



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

Papers downloaded from AgEcon Search may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

THE STATA JOURNAL

Editor

H. Joseph Newton
Department of Statistics
Texas A & M University
College Station, Texas 77843
979-845-3142; FAX 979-845-3144
jnewton@stata-journal.com

Editor

Nicholas J. Cox
Geography Department
Durham University
South Road
Durham City DH1 3LE UK
n.j.cox@stata-journal.com

Associate Editors

Christopher Baum
Boston College
Rino Bellocchio
Karolinska Institutet, Sweden and
Univ. degli Studi di Milano-Bicocca, Italy
David Clayton
Cambridge Inst. for Medical Research
Mario A. Cleves
Univ. of Arkansas for Medical Sciences
William D. Dupont
Vanderbilt University
Charles Franklin
University of Wisconsin, Madison
Joanne M. Garrett
University of North Carolina
Allan Gregory
Queen's University
James Hardin
University of South Carolina
Ben Jann
ETH Zurich, Switzerland
Stephen Jenkins
University of Essex
Ulrich Kohler
WZB, Berlin
Jens Lauritsen
Odense University Hospital

Stanley Lemeshow
Ohio State University
J. Scott Long
Indiana University
Thomas Lumley
University of Washington, Seattle
Roger Newson
Imperial College, London
Marcello Pagano
Harvard School of Public Health
Sophia Rabe-Hesketh
University of California, Berkeley
J. Patrick Royston
MRC Clinical Trials Unit, London
Philip Ryan
University of Adelaide
Mark E. Schaffer
Heriot-Watt University, Edinburgh
Jeroen Weesie
Utrecht University
Nicholas J. G. Winter
Cornell University
Jeffrey Wooldridge
Michigan State University

Stata Press Production Manager
Stata Press Copy Editors

Lisa Gilmore
Gabe Waggoner, John Williams

Copyright Statement: The Stata Journal and the contents of the supporting files (programs, datasets, and help files) are copyright © by StataCorp LP. The contents of the supporting files (programs, datasets, and help files) may be copied or reproduced by any means whatsoever, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

The articles appearing in the Stata Journal may be copied or reproduced as printed copies, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

Written permission must be obtained from StataCorp if you wish to make electronic copies of the insertions. This precludes placing electronic copies of the Stata Journal, in whole or in part, on publicly accessible web sites, fileservers, or other locations where the copy may be accessed by anyone other than the subscriber.

Users of any of the software, ideas, data, or other materials published in the Stata Journal or the supporting files understand that such use is made without warranty of any kind, by either the Stata Journal, the author, or StataCorp. In particular, there is no warranty of fitness of purpose or merchantability, nor for special, incidental, or consequential damages such as loss of profits. The purpose of the Stata Journal is to promote free communication among Stata users.

The *Stata Journal*, electronic version (ISSN 1536-8734) is a publication of Stata Press, and Stata is a registered trademark of StataCorp LP.

Decomposing inequality and obtaining marginal effects

Alejandro López-Feldman
Department of Agricultural and Resource Economics
University of California–Davis
Davis, CA
feldman@primal.ucdavis.edu

Abstract. This article describes a user-written command, `descogini`, that decomposes the Gini coefficient by income source and allows the calculation of the impact that a marginal change in a particular income source will have on inequality. `descogini` can be used with `bootstrap` to obtain standard errors and confidence intervals.

Keywords: st0100, `descogini`, Gini, Gini decomposition by income source, inequality

1 Introduction

The Gini coefficient is widely used to measure inequality in the distribution of income, consumption, and other welfare proxies. Decomposing this measure can help you understand the determinants of inequality. The techniques used more often decompose inequality either by subpopulations or by income source. This article describes a user-written command, `descogini`, that decomposes the Gini coefficient by income source and allows the calculation of the impact that a marginal change in a particular income source will have on inequality.

There are other user-written commands available for Stata that decompose inequality. For example, `ineqdeco` and `ginidesc` decompose inequality by subpopulation groups, whereas `ineqfac` does so by factor components of total income. `descogini` complements the commands that do subpopulation decompositions and is an alternative to `ineqfac`. The main difference between `ineqfac` and `descogini` is that `descogini` allows the estimation of the marginal effects that every income source has on inequality by using the approach proposed by Lerman and Yitzhaki (1985). Furthermore, this command can be used with `bootstrap` to obtain standard errors and confidence intervals.

