Rolf Aaberge, Steinar Bjerve and
Kjell Doksum

Decomposition of Rank-Dependent Measures of Inequality by Subgroups

Abstract:

The purpose of additive subgroup decomposition is to study the relationship between overall inequality and inequality within and between population subgroups defined by variables like gender, age, education and region of residence. As opposed to the inequality measures that are additively decomposable, the so-called generalized entropy family of inequality measures, the Gini coefficient does not admit decomposition into within- and between-group components but does also require an interaction (overlapping) term. The purpose of this paper is to introduce an alternative decomposition method that can be considered to be a parallel to Lerman and Yitzhaki's (1985) elasticity approach for decomposing the Gini coefficient by income sources, which means that the elasticity of the Gini coefficient with respect to various income components is treated as the basic quantities of the decomposition method. Thus, rather than decomposing the Gini coefficient or any other inequality measure into a within-inequality term, a between-inequality term and eventually an interaction term, the basic quantities of the introduced method are the effects of marginal changes in variables that are used to specify the population subgroups.

Address: Rolf Aaberge, Statistics Norway, Research Department. E-mail: rolf.aaberge@ssb.no

Steinar Bjerve, Department of Mathematics, University of Oslo.
E-mail: steinar@math.uio.no

Kjell Doksum, Department of Statistics, University of Wisconsin-Madison, USA.
E-mail: doksum@stat.wisc.edu

1. Introduction

The most widely used measure of income inequality is the Gini coefficient, which is defined equal to twice the area between the Lorenz curve and its equality reference[1]. The simple and direct relationship between the Gini coefficient and the Lorenz curve appears to be the major reason for its popularity in applied work. However, since empirical analyses of income inequality normally deals with issues that require use of decomposition methods numerous proposals on how to decompose the Gini coefficient by income sources as well as by subgroups has occurred in the literature[2]. The purpose of subgroup decomposition is to study the relationship between overall inequality and inequality within and between population subgroups defined by variables like gender, age, education and region of residence[3]. As opposed to the inequality measures that are additively decomposable, the so-called generalized entropy family of inequality measures, the Gini coefficient does not admit decomposition into within- and between-group components. However, by adding an extra term that captures the overlap between the marginal income distributions of subgroups it can be demonstrated that the Gini coefficient can be decomposed into three terms, the within-group term, the between-group term and an interaction term[4]. Note that the interaction term vanishes when there is no overlapping of income ranks between income units belonging to different subgroups; i.e. when the income distributions of subgroups do not overlap. However, a number of alternative approaches for decomposing the Gini coefficient and other measures of inequality by subgroups could be defined, see Shorrocks (1984). The purpose of this paper is to introduce a new method that can be considered to be a parallel to Lerman and Yitzhaki's (1985) elasticity approach for decomposing the Gini coefficient by income sources, which means that the elasticity of the Gini coefficient with respect to various income components is treated as the basic quantities of the decomposition method. Thus, we turn the focus from decomposing the Gini coefficient or any other inequality measure into a within-inequality term, a between-inequality term and eventually an interaction (overlapping) term to the effects of marginal changes in the variables that are used to specify the population subgroups.

---

[1] See Giorgi (1990) for a bibliographical portrait of the Gini coefficient.

[2] See e.g. Rao (1969), Kakwani (1977, 1980), Lerman and Yitzhaki (1985), Chakravarty (1990) and Silber (1993) for useful discussions on decomposing the Gini coefficient by income sources.

[3] See e.g Shorrocks (1984).

[4] More on the derivation and interpretation of the subgroup decomposition of the Gini coefficient, see Bhatacharya and Mahalanonis (1967), Piesch (1975), Silber (1989), Yitzhaki (1994), Yitzhaki and Lerman (1991), Lambert and Aronson (1993) and Dagum (1997).

2. Decomposition of Lorenz curves and rank-dependent measures of inequality

Let Y be a positive, continuous random variable representing wage or income, $\mathbf{X}$ a random covariate vector. Leaving out the influence of $\mathbf{X}$, the overall Lorenz curve is

$$L(u) = \frac{1}{\mu} \int_0^1 I[t \le u] F^{-1}(t) dt = \frac{1}{\mu} \int_0^\infty I\left[y \le F^{-1}(u) \right] y \, dF(y)$$

where $\mu = \int_0^\infty y \, dF(y)$ is the mean of Y, F^{-1} denotes the left inverse of the distribution function F of Y, and I is the indicator function. L(u) gives the proportion of the total amount of income that is owned by the 100 u poorest percent of the population. We extend this definition to include the influence of covariates by considering the proportion of the total amount of income that is owned by the subpopulation with covariate values $\mathbf{x}$ and with income below the u'th quantile in the entire population. To this end we define the *pseudo-Lorenz regression curve* as

