The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
1-year subscription	$ 79	1-year subscription	$115
2-year subscription	$155	2-year subscription	$225
3-year subscription	$225	3-year subscription	$329
3-year subscription (electronic only)	$210	3-year subscription (electronic only)	$210
1-year student subscription	$ 48	1-year student subscription	$ 79
1-year university library subscription	$ 99	1-year university library subscription	$135
2-year university library subscription	$195	2-year university library subscription	$265
3-year university library subscription	$289	3-year university library subscription	$395
1-year institutional subscription	$225	1-year institutional subscription	$259
2-year institutional subscription	$445	2-year institutional subscription	$510
3-year institutional subscription	$650	3-year institutional subscription	$750

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The Stata Journal (2012)
12, Number 3, pp. 393–405

Adjusting for age effects in cross-sectional distributions

Ingvild Almås
Norwegian School of Economics and Business Administration and
University of Oslo
Bergen/Oslo, Norway
ingvild.almas@nhh.no

Tarjei Havnes
University of Oslo
Oslo, Norway
tarjei.havnes@econ.uio.no

Magne Mogstad
University College London
London, United Kingdom
m.mogstad@ucl.ac.uk

Abstract. Income and wealth differ over the life cycle. In cross-sectional distributions of income or wealth, classical inequality measures such as the Gini could therefore find substantial inequality even if everyone has the same lifetime income or wealth. We describe the adjusted Gini index (Almås and Mogstad, 2012, *Scandinavian Journal of Economics* 114: 24–54), which is a generalization of the classical Gini index with attractive properties, and we describe the `adgini` command, which provides the adjusted Gini index and the classical Gini index. The `adgini` command also provides options to produce other well-known age-adjusted inequality measures, such as the Paglin–Gini (Paglin, 1975, *American Economic Review* 65: 598–609) and the Wertz–Gini (Wertz, 1979, *American Economic Review* 69: 670–672), and provides efficient estimation of the classical Gini coefficient.

Keywords: st0266, adgini, inequality, life cycle, age adjustments, Gini coefficient, Paglin–Gini, Wertz–Gini

1 Introduction

Because of data availability, many researchers are forced to work on cross-sectional distributions of income and wealth. For example, all the frequently used datasets in both the Luxembourg Income Surveys and the Luxembourg Wealth Surveys are cross-sectional. This is problematic because both theoretical models and empirical results suggest a strong relationship between age and income and age and wealth holdings (see, for example, Davies and Shorrocks [2000]). Both relationships are firmly established as increasing to a certain midlife age and then decreasing thereafter.[1] Hence a snapshot of inequality within a country or other geographical area runs the risk of providing a misleading picture of the differences in lifetime wealth or income of its citizens. Because the income and wealth profiles differ across countries, the inequality ranking of countries may also be affected by differences in transitory income or wealth attributable to

1. The income profile is likely to have its peak earlier than the wealth profile.

life-cycle factors. For these reasons, it has long been argued that age adjustments of inequality measures based on cross-section data are necessary (see, for example, Atkinson [1971]).[2]

Almås and Mogstad (2012) propose the adjusted Gini (AG) index, a new method to adjust for age effects, which unlike existing methods considers that individuals differ both in age and in other wealth-generating factors. For example, an individual's education is strongly correlated not only with wealth but also with age. Existing methods (such as the Paglin and Wertz–Gini [WG]) assume that differences between age groups in the unconditional distribution represent age effects and will, therefore, eliminate not only wealth inequality attributable to age but also differences owing to wealth-generating factors correlated with age, such as education. By contrast, the AG index eliminates inequality due to age, yet preserves inequality arising from other factors. To this end, a multivariate regression model is used, allowing isolation of the net age effects and holding other determinants of wealth constant. Perfect equality for the AG measure requires that each individual receive a share of total wealth equal to the proportion that the individual would hold if all wealth-generating factors except age were the same for everyone in the population.

Similar procedures have been developed and used by Almås (2008), Almås et al. (2011), and Almås, Havnes, and Mogstad (2011). The first two articles focus on fairness and allow the isolation of the effect from factors other than age. The third article describes how age adjustments may influence trends in earnings inequality, focusing on Norway from 1967 to 2000. Note that the `adgini` command is general in the sense that it can be used to isolate the effects of any factor influencing income or wealth, not just the age factor.

