entropy index (the parameter theta = 0). Do this for each of the three periods.

**A: /\*Decomposing entropy index\*/**

**dentropyg inc1, hgroup(Group) theta(0)**

**dentropyg inc2, hgroup(Group) theta(0)**

**dentropyg inc3, hgroup(Group) theta(0)**

1.3 Estimate the Gini inequality for each of the three distributions with the ***igini*** DASP command and discuss the results.

**A: igini inc\***

# Exercise 2 (5.5%)

Assume that the population is composed of eight households.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *identifier* | *pre\_tax\_income* | *hhsize* | *nchild* | *nelderly* |
| 1 | 240 | 4 | 2 | 1 |
| 2 | 600 | 5 | 3 | 1 |
| 3 | 230 | 3 | 2 | 0 |
| 4 | 1250 | 3 | 1 | 1 |
| 5 | 1900 | 4 | 1 | 1 |
| 6 | 280 | 4 | 2 | 0 |
| 7 | 620 | 3 | 1 | 1 |
| 8 | 880 | 4 | 3 | 0 |
| **Total** | **6000** | **30** | **15** | **5** |

The disposable income of the household is composed of three income sources:

1. The post tax income = pre-tax income – income tax;
2. The received child allowances
3. The received elderly pension

The government disposes two potential scenarios (A and B).

1. ***Scenario A:*** apply a proportional income tax of 10%. Then, 20% of the total collected taxes are equally distributed on the elderly population as pensions. The rest of the budget is equally distributed across the population of children, as allowances.
2. ***Scenario B:*** apply a proportional income tax of 10%, and then redistribute the generated revenue equally across the population of children. In that case, the universal elderly pension is equal to zero.

2.1 Using Stata, input the data (the eight observations), and then generate the variables:

* *pcincatA:* per capita post tax income with the scenario A;
* *pcincatB:* per capita post tax income with the scenario B;
* *pceldA:* per capita elderly pension with the scenario A;
* *pceldB:*  per capita elderly pension with the scenario B;
* *pcallowA:* per capita child allowances with the scenario A;
* *pcallowB:*  per capita child allowances with the scenario B;
* *dpcincA:* per capita disposable income with the scenario A (*pcincatA+ pceldA+ pcallowA*);
* *dpcincB:* per capita disposable income with the scenario B (*pcincatB+ pceldB + pcallowB*).

**A:**

**/\***

**For a given household, the post\_tax\_income in Scenario A is equal to its pre\_tax\_income minus 10% of the pre\_tax\_income.**

**Thus, post\_tax\_income= pre\_tax\_income\*(1-0.1)**

**Now, to compute the per capita post\_tax\_income (pcincatA), we have to divide by the household size.**

**This explains the form of the two following command lines.**

**\*/**

**gen pcincatA = pre\_tax\_income \* (1.00-0.1)/hhsize**

**gen pcincatB = pre\_tax\_income \* (1.00-0.1)/hhsize /\*we also do the same for post\_tax\_income inscenario B)\*/**

**/\***

**The collected tax revenue in scenario A is 10% of the total incomes : 0.1\*6000=600.**

**-20% of 600 (tax revenue) is equally distributed among 5 elderly in the population: ((0.1\*6000)\*0.2)/8 =120/5.**

**-the remaining tax income is distributed equally among 15 children allowance: (600-120)/15**

**-We would like to store the value of elderly allowance and child allowance in scenario A in the memory to be used latter.**

**-The command scalar can be used to the elderly allowance in the instance elder\_all\_A.**

**-The command scalar can be used to the child allowance in the instance child\_all\_A.**

**Thus for Stata, the component child\_all\_A is equal to (600-120)/15.**

**\*/**

**scalar elder\_Pens\_A= 120/5**

**scalar child\_all\_A = (600-120)/15**

**/\***

**In scenario B 10 % of the total income is taxed and redistributed to children with none given to the eldery pension.**

**-Thus allowance to children in scenario B is (0.1\*6000)/15 while for elderly its zero**

**using the same scalar principle, we keep these components as below\*/**

**scalar elder\_Pens\_B=0**

**scalar child\_all\_B=(0.1\*6000)/15**

**/\*we now generate percapita eldery pension and percapita child allowance in both scenarios A and B.**

**By household, the total received allowances:**

**-In scenario A: nchild\*child\_all\_A + nelderly\*elder\_Pens\_A**

**-In scenario B: nchild\*child\_all\_B**

**Thus we compute percapita eldery pension and child allowance for the household:**

**\*/**

**gen pceldA= (nelderly\*elder\_Pens\_A)/hhsize**

**gen pceldB= (nelderly\*elder\_Pens\_B)/hhsize**

**gen pcallowA= (nchild\*child\_all\_A)/hhsize**

**gen pcallowB= (nchild\*child\_all\_B)/hhsize**

**/\***

**Recall that the per capita disposable income is equal to the per capita pre\_tax\_income plus the per capita child allowances and plus the percapita eldery pension. It is the same as addition zero for the for eldery pension in scenario B.**

**\*/**

**gen dpcincA= pcincatA+ pcallowA + pceldA**

**gen dpcincB= pcincatB+ pcallowB + pceldB**

2.2 Using the DASP command *igini*, estimate the inequality in the distribution of the per capita disposable income for each of the two scenarios and discuss the results.