(2.1)
$$\Lambda(u|\mathbf{x}) = \frac{1}{\mu} E\left\{ Y I\left[Y \le F^{-1}(u) \right] \middle| \mathbf{X} = \mathbf{x} \right\} = \frac{1}{\mu} \int_0^\infty y I\left[y \le F^{-1}(u) \right] dF(y|\mathbf{x})$$
$$= \int y \left\{ I\left[y \le F^{-1}(u) \right] \middle/ \int_0^1 F^{-1}(u) du \right\} dF(y|\mathbf{x}), \quad 0 \le u \le 1.$$

where $F(y|\mathbf{x})$ denotes the distribution function of Y given $\mathbf{X} = \mathbf{x}$. Although this curve differs from the standard Lorenz curve it has the nice property that it is a decomposition of the Lorenz curve in the sense that its expected value equals the Lorenz curve for the total population, i.e. by using the iterated expectation theorem, see Bickel and Doksum (2001), we find

(2.2)
$$E\left[\Lambda(u|\mathbf{X}) \right] = L(u).$$

As (2.1) shows, this definition of Lorenz regression aggregates incomes from the subgroup with covariate vector $\mathbf{x}$, but uses $F^{-1}(u)$ as a *common* reference when computing proportions[5]. This reference quantile $F^{-1}(u)$ is the u'th quantile of the overall income distribution F(y) which is obtained by averaging out $\mathbf{x}$, that is, $F(y) = E\left[F(y|\mathbf{X}) \right]$. If $\mathbf{X}$ is used to partition the sample space into distinct categories $C_1, ..., C_s$ with probabilities $P(C_j) = P(\mathbf{X} \in C_j)$, $j = 1, ..., s$, then (2.2) becomes

---

[5] See Aaberge, Bjerve and Doksum (2005) who have used conditional Lorenz curves for deriving a regression framework for the Lorenz curve and the Gini coefficient.

$$L(u) = \sum_{j=1}^{s} P(C_j) \Lambda(u|C_j)$$

where

(2.3)
$$\Lambda(u|C_j) = \frac{1}{\mu} E\{ Y I[Y \le F^{-1}(u)] | \mathbf{x} \in C_j\} = \frac{\mu(C_j)}{\mu} L(F_j(F^{-1}(u))|C_j),$$

and $L(\cdot|C_j)$ is the Lorenz curve for sub-population C_j. Note that $\sum P(C_j) \Lambda(1|C_j) = 1$, but

$\Lambda(1|C_j) = \frac{\mu(C_j)}{\mu} \ne 1$ except when $\mu(C_j) = \mu$. Thus, the above decomposition of the Lorenz curve

gives a method for identifying the contribution to overall inequality from each subgroup, where the subgroup contributions can be expressed as the product of three components; the proportion of the population that belong to the subgroup, the ratio between the subgroup mean income and the overall mean income and an interaction component that depends on income inequality within the subgroup as well as the relative location of the subgroup distribution.

Similar to (2.3) for the discrete case we get the following expression for the continuous case,

(2.4)
$$\Lambda(u|\mathbf{x}) = \frac{\mu(\mathbf{x})}{\mu} L(g(u)|\mathbf{x})$$

where $g(u) = F(F^{-1}(u)|\mathbf{x})$ and $L(\cdot|\mathbf{x})$ is the (conditional) Lorenz curve for $F(y|\mathbf{x})$.

To summarize the information provided by the pseudo Lorenz curve $\Lambda(u|\mathbf{x})$ we may use the *pseudo-Gini coefficient*[6] defined by

(2.5)
$$\Gamma(\mathbf{x}) = 2 \int_0^1 [u - \Lambda(u|\mathbf{x})] du = \frac{1}{\mu} E\{ Y[2F(Y) - 1] | \mathbf{x}\},$$

or alternatively any member of the following family of pseudo inequality measures

(2.6)
$$\Psi_P(\mathbf{x}) = 1 + \int_0^1 P''(u) \Lambda(|\mathbf{x}) du = E\{ Y[1 - P'(F(Y))] | \mathbf{x}\},$$

---

[6] Kakwani (1980) introduced a similar definition in cases where $\mathbf{x}$ is a vector of discrete variable. See also Mahalanobis (1960).

where the weight-function P' is the derivative of a concave function P defined on the unit interval that satisfies the conditions $P(0) = 0$, $P(1) = 1$ and $P'(1) = 0$. Note that the unconditional counterpart of (2.6) is the family of rank-dependent measures of inequality introduced by Mehran (1976)[7]. By inserting for $P(u) = 2u - u^2$ in (2.5) we find that $\Psi_P(\mathbf{x}) = \Gamma(\mathbf{x})$. As for the pseudo-Lorenz curve we find the following convenient aggregation property for the pseudo inequality measures $\Psi_P(\mathbf{x})$

$$(2.7) \qquad J_P = E\left[\Psi_P(\mathbf{X})\right] = EY\left[1 - P'(F(Y))\right] = 1 - \int_0^1 P'(u)dL(u) = 1 - \frac{1}{\mu}\int_0^1 P'(u)F^{-1}(u)du \;.$$