The idea of an age-adjusted Gini index was first put forward in the seminal work of Paglin (1975). Numerous comments were written as responses to his article, among them the comment by Wertz (1979). While the Paglin–Gini (PG) is easy to implement, it fails to meet some attractive conditions met by Wertz's suggested measure (WG). However, WG fails to control for the correlation of other variables with age, because it takes the differences in mean wealth by age to represent the age effect.

Section 2 describes different age-inequality measures with specific focus on the AG index. Section 3 describes the `adgini` command, and section 4 provides examples of how the `adgini` command can be used and how age adjustment affects inequality results.

2 Age-adjusted inequality measures

The method underlying the AG index may be described as a three-step procedure. First, a generalization of the Gini formula is derived. Second, a multivariate regression model is used, allowing us to isolate the net age effects while holding other determinants of

2. For expositional convenience, we will from here consider inequality in wealth only. However, the method applies equally to income, earnings, or any other variable for which one is estimating inequality. For an application to earnings, see Almås, Havnes, and Mogstad (2011).

income or wealth constant (hereafter, just wealth). Third, the wealth distribution that characterizes perfect equality in age-adjusted wealth is determined. We describe the three steps below before showing that the AG index can be viewed as a generalization of the classical Gini coefficient (G).[3]

2.1 AG—a generalization of the Gini formula

Consider a society consisting of n individuals where every individual i is characterized by the pair $(w_i, \widetilde{w}_i)$. w_i denotes the actual wealth level, and $\widetilde{w}_i$ is an equalizing wealth level. If actual and equalizing wealth are the same for all individuals and if all individuals live equally long, then there is perfect equality of lifetime wealth in this society. As will be clear when we formally define the equalizing wealth level in section 2.3, the equalizing wealth is the same for all individuals belonging to the same age group in this society; it is a function of individual i's age but not of any other individual characteristics. If none of the wealth-generating factors (except age) are correlated with age, then the equalizing wealth is simply the mean wealth of each age group. Further, if there are no age effects on wealth, the equalizing wealth will be equal to the mean wealth for all individuals in the society.

The joint cross-sectional distribution Y of actual and equalizing wealth is given by

$$Y = \{(w_1, \widetilde{w}_1), (w_2, \widetilde{w}_2), \ldots, (w_n, \widetilde{w}_n)\}$$

Let Ξ denote the set of all possible joint distributions of actual and equalizing wealth such that the sum of actual wealth equals the sum of equalizing wealth. Suppose that the social planner imposes the following modified versions of the standard conditions on an inequality-partial ordering defined on the alternatives in Ξ, where $A \preceq B$ represents that there is at least as much age-adjusted inequality in B as in A.[4] Let μ denote the mean wealth of the population as a whole. Let the distributions of differences (Δ_i's) between actual wealth w_i and equalizing wealth $\widetilde{w}_i$ for the two distributions $[\Delta_i(A) = w_i(A) - \widetilde{w}_i(A)$ and $\Delta_i(B) = w_i(B) - \widetilde{w}_i(B)]$ be sorted in ascending order such that $\Delta_i \leq \Delta_{i+1}$.

Condition 1 ***Scale Invariance:*** *For any $a > 0$ and $A, B \in \Xi$, if $A = aB$, then $A \sim B$.*

Condition 2 ***Anonymity:*** *For any permutation function $\rho\colon n \to n$ and for $A, B \in \Xi$, if $\{w_i(A), \widetilde{w}_i(A)\} = \{w_{\rho(i)}(B), \widetilde{w}_{\rho(i)}(B)\}$ for all $i \in n$, then $A \sim B$.*

Condition 3 ***Unequalism:*** *For any $A, B \in \Xi$ such that $\mu(A) = \mu(B)$, if $\Delta_i(A) = \Delta_i(B)$ for every $i \in n$, then $A \sim B$.*

Condition 4 ***Generalized Pigou–Dalton:*** *For any $A, B \in \Xi$, if there exist two individuals s and k such that $\Delta_s(A) < \Delta_s(B) \leq \Delta_k(B) < \Delta_k(A)$, $\Delta_i(A) = \Delta_i(B)$ for all $i \neq s, k$, and $\Delta_s(B) - \Delta_s(A) = \Delta_k(A) - \Delta_k(B)$, then $A \succ B$.*

3. This section relies heavily on Almås and Mogstad (2012), and we refer interested readers to that work for further details.
4. See Almås et al. (2011) for analogous conditions imposed to study equality of opportunity.

