# Mathematical Analysis of Distribution and Redistribution of Income

Johan Fellman


Once I proved a theorem that has later inspired numerous scientists.

SciencePG
Science Publishing Group

Science Publishing Group

# Mathematical Analysis of Distribution and Redistribution of Income

$$L_X(p) = \frac{1}{\mu_X}\int_0^{x_p} x f_X(x)\,dx$$

Johan Fellman

SciencePG
Science Publishing Group

Science Publishing Group

548 Fashion Avenue
New York, NY 10018
www.sciencepublishinggroup.com

Published by Science Publishing Group 2015



First Edition

**ISBN: 978-1-940366-25-8**



To order additional copies of this book, please contact:
Science Publishing Group
book@sciencepublishinggroup.com
www.sciencepublishinggroup.com

Printed and bound in India

# Acknowledgements

This work was supported in part

by a grant from the

Magnus Ehrnrooths Stiftelse foundation.

# Preface

Different skew models such as the lognormal and the Pareto have been proposed as suitable descriptions of the income distribution, but such specific distributions are usually applied in empirical investigations. For general studies more wide-ranging tools have been considered. The central and most commonly applied theory is connected to the Lorenz curve. Without any assumptions concerning specific distributions, this theory enables analyses of temporal and regional variations in the income inequalities. Particularly, it is a valuable tool for studies of the effect of taxes and transfers to the redistribution of income. Taxation and transferring may have similar effects, but some marked differences with respect to their applications can be identified and therefore, both will usually be given individual presentations.

In this study I have collected the central parts of my contributions to the theory of income distributions and furthermore, I have tried to locate my results within the framework of the general literature.

# Contents

# 1

# Introduction

## 1.1 Historical Background

It is a well-known fact that the income distributions are commonly unimodal and skew with a heavy right tail. Therefore, different skew models such as the lognormal and the Pareto have been proposed as suitable descriptions of the income distribution, but they are usually applied in specific empirical situations.

For general studies, more wide-ranging tools have been considered. The most commonly used theory is based on the Lorenz curve. Lorenz (1905) developed it in order to analyse the distribution of income and wealth within populations. He described the Lorenz curve in the following way:

> *"Plot along one axis accumulated per cents of the population from poorest to richest, and along the other, wealth held by these per cents of the population".*

Consequently, the Lorenz curve $L(p)$ is defined as a function of the proportion $p$ of the population. It is convex and satisfies the condition $L(p) \leq p$ because the income share of the poor is less than their proportion of the population. A sketch of a Lorenz curve is given in Figure 1.1.1.

The theoretical Lorenz curve $L_X(p)$ for the income distribution $F_X(x)$ of a non-negative variable $X$ can be described in the following way: Let $f_X(x)$ be the corresponding frequency distribution,

$$\mu_X = \int_0^\infty x\, f_X(x)dx \qquad (1.1.1)$$

be the mean of $X$ and let $x_p$ be the $p$ quantile, that is $F_X(x_p) = p$. Then

$$L_X(p) = \frac{1}{\mu_X} \int_0^{x_p} x f_X(x) dx, \qquad (1.1.2)$$

is the Lorenz curve. The Lorenz curve is not defined if the mean is zero or infinite.

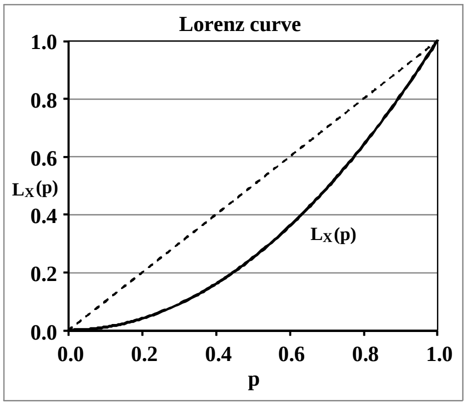



***Figure 1.1.1*** *A sketch of a Lorenz curve* $L_X(p)$.

Consider a transformed variable $Y = g(X)$, where $g(\cdot)$ is positive and monotone increasing. Define the inverse transformation $X = \gamma(Y)$. Then

$$F_Y(y) = P(Y \le y) = P(g(X) \le g(x)) = P(X \le x) = F_X(x).$$

For the transformed variable $Y$ the $p$ quantile is $F_Y(y_p) = p$, that is $y_p = g(x_p)$.

Now

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(x)}{dx}\frac{dx}{dy} = f_X(x)\frac{dx}{dy} = f_X(x)\frac{d\gamma}{dy}. \qquad (1.1.3)$$

Hence

$$\mu_Y = \int_0^\infty y f_Y(y)dy = \int_0^\infty g(x) f_X(x)dx \tag{1.1.4}$$

and

$$L_Y(p) = \frac{1}{\mu_Y}\int_0^{x_p} g(x) f_X(x)dx. \tag{1.1.5}$$

If the transformation is linear $g(x) = \theta x$, then $Y = \theta X$, $\mu_Y = \theta \mu_X$,

$$L_Y(p) = \frac{1}{\theta \mu_X}\int_0^{x_p} \theta x f_X(x)dx = L_X(p) \tag{1.1.6}$$

and consequently, the Lorenz curve is invariant under linear transformations.

A simple example of this property is that the Lorenz curve of the income distribution is independent of the currency used.

Consequently, the Lorenz curve satisfies the general rules:

*To every distribution $F(x)$ corresponds a unique Lorenz curve, $L_X(p)$. The contrary does not hold because every Lorenz curve $L_X(p)$ is a common curve for a whole class of distributions $F(\theta x)$ where $\theta$ is an arbitrary positive constant.*

A Lorenz curve always starts at $(0,0)$ and ends at $(1,1)$. The higher Lorenz curve the lesser is the inequality of the income distribution. The diagonal $L(p) = p$ is commonly interpreted as the Lorenz curve for complete equality between the income receivers, but according to Wang et al. (2011), $L(p) = p$ is strictly speaking not a Lorenz curve associated with complete inequality. They discuss the possibility how to identify this Lorenz curve with the situation that all individuals receive the same income. Mathematically this result can be obtained as a limiting curve when the inequality of the income distribution converges towards zero. Increasing inequality lowers the Lorenz curve and theoretically, it can converge towards the lower right corner of the square.

Consider two variables *X* and *Y*, their distributions $F_X(x)$ and $F_y(y)$, and their Lorenz curves $L_X(p)$ and $L_Y(p)$. If $L_X(p) \geq L_Y(p)$ for all *p*, then measured by the Lorenz curves, the distribution $F_X(x)$ has lower inequality than the distribution $F_y(y)$ and $F_X(x)$ is said to *Lorenz dominate* $F_y(y)$. We denote this relation $F_X(x) \underset{L}{\succ} F_Y(y)$. An example of Lorenz dominance is given in Figure 1.1.2. This is the common definition of the Lorenz dominance although that some define the dominance in the opposite way.

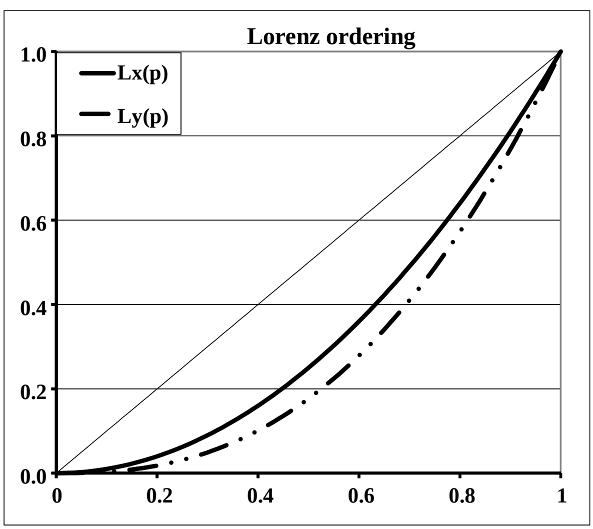



***Figure 1.1.2*** *Lorenz curves with Lorenz ordering, that is* $L_X(p) \underset{L}{\succ} L_Y(p)$.

Income inequalities can be of different type and the corresponding Lorenz curves may intersect and for these no Lorenz ordering can be identified (c.f. Figure 1.1.3). The Lorenz curve $L_2(p)$ corresponds to a population where the poor are relatively not so poor and the rich are relatively rich. On the other hand the Lorenz curve $L_1(p)$ corresponds to a population with very poor among the poor and the rich are not so rich.

For intersecting Lorenz curves alternative inequality measures have to be defined. The most frequently used is the Gini coefficient, $G$ (Gini, 1914). Using the Lorenz curves, this coefficient is the ratio between the area between the diagonal and the Lorenz curve and the whole area under the diagonal. The formula is

$$G = 1 - 2\int_0^1 L(p)dp. \tag{1.1.7}$$

This definition yields Gini coefficients satisfying the inequalities $0 < G < 1$. The higher $G$ value the stronger inequality. If $G_X < G_Y$, then the distribution $F_X(x)$, measured by the Gini coefficient, has lower inequality than the distribution $F_y(y)$ and we say that $F_X(x)$ *Gini dominates* $F_y(y)$. We denote this relation $F_X(x) \underset{G}{\succ} F_Y(y)$.