As demonstrated by Aaberge (2000) the Gini coefficient attaches an equal weight to a given transfer of income irrespective of where it takes place in the income distribution, as long as the income transfer occurs between individuals with the same difference in ranks. Thus, in general the Gini coefficient favors neither the lower nor the upper part of the Lorenz curve. To supplement the information provided by the Gini coefficient it might be relevant to use the Bonferroni coefficient[8] defined by

$$(2.8) \qquad B = \int_0^1 \left[1 - u^{-1}L(u)\right]du = 1 + \frac{1}{\mu}E\left[Y\log F(Y)\right]$$

and the pseudo-Bonforroni coefficient defined by

$$(2.9) \qquad B(\mathbf{x}) = \int_0^1 \left[1 - u^{-1}\Lambda(u|\mathbf{x})\right]du = 1 + \frac{1}{\mu}E\left\{\left[Y\log F(Y)\right]|\mathbf{x}\right\}$$

Note that B and Б($\mathbf{x}$) corresponds to J_P and Ψ_P for $P(u) = u(1 - \log u)$. As demonstrated by Aaberge (2000) the Bonferroni coefficient B satisfies Mehran's principle of positional transfer sensitivity[9] for any distribution function F and Kolm's principle of diminishing transfers for all F for which $\log F(x)$ is strictly concave. Thus, B is particular sensitive to transfers that occur in the lower part of the income distribution for logconcave distribution functions.

As suggested in Section 1 the main purpose of this paper is not to focus attention on the various components defined by the covariable vector $\mathbf{x}$ in cases where $\mathbf{x}$ is a vector of discrete variables, but to treat $\mathbf{x}$ as a vector of continuous variables and develop a framework that can be

---

[7] Mehran (1976) introduced the J_P-family by relying on descriptive arguments, whereas alternative normative motivations of the J_P-family and various subfamilies of the J_P-family have been provided by Donaldson and Weymark (1980, 1983), Weymark (1981), Yaari (1987,1988), Ben Porath and Gilboa (1994) and Aaberge (2001).

[8] For a discussion of the Bonferroni coefficient see D'Addario (1936), Nygård and Sandström (1981), Aaberge (1982, 2000) and Giorgi (1998). A poverty measure derived from the Bonferroni coefficient has been introduced by Giorgi (2001).

[9] See also Nygård and Sandstrøm (1981) and Giorgi (1998).

considered to provide similar information as the decomposition method in a situation with discrete variables. To this end we introduce the regression coefficients of the regression functions (2.1), (2.5) and (2.6) as quantities that provide information on the influence of covariates on overall inequality.

3. Measuring the effect of covariates on rank-dependent measures of inequality

By exploiting the parallel with the quantile regression approach, Aaberge, Bjerve and Doksum (2005) developed a regression framework for the conditional Lorenz curve, the conditional Gini coefficient and conditional rank-dependent measures of inequality, which can be used to examine the influence of covariates $\mathbf{x}$ on income inequality in the conditional distribution ($F(y|\mathbf{x})$) of Y given given $\mathbf{X} = \mathbf{x}$. However, sine the overall Lorenz curve and the overall Gini coefficient will not be attained by averaging out the covariates in the conditional Lorenz curve and the conditional Gini coefficient, the effects of covariates on the conditional Lorenz curve and the conditional Gini coefficient do not immediately carry over to the overall Lorenz curve and the overall Gini coefficient. Thus, the (conditional) Lorenz and Gini regression coefficients are not the appropriate quantities when focus is turned to the effects of covariates on overall inequality. To this end it appears more relevant to consider the regression coefficients of the pseudo-Lorenz curve and the pseudo-Gini coefficient introduced in Section 2. The pseudo-Lorenz regression coefficient curves are defined by

$$(3.1) \qquad \lambda_j(u; \mathbf{x}) = \frac{\partial \Lambda(u|\mathbf{x})}{\partial x_j}, \, 0 \le u \le 1, \, j = 1, 2, ..., s,$$

and can be considered as measures of the relative importance of the covariate x_j on income inequality[10]. They show how much a small perturbation of x_j for j=1, 2, …s changes the pseudo-Lorenz curves and allows the effects of the covariates to depend on whether the response is located in the lower, the central or the upper segment of the income distribution. Similarly as for the quantile regression coefficients curves it may be useful to summarize the pseudo-Lorenz regression coefficient curves across the covariates by

$$(3.2) \qquad \lambda_j(u) = E\lambda_j\left(u; \mathbf{X}\right), \, 0 \le u \le 1, \, j = 1, 2, ..., s.$$

---

[10] A similar approach for quantile regression was introduced by Chaudhury et al. (1997).

Note that $\lambda_j(u)$ gives the average change of the pseudo-Lorenz curves due to small change in the j'th covariate when the remaining covariates are first kept fixed, then averaged out. We call $\lambda_j(\cdot)$ the j'th marginal pseudo-Lorenz curve.