2 Framework

Extending the results of Shorrocks (1982), Lerman and Yitzhaki (1985) show that the Gini coefficient for total income inequality, G , can be represented as

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

where S_k represents the share of source k in total income, G_k is the source Gini corresponding to the distribution of income from source k , and R_k is the Gini correlation of income from source k with the distribution of total income ($R_k = \text{Cov}\{y_k, F(y)\}/\text{Cov}\{y_k, F(y_k)\}$, where $F(y)$ and $F(y_k)$ are the cumulative distributions of total income and of income from source k).

As noted by Stark, Taylor, and Yitzhaki (1986), the relation among these three terms has a clear and intuitive interpretation; the influence of any income component upon total income inequality depends on

- how important the income source is with respect to total income (S_k);
- how equally or unequally distributed the income source is (G_k); and
- how the income source and the distribution of total income are correlated (R_k).

If an income source represents a large share of total income, it may potentially have a large impact on inequality. However, if income is equally distributed ($G_k = 0$), it cannot influence inequality, even if its magnitude is large. On the other hand, if this income source is large and unequally distributed (S_k and G_k are large), it may either increase or decrease inequality, depending on which households (individuals), at which points in the income distribution, earn it. If the income source is unequally distributed and flows disproportionately toward those at the top of the income distribution (R_k is positive and large), its contribution to inequality will be positive. However, if it is unequally distributed but targets poor households (individuals), the income source may have an equalizing effect on the income distribution.

Lerman and Yitzhaki (1985) show that by using this particular method of Gini decomposition, you can estimate the effect of small changes in a specific income source on inequality, holding income from all other sources constant. Consider a small change in income from source k equal to ey_k , where e is close to 1 and y_k represents income from source k . It can be shown (see Stark, Taylor, and Yitzhaki [1986]) that the partial derivative of the Gini coefficient with respect to a percent change e in source k is equal to

$$\frac{\partial G}{\partial e} = S_k(G_k R_k - G)$$

where G is the Gini coefficient of total income inequality prior to the income change. The percent change in inequality resulting from a small percent change in income from

source k equals the original contribution of source k to income inequality minus source k 's share of total income:

$$\frac{\partial G/\partial e}{G} = \frac{S_k G_k R_k}{G} - S_k \quad (1)$$

3 The descogini program

3.1 Syntax

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

3.2 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` includes a vertical bar before each field to facilitate the creation of a table of results in Microsoft Word.

3.3 Description

For the command to work properly, the variable that captures total income must be included first, followed by the income source variables. The order in which the income source variables are included in the syntax of the command is not important for the consistency of the results, and in fact, as shown in the examples below, you may include only a partial list of the income source variables.

The estimates of S_k , G_k , R_k , G , and the marginal effects obtained from (1) are all available after `descogini` is run. The marginal effects are saved as a vector `e(b)`, whereas the rest of the results are scalars under the following names: $S_k = \mathbf{sname}$, $G_k = \mathbf{gname}$, $R_k = \mathbf{rname}$, and $G = \mathbf{gtotal}$, where *name* is the name of each income source variable. `bootstrap` can be easily applied to obtain standard errors and confidence intervals for any of these estimates.

4 An example of Gini decomposition by income source

The following example illustrates the use of `descogini` and presents an interpretation of the results. Total income can be disaggregated in multiple ways, depending on the characteristics of the information and on the question that you are trying to answer. In this example, total income from household data is divided into three sources: farm income (`farminc`), off-farm income (`ofarminc`), and income from remittances and government transfers (`transinc`). The output obtained from using `descogini` is the following:

```
. use desco
(Data for examples of descogini)
```

```
. descogini totinc farminc ofarminc transinc
```

Gini Decomposition by Income Source

Total Income Variable: totinc

Source	Sk	Gk	Rk	Share	% Change
farminc	0.6225	0.7374	0.9428	0.8207	0.1982
ofarminc	0.1929	0.6577	0.5502	0.1324	-0.0605
transinc	0.1846	0.3793	0.3535	0.0469	-0.1377
Totalincome		0.5273			