Yitzhaki (1983) proposed the generalized Gini coefficient

$$G(\nu) = 1 - \nu(1-\nu)\int_0^1 (1-p)^{\nu-2} L(p)dp, \tag{1.1.8}$$

where $\nu > 1$. Different $\nu'$s are used in order to identify different inequality properties. For low $\nu'$s greater weights are associated with the rich and for high

$\nu'$s greater weights are associated with the poor. Using the mean income ( $\mu$ ) and the Gini coefficient ( $G$ ), Sen (1973) proposed a welfare index

$$W = \mu(1 - G). \tag{1.1.9}$$

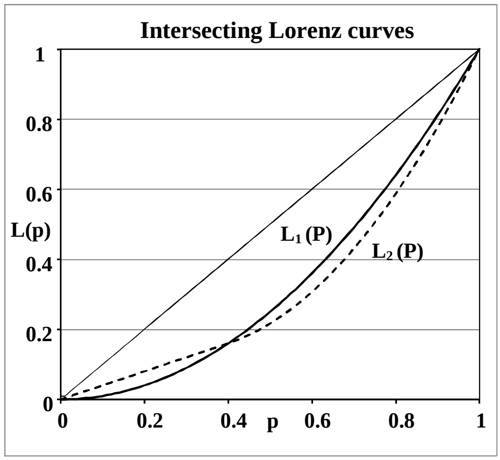



***Figure 1.1.3*** *Two intersecting Lorenz curves. Using the Gini coefficient presented in the text,* $L_1(p)$ *has less inequality (* $G_1 = 0.3333$*) than* $L_2(p)$ *(* $G_2 = 0.3600$*). The Pietra coefficients, presented below, are* $P_1 = 0.2500$ *and* $P_2 = 0.2940$*.*

Alternative inequality measures have been defined and such measures are discussed later in section 1.3.

## 1.2 Income Distributions

According to Aichison and Brown (1954) general description of an income distribution may be defined as a rule which gives for each value of income $x$ the proportion $F(x)$ of persons in a given population who have an income not greater than $x$. Such a description is a useful analytical tool if it requires that $F(x)$ has to be given a precise mathematical expression involving known, or more frequently unknown, parameters. It is interesting to recall that Pareto (1897), when he first presented his law, emphasised its empirical basis, but on the other hand the process of reasoning by Gibrat (1931) started from theory to observations.

Aichison and Brown (1954) gave four criteria on which the success of a particular description may be assessed.

- How closely the description approximate to the observed distribution of incomes when specific values are assigned to the parameters? These values will usually be estimated from the data.
- To what extent may the statistical description be shown to rest on assumptions which are consistent with our knowledge of the way in which incomes are generated?
- What facilities does the description provide in the statistical analysis of the data?
- What economic meaning or significance can be attached to the parameters of the description?

Furthermore, Aichison and Brown gave a thorough presentation of studies of income distributions presented during the first half of the $20^{th}$ century. They stated that it is well known that income distributions almost invariably possess a single mode and are positively skewed. Many statistical descriptions satisfying these rather general conditions have been proposed in the past as applicable to the distribution of incomes, among which one may note the frequency curves of Pareto (1897), Kapteyn (1903), Gibrat (1931) and Champernowne (1953).

Already Quensel (1944) stated that the lognormal curve agrees fairly well with the actual distribution of the lower incomes, although the Pareto curve often provides a more adequate description of the higher incomes.

Champernowne (1953) described an ingenious model which under realistic assumptions generates exactly or approximately a distribution of incomes obeying Pareto ́s law. Champernowne ́s model provides a basis for the comparison of processes of generating the Pareto and the lognormal descriptions of income distributions. Before Champernowne ́s article Rhodes (1944) and Castellani (1950) presented attempts to derive Pareto ́s distribution.

Furthermore, Aichison and Brown (1954) noted that the law of proportionate effect, postulated by models predicting lognormality, is less appropriate when we are considering a heterogeneous group of income receivers than if the population is divided in sectors, within each the postulate applies. Under the assumptions which are necessary for the application of the central limit theorem, the multiplicative form of the central limit theorem leads us to expect that the distribution of incomes will eventually be described by a lognormal curve. If the population is divided into a large number of sectors and that in each sector the basic postulate of proportionate effect may be assumed to apply, means that a lognormal description of incomes will be valid in each sector, though the parameters of the description may take on different numerical values in each sector.

Finally, Aichison and Brown (1954) stressed that it is useless to posit a statistical description of income distribution unless it is possible with the help of this description to derive analytical tools for any investigation that is likely to be required. To take an extreme example, there would be little point in giving $F(x)$ an explicit mathematical form involving unknown parameters if no method of estimating these parameters from data were available. It is, however, comforting in statistical work to be sure that one is not wasting any of the information available and this is always possible with the lognormal description.

An example of a skewed lognormal distribution can be seen in Figure 1.2.1. Note that the income receivers in this example are a homogeneous group from the upper part of the hierarchy (c.f. Aichison and Brown, 1954).

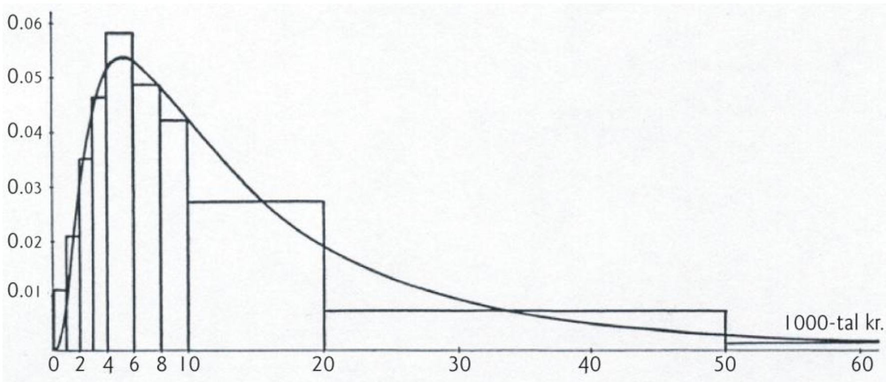



***Figure 1.2.1*** *Distribution of the income among 4103 industrial managers compared to a lognormal distribution (Cram ér, 1949).*

McDonald and Ransom (1979) compared alternative income distribution models and applied them on US family income data. The interesting models were the lognormal, the gamma, the beta and the Sing-Maddala functions. They applied the models on family income for 1960 and 1969 through 1975 and compared the estimation methods: the method of scoring, the Pearson minimum chi-squared method and the least squares estimation. The estimation of the mean income and the Gini coefficient were directly obtained by substituting estimates of the parameters characterizing the associated distribution functions into the appropriate theoretical expressions of the coefficients. They noted that even though they observed situations in which parameter estimates change significantly from one time period to another, the associated population characteristics such as the mean and the Gini coefficients are much more stable. However, the estimated Gini coefficients associated with the scoring and the minimum chi-squared estimates of the lognormal density are much larger than for any other case considered. A general observation was that the scoring and the minimum chi-squared results were very similar for the three parameter functions, with greater differences for the gamma, and still greater for the lognormal.

Summing up, McDonald and Ransom concluded that the gamma provided a better fit than the lognormal, regardless of the estimation technique used. The three parameter functions (beta and Singh-Maddala) provided a better fit to the data than did the gamma density function. This finding is obviously due to the number of distribution parameters.

Over time has come the realization that only the upper tail of the distribution is Pareto in form. Proceeding from the observation that the distribution has a Pareto tail for the top 15-20% of employees. Lydall (1968) advances a model of hierarchal earnings based on the notation that large organisations are organised on hierarchical principle.

Harrison (1981) noted that a number of observed earnings distributions were well described by the Pareto distribution

$$F(y)=\begin{cases}0 & y\le 1\\ 1-y^{-\alpha} & y>1\end{cases}, \qquad (1.2.1)$$

where $\alpha>0$ and $y=Y/Y_L$, $Y_L$ being the minimum income. For $\alpha>1$, the mean is $E(Y)=\dfrac{\alpha}{\alpha-1}$. Furthermore, the Lorenz curve is $L(p)=1-(1-p)^{\frac{\alpha-1}{\alpha}}$ and the Gini coefficient is $G=\dfrac{1}{2\alpha-1}$. It may perhaps be convenient to remark here that for commonly occurring values of the parameter $\alpha$ a second moment of the Pareto distribution does not exist unless $\alpha>2$. Furthermore, Harrison stressed that equally compelling reasons supporting the use of disaggregated data can be found in the case of the lognormal function.

Dagum (1977, 1980, 1987) has paid continuous attention to alternative income distributions.

A common technique for estimating the Pareto constant, $\alpha$, is to linearize the survival function by taking logarithms, and apply ordinary least squares. The survival function is

$$S(y) = 1 - F(y) = \left(\frac{Y}{Y_L}\right)^{-\alpha}.$$

After taking natural logarithms one obtains the linear model

$$ln(S(y)) = -\alpha\, ln(Y) + \alpha\, ln(Y_L) = C - \alpha\, ln(Y).$$

This model indicates a linear, decreasing association between $\ln(S(y))$ and $\ln(Y)$. A regression analysis gives an estimate of $\alpha$ and the coefficient of determination, $R^2$, measures the linearity in the model and the goodness of fit of the Pareto model.

We apply this analysis on annual taxable incomes in Finland for 2009 (http://pxweb2.stat.fi/Database/StatFin/tul/tvt/2009/2009_en.asp).

The data are presented in a grouped table (Table 1.2.1). We assume that the Pareto model may start from ca. $Y = 25000$€. For values equal to or greater than that we obtain the estimate $\hat{\alpha} = 2.637$ and in addition, the coefficient of determination is $R^2 = 0.99241$. For the income distribution for incomes greater than 25000 the Gini coefficient is $G = \dfrac{1}{2\alpha - 1} = 0.234$.

***Table 1.2.1*** *Taxable income receivers in Finland 2009.*

| Classes of annual income (€) | Number of income recipients |
|---|---|
| - 1000 | 182281 |
| 1000 - 2000 | 96836 |
| 2000 - 3000 | 80056 |
| 3000 - 4000 | 65800 |
| 4000 - 5000 | 59595 |
| 5000 - 6000 | 62171 |
| 6000 - 7000 | 107558 |
| 7000 - 8000 | 146526 |
| 8000 - 9000 | 114602 |
| 9000 - 10000 | 121555 |
| 10000 - 12500 | 319042 |
| 12500 - 15000 | 329083 |
| 15000 - 17500 | 259979 |
| 17500 - 20000 | 243284 |
| 20000 - 25000 | 481753 |
| 25000 - 30000 | 487376 |
| 30000 - 35000 | 385672 |
| 35000 - 40000 | 266075 |
| 40000 - 50000 | 307810 |
| 50000 - 60000 | 152714 |
| 60000 - 80000 | 120327 |
| 80000 - | 88488 |
| All | 4478583 |

In Figure 1.2.2 we sketch the result.