To complete the summarization of the pseudo-Lorenz regression coefficients provided by $\lambda_j(u)$ a summary measure that captures the variation across quantiles will be introduced. To this end we may use the pseudo-Gini coefficient as a summary measure of the information content of the pseudo-Lorenz curve. The pseudo-Gini regression coefficients that correspond to (3.1) are defined by

$$(3.3) \qquad \gamma_j(\mathbf{x}) = \frac{\partial \Gamma(\mathbf{x})}{\partial x_j} = -2 \int_0^1 \lambda_j(u, \mathbf{x}) \, du, \quad j = 1, 2, \ldots, s.$$

Moreover, by summarizing over $\mathbf{x}$ we get

$$(3.4) \qquad \gamma_j = E \gamma_j(\mathbf{X}) = -2 \int_0^1 \lambda_j(u) \, du, \quad j = 1, 2, \ldots, s.$$

The corresponding pseudo-Bonferroni summary measures are given by

$$(3.5) \qquad b_j(\mathbf{x}) = \frac{\partial B(\mathbf{x})}{\partial x_j}, \quad b_j = E b_j(\mathbf{X}).$$

Since alternative methods for summarizing the pseudo-Lorenz regression coefficients may be called for, we introduce the Ψ_P-regression coefficients derived from the pseudo-inequality measures defined by (2.5),

$$(3.6) \qquad \xi_{jP}(\mathbf{x}) = \frac{\partial \Psi_P(\mathbf{x})}{\partial x_j} = \int_0^1 P''(u) \lambda_j(u, \mathbf{x}) \, du, \quad j = 1, 2, \ldots, s,$$

where P'' is the second derivative of the weight-function P. By summarizing over $\mathbf{x}$ we get

$$(3.7) \qquad \xi_{jP} = E \xi_{iP}(\mathbf{X}) = \int_0^1 P''(u) \lambda_j(u) \, du, \quad j = 1, 2, \ldots, s.$$

Note that $P(u) = 2u - u^2$ is the P-function that corresponds to the Gini coefficient, whilst $P(u) = u(1 - \log u)$ corresponds to the Bonferroni coefficient $(P''(u) = -1/u)$.

4. Estimation

We have considered a variety of maps $m : R^s \to R$ that measure inequality in income Y as a function of covariates $\mathbf{x} \in R^s$. These surfaces m(·), which are referred to as "curves" in the literature and this paper, can not be displayed effectively, nor estimated efficiently unless the sample sizes are enormous. For this reason we turn to summary measures: The average derivative nonparametric parameter is the gradient vector

$$\text{(4.1)} \qquad \nabla m(\mathbf{x}) = \left(\left[\frac{\partial}{\partial x_j} m(\mathbf{x}) \right], j = 1, \ldots, s \right)^T$$

In the case of single index models, $\nabla m(\mathbf{x})$ is proportional to the single index parameter vector.

Average Derivative Estimates (ADE's) have been proposed and analysed by Stoker (1986), Härdle and Stoker (1989), Härdle et al. (1993), Chaudhury et al. (1997), and Hristache et al. (2001), among others. Related work on projection pursuit regression appears in Friedman and Stuetzle (1981) and Hall (1989). One basic idea is to estimate the gradient ∇m locally near a sample point $\mathbf{x}_i$ by using locally weighted least squares. That is, use $\hat{\nabla} m$, where

$$
\text{(4.2)} \quad
\begin{aligned}
\begin{pmatrix} \hat{m}(\mathbf{X}_i) \\ \hat{\nabla} m(\mathbf{X}_i) \end{pmatrix} &= \operatorname*{arg\,min}_{a \in R,\ \boldsymbol{\beta} \in R^s} \sum_{j=1}^{n} \left\{ V_j - \left[a + \boldsymbol{\beta}^T \left(\mathbf{X}_j - \mathbf{X}_i \right) \right] \right\}^2 K\left(\frac{\left| \mathbf{X}_j - \mathbf{X}_i \right|}{h^2} \right) \\
&= \left\{ \sum_{j=1}^{n} \begin{pmatrix} 1 \\ \mathbf{X}_{ij} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{X}_{ij} \end{pmatrix}^T K\left(\frac{\left| \mathbf{X}_{ij} \right|^2}{h^2} \right) \right\}^{-1} \sum_{j=1}^{n} V_j \begin{pmatrix} 1 \\ \mathbf{X}_{ij} \end{pmatrix} K\left(\frac{\left| \mathbf{X}_{ij} \right|^2}{h^2} \right).
\end{aligned}
$$

Here h is a tuning parameter selected using the data, $\mathbf{X}_{ij} = \mathbf{X}_j - \mathbf{X}_i, |\cdot|$, is Euclidean distance, and the basic data $\left\{ (\mathbf{X}_1, V_1), \ldots, (\mathbf{X}_n, V_n) \right\}$ is assumed to be i.i.d. The proceeding references give various modifications of this basic formula in order to deal with regions with sparse data.