Scale invariance states that if all actual and equalizing wealth levels are rescaled by the same factor, then the level of age-adjusted inequality remains the same. Anonymity implies that the ranking of alternatives should be unaffected by a permutation of the identity of individuals. Unequalism entails that the social planner is only concerned with how unequally each individual is treated, with "unequal" defined as the difference between actual and equalizing wealth. Finally, the generalized version of the Pigou–Dalton criterion states that any fixed transfer of wealth from an individual i to an individual j, where $\Delta_i > \Delta_j$, reduces age-adjusted inequality.

The generalized Gini formula is based on a comparison of the absolute values of the differences in actual and equalizing wealth between all pairs of individuals. It is defined as follows:

$$\text{AG}(Y) = \frac{\sum_j \sum_i |\,(w_i - \widetilde{w}_i) - (w_j - \widetilde{w}_j)\,|}{2\mu n^2} \tag{1}$$

The AG index satisfies conditions 1–4. These conditions are similar to those underlying G in all respects but one: the equalizing wealth is not given by the mean wealth in the society as a whole but instead depends on the age of the individuals.

2.2 Identifying the net age effects

Suppose that the wealth level of individual i at a given point in time depends on the age group a that i belongs to and on i's lifetime resources given as a function h of a vector X of individual characteristics:

$$w_i = f(a_i)h(X_i)$$

The functional form of f depends on the underlying model of wealth accumulation. In the simplest life-cycle model, there is no uncertainty: individuals earn a constant income until retirement age, and the interest rate, as well as the rate of time preference, is 0. In this model, the wealth of an individual increases until retirement and decreases afterward. If the earnings profile slopes upward, the model predicts borrowing in the early part of the life cycle. The fact that this is not always observed could be explained by credit market imperfections. Introducing lifetime uncertainty and noninsurable health hazard induces the elderly to hold assets for precautionary purposes, which reduces the rate at which wealth decreases during retirement. If the sole purpose of saving is to leave a bequest to children, individuals behave as if their horizons were infinite and wealth does not decline with age.

Empirically, we can specify a flexible, functional form of f, yielding the wealth-generating function[5]

$$\ln w_i = \ln f(a_i) + \ln h(X_i) = \delta_i + X_i' B \tag{2}$$

5. The default of the `adgini` command is the log-linear distribution, though `adgini` provides other distributions as options.

where δ_i gives the percentage wealth difference of being in the age group of individual i relative to some reference age group, holding all other variables constant. The `adgini` command will give an error message if negative values are used and will add one unit to observations with 0 values in the dependent variable. Because wealth may be negative, it is possible to adjust the location of the distribution by adding to each wealth observation a constant equal to the absolute value of the minimum wealth observation when estimating the log-linear specification.

We must emphasize that the objective of the estimation of (2) is not to explain as much variation as possible in wealth holdings, but simply to attain an empirically sound estimate of the effects of age on wealth, δ_i.

2.3 Defining equalizing wealth

To eliminate wealth differences attributable to age but preserve inequality arising from all other factors, the `adgini` command uses the so-called general proportionality principle proposed by Bossert (1995) and Konow (1996) and further studied by Cappelen and Tungodden (2010). The absence of age-adjusted inequality requires that any two individuals belonging to a given age group have the same wealth level. Moreover, in any situation where everyone has the same wealth-generating factors except age, there should be no lifetime wealth inequality.[6]

More formally, the equalizing wealth level of individual i depends on age and every other wealth-generating factor of all individuals in the society; it is formally defined as

$$\widetilde{w}_i = \frac{\mu n \sum_j f(a_i)h(X_j)}{\sum_k \sum_j f(a_k)h(X_j)} = \frac{\mu n e^{\delta_i}}{\sum_k e^{\delta_k}}$$

where e^{δ_k} gives the net age effect of belonging to the age group of individual k after integrating out the effects of other wealth-generating factors correlated with age. No age-adjusted inequality corresponds to every individual i receiving $\widetilde{w}_i$, which is the share of total wealth equal to the proportion of wealth that an individual from i's age group would hold if all wealth-generating factors except age were the same for everyone in the population. If there is no age effect on wealth, the equalizing wealth level is equal to the mean wealth level in the society.