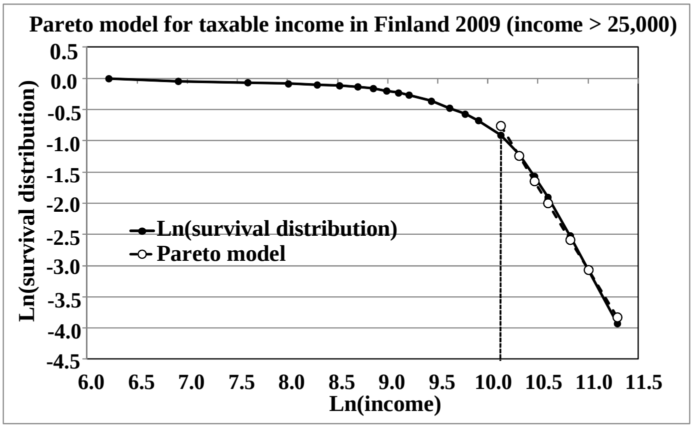



***Figure 1.2.2*** *Graphical sketch of the distribution of taxable income in Finland (2009) and a Pareto model for annual incomes greater than* $Y = 25000$ *€.*

## 1.3 Lorenz Curves and Concentration of Incomes

A central topic in the analyses of income distributions is the concept of *concentration of incomes*, which is defined in the literature (Lorenz, 1905) in such a way as to be free of any particular hypothesis concerning the genesis of the description of the income distribution.

In Section 1.1 we introduced the Lorenz curve $L(p)$ defined by $L(p) = \frac{1}{\mu}\int_0^{x_p} x\, f(x)dx$, where $\mu_X = \int_0^{\infty} x\, f_X(x)dx$ is the mean and $F(x_p) = p$. Lorenz curves were presented in the Figures 1.1.1, 1.1.2 and 1.1.3.

The Lorenz curve has the following general properties:

i. $L(p)$ is monotone increasing.

ii. $L(p) \le p$.

iii. $L(p)$ is convex.

iv. $L(0)=0$ and $L(1)=1$.

The Lorenz curve $L(p)$ is convex because the income share of the poor is less than their proportion of the population. The higher Lorenz curve the lesser inequality in the income distribution (c.f. Section 1.1).

The Lorenz curve for a probability distribution is a continuous function. However, Lorenz curves representing discontinuous functions can be constructed as the limit of Lorenz curves of probability distributions, the line of perfect inequality being an example.

If the Lorenz curve is differentiable the derivatives have the following properties. Let $L_X(p)=\frac{1}{\mu_X}\int_0^{x_p} x\, f_X(x)dx$, $F_X(x_p)=p$ and the density function $f_X(x)$ . When we differentiate the equation $F_X(x_p)=p$ we obtain $\frac{dF_X(x_p)}{dp}=\frac{dF_X(x_p)}{dx_p}\frac{dx_p}{dp}=1$,

$$f_X(x_p)\frac{dx_p}{dp}=1$$

and

$$\frac{dx_p}{dp}=\frac{1}{f_X(x_p)}.$$

The derivation of $L_X(p)=\frac{1}{\mu_X}\int_0^{x_p} x\, f_X(x)dx$ yields

$$\frac{dL_X(p)}{dp}=\frac{1}{\mu_X}\frac{d\int_0^{x_p} x\, f_X(x)dx}{dx_p}\frac{dx_p}{dp}=\frac{1}{\mu_X}x_p f_X(x_p)\frac{dx_p}{dp}=\frac{x_p}{\mu_X}$$

and consequently,

$$\frac{dL_X(p)}{dp}=\frac{x_p}{\mu_X} \tag{1.3.1}$$

If the Lorenz curve is differentiable twice, then the second derivative is

$$\frac{d^2L_X(p)}{dp^2}=\frac{1}{\mu_X}\frac{dx_p}{dp}=\frac{1}{\mu_X}\frac{1}{f_X(x_p)}.$$

Hence,

$$\frac{d^2L(p)}{dp^2}=\frac{1}{\mu_X f_X\left(x_p\right)} \tag{1.3.2}$$

The difference between the diagonal and the Lorenz curve

$$D=p-L_X(p)$$

$$\frac{dD}{dp}=1-L_X'(p)=1-\frac{x_p}{\mu_X}$$

$$\frac{d^2D}{dp^2}=-L_X''(p)=-\frac{1}{\mu_X}\frac{dx_p}{dp}=-\frac{1}{\mu_X f_X(x)}<0.$$

The maximum of $D$ implies $1-\frac{x_p}{\mu_X}=0$, that is $x_p=\mu$.

For $x_p=\mu_X$, $L_X'(p)=\frac{\mu_X}{\mu_X}=1$ and at the point $p_\mu=F_X(\mu_X)$ the tangent is parallel to the line of perfect equality. This is also the point at which the vertical distance between the Lorenz curve and the egalitarian line attains its maximum $P = p_\mu - L_X(p_\mu)$. This maximum is defined as the Pietra index (Lee, 1999). According to this definition $0 < P < 1$. The lower bound is obtained when there is total income-equality that is the Lorenz curve coincides with the diagonal. The upper bound can be obtained when the Lorenz curve converges towards the lower right corner. The Pietra index can be interpreted as income of the rich that should be redistributed to the poor in order to obtain total income equality. Therefore, the index is sometimes named the Robin Hood index. Lee (1999) used the Lorenz curve and the summary measures based on it for diagnostic tests medical studies. He associated the Gini and the Pietra indices with the receiver operating characteristic curve (ROC). He also gave in his reference list additional papers where these summary statistics were applied.

An alternative definition has also been given. The Pietra index can be defined as twice the area of the largest triangle inscribed in the area between the Lorenz curve and the diagonal line (Lee, 1999). In Figure 1.3.1 one observes that the triangle obtains its maximum when the corner lies on the Lorenz curve where the tangent is parallel to the diagonal. The height of the triangle is $h = \frac{P}{\sqrt{2}}$ and the base is the diagonal $b = \sqrt{2}$. The double of the area is

$$2\text{area} = 2\frac{h\sqrt{2}}{2} = 2\frac{P\sqrt{2}}{2\sqrt{2}} = P.$$

Compared to the Gini coefficient we obtain that $G > P$ (see, Lee, 1999).

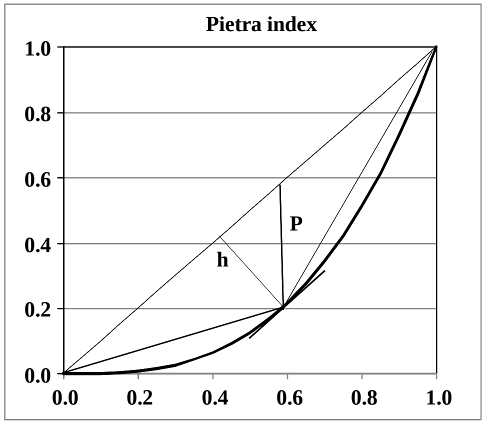



***Figure 1.3.1*** *The Lorenz curve and the geometric interpretations of the Pietra index.*

The definition yields Pietra coefficients satisfying the inequality $0 \le P \le 1$. If $P_X < P_Y$, then the distribution $F_X(x)$ measured by the Pietra index has lower inequality than the distribution $F_y(y)$ and we say that $F_X(x)$ *Pietra dominates* $F_y(y)$. We denote this relation $F_X(x) \succ_P F_Y(y)$. For the Lorenz curves in Figure 1.1.3, $P_1 \approx 0.2500$ and $P_2 \approx 0.2940$. According to the Pietra index, $L_1(p)$ is less unequal than $L_2(p)$.

In general, the Pietra and the Gini orderings are not identical. The following simple example supports this statement. Consider the situation described in Figure 1.3.2. There are two polygonal Lorenz curves, OABC ( $L_1(p)$ ) and ODC ( $L_2(p)$ ). [1]For $L_1(p)$ we obtain $P_1 < G_1$ and for $L_2(p)$ we obtain $P_2 = G_2$ because ODC is a triangle yielding identical indices. Furthermore if the point D

[1]The Lorenz curves in this example are not continuously differentiable, but slight modifications yield differentiable Lorenz curves. One has only to modify the edges to mini curves. If these modifications are minute, the inequalities given above still hold.

is close to the line AB, we observe that $G_1 > G_2$ and $P_1 < P_2$. Combining these inequalities we obtain $P_1 < P_2 = G_2 < G_1$. Consequently, $L_1(p)$ Pietra dominates $L_2(p)$, but $L_2(p)$ Gini dominates $L_1(p)$.

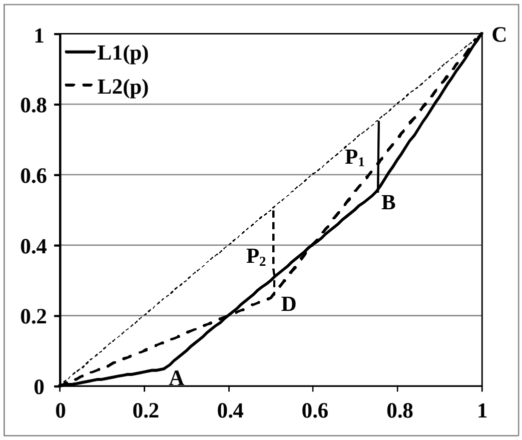



***Figure 1.3.2*** *Comparisons between Gini and Pietra indices. For the Lorenz curve $L_1(p)$ the Pietra index is $P = 0.20$ and the Gini coefficient is $G = 0.30$ and for Lorenz curve $L_2(p)$ the Pietra index is $P = 0.25$ and the Gini coefficient is $P = 0.25$.*

Above we obtained the inequality $0 < P < 1$. The limits in the inequalities can be obtained and this can be explained by the following example and Figure 1.3.3.