To use these methods we need further specification. We have considered the following three m's:

$$
\text{(4.3)} \quad
\begin{aligned}
m_1(u; \mathbf{x}) &= \Lambda(u | \mathbf{x}) = \mu^{-1} E\left[Y I[F(Y) \le u] | \mathbf{x} \right], \\
m_2(\mathbf{x}) &= \Gamma(\mathbf{x}) = 1 - 2\mu^{-1} \left[E(Y | \mathbf{x}) - E(Y F(Y) | \mathbf{x}) \right], \\
m_3(\mathbf{x}) &= B(\mathbf{x}) = 1 + \mu^{-1} E\left[Y \log F(Y) | \mathbf{x} \right].
\end{aligned}
$$

Thus, we need ADE's for the four cases where $V = Y\,I\big[F(Y) \le u\big]$, $V = Y$, $V = Y\,F(Y)$ and $V = Y\log F(Y)$.

Because F is unknown, we need to replace $F(Y_i)$ by its empirical version $\hat{F}(Y_i) = i/n$, where the incomes $\{Y_i\}$ have been arranged in increasing order and $\mathbf{X}_i$ now denotes the covariate vector that belongs with the i'th ordered Y. For ease of interpretation and display the ADE algorithms require that each X_{ij} in the sample have the sample mean $\bar{X}_j$ subtracted and be divided by the sample standard deviation s_j, $j = 1,...,s$. Our curves require an estimate of $\mu = E(Y)$, which we take as $\hat{\mu} = \bar{Y}$.

We label the outputs from the ADE algorithms as $\nabla m_{kj}(\mathbf{X}_i)$, $k = 1,2,3$, $j = 1,...,s$, $i = 1,...,n$. Then our estimates are

$$\hat{\lambda}_j(u) = n^{-1}\sum_{i=1}^{n} \nabla m_{1j}(u;\mathbf{X}_i) \text{ (Lorenz curve in direction } X_j)$$

(4.4)
$$\hat{\gamma}_j = n^{-1}\sum_{i=1}^{n} \nabla m_{2j}(\mathbf{X}_i) \text{ (Gini coefficient in direction } X_j)$$
(4.4)

$$\hat{b}_j = n^{-1}\sum_{i=1}^{n} \nabla m_{3j}(\mathbf{X}_i) \text{ (Bonferroni coefficient in direction } X_j)$$

When there is only one covariate X, estimation is more straightforward. In the case of $\Lambda(u|x)$ we can apply any nonparametric regression estimator to the data $(X_1, V_1(u),...,X_n, V_n(u))$ where

(4.5)
$$V_i(u) = I\big(i \le [un]\big)Y_i$$

where $[\cdot]$ is the greatest integer function. One simple such estimator would be

(4.6)
$$\hat{\Lambda}(u|x) = \frac{\sum_{i=1}^{n} V_i(u)K_h(X_i - x)}{\bar{Y}\sum_{i=1}^{n} K_h(X_i - x)}$$

where $K_h(u) = h^{-1}K(u/h)$, $K(u)$ is a kernel on R with $\int K(u)du = 1$ and h>0 a tuning parameter.

The Gini regression index can be estimated as

(4.7)
$$\hat{\Gamma}(x) = 1 - 2(\bar{Y})^{-1}\hat{\mu}_G(x)$$

where

$$(4.8) \qquad \hat{\mu}_G(x) = \frac{\displaystyle\sum_{i=1}^{n} \left(1 - \frac{i}{n+1}\right) Y_i\, K_h\left(X_i - x\right)}{\displaystyle\sum_{i=1}^{n} K_h\left(X_i - x\right)}$$

Here the $\{Y_i\}$ are in increasing order and X_i is the covariate value the case with ordered response Y_i. Similarly, the Bonferroni regression index can be estimated as

$$(4.9) \qquad \hat{B}(x) = 1 + \left(\overline{Y}\right)^{-1} \hat{\mu}_B(x)$$

where

$$(4.10) \qquad \hat{\mu}_B(x) = \frac{\displaystyle\sum \log\left(\frac{i}{n+1}\right) Y_i\, K_L\left(X_i - x\right)}{\displaystyle\sum K_L\left(X_i - x\right)}$$

References

Aaberge, R. (1982): Om måling av ulikskap. Rapporter 82/9, Statistisk sentralbyrå (In Norwegian).

Aaberge, R. (2000): Characterizations of Lorenz Curves and Income Distributions. *Social Choice and Welfare* **17**, 639-653.

Aaberge, R. (2001): Axiomatic Characterization of the Gini Coefficient and Lorenz Curve Orderings. *Journal of Economic Theory* **101**, 115-132.

Aaberge, R, S. Bjerve and K. Doksum (2005): Lorenz and Gini Regression. Mimeo, Research Department, Statistics Norway.

Ben Porath, E. and I. Gilboa (1994): Linear Measures, the Gini Index, and the Income-Equality Trade-off. *Journal of Economic Theory* **64**, 443-467.

Bhattacharya, N. and B. Mahalanobis (1967): Regional Disparities in Household Consumption in India. *JASA* **62**, 143-161.

Bickel, P.J. and K.D. Doksum (2001): *Mathematical Statistics. Basic Ideas and Selected Topics.* 2nd Ed., Vol. I. New Jersey: Prentice Hall.

Chaudhuri, P., K.A. Doksum and A. Samarov (1997): On Average Derivative Quantile Regression. *Journal of the American Statistical Association* **74**, 829-836.

