

Decomposition of the Gini coefficient using Stata

Alejandro López-Feldman
Universidad de Guanajuato
(Moving to CIDE in June)

Motivation

- The Gini coefficient is widely used to measure inequality in the distribution of income, wealth, expenditures, etc.
- By decomposing this measure you can better understand the determinants of inequality.
- Today I will show the use of a user written command *descogini* to decompose the Gini by income source.

Background

- Lerman and Yitzhaki (1985) show that the Gini coefficient for total income, G , can be represented as:

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

S_k = share of source k in total income

G_k = the source Gini

R_k = Gini correlation of income from source k with the distribution of total income

Background

- Stark, Taylor and Yitzhaki (1986) note that the influence of any income component upon total income inequality depends on:
 - how important the income source is with respect to total income;
 - how equally or unequally distributed the income source is; and
 - how the income source and the distribution of total income are correlated.

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

Background

- By using this method of Gini decomposition you can estimate the effect that a 1% change in income from source k will have on total income inequality. This effect is given by:

$$\frac{S_k G_k R_k}{G} - S_k$$

Program syntax

- Syntax

descogini *varlist* [*if*] [*in*] [, **d(#)** **bar**]

- Options

d(#) allows the user to specify the number of decimal places to be reported in the table of results. The default is d(4)

bar specifies that a vertical bar should be included before each output field.

Program syntax

- Remarks

The first variable in `varlist` should always be the variable that captures total income. The order in which the income sources are included in the syntax does not matter.

Examples

- Example 1
 - Use data from Chiapas Mexico.
 - 3 income sources (family income, remittances and government transfers).

descogini output

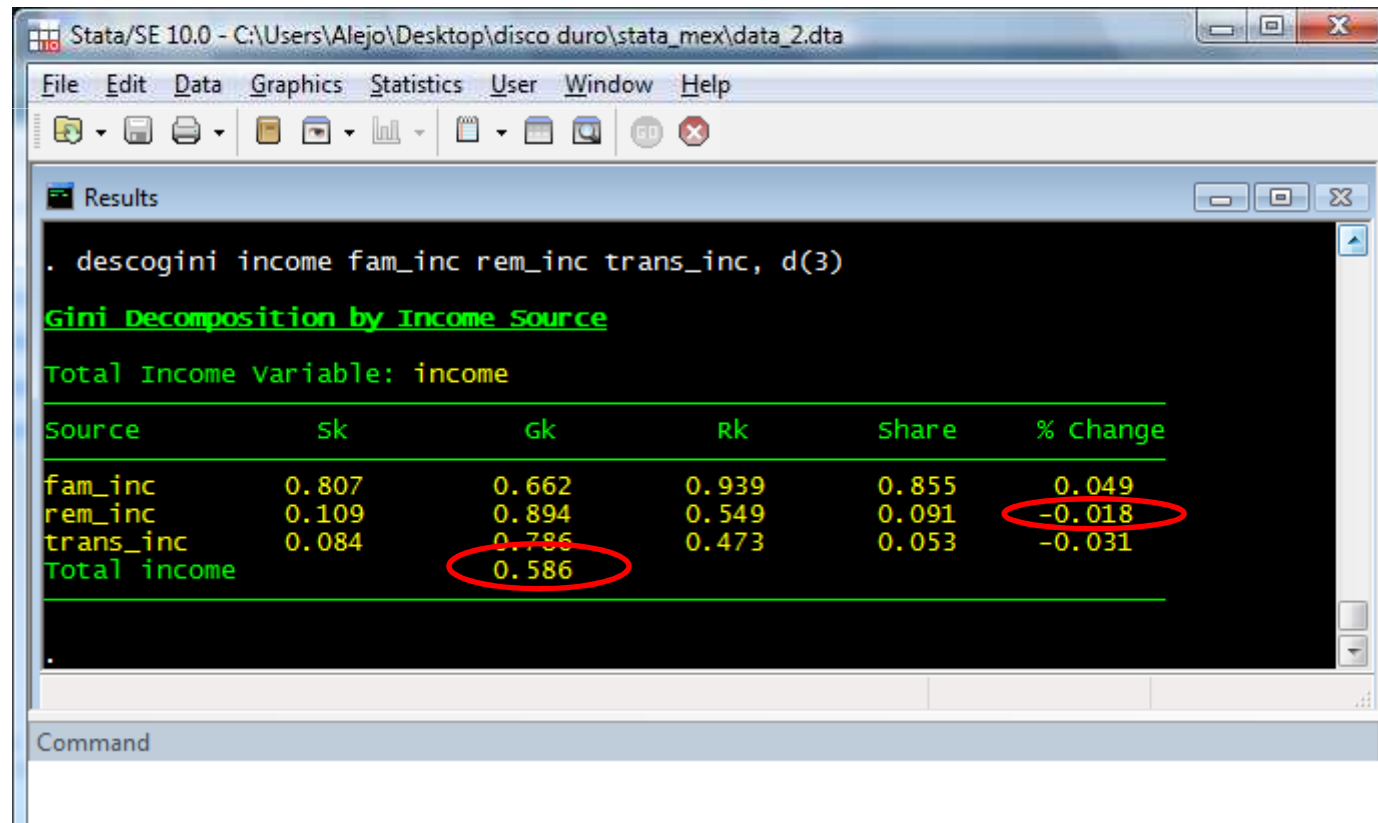
```
Stata/SE 10.0 - C:\Users\Alejo\Desktop\disco duro\stata_mex\data_1.dta
File Edit Data Graphics Statistics User Window Help
Results
. descogini income fam_inc rem_inc trans_inc, d(3)
Gini Decomposition by Income Source
Total Income Variable: income
Source      sk      Gk      Rk      share  % change
fam_inc     0.744   0.433   0.960   0.893   0.149
rem_inc     0.011   0.981   0.202   0.006  -0.005
trans_inc   0.245   0.341   0.418   0.101  -0.144
Total income      0.346
```