Consider the simplified RT model defined in (1.4.5)

$$L(p) = p^{\alpha} \quad \alpha \geq 1.$$

Examples of these Lorenz curves are sketched in Figure 1.3.3. The Gini coefficient is $G = \dfrac{\alpha - 1}{\alpha + 1}$. When $\alpha \to 1$ then $G \to 0$ and when $\alpha \to \infty$ then $G \to 1$. The Pietra index is $P = \left(\dfrac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\dfrac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$. We select a sequence of $\alpha$

values $\alpha = 1 + \frac{1}{n}$, for $n = 1, 2, \ldots$. The $P$ values are

$$P = \left(1 - \frac{1}{n+1}\right)^n - \left(\left(1 - \frac{1}{n+1}\right)^n\right)^{\left(1+\frac{1}{n}\right)}.$$

When $n \to \infty$, both terms converge towards $e^{-1}$ and $P \to 0$. According to the definition of the $P$ index

$$P = \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}} \geq p - p^{\alpha} \text{ for all } p < 1.$$

For increasing $\alpha$ values the supremum of $p - p^{\alpha}$ is one. This must also be the supremum of $P = \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$. Consequently, the interval $0 < P < 1$ cannot be shortened.

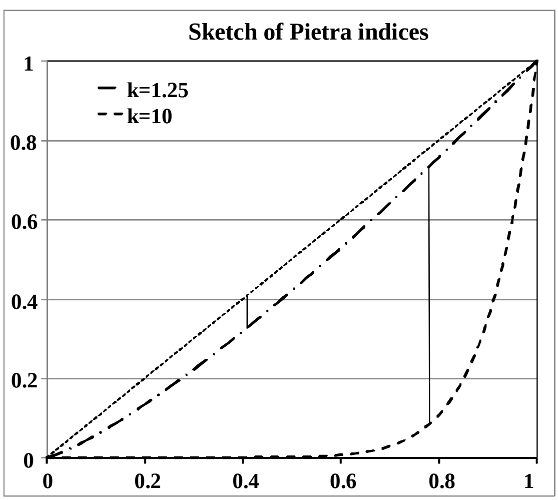



***Figure 1.3.3*** *Sketches of extreme Lorenz curves with corresponding P indices. For the Lorenz curve* $k = 1.25$ *the Pietra index is 0.0819 and for the Lorenz curve* $k = 10$ *the index is 0.6966.*

We can prove ( $\lim\limits_{p\uparrow 1}$ denotes limit from the left).

**Theorem 1.3.1.** If $\mu_X$ exists, then $\lim\limits_{p\uparrow 1} L'(p)(1-p)=0$.

**Proof.** Consider the integral $\int_x^\infty t\, f_X(t)dt$. If $\mu_X$ exists, then $\int_0^\infty t\, f_X(t)dt=\mu_X$ and for every $\varepsilon>0$ there exists an $x'$ such that $\int_x^\infty t\, f_X(t)dt<\varepsilon$ if $x>x'$. Choose p so that $x_p > x'$, then

$$\varepsilon > \int_{x_p}^\infty t\, f_X(t)dt \geq x_p \int_{x_p}^\infty f_X(t)dt = x_p(1-p)\,. \tag{1.3.3}$$

As a consequence of (1.3.3),

$$\lim_{p\uparrow 1} L'_X(p)(1-p) = \lim_{p\uparrow 1} \frac{x_p}{\mu_X}(1-p) = \frac{1}{\mu_X}\lim_{p\uparrow 1} x_p(1-p)=0\,.$$

Consider an one-parametric class of cumulative distribution functions $F(x,\theta)$, defined on the positive x-axis. If we assume that $F(x,\theta)=F(\theta x)$, i.e. it depends only on the product $\theta x$, then the following theorem holds:

**Theorem 1.3.2.** Let $F(x,\theta)$ be a one-parametric class of distributions with the properties:

i. $F(x,\theta)=F(\theta x)$.

ii. $F(\theta x)$ is defined on the positive x-axis.

iii. $F(\theta x)$ and its derivative are continuous.

iv. $\mu_X = E(X)$ exists.

Let $T = \theta X$, then

$$x_p(\theta) = \frac{t_p}{\theta} \tag{1.3.3}$$

and

$$\mu_X(\theta) = \frac{c}{\theta}, \tag{1.3.5}$$

where $t_p$ and c are independent of $\theta$.

**Proof.** Let $\theta$ be an arbitrary, positive parameter. Then the quantile $x_p(\theta)$ is defined by the equation $F(\theta x_p) = p$. If we define $t_p$ by the equation $F(t_p) = p$ then $t_p$ does not depend on $\theta$ and $\theta x_p(\theta) = t_p$ and (1.3.3) is proved. The formula (1.3.5) and the statement that $L(p) = \frac{1}{\mu(\theta)} \int_0^{x_p(\theta)} x\,dF(\theta x)$ is independent of $\theta$ is proved by using the substitution $t = \theta x$ in the integrals $E(X) = \int_0^{\infty} x\,dF(\theta x)$ and $L(p) = \frac{1}{\mu(\theta)} \int_0^{x_p(\theta)} x\,dF(\theta x)$.

Furthermore, we can prove:

**Theorem 1.3.3.** Consider a function $L(p)$ defined on the interval [0, 1] with the properties:

i. $L(p)$ is monotone increasing and convex.

ii. $L(0) = 0$ and $L(1) = 1$.

iii. $L(p)$ is differentiable twice.

iv. $\lim_{p\uparrow 1} L'(p)(1-p)=0$.

then $L(p)$ is a Lorenz curve of a distribution with finite mean.

**Proof.** If we denote the unknown distribution $F(x)$ and its derivative $f(x)$, then necessarily $L'(p)=\frac{x_p}{\mu}$. The derivative $L'(p)$ is a monotone-increasing function. If its inverse is denoted $M(p)$, we get the necessary relation

$$F(x_p)=p=M(\frac{x_p}{\mu}).$$

If $\theta=\frac{1}{\mu}$, then $F(x)=M(\theta x)$. Now we shall prove the sufficiency, that is, that $M(\theta x)$ is a distribution, whose mean is $\mu=\frac{1}{\theta}$ and whose Lorenz curve is $L(p)$. We denote $M(\theta x)=F(x)$ then $f(x)=F'(x)=\theta M'(\theta x)$. After observing that the property (iv) indicates that $L'(p)$ is integrable from *0* to *1*, we introduce the variable transformation

$$y=M(\theta x)$$

$$dy=\theta M'(\theta x)dx$$

$$x=\frac{1}{\theta}L'(y)$$

We obtain

$$\mu=\lim_{t\to\infty}\int_0^t x\theta M'(\theta x)dx=\lim_{p\uparrow 1}\int_0^p \frac{1}{\theta}L'(y)\,dy=\frac{1}{\theta}\lim_{p\uparrow 1}\int_0^p L'(y)\,dy=\frac{1}{\theta}$$

The given function $L'(p)$ has a monotone-increasing inverse function, $M(\theta x)$ giving a corresponding distribution function $F(x) = M(\theta x)$ whose mean is $\mu$.

Using the same transformation we obtain that the Lorenz curve $\tilde{L}(p)$ of $F(x) = M(\theta x)$ is

$$\tilde{L}(p) = \theta \int_0^{x_p} x\theta M'(\theta x) dx = \int_0^p L'(v)\, dv = \int_0^p L'(v)\, dv$$

and the theorem is proved.

These results have been collected in the following theorem (Fellman, 1976, 1980).

**Theorem 1.3.4.** Consider a given function $L(p)$ with the properties:

i. $L(p)$ is monotone increasing and convex to the p-axis.

ii. $L(0) = 0$ and $L(1) = 1$.

iii. $L(p)$ is differentiable.

iv. $\lim_{p \uparrow 1} L'(p)(1-p) = 0$.

Then $L(p)$ is the Lorenz curve of a whole class of distribution functions $F(\theta x)$, where $\theta$ is an arbitrary positive constant and the function $F(\cdot)$ is the inverse function to $L'(p)$.

In Fellman (1976) the result was presented and later Fellman (1980) presented the following theorem.

**Theorem 1.3.5.** A class of continuous distributions $F(x, \theta)$ with finite mean has a common Lorenz curve if and only if $F(x, \theta) = F(\theta x)$.

The formula for Gini coefficient $G$ is given in (1.1.7) and $G$ satisfies the inequality $0 \leq G \leq 1$. The higher $G$ value the stronger inequality in the income distribution. Later, alternative inequality indices have been defined and introduced. The generalized Gini coefficient $G(\nu) = 1 - \nu(1-\nu)\int_0^1 (1-p)^{\nu-2} L(p)dp$, where $\nu > 1$, is given in (1.1.3) and has been proposed by Yitzhaki (1983) in order to identify different distribution properties. The welfare index $W = \mu(1-G)$, given in (1.1.8) and proposed by Sen (1973) is based on the mean income ( $\mu$ ) and the Gini coefficient ( $G$ ).

Kleiber and Kotz (2001, 2002) have outlined how the income distributions can be characterised by their Lorenz curves:

## 1.4 Modelling Lorenz Curves

As an alternative to income distributions some scientists have built models for the Lorenz curve. Among these we may list the following studies: Kakwani & Podder (1973, 1976), Kakwani (1980), Rasche et al. (1980), Gupta (1984), Rao & Tam (1987), Chotikabanich (1993), Ogwang & Rao (2000), Cheong (2002), Rohde (2009) and Fellman (2012). The theoretical step from Lorenz curve to distribution function is more difficult than that from distribution function to Lorenz curve. Fellman (2012) noted that there is a difference between advanced and simple Lorenz models. Advanced Lorenz models yield a better fit to data, but are difficult to exactly connect to income distributions. Simple one-parameter models can more easily be associated with the corresponding income distribution, but when statistical analyses are performed the goodness of fit is often poor.