6. In a study of income inequality in the United States, Bishop, Formby, and Smith (1997) use a method to make age adjustments that disregards that the underlying income function is not additively separable. First, they estimate a multiplicative separable income function, which can be expressed as $\ln Y = \alpha_0 + \beta \text{Age} + Z'\gamma + \epsilon$, where α_0 is a constant, Age is the age, and Z is a set of controls. Second, they use the prediction $\ln Y^* = \ln Y - \beta \text{Age}$ as their age-adjusted income measure. However, the net age effect is given by $dY/d\text{Age}$, which is generally different from $\beta = d \ln Y/d\text{Age} = (dY/Y)(1/d\text{Age})$, because Y is a function of Z. If Z is correlated with Age, then Bishop, Formby, and Smith's approach will fail to capture the net age effects.

2.4 Relationship to the classical Gini coefficient

From (1), we can see that the AG index is closely linked to G. Both measures are based on a comparison of the absolute values of the differences in the actual and equalizing wealth levels between all pairs of individuals. The distinguishing feature is how equalizing wealth is defined. For G, the equalizing wealth level is assumed to be μ. Perfect equality requires not only that individuals have equal lifetime wealth but also that individuals of all ages have the same wealth holding in any given year, which can be realized only if there is a flat age–wealth profile.

However, a flat age–wealth profile runs counter to consumption needs over the life cycle as well as to productivity variation depending on human capital investment and experience. Indeed, the relationship between wealth and age can produce wealth inequality at a given point in time even if everyone is completely equal in all respects except age. Because differences in transitory wealth even out over time, a snapshot of inequality produced by G runs the risk of producing a misleading picture of actual variation in lifetime wealth. In comparison, the AG index abandons the assumption of a flat age–wealth profile and allows equalizing wealth to depend on the age of the individuals. By doing this, the AG index purges the cross-sectional measure of inequality of its interage or life-cycle component. If $\widetilde{w}_i = \mu$ for all individuals in every age group, the age–wealth profile is flat and the AG index coincides with G. If there is a relationship between age and wealth, the AG index will, in general, differ from G.

To get further intuition on the similarities and differences between G and the AG index, we should see the correspondence between the standard representation of the Lorenz curve and a Lorenz curve expressed in differences between actual wealth and mean wealth in the society as a whole. Figure 1 displays standard and difference-based Lorenz curves for the same wealth distribution. The area between the standard Lorenz curve and the diagonal of the upper diagram (the line of equality) is identical to the area between the difference-based Lorenz curve and the horizontal axis (the line of equality) in the lower diagram. The classical Gini coefficient is in both cases equal to twice the area A between the Lorenz curve and the line of equality.

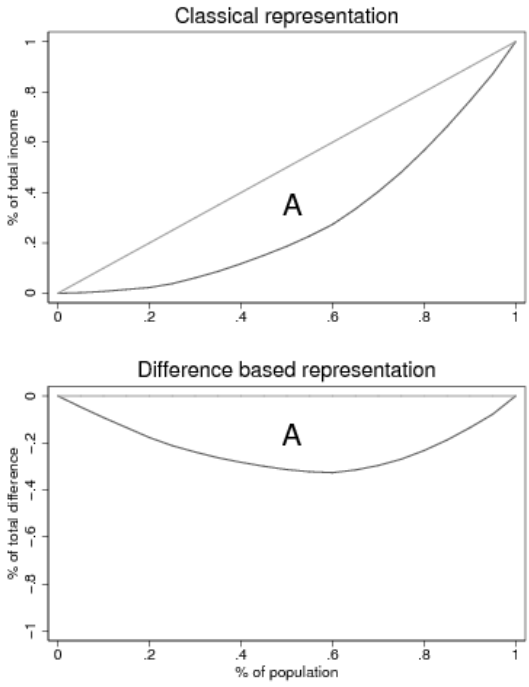

Figure 1. *Two representations of the standard Lorenz curve.* The figure displays two representations of the standard Lorenz curve: the classical representation relies on cumulative income shares, and the difference-based representation relies on cumulative shares of the difference between the average income and the actual income. The area A is the same in both panels.