**A: /\***

**Estimating per capita disposable income for both scenario A and B. We are using the DASP command igini that enables us to compute the Gini inequality for a given variable or a set of variables.**

**\*/**

**igini dpcincA dpcincB , hsize(hhsize)**

**/\*The results show that scenario B has the highest reduction in inequality in disposable income as the gini estimate is 0.348667compared to 0.352933 in scenario A. This implies that giving allowance to children is a much better policy ti help reduce inequality than sharing the tax income allowance for children and the elderly\*/**

2.3 Using the DASP command *diginis*, decompose the inequality in the distribution of the per capita disposable income for each of the two scenarios (remember that the three income sources are *pcincatA, pceldA and pcallowA* for the scenario A and *pcincatB, pceldB and pcallowB* for the scenario B)*.*

**A: /\*We decompose inequality by using the diginis DASP command that enables to compute the total Gini inequality, as well as, to decompose the total inequality by -disposable-income sources. In our case the income sources for scenario A are pcincatA pceldA pcallowA and for scenario B are vpcincatB pceldB pcallowB \*/**

**diginis pcincatA pceldA pcallowA, hsize(hhsize)**

**diginis pcincatB pceldB pcallowB, hsize(hhsize)**

**/\***

2.4 Based on the results of 2.2 and 2.3, in which case will the set of transfer programs reduce inequality in disposable incomes the most? Why?

**A: /\***

**Scenario B is the one with the highest reduction in inequality in disposable incomes.**

**This is because this program effectively targets children who may be in need morre tha the eldery.**

**This also makes the contribution of the source Child Allowances more effective in reducing inequality.**

**\*/**

2.5 Estimate the change in the headcount when the scenario B is enacted (with respect to the initial distribution), and the poverty line is 100 (use the DASP command *difgt*).

**A: /\* generating the per capita income for the initial distribution without applying any program \*/**

**gen pcinc = pre\_tax\_income/hhsize**

**difgt dpcincB pcinc, alpha(0) hsize1(hhsize) hsize2(hhsize) pline1(100) pline2(100)**

**Without child allowances, the poverty head count is : .3666667**

**With child allowances, the poverty headcount is : .3666667**

**Child allowances does not reduce the poverty head count as the difference between previous headcount and current head ciunt is just equal to 0. This difference is significant by about 10% (i.e. P>|t| = 0.1)**

2.6 Estimate the change in the poverty gap related to the scenario B (with respect to the initial distribution) and when the poverty line is 100 (use the DASP command *difgt*). Compare the results found here with those found in the previous point (2.5).

**A: /\* Estimating Poverty gap\*/**

**difgt dpcincB pcinc, alpha(1) hsize1(hhsize) hsize2(hhsize) pline1(100) pline2(100)**

**Without child allowances, the poverty gap is : .1166667**

**With child allowances, the poverty gap is : 0.0616667**

**Child allowances reduce the poverty gap from .1166667 to 0.0616667 or by .055.**

**This difference is significant by about 10% (i.e. P>|t| = 0.1)**

**The results are showing the difference in poverty measures. Poverty headcount is not sensitive to income transfers among the poor as it still recodes no difference in terms of poverty while the poverty gap has shown that even though people have not moved above the poverty line, the tax policy has helped those who were so poor to catch up close to the poverty line.**

**\*/**

# Exercise 3 (3%)

* 1. Load the file data\_2, then initialize the sampling design with the variables *strata, psu* and *sweight*.

**A: clear**

**use "data\_2.dta" , replace**

**//Initializing sampling desgin of the data file using strata, psu and sweight**

**svyset psu [pweight=sweight], strata(strata)**

* 1. Using the DASP ***ifgt*** command, estimate the headcount when the measurement of well-being is the adult equivalent expenditures and the poverty line is equal to 21 000.

**A: /\*Estimating the Poverty Head count (theta=0), wellfare variable= a\_e and poverty line=21000\*/**

**ifgt ae\_exp, alpha(0) hsize(hsize) pline(21000)**

* 1. Now, estimate the headcount poverty by population groups (defined by the sex of the household head) and discuss the results.

**A: /\*Estimating the Poverty Head count (theta=0), wellfare variable= a\_e,**

**poverty line=21000 and subgrouped by sex\*/**

**ifgt ae\_exp, alpha(0) hsize(hsize) hgroup(sex) pline(21000)**

**/\*Discussion: Total population poverty head count is 33.7% implying that these are total number of individuals below the poverty line. The results show that more females are are poor compared to males. About 37.9% of females are poor while 32.5% of males are poor. The imolication is that there is income differentials between males and females in this country\*/**