Rao and Tam (1987) compared five different models. The first was the three-parameter Kakwani & Podder (KP) model (1973),

$$\eta = a\pi^b\left(\sqrt{2}-\pi\right)^c \quad \begin{matrix} a>0 \\ 0\le b\le 1 \\ 0\le c\le 1 \end{matrix}, \tag{1.4.1}$$

where $\pi = \frac{L+p}{\sqrt{2}}$ and $\eta = \frac{L-p}{\sqrt{2}}$.

The second is the two-parameter generalised Pareto model (GP) analysed by Rasche et al. (1980)

$$L_{GP} = \left(1-(1-p)^a\right)^{1/b} \qquad \begin{matrix} 0\le a\le 1 \\ 0\le b\le 1 \end{matrix}, \tag{1.4.2}$$

and the third is the one-parameter Gupta (G) model (1984)

$$L_G = p\beta^{p-1}, \quad \beta > 1. \tag{1.4.3}$$

In addition, Rao and Tam constructed a generalized two-parameter Gupta model (RT)

$$L_{RT} = p^a\beta^{p-1}, \quad a, \beta > 1. \tag{1.4.4}$$

Finally, they introduced a simplified one-parameter version (S) of the RT model ( $\beta = 1$)

$$L_S = p^a \quad \alpha > 1 \tag{1.4.5}$$

Chotikabanich (1993) defined an alternative one-parameter Lorenz curve (C):

$$L_C(p) = \frac{e^{kp}-1}{e^k-1} \qquad k > 0. \tag{1.4.6}$$

The models G, S and C contain only one parameter. They are so simple that it is impossible to distinguish between the estimated length of the range for the income distribution function and the Gini coefficient. If one of these properties is estimated the other is fixed. Therefore, Fellman (2012) paid these models special attention and analysed them in more detail.

In general, the step from the Lorenz curve to the income distribution starts from the formula

$$L'(p) = \frac{x_p}{\mu}, \tag{1.4.7}$$

where $x_p$ is the $p$-percentile and $\mu$ is the mean of the corresponding distribution $F(x)$. We define $M(\cdot)$ as the inverse function of $L'(\cdot)$. From (1.4.7) we obtain

$$p = M\left(\frac{x_p}{\mu}\right). \tag{1.4.8}$$

Equation (1.4.8) indicates that $M(\cdot)$ is the income distribution function corresponding to the given Lorenz curve, that is, $F(x) = M\left(\frac{x}{\mu}\right)$. This connection between the Lorenz curve and the distribution function is easily defined, but for most of the exact Lorenz curves it is difficult or even impossible to obtain the income distribution mathematically.

**The Gupta model.** Examples of Lorenz curves for the Gupta model (1.4.3) are given in Figure 1.4.3.

Following Gupta, we observe that

$$L_G'(p) = p\beta^{p-1}\log\beta + \beta^{p-1} = \frac{x_p}{\mu}. \tag{1.4.9}$$

Consequently,

$$\lim_{p\to 0^+} x_p = \mu \lim_{p\to 0^+} L'_G = \frac{\mu}{\beta} \tag{1.4.10}$$

and

$$\lim_{p\to 1^-} x_p = \mu \lim_{p\to 1^-} L'_G = \mu(1+\log\beta) \tag{1.4.11}$$

From this it follows that Gupta ́s model corresponds to distributions defined on a finite interval $\left(\mu\beta^{-1}, \mu(1+\log\beta)\right)$. In spite of the fact that the Gupta model is relatively simple, the corresponding income distribution is not attainable. The equation (1.4.9) cannot be solved exactly with respect to variable $p$ because the variable $p$ can be found both as a factor and in the exponent.

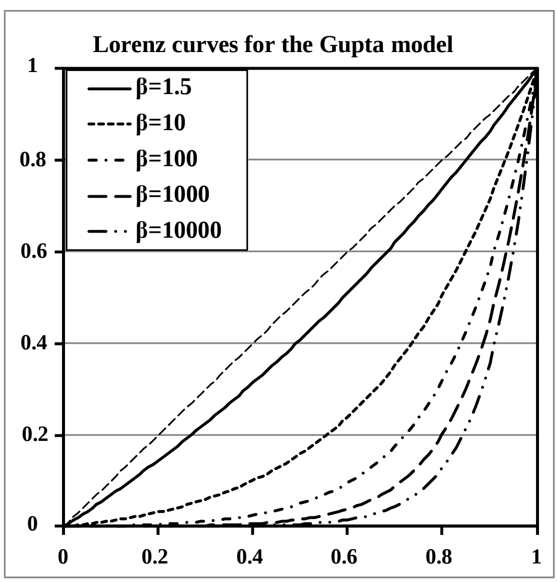



***Figure 1.4.3*** *The Lorenz curves for the Gupta model for various β values (Fellman, 2012).*

For the Gupta model, the Gini coefficient is

$$G_G = 1 - 2\int_0^1 p\beta^{p-1}dp = 1 - \frac{2}{\log\beta}\left(1 - \frac{\beta - 1}{\beta\log\beta}\right). \quad (1.4.12)$$

Figure 1.4.3 shows that the Gini coefficient tends towards 1 when $\beta \to \infty$.[2]

Following Gupta, the variable $log\,\beta$ can be estimated by using the logarithm of the model in (1.4.4), that is, from the equation $\log\left(\frac{L}{p}\right) = (p-1)\log(\beta)$.

**The generalized Gupta model (RT).** For the generalized Gupta model, we obtain.

$$\frac{x_p}{\mu} = L'_G(p) = p^{\alpha-1}\left(p\beta^{p-1}\log\beta + \alpha\beta^{p-1}\right). \quad (1.4.13)$$

The income distribution is defined on the interval $\left(0, \mu(\alpha + \log\beta)\right)$. It can be observed that if $\beta \to \infty$ the range of the income distribution then tends towards $(0, \infty)$ for both the Gupta and the generalized Gupta models.

Following Gradsheteyn and Ryshnik (1965), Rao and Tam give for the generalised Gupta model the Gini coefficient

$$G_{RT} = 1 - 2e^{-\frac{\log\beta}{(1+\alpha)}}\,{}_1F_1\left(1+\alpha; 2+\alpha; \log\beta\right), \quad (1.4.14)$$

where ${}_1F_1$ denotes the confluent hyper-geometric function with the parameters indicated in the parentheses.

[2]In Rao and Tam (1987), the formula for the Gini coefficient based on the Gupta model contains a misprint, but a numerical check of the Rao and Tam results indicates that the authors have used the correct formula in their calculations.

**The simplified RT model (S).** The simplified RT model is obtained for $\beta = 1$ and is given in (1.4.5). The Lorenz curves for various $\alpha$ values are given in Figure 1.4.4.

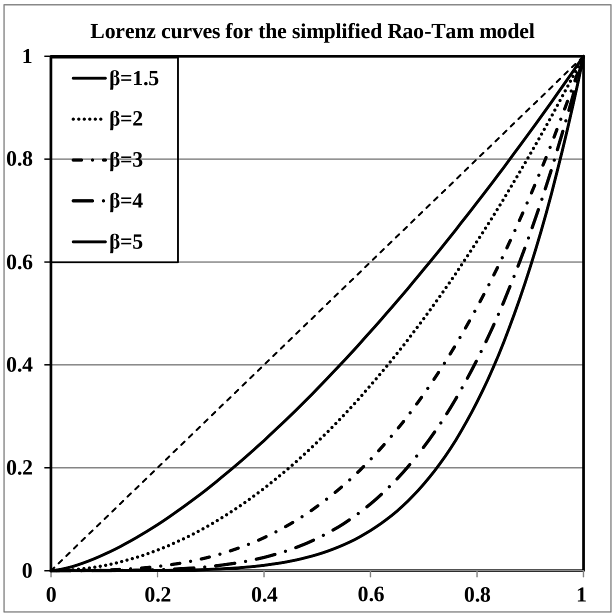



***Figure 1.4.4*** *Rao-Tam simplified Lorenz curves (Fellman, 2012).*

The Gini coefficient is $G_S = \dfrac{\alpha - 1}{\alpha + 1}$. The income distribution corresponding to the S model can be found. The derivative of $L_S(p) = p^{\alpha}$ is $L_S'(p) = \alpha p^{\alpha-1}$.

We obtain $\dfrac{x_p}{\mu} = L_S'(p) = \alpha p^{\alpha-1}$, $p^{\alpha-1} = \dfrac{x_p}{\alpha \mu}$ and $p = \left(\dfrac{x_p}{\alpha \mu}\right)^{\frac{1}{\alpha-1}}$.

Hence, the income distribution is $F(x) = \left(\dfrac{x}{\alpha \mu}\right)^{\frac{1}{\alpha-1}}$ defined on the interval $(0, \alpha \mu)$. Income distributions are given in Figure 1.4.5 for various α values.

