lerman: A Stata module to decompose inequality using sampling weights

Bruno Damásio and David Neves

Portuguese Stata Users Group meeting

2015

Outline

Motivation

- The impact of economic trends and government policies on the distribution of income is a central topic in economic analysis
- Income inequality is, at last, becoming a major concern for public policy in the western economies
- Thus, it is crucial to have rigorous tools to evaluate the sources of income inequality and the efectiveness of the goverment's social policies

Motivation

- Lerman and Yitzhaki (1985) proposed a methodology to decompose the Gini coefficient
- López-Feldman (2006) presented a Stata module to operationalize Lerman and Yitzhaki's method (`descogini`)
- However it does not allow the usage of sampling weights
- That considerably narrows its application to surveys that collect information on household income dynamics, such as the PSID in the USA and the EU-SILC in the EU.
- Due to the sampling design of these surveys, the estimates come biased whenever sampling weights are excluded from the statistical analysis

Gini index decomposition

- Lerman and Yitzhaki shown that the Gini index can be written as follows:

$$G = \sum_{k=1}^{K} S_k G_k R_k$$

  - G is the Gini index
  - S_k is the the share of the $k - th$ source in total income
  - G_k is the source Gini corresponding to the distribution of income from source k
  - R_k represents the correlation between income from source k with the distribution of total income

Gini index decomposition

- S_k represents the relative importance of the $k - th$ income source
- G_k gives us an idea of the (in)equality of the distribution of income from source k
- R_k is a measure of the strength and direction of the linear relationship between the source k and the distribution of total income

Gini index decomposition

Advantages of this approach:

- It has an intuitive interpretation
- It allows us to identify which source of income mostly contributes to household income inequality
- It allows to understand how changes in particular income sources will affect overall income inequality
- The latter is particularly useful to evaluate the efectiveness of public transfers in reducing income inequality

Gini index decomposition

- Consider a change in the household income from source k equal to εY_k (Y_k is the income form $k-th$ source)
- It can be shown that the partial derivative of the overall Gini with respect to a percentage change in source k is

$$\frac{\delta G}{\delta \varepsilon} = S_k \left(G_k R_k - G \right) \tag{1}$$

Gini index decomposition

- Dividing (1) by G gives the source's marginal effect relative to the overall Gini:

$$\frac{\delta G/\delta \varepsilon}{G} = \frac{S_k G_k R_k}{G} - S_k \tag{2}$$

- Which can be written as the source's inequality contribution as a percentage of the overall Gini minus the source's share of total income
- The sum of relative marginal effects is zero
- Multiplying all sources by ε leaves the overall Gini unchanged
- We can estimate the impact that a 1% change in income from source k will have on total income inequality

Program syntax

- Syntax

```
lerman country year varlist [if] [in] [pw=weight]
```

  - country = a string variable that includes one or more geographic units (countries, states, divisions, etc).
  - year = a numeric variable that includes one or more time units
  - varlist = total income followed by its components

US Panel Study of Income Dynamics

- To illustrate the program, we used the The Panel Study of Income Dynamics (PSID):
  - Three waves: 1994, 2001 and 2013
  - 13 US divisions and states
  - 24.825 families
  - Cross-secional sampling weights

US Panel Study of Income Dynamics

- Five equivalized income sources:
  - employment (laboreq)
  - self employment (laborselfeq)
  - pensions (pensionseq)
  - property (propertyeq)
  - transfers (transferseq)

lerman by year ignoring sampling weights

Figure : ignoring sampling weights

```
. lerman USasawhole year totaleq laboreq laborselfeq propertyeq pensionseq transferseq
US
1994 2001 2013

table_1994_US[5,9]
               year_US   source_US       Rk_US       Gk_US    Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq       1994           1    .84401059    .56882647    .45343775    .56936131    .61007044    .44805947   -.12130184
laborselfeq       1994           2    .71626324    .98913374    .58853577    .01977884    .61007044    .02296934     .0031905
 propertyeq       1994           3    .84576306    .98889716    .86807204    .10243482    .61007044    .14043245    .03799763
 pensionseq       1994           4    .74604999    .94852981    .78945434    .22039411    .61007044    .25564595    .03525184
transferseq       1994           5    .92782517    .99261316    .96498348    .08803091    .61007044    .13289279    .04486187

table_2001_US[5,9]
               year_US  source_US       Rk_US       Gk_US    Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq       2001          1    .8433403    .53632865    .44436641    .6288769    .57023961    .49881799   -.1300589
laborselfeq       2001          2    .73296995    .9895016    .65124365    .02021464   .57023961    .02571055    .00549591
 propertyeq       2001          3    .77637384    .99769785   .82746423    .07831979   .57023961    .10638589    .02806611
 pensionseq       2001          4    .79778914    .96102607   .83280868    .23992811   .57023961    .32258714    .08265903
transferseq       2001          5    .82032017    .9896659    .94163206    .03266056   .57023961    .04649843    .01383787

table_2013_US[5,9]
               year_US   source_US       Rk_US       Gk_US    Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq       2013           1    .91050144    .57372015    .49050199    .80558746    .48434464    .86883828    .06325082
laborselfeq       2013           2    .70155329    1.0166361    .70366458    .01521114    .48434464    .02239925    .00718811
 propertyeq       2013           3    .67931407    .93886415    .84648213    .06280746    .48434464    .08270469    .01989722
 pensionseq       2013           4    .19129855    .85036711    .46041918    .09836782    .48434464    .03303822   -.0653296
transferseq       2013           5   -.20722359    .90509821     .497503     .01802612    .48434464   -.00698044   -.02500655
```