For brevity, let us look at the results of just two of the three income sources. The results of farm income show that a 1% increase in that income source, all else being equal, increases the Gini coefficient of total income by 0.198%. Farm income is unequally distributed (0.737), and the Gini correlation between farm income and total income is high (0.943), indicating that farm income favors the rich more than any other income source. On the other hand, off-farm income has a slight equalizing effect on the distribution of total income. This finding shows that a relatively high source Gini (0.658) does not imply that an income source has an unequalizing effect on total-income inequality. An income source may be unequally distributed yet favor the poor, as is the case for off-farm income. For detailed examples of Gini decomposition by income source, see Stark, Taylor, and Yitzhaki (1986); López-Feldman, Mora, and Taylor (2005); and Taylor et al. (2005).

As mentioned earlier, `descogini` allows the estimation of bootstrapped standard errors and confidence intervals. Assume that we are interested only in the standard errors of the farm income Gini and in the marginal change in inequality due to a 1% change in farm income. These answers can be obtained in the following way:

```
. set seed 1234567
. bootstrap "descogini totinc farminc" Gfarm=gfarminc Mgfarminc=_b[mgfarminc]
  command: descogini totinc farminc
  statistics: Gfarm      = gfarminc
               Mgfarminc = _b[mgfarminc]
Bootstrap statistics
                                         Number of obs      =      185
                                         Replications      =       50

```

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf. Interval]
Gfarm	50	.7373723	-.0124623	.0359332	.6651617 .8095828 (N) .6646934 .7848438 (P) .6690543 .7929884 (BC)
Mgfarminc	50	.1981957	-.0019509	.0193368	.159337 .2370545 (N) .1618218 .2267272 (P) .1707874 .2476358 (BC)

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

To guarantee reproducibility of the results, `set seed` must be used. All the options of `bootstrap` (see [R] `bootstrap`) are available when bootstrapping `descogini`.

Finally, since the estimates of marginal effects are saved in the vector `e(b)`, we can easily bootstrap them all at the same time:

```
. set seed 1234567
. bootstrap "descogini totinc farminc ofarminc transinc" _b
  command: descogini totinc farminc ofarminc transinc
  statistics: b_mgfarm-c = _b[mgfarminc]
               b_mgofar-c = _b[mgofarminc]
               b_mgtran-c = _b[mgtransinc]
  Bootstrap statistics
                                         Number of obs      =      185
                                         Replications      =       50
  Variable      Reps  Observed      Bias  Std. Err.  [95% Conf. Interval]
  b_mgfarminc   50  .1981957  -.0019509  .0193368   .159337  .2370545  (N)
                           .1618218  .2267272  (P)
                           .1707874  .2476358  (BC)
  b_mgofarminc  50  -.0605021  .0006287  .0168256  -.0943144  -.0266899  (N)
                           -.0906175  -.0285774  (P)
                           -.0906175  -.0069495  (BC)
  b_mgtransinc  50  -.1376936  .0013222  .0158715  -.1695885  -.1057987  (N)
                           -.1674831  -.1117099  (P)
                           -.1701275  -.1168307  (BC)
  Note: N = normal
        P = percentile
        BC = bias-corrected
```

5 Acknowledgment

I thank an anonymous reviewer for helpful comments and suggestions.

6 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.

López-Feldman, A., J. Mora, and J. E. Taylor. 2005. Does natural resource extraction mitigate poverty and inequality? Evidence from rural Mexico. Working paper.

Shorrocks, A. F. 1982. Inequality decomposition by factor components. *Econometrica* 50: 193–212.

Stark, O., J. E. Taylor, and S. Yitzhaki. 1986. Remittances and inequality. *Economic Journal* 96: 722–740.

Taylor, J. E., J. Mora, R. Adams, and A. López-Feldman. 2005. Remittances, inequality, and poverty. Evidence from rural Mexico. Working paper.
<http://repositories.cdlib.org/are/arewp/05-003>.

About the author

Alejandro López-Feldman is a Ph.D. student at the University of California–Davis in the Department of Agricultural and Resource Economics. His main research interests are in the areas of environmental, resource, and development economics, as well as in applied microeconomics.