D'Addario, R. (1936): Curve di concentrazione, elasticitá, flessibilitá, densitá media e densitá marginale dei redditi, Cressati, Bari.

Dagum, C. (1997): A new Approach to the Decomposition of the Gini Income Inequality Ration. *Empirical Economics* **22**, 515-531.

Donaldson, D. and J.A. Weymark (1980): A Single Parameter Generalization of the Gini Indices of Inequality. *Journal of Economic Theory* **22**, 67-86.

Donaldson, D. and J.A. Weymark (1983): Ethically Flexible Indices for Income Distribution in the Continuum. *Journal of Economic Theory* **29**, 353-358.

Friedman, F.H. and W. Stuetzle (1981): Projection Pursuit Regression. *Journal of the American Statistical Association* **76**, 817-823.

Giorgi, G.M. (1990): Bibliographic portrait of the Gini concentration ratio. *Metron* **48**, 183-221.

Giorgi, G.M. (1998): Concentration index, Bonferroni; in Encyclopedia of Statistical Sciences. *Update* **2**, 141-146.

Giorgi, G.M. (2001): A proposal of poverty measures based on the Bonferroni inequality index. *Metron* **59** (3-4), 2-1.

Hall, P. (1989): On Projection Pursuit Regression. *Annals of Statistics* **17**, 573-588.

Hristache, M., A. Juditsky and V. Spokoiny (2001): Direct Estimation of the Index Coefficient in a Single-Index Model. *Annals of Statistics* **17**, 1009-1052.

Härdle, W. and T. Stoker (1989): Investigating Smooth Multiple Regression by the Method of Average Derivatives. *Journal of the American Statistical Association* **84**, 986-995.

Härdle, W., P. Hall and H. Ichimura (1993): Optimal Smoothing in Single-Index Models. *Annals of Statistics* **21**, 157-178.

Kakwani, N.C. (1980): *Income Inequality and Poverty: Methods of Estimation and Policy Applications.* Oxford University Press.

Lambert, P. and R. Aronson (1993): Inequality Decomposition Analysis and the Gini Coefficient Revisited. *The Economic Journal* **103**, 1221-1227.

Mahalanobis, P.C. (1960): A Method of Fractile Graphical Analysis. *Econometrica* **28**, 325-351.

Mehran, F. (1976): Linear measures of inequality. *Econometrica* **44**, 805-809.

Nygård, F. and A. Sandstrøm (1981): *Measauring Income Inequality.* Almqvist and Wiksell, Stockholm.

Piesch, W. (1975): *Statistische Konzentrationsmasse.* Mohr (Siebeck), Tübingen.

Silber, J. (1989): Factor Components, Population Subgroups, and Computation of the Gini Index of Inequality. *Review of Economics and Statistics* **71**, 107-115.

Shorrocks, A.F. (1984): Inequality Decomposition by Population Subgroups. *Econometrica* **52**, 1359-1385.

Stoker, T. (1986): Consistent Estimation of Scaled Coefficients. *Econometrica* **54**, 1461-1481.

Weymark, J. (1981): Generalized Gini Inequality Indices. *Mathematical Social Science* **1**, 409-430.

Yitzhaki, S. (1994): Economic Distance and Overlapping of Distribution. *Journal of Economics* **61**, 147-159.

Yitzhaki, S. and R.I. Lerman (1991): Income Stratification and Income Inequality. *Review of Income and Wealth* **37**, 313-329.

Yaari, M.E. (1987): The Dual Theory of Choice under Risk. *Econometrica* **55**, 95-115.

Yaari, M.E. (1988): A Controversial Proposal Concerning Inequality Measurement. *Journal of Economic Theory* **44**, 381-397.

Recent publications in the series Discussion Papers

346   B.M. Larsen and R.Nesbakken (2003): How to quantify household electricity end-use consumption

347   B. Halvorsen, B. M. Larsen and R. Nesbakken (2003): Possibility for hedging from price increases in residential energy demand

348   S. Johansen and A. R. Swensen (2003): More on Testing Exact Rational Expectations in Cointegrated Vector Autoregressive Models: Restricted Drift Terms

349   B. Holtsmark (2003): The Kyoto Protocol without USA and Australia - with the Russian Federation as a strategic permit seller

350   J. Larsson (2003): Testing the Multiproduct Hypothesis on Norwegian Aluminium Industry Plants

351   T. Bye (2003): On the Price and Volume Effects from Green Certificates in the Energy Market

352   E. Holmøy (2003): Aggregate Industry Behaviour in a Monopolistic Competition Model with Heterogeneous Firms

353   A. O. Ervik, E.Holmøy and T. Hægeland (2003): A Theory-Based Measure of the Output of the Education Sector

354   E. Halvorsen (2003): A Cohort Analysis of Household Saving in Norway

355   I. Aslaksen and T. Synnestvedt (2003): Corporate environmental protection under uncertainty

356   S. Glomsrød and W. Taoyuan (2003): Coal cleaning: A viable strategy for reduced carbon emissions and improved environment in China?