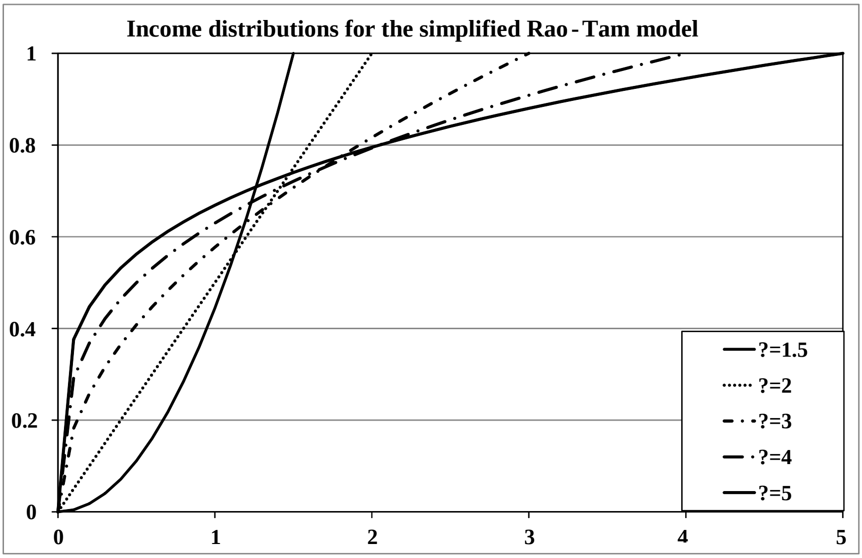



*Figure 1.4.5 Income distributions corresponding to the Rao-Tam simplified Lorenz curve (Fellman, 2012).*

**The Chotikabanich model.** Chotikabanich (1993) introduced an alternative one-parameter Lorenz curve (cf. 1.4.6)

$$L_C(p) = \frac{e^{kp} - 1}{e^k - 1} \qquad k > 0.$$

It is easily found that

$$L_C(0) = 0, \; L_C(1) = 1, \; \frac{dL_C(p)}{dp} = \frac{ke^{kp}}{e^k - 1} > 0$$

and

$$\frac{d^2L_C(p)}{dp^2} = \frac{k^2e^{kp}}{e^k - 1} > 0.$$

The second derivative is positive and hence the Lorenz curve is convex. Consequently, the first derivative is increasing from the minimum $\frac{dL_C(0)}{dp} = \frac{k}{e^k - 1} > 0$ to $\frac{dL_C(1)}{dp} = \frac{ke^k}{e^k - 1}$.

If we consider an income distribution with the mean $\mu$, then income is distributed over the interval $\left( \frac{\mu k}{e^k - 1}, \frac{\mu k e^k}{e^k - 1} \right)$. When $k \to \infty$, this interval converges towards $(0, \infty)$

Lorenz curves as functions of parameter *k* are given in Figure 1.4.6.

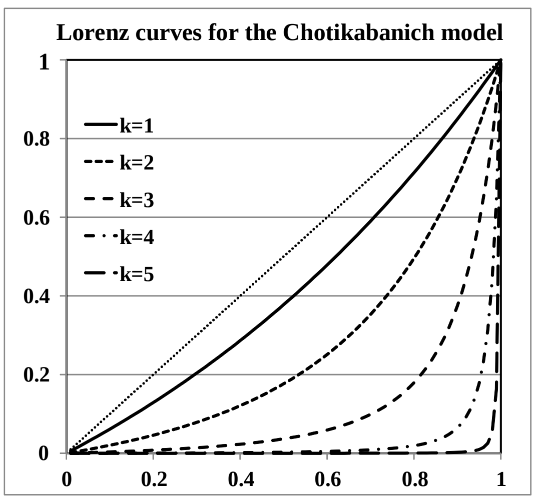



***Figure 1.4.6*** *Lorenz curves for the Chotikabanich models (c.f. Fellman, 2012).*

The Gini coefficient is.

$$G_C = 1 - 2\int_0^1 L_C(p)dp = 1 - 2\int_0^1 \frac{e^{kp} - 1}{e^k - 1} dp$$

$$=1-2\left(\left(\frac{\frac{1}{k}e^{kp}-1}{e^k-1}\right)_{p=1}-\left(\frac{\frac{1}{k}e^{kp}-1}{e^k-1}\right)_{p=0}\right)$$

$$=1-2\left(\left(\frac{\frac{1}{k}e^{k}-1}{e^k-1}\right)-\left(\frac{\frac{1}{k}-1}{e^k-1}\right)\right)=1-2\left(\frac{e^k-1}{k(e^k-1)}\right)$$

$$=\frac{k(e^k-1)-2e^k+2}{k(e^k-1)}=\frac{(k-2)e^k+k+2}{k(e^k-1)}$$

The Gini coefficient increases toward *1* when $k\to\infty$ .

If we assume an arbitrary $\mu$ , then $x_p=\frac{\mu dL_C(p)}{dp}=\frac{\mu ke^{kp}}{e^k-1}$ and we get $\frac{\mu ke^{kp}}{e^k-1}=x_p$ . Hence, $F(x)=\frac{1}{k}log\left(\frac{x(e^k-1)}{\mu k}\right)$ and the theoretical income distribution is obtained.

Figure 1.4.7 presents income distributions for various *k* values.

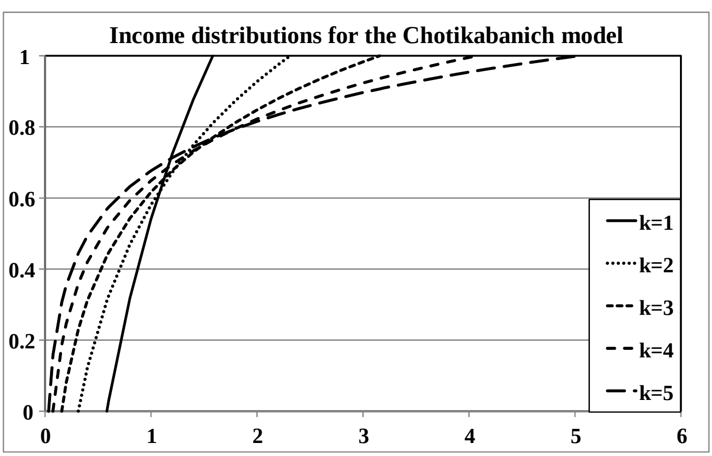



***Figure 1.4.7*** *Income distributions for the Chotikabanich models (Fellman, 2012).*

Kakwani and Podder (1976) applied their Lorenz model to Australian data, comparing four alternatives, of which all resulted in accurate estimates. The estimates varied between 0.3195 and 0.3208 when the actual value was 0.3196. Rao and Tam (1987) applied the Kakwani-Podder, the generalised Pareto, the RT, the Gupta and the simplified RT models to the same data. Their comparison of the models indicates that the Kakwani-Podder, the generalised Pareto and the RT model yielded the best estimates. The G and the S models resulted in estimates with the largest errors. For the Gupta model, the estimate was too high (0.3691) and for the simplified RT model it was too low (0.2508). The magnitudes of these errors were comparable. These findings support the criticism of the estimation based on simple one-parameter Lorenz models.

Fellman (2012) applied the Chotikabanich model and obtained the following results. He considered $\min_k \sum \left( f_{obs} - f(k) \right)^2$ and estimated the parameter $k$ and performed the minimization by using $f = L$ and $f = log(L)$. Fellman fitted the model to the Kakwani & Podder data obtained, $k = 0.2095$ and $G = 0.3262$, and $k = 0.2097$ and $G = 0.3263$, respectively. He observed that the one-parameter Chotikabanich model yields slightly better but still less exact results. As a comparison, he presented Lorenz models fitted to the Australian data graphically in his Figure 6, which we reprint in Figure 1.4.8. One observes that the Chotikabanich model is closest to the empirical curve. The simplified RT and the Gupta models show larger but comparable discrepancies. These findings support the results obtained by Rao and Tam. In Figure 1.4.8, we also observe that Gupta model yields too high an estimate of $G$ and the simplified model too low an estimate.

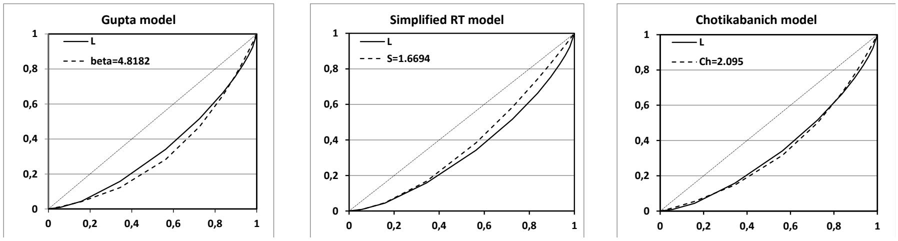



***Figure 1.4.8*** *Graphical presentation of the goodness of fit obtained by the Gupta, RT and Chotikabanich models. Note that the Chotikabanich gives the best fit (Fellman, 2012).*

Fellman (2012) studied the numerical estimation of the Gini coefficient based on Lorenz curves and this is discussed more in detail in Section 2.4. The methods were the trapezium rule, Simpson ́s rule, a modified version of Golden ́s method (2008) and the Lagrange method. In Fellman (2012) the Simpson rule could not be performed because it demands equidistant points. In general, the trapezium rule yields Gini coefficients which are too low. For the Australian data, the trapezium rule yielded the result *0.3134*, which is slightly below the correct value. Since the Lagrange method demands an even number of sub-intervals, Fellman (2012) had to modify the method slightly. He applied the Lagrange method for the ten last sub-intervals and added a small (triangular) correction from the first sub-interval. The estimate obtained was *0.3199*, a result which is closest to the correct value. Fellman (2012) presented a modified version of Golden ́s method. When he applied this method to the Australian data, he obtained the estimate of *0.3075*. This is too low, but still greater than the extremely low value obtained by the simplified RT model. Summing up, one has to choose the Lorenz model with due consideration. This is especially important if the selection should be performed among simple one-parameter models.

# References

[1] Aaberge, R. (2000). Characterizations of Lorenz curves and income distributions. *Social choice and welfare 17*:639-653.

[2] Aichison, J., Brown, J. A. C. (1954). On criteria for description of income distribution. *Metroeconomica 6*:88-107.

[3] Castellani, M. (1950). On Multinomial Distributions with Limited Freedom: A Stochastic Genesis of Pareto's and Pearson's Curves. *Ann. Math. Statist. 21*:289-293.

[4] Cheong, K. S. (2002). An empirical comparison of alternative functional forms for the Lorenz curve. *Applied Economics Letters 9*:171-176

[5] Champernowne, (1953). A model of income distribution. *The Economic Journal 63*:318-351.

[6] Chotikabanich, D. (1993). A comparison of alternative functional forms for the Lorenz curve. *Economics Letters 41*:129-138.

[7] Cramér, H. (1949). *Sannolikhetskalkylen och några av dess användningar*. Almqvist & Wiksell, Uppsala. 255pp.