In a similar vein, we can draw the age-adjusted Lorenz curve underlying the AG index, expressing the differences between actual wealth and the equalizing wealth in the population. And just as for G, the AG index is equal to twice the area between this difference-based Lorenz curve and the horizontal axis (line of equality). When drawing age-adjusted Lorenz curves, however, one orders individuals not by their wealth per se, as in figure 1, but by the difference between their actual wealth holdings and the equalizing wealth in their age group. Both G and the AG index reach their minimum value of 0 if all individuals receive their equalizing wealth. Moreover, both measures reach their maximum when the difference between actual and equalizing wealth is at its highest possible level. Specifically, G reaches its maximum value of 1 if one individual holds all the wealth. In comparison, the AG index reaches its maximum of 2 in the hypothetical situation where the equalizing wealth of the individual who has all the wealth is 0 and where the equalizing wealth of one of the individuals with no wealth is equal to the aggregate wealth in the economy. That G and the AG index range over

different intervals is therefore a direct result of their different views of perfect equality: age-adjusted inequality is not only due to differences in individual wealth holdings but also due to differences in equalizing wealth across individuals in different age groups.

2.5 Relationship to WG and PG

There are two distinguishing aspects of age-adjusted inequality measures. First, they hold different views on how equalizing wealth should be measured. Second, the formulas for calculating the differences between individuals' actual and equalizing wealth levels differ. The `adgini` command gives two alternative age-adjusted inequality measures as options: PG and WG. They both have the same objective as the AG index, namely, to purge the classical Gini coefficient applied to snapshots of wealth inequality of its interage or life-cycle component. In particular, the condition of a flat age–wealth profile is abandoned. We use the above conditions below to assess the properties of PG and WG and to characterize their relationship to the AG index.

Because of its close relationship to the AG index, we will first consider WG, which was proposed by Wertz (1979). WG can be expressed as follows:

$$\text{WG}(Y) = \frac{\sum_j \sum_i |(w_i - \mu_i) - (w_j - \mu_j)|}{2\mu n^2}$$

where μ_i and μ_j denote the mean wealth level of all individuals belonging to the age group of individuals i and j, respectively. Like the AG index, WG is based on a comparison of the absolute values of the differences in actual and equalizing wealth levels between all pairs of individuals and ranges over the interval $[0, 2]$. It also satisfies conditions 1–4. However, WG defines the equalizing wealth of an individual i as the unconditional mean wealth levels in i's age group, μ_i, and will therefore eliminate not only wealth inequality due to age but also differences due to wealth-generating factors correlated with age, such as education. The standard omitted-variables-bias formula tells us that WG will be equal to AG whenever age is uncorrelated with omitted wealth-generating factors. Hence, AG may be viewed as a generalization of WG, important in situations where omitted variables bias is a major concern.

Next consider the much-used PG, which can be expressed as

$$\text{PG}(Y) = \frac{\sum_j \sum_i (|w_i - w_j| - |\mu_i - \mu_j|)}{2\mu n^2}$$

where μ_i and μ_j denote the mean wealth level of all individuals belonging to the age group of individuals i and j, respectively. Applying the standard Gini decomposition, we can rewrite PG as

$$\text{PG} = \text{G} - G_b = \sum_i \theta_i G_i + R$$

where G_b represents the Gini coefficient that would be obtained if the earnings of each individual in every age group were replaced by the relevant age group mean μ_i; G_i is the Gini coefficient of earnings within the age group of individual i; θ_i is the weight

given by the product of this group's earnings share $n_i\mu_i/\mu n$ and population share n_i/n (n_i being the number of individuals in the age group of individual i); and R captures the degree of overlap in the earnings distributions across age groups (see, for example, Lambert and Aronson [1993]).[7]

Similarly to the case of WG, PG also defines the equalizing wealth of an individual i as the unconditional mean wealth level in i's age group, μ_i, disregarding that other wealth-generating factors are correlated with age.

In addition, PG is based on a comparison of differences in the absolute values of actual and equalizing wealth levels between all pairs of individuals, $|(w_i - w_j)| - |(\mu_i - \mu_j)|$. This violates the unequalism condition because $|(w_i - w_j)| - |(\mu_i - \mu_j)| = 0$ does not necessarily imply that $|(w_i - \mu_i) - (w_j - \mu_j)| = 0$.[8]

Because $|(w_i - w_j) - (\mu_i - \mu_j)|$ provides an upper bound for $|(w_i - w_j)| - |(\mu_i - \mu_j)|$, it follows that WG $\geq$ PG. As stated in proposition 1 in Almås and Mogstad (2012), PG will differ from WG if there is any age effect on wealth, provided that there is some within-age-group wealth variation. Moreover, overlap in the wealth distributions across age groups, that is, $R > 0$, is a sufficient condition for WG $>$ PG. A corollary is therefore that PG is likely to yield a different ranking than WG in situations where countries differ substantially in the degree of overlap.