357   A. Bruvoll T. Bye, J. Larsson og K. Telle (2003): Technological changes in the pulp and paper industry and the role of uniform versus selective environmental policy.

358   J.K. Dagsvik, S. Strøm and Z. Jia (2003): A Stochastic Model for the Utility of Income.

359   M. Rege and K. Telle (2003): Indirect Social Sanctions from Monetarily Unaffected Strangers in a Public Good Game.

360   R. Aaberge (2003): Mean-Spread-Preserving Transformation.

361   E. Halvorsen (2003): Financial Deregulation and Household Saving. The Norwegian Experience Revisited

362   E. Røed Larsen (2003): Are Rich Countries Immune to the Resource Curse? Evidence from Norway's Management of Its Oil Riches

363   E. Røed Larsen and Dag Einar Sommervoll (2003): Rising Inequality of Housing? Evidence from Segmented Housing Price Indices

364   R. Bjørnstad and T. Skjerpen (2003): Technology, Trade and Inequality

365   A. Raknerud, D. Rønningen and T. Skjerpen (2003): A method for improved capital measurement by combining accounts and firm investment data

366   B.J. Holtsmark and K.H. Alfsen (2004): PPP-correction of the IPCC emission scenarios - does it matter?

367   R. Aaberge, U. Colombino, E. Holmøy, B. Strøm and T. Wennemo (2004): Population ageing and fiscal sustainability: An integrated micro-macro analysis of required tax changes

368   E. Røed Larsen (2004): Does the CPI Mirror Costs.of.Living? Engel's Law Suggests Not in Norway

369   T. Skjerpen (2004): The dynamic factor model revisited: the identification problem remains

370   J.K. Dagsvik and A.L. Mathiassen (2004): Agricultural Production with Uncertain Water Supply

371   M. Greaker (2004): Industrial Competitiveness and Diffusion of New Pollution Abatement Technology – a new look at the Porter-hypothesis

372   G. Børnes Ringlund, K.E. Rosendahl and T. Skjerpen (2004): Does oilrig activity react to oil price changes? An empirical investigation

373   G. Liu (2004) Estimating Energy Demand Elasticities for OECD Countries. A Dynamic Panel Data Approach

374   K. Telle and J. Larsson (2004): Do environmental regulations hamper productivity growth? How accounting for improvements of firms' environmental performance can change the conclusion

375   K.R. Wangen (2004): Some Fundamental Problems in Becker, Grossman and Murphy's Implementation of Rational Addiction Theory

376   B.J. Holtsmark and K.H. Alfsen (2004): Implementation of the Kyoto Protocol without Russian participation

377   E. Røed Larsen (2004): Escaping the Resource Curse and the Dutch Disease? When and Why Norway Caught up with and Forged ahead of Its Neughbors

378   L. Andreassen (2004): Mortality, fertility and old age care in a two-sex growth model

379   E. Lund Sagen and F. R. Aune (2004): The Future European Natural Gas Market - are lower gas prices attainable?

380   A. Langørgen and D. Rønningen (2004): Local government preferences, individual needs, and the allocation of social assistance

381   K. Telle (2004): Effects of inspections on plants' regulatory and environmental performance - evidence from Norwegian manufacturing industries

382   T. A. Galloway (2004): To What Extent Is a Transition into Employment Associated with an Exit from Poverty

383   J. F. Bjørnstad and E.Ytterstad (2004): Two-Stage Sampling from a Prediction Point of View

384   A. Bruvoll and T. Fæhn (2004): Transboundary environmental policy effects: Markets and emission leakages

385   P.V. Hansen and L. Lindholt (2004): The market power of OPEC 1973-2001

386   N. Keilman and D. Q. Pham (2004): Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area

387   G. H. Bjertnæs and T. Fæhn (2004): Energy Taxation in a Small, Open Economy: Efficiency Gains under Political Restraints

388   J.K. Dagsvik and S. Strøm (2004): Sectoral Labor Supply, Choice Restrictions and Functional Form

389   B. Halvorsen (2004): Effects of norms, warm-glow and time use on household recycling

390   I. Aslaksen and T. Synnestvedt (2004): Are the Dixit-Pindyck and the Arrow-Fisher-Henry-Hanemann Option Values Equivalent?

391   G. H. Bjønnes, D. Rime and H. O.Aa. Solheim (2004): Liquidity provision in the overnight foreign exchange market

392   T. Åvitsland and J. Aasness (2004): Combining CGE and microsimulation models: Effects on equality of VAT reforms

393   M. Greaker and Eirik. Sagen (2004): Explaining experience curves for LNG liquefaction costs: Competition matter more than learning

394   K. Telle, I. Aslaksen and T. Synnestvedt (2004): "It pays to be green" - a premature conclusion?

395   T. Harding, H. O. Aa. Solheim and A. Benedictow (2004). House ownership and taxes

396   E. Holmøy and B. Strøm (2004): The Social Cost of Government Spending in an Economy with Large Tax Distortions: A CGE Decomposition for Norway