[8] Dagum, C. (1977). A new model of personal income distribution: specification and estimation. *Economie Appliquee 30*:413-436.

[9] Dagum, C. (1980). Inequality measures between income distributions with applications. *Econometrica 48*:1791-1803.

[10] Dagum, C. (1987). Measuring the economic affluence between populations of income receivers. *Journal of Business & Economic Statistic, 5*:5-12.

[11] Fellman, J. (1976). The effect of transformations on Lorenz curves. *Econometrica 44*:823-824.

[12] Fellman, J. (1980). *Transformations and Lorenz curves.* Swedish School of Economics and Business Administration Working Papers: 48, 18 pp.

[13] Fellman. J. (2012). Modelling Lorenz curves. *Journal of Statistical and Econometric Methods,* 1(3):53-62.

[14] Gibrat, (1931). *Les Inégalités Economiques*. Librairie du Recueil Sirey, Paris.

[15] Gini, C. (1914). Sulla misura della concentrazione e della variabilita dei caratteri. Atti del R. Instituto veneto 73c:1203-1248.

[16] Giorgi, G. M. & Pallini, A. (1987). About a general method for the lower and upper distribution-free bounds on Gini ́s concentration ratio from grouped data. *Statistica 47*:171-184.

[17] Golden, J. (2008). A simple geometric approach to approximating the Gini coefficient. *J. Economic Education 39(1)*:68-77.

[18] Gradshteyn, I. S. & Ryshnik, I. M. (1965). *Tables of Integral Series and Products*. New York Academic press. 1086 pp.

[19] Gupta, M. R. (1984). Functional form for estimating the Lorenz curve. *Econometrica 52*:1313-1314.

[20] Harrison, A. (1981). Earnings by size: A tale of two distributions. *Review of Economic Studies 48*:621-631.

[21] Kakwani, N. (1980). On a Class of Poverty Measures *Econometrica 4*: 437-446.

[22] Kakwani N. C. & Podder N. (1973). On the Estimation of Lorenz Curves from Grouped Observations. *International Economic Review 14*: 278-292.

[23] Kakwani N. C. & Podder N. (1976). Efficient estimation of the Lorenz curve and the associated inequality measures from grouped observations. *Econometrica 44*:137-148.

[24] Kapteyn J. C. (1903). *Skew Frequency Curves in Biology and Statistics*. Vol. 1-2. 45 pp.

[25] Kleiber, Ch., Kotz, S. (2001). Characterizations of income distributions and the moment problem of order statistics. The 53rd Session of the International Statistical Institute in Seoul, Korea, Aug 22-29, Contributed Papers.

[26] Kleiber, Ch., Kotz, S. (2002). A characterization of income distributions in terms of generalized Gini coefficients. *Social Choice and Welfare 19*:789-794.

[27] Lee, W.-C. (1999). Probabilistic analysis of global performances of diagnostic tests: Interpreting the Lorenz curve based summary measures. *Statistics in Medicine 18*:455-471.

[28] Lorenz, M. O. (1905). Methods for measuring concentration of wealth. *J. Amer. Statist. Assoc. New Series, No. 70*: 209-219.

[29] Lydall, H. F. (1968). *The Structure of Earnings*. (London: Oxford University press).

[30] McDonald, J. B., Ransom, M. R. (1979). Functional Forms, estimation techniques and the distribution of income. *Econometrica 47*:1513-1525.

[31] Ogwang, T. & Rao, U. L. G, (2000). Hybrid models of the Lorenz curve, *Economics Letters, 69*:39-44.

[32] Pareto, V. (1897). *Cours d'Economie Politique*. Lausanne, Suisse.

[33] Quensel, C. E. (1944). *Inkomstfördelning och skattetryck*. Sveriges industriförbund, Lund.

[34] Rao, U. L. G. & Tam, A. Y.-P. (1987). An empirical study of selection and estimation of alternative models of the Lorenz curve. *J. of Applied Statistics 14*:275-280.

[35] Rasche, R. H., Gaffney, J., Koo A. Y. C. & Obst, N. (1980). Functional Forms for Estimating the Lorenz Curve. *Econometrica 48*:1061-1062.

[36] Rhodes, E. C. (1944). The Pareto distribution of incomes. *Economica New Series, XI 41*:1-11.

[37] Rohde, N. (2009). An alternative functional form for estimating the Lorenz curve. *Economics Letters 105*:61-63.

[38] Sen, A. (1973). *On Economic Inequality*. Clarendon Press, Oxford.

[39] Yitzhaki, S. (1983). On an extension of the Gini index. *International Economic Review 24*:617-628.

# 2

# Income Transformations

Redistributions of income according to tax or transfer policies can be considered as variable transformations of the initial income. The transformation is usually assumed to be positive, monotone-increasing and continuous. Recently, Fellman (2009, 2011) has also discussed discontinuous transformations. If the transformation is considered as a tax or a transfer policy, the transformed variable is either the post-tax or the post-transfer income. A central problem in the literature has been the Lorenz dominance, defined above, between the initial and the transformed income (c.f. Fellman, 1976; Jakobsson, 1976; Kakwani, 1977) (see also Theorem 2.1.1 below). Under the assumption that the theorems should hold for all income distributions, the conditions are both necessary and sufficient (Jakobsson, 1976; Fellman, 2009).

## 2.1 Income Redistributions

Variable transformations are valuable when one studies the effect of tax and transfer policies on the income inequality. If the transformation should result in an increasing transformed variable with finite mean then discontinuities can only consist of finite positive jumps and the number of jumps has to be finite or countable. In this study we reconsider the effect of variable transformations on the redistribution of income (Fellman, 1976, Jakobsson, 1976, Kakwani, 1977 and Hemming & Keen, 1983). The continuity of the transformations can be implicitly included in the necessary and sufficient conditions. One main result is that continuity is a necessary condition if one pursues that the income inequality should remain or be reduced.

Consider the income $X$ with the distribution function $F_X(x)$, the mean $\mu_X$, and the Lorenz curve $L_X(p)$. We assume that $X$ is defined for $x \geq 0$. If we assume that the density function $f_X(x)$ exists, we follow Section 1.1 and obtain the formulae

$$\mu_X = \int_0^\infty x\, f_X(x)dx \tag{2.1.1}$$

and

$$L_X(p) = \frac{1}{\mu_X}\int_0^{x_p} x\, f_X(x)dx. \tag{2.1.2}$$

We consider the transformation $Y = u(X)$, where $u(\cdot)$ is non-negative and monotone increasing. The transformation can be considered as a tax or a transfer policy and consequently, the transformed variable is the post-tax or post-transfer income, respectively.

For the transformed variable $Y$ we obtain the distribution function

$$F_Y(y) = P(Y \le y) = P(u(X) \le y) = P\left(X \le u^{-1}(y)\right) = F_X\left(u^{-1}(y)\right) \tag{2.1.3}$$

Using this result we obtain the mean and the Lorenz curve for the variable $Y$.

$$\mu_Y = \int_0^1 u(x) f_X(x)dx \tag{2.1.4}$$

and

$$L_Y(p) = \frac{1}{\mu_Y}\int_0^{x_p} u(x) f_X(x)dx. \tag{2.1.5}$$

The fundamental theorem is:

**Theorem 2.1.1.** (Fellman, 1976, Jakobsson, 1976 and Kakwani, 1977). Let X be an arbitrary non-negative, random variable with the distribution $F_X(x)$, mean $\mu_X$ and Lorenz curve $L_X(p)$. Let $u(x)$ be non-negative, continuous and monotone-increasing and let $\mu_Y = E(u(X))$ exist. Then the Lorenz curve $L_Y(p)$ of $Y = u(X)$ exists and the following results hold:

(i) $L_Y(p) \geq L_X(p)$ if $\dfrac{u(x)}{x}$ is monotone decreasing.

(ii) $L_Y(p) = L_X(p)$ if $\dfrac{u(x)}{x}$ is constant and

(iii) $L_Y(p) \leq L_X(p)$ if $\dfrac{u(x)}{x}$ is monotone increasing.

According to this theorem we obtain in (i) a sufficient condition that the transformation $u(x)$ generates a new income distribution which Lorenz dominates the initial one. If we analyse the proof of the case (i) in Fellman (1976, Theorem 1) we observe that the difference $L_Y(p) - L_X(p)$ can be written

$$D(p) = L_Y(p) - L_X(p) = \int_0^{x_p} \left( \frac{u(x)}{\mu_Y} - \frac{x}{\mu_X} \right) f_X(x) dx \qquad (2.1.6)$$

where $x_p = F_X^{-1}(p)$. In any case, $D(0) = D(1) = 0$. In order to obtain Lorenz dominance the difference $D(p)$ must start from zero and then attain positive values and after that decrease back to zero and the integrand in (2.1.6) must start from positive (non-negative) values and then change its sign and become negative. Consequently, $\dfrac{u(x)}{x}$ has to be a decreasing function.

The condition is necessary if the rule should hold for all income distributions $F_X(x)$ (Jakobsson, 1976). Otherwise we can find a transformation $u(x)$ for which the quotient $\dfrac{u(x)}{x}$ is not monotone decreasing for all $x > 0$, and a distribution $F_X(x)$ such that the result in the proof holds, i.e. dominance is obtained. Assume that the quotient $\frac{u(x)}{x}$ is both increasing and decreasing.[3] Let a transformation $u(x)$ satisfy the initial conditions (non-negative, continuous and monotone increasing) and let $\frac{u(x)}{x}$ be increasing within some interval ( $0 < a < x < b < \infty$ ). Now we present a distribution such that the transformed variable $Y = u(X)$ does not Lorenz dominate the initial variable *X*. Consider a distribution with a continuous density function,

$$f_X(x) = \begin{cases} 0 & 0 \le x < a \\ f_0(x) > 0 & a \le x \le b \\ 0 & x > b \end{cases} \tag{2.1.7}$$

For the pair ( $f_X(x), u(x)$) the formula (2.1.6) can be written

$$D(p) = \int_a^{x_p} \frac{x}{\mu_Y}\left(\frac{u(x)}{x} - \frac{\mu_Y}{\mu_X}\right) f_X(x)dx, \tag{2.1.8}$$

where $a \le x_p \le b$.