This result speaks to a main controversy surrounding the PG, namely, whether R should be treated as an interage or a within-age-groups component.[9] Until recently, the issue was unsettled simply because little was known about the overlap term. Shorrocks and Wan (2005), for example, refer to R as a "poorly specified" element of the Gini decomposition. However, Lambert and Decoster (2005) provide a novel characterization of the properties of R, showing first that R unambiguously falls as a result of a within-group progressive transfer and second that R increases when the wealth holding in the poorer group is scaled up, reaching a maximum when means coincide. Lambert and Decoster (2005, 378) conclude: "The overlap term in R is at once a between-groups and a within-groups effect: it measures a between-groups phenomenon, overlapping, that is generated by inequality within groups." Therefore, $R = 0$ is necessary for PG to net out the interage component—and nothing but the interage component—from cross-sectional inequality measures.

7. Overlap implies that the wealth holding of the richest person in an age group with a relatively low mean wealth level exceeds the wealth holding of the poorest person in an age group with a higher mean wealth level; that is, $w_i < w_j$ and $\mu_i > \mu_j$ for at least one pair of individuals i and j.
8. See Almås and Mogstad (2012) for further discussion and a simple numerical example.
9. Nelson (1977) and others argue that R is part of interage inequality and should thus be netted out when constructing age-adjusted inequality measures. Paglin (1975), however, maintains that R is capturing within-group inequality and that PG is accurately defined.

3 The adgini command

3.1 Syntax

`adgini` *depvar* [*effectvars*] [*if*] [*in*] [, `controls(`*varlist*`)` `estname(`*string*`)` `equalizing(`*varname*`)` `all` `paglin` *regress_options*]

3.2 Description

The `adgini` command estimates alternative Gini coefficients for *depvar*, adjusting for *effectvars*, while holding `controls()` constant. `adgini` always estimates the classical Gini coefficient. If one or more *effectvars* are specified but `controls()` are not specified, `adgini` also estimates WG by default (Wertz 1979) or PG if the option `paglin` is activated (Paglin 1975). If `paglin` is activated, the Between–Gini (BG) is stored but not reported. If both *effectvars* and `controls()` are specified, `adgini` also estimates the AG. If the `all` option is activated, `adgini` estimates G, WG, PG, BG, and AG (when relevant).

3.3 Options

`controls(`*varlist*`)` specifies a set of control variables (correlated with *effectvars*) that are not to be adjusted for when calculating AG.

`estname(`*string*`)` requests that regression results be stored in memory under the name specified.

`equalizing(`*varname*`)` specifies a variable in memory containing equalizing values to be used in calculating AG. In this case, *effectvars* and `controls()` are not used in the estimation even if specified.

`all` requests calculation of all relevant Ginis (G, WG, PG, BG, and AG).

`paglin` requests calculation of PG (the default is WG).

regress_options are any of the options documented in [R] **regress**.

3.4 Saved results

`adgini` saves the following in `r()`:

Scalars

`r(N)`	number of observations	`r(pg)`	PG
`r(gini)`	Gini	`r(bg)`	BG
`r(wg)`	WG	`r(ag)`	AG

Macros

`r(cmd)`	`adgini`	`r(equalizing)`	name of variable with equalizing values
`r(depvar)`	name of dependent variable		
`r(effectvars)`	list of variables in *effectvars*	`r(regoptions)`	`regress` options
`r(controls)`	list of variables in `controls()`		

4 Examples

We provide two examples. The first example is a straightforward application of the method and the `adgini` command, namely, to correct for age effects without simultaneously eliminating effects from education (which is likely to be correlated with age). The second example illustrates the generality of the procedure as it demonstrates how we can use the `adgini` command to show the dispersion of prices corrected for quality effects without simultaneously eliminating the effect from other variables correlated with quality.

