dsginideco

Decomposition of inequality change into pro-poor growth and mobility components

Stephen P. Jenkins
ISER, University of Essex[†]
stephenj@essex.ac.uk

Philippe Van Kerm
CEPS/INSTEAD, Luxembourg[‡]
ISER, University of Essex
philippe.vankerm@ceps.lu

February 2009

Abstract This note describes `dsginideco`, a user-written Stata package implementing the decomposition of a change in inequality (measured by the generalized Gini coefficient) into progressivity/pro-poor growth and reranking components. The decomposition was proposed by Jenkins & Van Kerm (*Oxford Economic Papers*, 2006).

Keywords `dsginideco` ; Stata ; generalized Gini; pro-poor growth ; reranking
JEL Classification: C88; D31

1 Introduction

This note describes `dsginideco`, a Stata command that we have written to calculate the decomposition of a change in inequality (measured by a generalized Gini coefficient) into progressivity/pro-poor growth and reranking components. The decomposition was proposed by Jenkins & Van Kerm (2006). The command is available online for installation in net-aware Stata.[1] At the command prompt, type

```
net install dsginideco , from(http://medim.ceps.lu/stata)
```

or

```
ssc install dsginideco
```

2 The decomposition

Jenkins & Van Kerm (2006) showed that the change in income inequality between two time periods could be expressed in terms of two components, one representing the progressivity (pro-poorness) of income growth, and the other representing reranking. Inequality is measured using the generalized Gini coefficient, also known as the S-Gini, $G(X;\upsilon)$, where X is the distribution of income at a point in time, and $\upsilon > 1$ is a parameter. $G(X;\upsilon)$ is a distributionally-sensitive inequality index, with larger values of υ placing greater weight on inequality differences among poorer (lower ranked) observations. The conventional Gini coefficient corresponds to the case $\upsilon = 2$.

The decomposition is of the form:

$$\Delta(\upsilon) = R(\upsilon) - P(\upsilon)$$

[†]Institute for Social and Economic Research, University of Essex, Colchester, Essex CO4 3SQ, U.K. `http://www.iser.essex.ac.uk`

[‡]Centre d'Etudes de Populations, de Pauvreté et de Politiques Socio-Economiques/International Networks for Studies in Technology, Environment, Alternatives, Development. B.P. 48, L-4501 Differdange, Luxembourg. `http://www.ceps.lu`

[1]The latest version of the `dsginideco` package is 1.0.2 (of 2009-02-20). Stata 8.2 or later is required.

where

$$\Delta(\upsilon) = \mathrm{G}(\mathrm{X}^1;\upsilon) - \mathrm{G}(\mathrm{X}^0;\upsilon)$$

is the growth of the generalized Gini coefficient between period 0 and period 1. $\mathsf{R}(\upsilon)$ is a measure of reranking, and $\mathsf{P}(\upsilon)$ is a measure of the progressivity of income growth defined, respectively, as

$$\mathsf{R}(\upsilon) = \mathrm{G}(\mathrm{X}^1;\upsilon) - \mathrm{C}(\mathrm{X}^0,\mathrm{X}^1;\upsilon).$$

and

$$\mathsf{P}(\upsilon) = \mathrm{G}(\mathrm{X}^0;\upsilon) - \mathrm{C}(\mathrm{X}^0,\mathrm{X}^1;\upsilon),$$

where $\mathrm{C}(\mathrm{X}^0,\mathrm{X}^1;\upsilon)$ is the generalized Concentration coefficient of period 1 incomes against period 0 ranking. $\mathsf{P}(\upsilon)$ can be interpreted as an indicator of how much growth has benefited disproportionately to individuals towards the bottom of the distribution in the initial time period. $\mathsf{R}(\upsilon)$ captures how much a progressive income growth has lead to reranking between individuals, so that the net reduction in inequality is the difference between $\mathsf{P}(\upsilon)$ and $\mathsf{R}(\upsilon)$. $\mathsf{R}(\upsilon)$ may be interpreted as a measure of mobility (in the form of reranking) in its own right (Yitzhaki & Wodon, 2004). Realize that this decomposition is a panel data technique: computation of the $\mathsf{P}(\upsilon)$ and $\mathsf{R}(\upsilon)$ components requires two observations on income for a sample of individuals or households.

In an analysis of cross-country convergence in GDP, O'Neill & Van Kerm (2008) interpreted $\Delta(\upsilon)$ as a measure of 'σ-convergence' and $\mathsf{P}(\upsilon)$ as a measure of 'β-convergence', thereby reconciling the two concepts within a single framework.

3 The `dsginideco` command

3.1 Syntax

`dsginideco` *var0 var1* [*if*] [*in*] [*weight*] [`,` `parameters(`*numlist*`)` `format(`*string*`)`
`percentage` `percformat(`*string*`)` `kakwani`]

`aweight` and `fweight` are allowed; see [U] **11.1.6 weight − Weights**.
`by`, `bootstrap`, `jackknife` are allowed; see [U] **11.1.10 Prefix commands**.