lerman by year with sampling weights

Figure : considering sampling weights

```
. lerman USasawhole year totaleq laboreq laborselfeq propertyeq pensionseq transferseq [pw=weight]
US
1994 2001 2013

table_1994_US[5,9]
                year_US   source_US       Rk_US       Gk_US     Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq        1994           1    .80786131    .58988698    .45284182    .50389404    .61104224    .39298286   -.11091117
laborselfeq        1994           2    .68039861    .98542972    .57117286    .02072377    .61104224     .0227398    .00201603
 propertyeq        1994           3    .82532097    .94746728     .8523925    .11809904    .61104224    .15113402    .03303498
 pensionseq        1994           4    .75510845    .92936316    .77713907    .29030802    .61104224    .33341304    .04310502
transferseq        1994           5    .91555553    .99380219    .96170772    .06697514    .61104224    .09973028    .03275514

table_2001_US[5,9]
                year_US   source_US       Rk_US       Gk_US     Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq        2001           1    .81713498    .56049799     .4409935    .56361774    .58192654    .44359266   -.12002508
laborselfeq        2001           2    .70713755    .99679036    .65255626    .02143427    .58192654     .0259626    .00452833
 propertyeq        2001           3    .73740652     .9462024     .8100832    .08405958    .58192654    .10078829    .01672871
 pensionseq        2001           4    .79140708    .94371761    .81422127    .29963093    .58192654    .38455678    .08492585
transferseq        2001           5    .84553762    .99301222    .94798629    .03125748    .58192654    .04509967    .01384219

table_2013_US[5,9]
                year_US   source_US       Rk_US       Gk_US     Gkadj_US       Sk_US     gini_US    share_US    delta_US
    laboreq        2013           1    .87827637    .61000913    .48854786    .74429226    .47230659    .84428104    .09998877
laborselfeq        2013           2     .6510064   1.0330151    .67684569    .01629988    .47230659    .02320879     .0069089
 propertyeq        2013           3    .66420227    .91932516    .83593665    .08035729    .47230659    .10388929     .023532
 pensionseq        2013           4    .13937729    .77929625    .41241162    .14580692    .47230659    .03353117   -.11227575
transferseq        2013           5   -.1885103    .92894029    .49391843    .01324365    .47230659   -.00491028   -.01815393
```

Conclusions: US 2013

- Labor income increases inequality and Gini coefficient
  - 1% change increases Gini by 0.099%
- Pensions decreases inequality
  - 1% change decreases Gini by 0.112%
- Social transfers decreases inequality
  - 1% change decreases Gini by 0.018%
- Social transfers and pensions are unequally distributed but have a equalizing effect on income distribution

Conclusions: US over years

- The household income inequality of gross income have been decreasing (0.6110> 0.5819 > 0.4723)
- The social transfers mechanism is effective
- The social transfers (and pensions) have now an equalizing effect (0.0327 > 0.0138 > -0.01815)
- The labor share raised (0.5038 >0.5361 > 0.7443), but...
- It is more unequally distributed (0.5898 >0.6100)

References

1. Lerman, R. I., and Yitzhaki, S. 1985. Income inequality effects by income source: A new approach and applications to the United States. Review of Economics and Statistics 67: 151–156.
2. Lopez-Feldman, A. 2006. Decomposing Decomposing inequality inequality and obtaining marginal effects. The Stata Journal 6:106- 111.
3. Stark, O., Taylor, J. E. and Yitzhaki, S. 1986. Remittances and inequality. Economic Journal 96: 722- 740.