Command

C:\data

Examples

- Example 2
 - Use data from rural Mexico (ENHRUM).
 - 3 income sources (family income, remittances and government transfers).



```
. descogini income fam_inc rem_inc trans_inc, d(3)
```

Gini Decomposition by Income Source

Total Income Variable: income

Source	sk	Gk	Rk	share	% Change
fam_inc	0.807	0.662	0.939	0.855	0.049
rem_inc	0.109	0.894	0.549	0.091	-0.018
trans_inc	0.084	0.786	0.473	0.053	-0.031
Total income		0.586			

Command

Saved Results

- The % changes are saved as a vector $e(b)$
- The estimates of S_k, G_k, R_k, G are available as scalars.
- **bootstrap** can be easily applied to obtain standard errors and confidence intervals for any of these estimates.

bootstrap and descogini

```
Stata/SE 10.0 - C:\Users\Alejo\Desktop\disco duro\stata_mex\data_2.dta
File Edit Data Graphics Statistics User Window Help
Results
. set seed 34567
. bootstrap "descogini income fam_inc rem_inc trans_inc" _b
command: descogini income fam_inc rem_inc trans_inc
statistics: b_fam_inc = _b[fam_inc]
            b_rem_inc = _b[rem_inc]
            b_trans_~c = _b[trans_inc]

Bootstrap statistics                Number of obs    =    1745
                                   Replications      =     50

Variable      Reps  Observed      Bias  Std. Err.  [95% Conf. Interval]
-----
b_fam_inc     50  .0487089  .0016582  .007815   .0330041  .0644136  (N)
               .0377644  .0648405  (P)
               .0262083  .06232   (BC)
b_rem_inc     50  -.0177832 -.0013166  .0049589  -.0277484  -.0078179  (N)
               -.0277763  -.0078057  (P)
               -.0235923  -.0019931  (BC)
b_trans_inc   50  -.0309257  -.0003415  .0044945  -.0399577  -.0218937  (N)
               -.0387277  -.0232566  (P)
               -.0387277  -.0220648  (BC)

Note: N = normal
      P = percentile
      BC = bias-corrected
```

References

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

Gracias!

lopezfeldman@gmail.com