`dsginideco` requires panel data, in wide form, on income in two time periods. *var0* contains the measure of income in the initial period for each observation. *var1* contains the measure of income in the final period for each observation. If the data are held in long form, time-series operators ([U] **11.4.3 Time-series varlists**) may be used to define *var0* or *var1*; see Examples below.

From the balanced sample of observations with non-missing income for both *var0* and *var1*, the command computes the generalized Gini coefficient for the initial and final years, the change in the index, and the decomposition components. Optionally, these estimates may also expressed as a fraction of the initial period generalized Gini index.

3.2 Options

`parameters(`*numlist*`)` specifies a value or values for υ. The default is 2, leading to a decomposition of the standard Gini coefficient. Multiple parameters may be specified but each value must be greater than 1.

`format(`*string*`)` specifies a format for the displayed results. The default is `%5.3f`.

`percentage` requests that decomposition factors be reported as fractions of the initial period generalized Gini coefficient.

`percformat(`*string*`)` used in conjunction with `percentage` specifies a format for results expressed as a fraction of initial period generalized Gini coefficient. The default is `%4.1f`.

kakwani requests reporting of the Kakwani-type measure of progressivity of income growth. (See Jenkins & Van Kerm (2006) for the definition.) This statistic is meaningful only when average income growth is not close to zero.

3.3 Saved results

Scalars

r(sgini0)	Initial period inequality index, $G(X^0;\upsilon)$
r(sgini1)	Final period inequality index, $G(X^1;\upsilon)$
r(dsgini)	Change in inequality, $\Delta(\upsilon)$
r(pi)	Average income growth
r(P)	P-component, $P(\upsilon)$
r(R)	R-component, $R(\upsilon)$
r(K)	Kakwani index of progressivity, if requested
r(N)	Number of observations
r(sum_w)	Sum of weights

Macros

r(var0)	Name of variable *var0*
r(var1)	Name of variable *var1*
r(paramlist)	Values of υ

Matrices

r(coeffs)	All estimated statistics
r(parameters)	Vector containing the values of υ

When the percentage option is specified, an additional set of results is returned, each prefixed by rel, containing the estimates expressed as a fraction of the initial period generalized Gini, e.g. r(reldsgini). Type return list after dsginideco to ascertain precisely which results are returned.

When multiple parameters are specified in option parameters, returned scalars contain estimates for the smallest parameter only. The complete set of estimated coefficients is available in the matrix r(coeffs).

4 Examples

We illustrate dsginideco using data from the National Longitudinal Survey of Youth, available from the Stata Press website. These are panel data and stored in long form, so we tsset the data and use a time series operator to construct *var0*. We examine year-on-year changes in wage inequality, pro-poor growth and mobility, pooling observations from successive pairs of years.

```
. cap use http://www.stata-press.com/data/r9/nlswork , clear
. tsset idcode year
       panel variable:  idcode (unbalanced)
        time variable:  year, 68 to 88, but with gaps
                delta:  1 unit
. gen w = exp(ln_wage)


. dsginideco L.w w
Decomposition of change in S-Gini coefficient of inequality
Average growth rate = 0.077
```

Parameter:	v=2
Initial S-Gini	0.245
Final S-Gini	0.266
Change	0.021
R-component	0.062
P-component	0.041

```
. dsginideco L.w w , percentage parameters(1.5 2 3 4) kakwani
Decomposition of change in S-Gini coefficient of inequality
Average growth rate = 0.077
```

Parameter:	v=1.5	v=2	v=3	v=4
Initial S-Gini	0.163	0.245	0.333	0.383
Final S-Gini	0.182	0.266	0.353	0.402
Change	0.020	0.021	0.020	0.019
R-component	0.047	0.062	0.082	0.097
P-component	0.028	0.041	0.062	0.078
K-index	0.386	0.580	0.865	1.098

Change, P- and R-components as percentage of initial S-Gini:

Parameter:	v=1.5	v=2	v=3	v=4
Change	12.1	8.6	6.0	5.0
R-component	29.0	25.4	24.5	25.4
P-component	16.9	16.9	18.5	20.4
K-index	237.7	236.9	259.6	286.4

```
. return list
scalars:
               r(relK) =  237.653905548376
               r(relR) =  29.03050118340291
               r(relP) =  16.94475564992893
          r(reldsgini) =  12.08574553347398
                 r(K) =  .3864254001295567
                 r(R) =  .047203613211711
                 r(P) =  .0275521833609778
            r(dsgini) =  .0196514298507332
            r(sgini1) =  .1822514928448157
            r(sgini0) =  .1626000629940825
                r(pi) =  .0767741422048413
             r(sum_w) =  10891
                 r(N) =  10891
macros:
          r(paramlist) : "1.5 2 3 4"
               r(var1) : "w"
               r(var0) : "L.w"
matrices:
           r(relcoeffs) :  4 x 4
              r(coeffs) :  6 x 4
          r(parameters) :  1 x 4
. matrix list r(coeffs)
r(coeffs)[6,4]
           param1     param2     param3     param4
sgini0  .16260006  .24467002  .33315148  .38339245
sgini1  .18225149   .2656148  .35325426  .40243734
 dgini  .01965143  .02094478  .02010278  .01904489
     R  .04720361  .06226445  .08175698  .09733694
     P  .02755218  .04131966   .0616542  .07829205
     K   .3864254   .5795173  .86471359  1.0980632
```