397   T. Hægeland, O. Raaum and K.G. Salvanes (2004): Pupil achievement, school resources and family background

398   I. Aslaksen, B. Natvig and I. Nordal (2004): Environmental risk and the precautionary principle: "Late lessons from early warnings" applied to genetically modified plants

399   J. Møen (2004): When subsidized R&D-firms fail, do they still stimulate growth? Tracing knowledge by following employees across firms

400   B. Halvorsen and Runa Nesbakken (2004): Accounting for differences in choice opportunities in analyses of energy expenditure data

401   T.J. Klette and A. Raknerud (2004): Heterogeneity, productivity and selection: An empirical study of Norwegian manufacturing firms

402   R. Aaberge (2005): Asymptotic Distribution Theory of Empirical Rank-dependent Measures of Inequality

403   F.R. Aune, S. Kverndokk, L. Lindholt and K.E. Rosendahl (2005): Profitability of different instruments in international climate policies

404   Z. Jia (2005): Labor Supply of Retiring Couples and Heterogeneity in Household Decision-Making Structure

405   Z. Jia (2005): Retirement Behavior of Working Couples in Norway. A Dynamic Programming Approch

406   Z. Jia (2005): Spousal Influence on Early Retirement Behavior

407   P. Frenger (2005): The elasticity of substitution of superlative price indices

408   M. Mogstad, A. Langørgen and R. Aaberge (2005): Region-specific versus Country-specific Poverty Lines in Analysis of Poverty

409   J.K. Dagsvik (2005) Choice under Uncertainty and Bounded Rationality

410   T. Fæhn, A.G. Gómez-Plana and S. Kverndokk (2005): Can a carbon permit system reduce Spanish unemployment?

411   J. Larsson and K. Telle (2005): Consequences of the IPPC-directive's BAT requirements for abatement costs and emissions

412   R. Aaberge, S. Bjerve and K. Doksum (2005): Modeling Concentration and Dispersion in Multiple Regression

413   E. Holmøy and K.M. Heide (2005): Is Norway immune to Dutch Disease? CGE Estimates of Sustainable Wage Growth and De-industrialisation

414   K.R. Wangen (2005): An Expenditure Based Estimate of Britain's Black Economy Revisited

415   A. Mathiassen (2005): A Statistical Model for Simple, Fast and Reliable Measurement of Poverty

416   F.R. Aune, S. Glomsrød, L. Lindholt and K.E. Rosendahl: Are high oil prices profitable for OPEC in the long run?

417   D. Fredriksen, K.M. Heide, E. Holmøy and I.F. Solli (2005): Macroeconomic effects of proposed pension reforms in Norway

418   D. Fredriksen and N.M. Stølen (2005): Effects of demographic development, labour supply and pension reforms on the future pension burden

419   A. Alstadsæter, A-S. Kolm and B. Larsen (2005): Tax Effects on Unemployment and the Choice of Educational Type

420   E. Biørn (2005): Constructing Panel Data Estimators by Aggregation: A General Moment Estimator and a Suggested Synthesis

421   J. Bjørnstad (2005): Non-Bayesian Multiple Imputation

422   H. Hungnes (2005): Identifying Structural Breaks in Cointegrated VAR Models

423   H. C. Bjørnland and H. Hungnes (2005): The commodity currency puzzle

424   F. Carlsen, B. Langset and J. Rattsø (2005): The relationship between firm mobility and tax level: Empirical evidence of fiscal competition between local governments

425   T. Harding and J. Rattsø (2005): The barrier model of productivity growth: South Africa

426   E. Holmøy (2005): The Anatomy of Electricity Demand: A CGE Decomposition for Norway

427   T.K.M. Beatty, E. Røed Larsen and D.E. Sommervoll (2005): Measuring the Price of Housing Consumption for Owners in the CPI

428   E. Røed Larsen (2005): Distributional Effects of Environmental Taxes on Transportation: Evidence from Engel Curves in the United States

429   P. Boug, Å. Cappelen and T. Eika (2005): Exchange Rate Rass-through in a Small Open Economy: The Importance of the Distribution Sector

430   K. Gabrielsen, T. Bye and F.R. Aune (2005): Climate change- lower electricity prices and increasing demand. An application to the Nordic Countries

431   J.K. Dagsvik, S. Strøm and Z. Jia: Utility of Income as a Random Function: Behavioral Characterization and Empirical Evidence

432   G.H. Bjertnæs (2005): Avioding Adverse Employment Effects from Energy Taxation: What does it cost?

433.   T. Bye and E. Hope (2005): Deregulation of electricity markets—The Norwegian experience

434   P.J. Lambert and T.O. Thoresen (2005): Base independence in the analysis of tax policy effects: with an application to Norway 1992-2004

435   M. Rege, K. Telle and M. Votruba (2005): The Effect of Plant Downsizing on Disability Pension Utilization

436   J. Hovi and B. Holtsmark (2005): Cap-and-Trade or Carbon Taxes? The Effects of Non-Compliance and the Feasibility of Enforcement

437   R. Aaberge, S. Bjerve and K. Doksum (2005): Decomposition of Rank-Dependent Measures of Inequality by Subgroups