▷ Example: Income inequality (mother's labor income)

The most standard use of inequality indices concerns income distributions. In the current example, we use an instructional dataset from Wooldridge (2001) on the labor income of mothers. We are interested in the inequality of labor income when we adjust for the individual age. However, we do not want to remove the effect of education, which is likely correlated with age. To account for age and education in the most flexible way, we control for indicator variables for every value of age and education by using factor variables.

```
. use http://fmwww.bc.edu/ec-p/data/wooldridge2k/LABSUP
. adgini labinc i.age, controls(i.educ) all
==================================================
 Gini:   .654
 Between-Gini:   .114
 Paglin:   .539
 Wertz:   .666
 AG:   .654
==================================================
```

◁

▷ Example: Gini as a measure of dispersion

Gini coefficients may also be used as a measure of dispersion in contexts other than income or wealth. For instance, we may be interested in summarizing the dispersion of prices for comparable goods. However, we may not want our measure of price dispersion to reflect differences in the observable quality between goods. `adgini` can be used to calculate such a measure of dispersion: we enter price as *depvar*, quality variables as *effectvars*, and nonquality variables correlated with quality in `controls()`. The impact of quality variables on price should be properly identified in the empirical model.

```
. sysuse auto
(1978 Automobile Data)
. adgini price mpg length turn trunk, controls(foreign weight) all
==================================================
 Gini:   .232
 Between-Gini:   .133
 Paglin:   .099
 Wertz:   .215
 AG:   .305
==================================================
```

◁

5 Concluding remarks

We have provided a description of the method for age adjustment in cross-sectional distributions and of the `adgini` command, which provides corresponding inequality statistics in Stata. As a by-product, the `adgini` command provides a faster estimation of the classical Gini coefficient than do the existing algorithms, using Stata's built-in matrix language, Mata. We believe that the `adgini` command will serve as a useful tool for statistical bureaus and individual researchers studying wealth, earnings, or income distributions.

6 Acknowledgments

This article is part of the research activities at the ESOP Centre in the Department of Economics at the University of Oslo. ESOP is supported by the Research Council of Norway. We would like to thank ESOP and the Research Council for the opportunity to work on this article.

7 References

Almås, I. 2008. Equalizing income versus equalizing opportunity: A comparison of the United States and Germany. In *Research on Economic Inequality*, ed. J. Bishop and B. Zheng, vol. 16, 129–156. Bingley, UK: Emerald.

Almås, I., A. W. Cappelen, J. T. Lind, E. Ø. Sørensen, and B. Tungodden. 2011. Measuring unfair (in)equality. *Journal of Public Economics* 95: 488–499.

Almås, I., T. Havnes, and M. Mogstad. 2011. Baby booming inequality? Demographic change and earnings inequality in Norway, 1967–2000. *Journal of Economic Inequality* 9: 629–650.

Almås, I., and M. Mogstad. 2012. Older or wealthier? The impact of age adjustment on wealth inequality. *Scandinavian Journal of Economics* 114: 24–54.

Atkinson, A. B. 1971. The distribution of wealth and the individual life-cycle. *Oxford Economic Papers* 23: 239–254.

Bishop, J. A., J. P. Formby, and W. J. Smith. 1997. Demographic change and income inequality in the United States, 1976–1989. *Southern Economic Journal* 64: 34–44.

Bossert, W. 1995. Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences* 29: 1–17.

Cappelen, A. W., and B. Tungodden. 2010. Fairness and the proportionality principle. Discussion Paper 31/07, Department of Economics, Norwegian School of Economics.

Davies, J. B., and A. F. Shorrocks. 2000. The distribution of wealth. In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon, vol. 1, 605–675. Amsterdam: North-Holland.

Konow, J. 1996. A positive theory of economic fairness. *Journal of Economic Behavior & Organization* 31: 13–35.

Lambert, P. J., and J. R. Aronson. 1993. Inequality decomposition analysis and the Gini coefficient revisited. *Economic Journal* 103: 1221–1227.

Lambert, P. J., and A. Decoster. 2005. The Gini coefficient reveals more. *Metron* LXIII: 373–400.

Nelson, E. 1977. The measurement and trend of inequality: Comment. *American Economic Review* 67: 497–501.

Paglin, M. 1975. The measurement and trend of inequality: A basic revision. *American Economic Review* 65: 598–609.

Shorrocks, A., and G. Wan. 2005. Spatial decomposition of inequality. *Journal of Economic Geography* 5: 59–81.

Wertz, K. L. 1979. The measurement of inequality: Comment. *American Economic Review* 69: 670–672.

Wooldridge, J. M. 2001. Labsup. Instructional Stata datasets for econometrics labsup, Department of Economics, Boston College.

About the authors

Ingvild Almås is an associate professor in the Department of Economics at the Norwegian School of Economics.

Tarjei Havnes is a postdoctoral fellow in the Department of Economics at the University of Oslo.

Magne Mogstad is an assistant professor in the Department of Economics at the University College London.