dsginideco reports point estimates, but does not compute the sampling covariance matrix of the decomposition components. However resampling-based inference may be implemented using Stata's built-in bootstrap (or jackknife) prefix, as in the following examples (see Sanchez, 2007). It is the user's responsibility to ensure that the resampling technique implemented with the prefix command

correctly reflects the original survey sampling design.

```
. gen newid = idcode

. tsset newid year
       panel variable:  newid (unbalanced)
        time variable:  year, 68 to 88, but with gaps
                delta:  1 unit

. bootstrap dG=r(dsgini) R=r(R) P=r(P)                    ///
>     , cluster(idcode) idcluster(newid) reps(250) nodots: ///
>    dsginideco L.w w if !mi(L.w) & !mi(w)

Warning:  Since dsginideco is not an estimation command or does not set
          e(sample), bootstrap has no way to determine which observations are
          used in calculating the statistics and so assumes that all
          observations are used.  This means no observations will be excluded
          from the resampling because of missing values or other reasons.

          If the assumption is not true, press Break, save the data, and drop
          the observations that are to be excluded.  Be sure that the dataset
          in memory contains only the relevant data.

Bootstrap results                                 Number of obs      =     10891
                                                  Replications       =       250

      command:  dsginideco L.w w
           dG:  r(dsgini)
            R:  r(R)
            P:  r(P)

                                (Replications based on 3700 clusters in idcode)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          dG |   .0209448   .0037565     5.58   0.000     .0135822    .0283074
           R |   .0622644   .0061849    10.07   0.000     .0501423    .0743866
           P |   .0413197   .0039909    10.35   0.000     .0334977    .0491416
------------------------------------------------------------------------------

. jackknife dG=r(dsgini) R=r(R) P=r(P)                    ///
>     , cluster(idcode) idcluster(newid) rclass nodots:   ///
>    dsginideco L.w w if !mi(L.w) & !mi(w)

Jackknife results                                 Number of obs      =     10891
                                                  Replications       =      3700

      command:  dsginideco L.w w if !mi(L.w) & !mi(w)
           dG:  r(dsgini)
            R:  r(R)
            P:  r(P)
         n():  r(N)

                                (Replications based on 3700 clusters in idcode)
------------------------------------------------------------------------------
             |             Jackknife
             |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          dG |   .0209448   .0037447     5.59   0.000     .0136029    .0282867
           R |   .0622644   .0045442    13.70   0.000     .0533551    .0711738
           P |   .0413197   .0036637    11.28   0.000     .0341365    .0485028
------------------------------------------------------------------------------
```

References

Jenkins, S. P. & Van Kerm, P. (2006), 'Trends in income inequality, pro-poor income growth and income mobility', *Oxford Economic Papers*, **58**(3):531–48.

O'Neill, D. & Van Kerm, P. (2008), 'An integrated framework for analysing income convergence', *The*

Manchester School, **76**(1):1–20.

Sanchez, G. (2007), 'How do I obtain bootstrapped standard errors with panel data?', Stata FAQ. Available at http://www.stata.com/support/faqs/stat/xt_boot.html.

Yitzhaki, S. & Wodon, Q. (2004), 'Inequality, mobility, and horizontal equity', in Y. Amiel & J. A. Bishop (eds.), *Research on Economic Inequality (*Studies on Economic Well-Being: Essays in Honor of John P. Formby*)*, vol. 12, JAI Press, Elsevier Science, 177–198.

Citation, liability, conditions of use

The program should work as described, but it is freely offered 'as-is' to the research community. Use at your own risk! Of course, we appreciate bug reports, as well as comments and suggestions (email philippe.vankerm@ceps.lu).

Please cite as:
Jenkins, S. P. and Van Kerm, P. (2009), '`dsginideco` – Decomposition of inequality change into pro-poor growth and mobility components', v1.0, CEPS/INSTEAD, Differdange, Luxembourg.

Acknowledgements

This work is part of the MeDIM project (*Advances in the Measurement of Discrimination, Inequality and Mobility*) supported by the Luxembourg 'Fonds National de la Recherche' (contract FNR/06/15/08) and by core funding for CEPS/INSTEAD by the Ministry of Culture, Higher Education and Research of Luxembourg. Jenkins's research is supported by core funding from the University of Essex and the U.K. Economic and Social Research Council for the Research Centre on Micro-Social Change and the United Kingdom Longitudinal Studies Centre.