We observe that $D(0) = D(1) = 0$, that Theorem 1(iii) holds and that the transformation results in a new variable *Y* which is Lorenz dominated by the initial variable *X*. Hence, if we demand that the transformed variable $Y = u(X)$ shall Lorenz dominate *X* for all distributions $F_X(x)$, then the condition in Theorem 2.1.1 (i) is necessary (Jakobsson, 1976, Lambert, 2001; Chapter 8).

[3]If $\frac{u(x)}{x}$ is monotonously increasing for all $x > 0$ then the proposition (iii) holds and this case can be ignored.

Hemming and Keen (1983) gave a new condition for the Lorenz dominance. Their condition is, with our notations, that for a given distribution $F_X(x)$ the function $u(x)$ crosses the line $\frac{\mu_Y}{\mu_X}x$ once from above, that is that $\frac{u(x)}{x}$ crosses the level $\frac{\mu_Y}{\mu_X}$ once from above. We observe that if their condition holds then the integrand in (2.1.6) starts from positive values changes its sign once and ends up with negative values and their condition is equivalent with our condition. For the example considered above, the Hemming-Keen condition is not satisfied. The integrand is zero for $x < a$ and for $x > b$. For $a \le x \le b$ the ratio $\frac{u(x)}{x}$ is increasing and if it crosses $\frac{\mu_Y}{\mu_x}$ it cannot do it from above. Consequently, if $\frac{u(x)}{x}$ is not monotone decreasing then there are distributions for which the Hemming-Keen condition does not hold.

On the other hand if we assume that $\frac{u(x)}{x}$ is monotone decreasing then $\frac{u(x)}{x}$ satisfies the condition "crossing once from above for every distribution $F_X(x)$". Hence, our condition and Hemming-Keen condition are also equivalent as necessary conditions.

In a similar way we can prove that if the other results in Theorem 2.1.1 should hold for every income distribution the conditions in (ii) and in (iii) are also necessary.

The results obtained, indicate that if $\frac{u(x)}{x}$ is continuous and monotone increasing even in a short interval, then there are income distributions such that the transformation $u(x)$ cannot result in Lorenz dominance. What can be said if $u(x)$ is discontinuous? Assume that $u(x)$ is still positive and monotone increasing. Assume furthermore, that $E(u(X)) = \mu_Y$ exists for every stochastic variable $X$ with finite mean $\mu_X$. Above we stressed that the discontinuities of $u(x)$ can only consist of finite positive jumps and the number of jumps can be assumed to be finite or countable. Assume that elsewhere $u(x)$ satisfies all the other conditions including the condition in Theorem 2.1.1(i). We will prove that if $u(x)$ is discontinuous there exists a distribution $F_X(x)$ such that the transformation $Y = u(X)$ does not Lorenz dominate the initial variable $X$. Again we follow the arguments given by Jakobsson (1976). However, the discontinuity demands a more detailed reasoning.

Let $a > 0$ be a discontinuity point, such that $lim_{x \to a-} u(x) = u_0$ and $lim_{x \to a+} u(x) = u_0 + d$, where the jump $d > 0$. (The notation $lim_{x \to a-} u(x)$ indicates limit from the left and $lim_{x \to a+} u(x)$ limit from the right.) We do not assume anything about how $u(x)$ is defined in the point $a$. The following analyses are based on Fellman (2009). Choose $h > 0$ so small that the point $a$ is the only discontinuity point within the interval $(a - h, a + h)$. (Later we may reduce the interval even more). Let $t$ and $z$ be arbitrary values satisfying the inequalities

$$a - h < t \leq a \leq z < a + h.$$

If $u(x)$ is monotone increasing we have $u(t) \leq u_0 < u_0 + d \leq u(z)$ and

$$lim_{t\to a-}\left(\frac{u(t)}{t}\right)=\frac{u_0}{a}<\frac{u_0+d}{a}=lim_{z\to a+}\left(\frac{u(z)}{z}\right).$$

Hence, the quotient $\frac{u(x)}{x}$ cannot be monotone decreasing within the interval $(a-h,a+h)$. Consider a variable *X*, having the symmetric density function

$$f_X(x)=\begin{cases}0 & x<a-h\\ \frac{1}{h}\left(1-\frac{1}{h}|a-x|\right) & a-h\le x\le a+h\\ 0 & x>a-h\end{cases}. \tag{2.1.9}$$

The mean $E(X)=\mu_X=a$. For the transformed variable $Y=u(X)$ the mean is

$$\begin{aligned}\mu_Y=E(Y)&=\int_{a-h}^{a}u(x)f_X(x)dx+\int_{a}^{a+h}u(x)f_X(x)dx\\&=u(\alpha_1)\int_{a-h}^{a}f_X(x)dx+u(\alpha_2)\int_{a}^{a+h}f_X(x)dx\\&=\frac{1}{2}(u(\alpha_1)+u(\alpha_2))\end{aligned}, \tag{2.1.10}$$

where $a-h<\alpha_1<a$ and $a<\alpha_2<a+h$.

If $h\to 0$ then $\mu_Y\to u_0+\frac{1}{2}d$. Assume furthermore, that we have chosen *h* so small that $\mu_Y>u_0+\frac{1}{4}d$. Consider now

$$D(p)=L_Y(p)-L_X(p)=\int_{a-h}^{x_p}\frac{x}{\mu_Y}\left(\frac{u(x)}{x}-\frac{\mu_Y}{\mu_X}\right)f_X(x)dx, \tag{2.1.11}$$

where $F_X(x_p) = p$. In order to obtain Lorenz dominance the integrand must start from positive (non-negative) values and then change its sign and become negative in such a manner that the difference $D(p)$ starts from zero and then attains positive values and after that it decreases back to zero. Within the interval $(a-h, a+h)$ the sign of the integrand depends on the factor $\frac{u(x)}{x} - \frac{\mu_Y}{\mu_X}$, which starts from the value

$$\frac{u(a-h)}{a-h} - \frac{\mu_Y}{a} \le \frac{u_0}{a-h} - \frac{u_0 + \frac{1}{4}d}{a} \le \frac{-\frac{1}{4}ad + h(u_0 + \frac{1}{4}d)}{a(a-h)}.$$

If we assume that $h$ satisfies the earlier conditions and in addition, the condition $h < \frac{ad}{4u_0 + d}$, the parenthesis in (2.1.11) starts from negative values and consequently, the whole integrand is negative and $D(p)$ starts from negative values. For the corresponding income distribution the transformed variable $Y$ does not Lorenz dominate the initial variable $X$. Hence, the continuity of $u(x)$ is a necessary condition if we demand that the transformed variable should Lorenz dominate the initial variable for every distribution. From this it follows that if the condition in Theorem 2.1.1(i) has to be necessary it implies continuity and hence, an explicit statement of continuity can be dropped. If we study the condition in (ii) we observe that $u(x) = kx$ and consequently, $u(x)$ has to be continuous.

However, in the case (iii) the discontinuity does not jeopardize the monotone increasing property of the quotient $\frac{u(x)}{x}$ and the result in Theorem 2.1.1 (iii) holds even if the function is discontinuous. Therefore, also in this case we can drop the explicit continuity assumption.

Summing up, for arbitrary distributions, $F_X(x)$, the conditions (i), (ii), and (iii) in Theorem 2.1.1 are both necessary and sufficient for the dominance relations and an additional assumption about the continuity of the transformation $u(x)$ can be dropped. We obtain the more general theorem (Jakobsson, 1976; Fellman, 2009).

**Theorem 2.1.2.** Let X be an arbitrary non-negative, random variable with the distribution $F_X(x)$, mean $\mu_X$ and the Lorenz curve $L_X(p)$, let $u(x)$ be a non-negative, monotone increasing function and let $Y = u(X)$ and $E(Y) = \mu_Y$ exist. Then the Lorenz curve $L_Y(p)$ of Y exists and the following results hold:

(i) $L_Y(p) \geq L_X(p)$ if and only if $\dfrac{u(x)}{x}$ is monotone-decreasing.

(ii) $L_Y(p) = L_X(p)$ if and only if $\dfrac{u(x)}{x}$ is constant.

(iii) $L_Y(p) \leq L_X(p)$ if and only if $\dfrac{u(x)}{x}$ is monotone-increasing.

**Remark.** From the discussion above it follows that only in the case (iii) the transformation $u(x)$ can be discontinuous.

Now, we analyse the effect of a finite step in $u(x)$ on the Lorenz curve. We use the notations presented above.

Let $t \leq a \leq z$, $r = F_X(t)$, $q = F_X(a)$ and $s = F_X(z)$.

Consider the difference

$$\Delta L_Y = L_Y\left(F_X(z)\right) - L_Y\left(F_X(t)\right) = \frac{1}{\mu_Y}\int_t^z u(x) f_X(x)\, dx$$

$$= \frac{1}{\mu_Y}\int_t^a u(x) f_X(x) dx + \frac{1}{\mu_Y}\int_a^z u(x) f_X(x) dx = \frac{u(\alpha_1)}{\mu_Y}(q-r) + \frac{u(\beta_1)}{\mu_Y}(s-q)$$

where $t \le \alpha_1 \le a$ and $a \le \beta_1 \le z$.

When $t \to a-$ and $z \to a+$, then $q-r \to 0$, $s-q \to 0$ and $\Delta L_Y \to 0$. Hence, although the transformation $u(x)$ is discontinuous in the point $a$, the Lorenz curve is continuous. However, it is not differentiable. For every $t < a$ we obtain

$$\Delta L_Y = L_Y(q) - L_Y(r) = \frac{1}{\mu_Y}\int_t^a u(x) f_X(x) dp = \frac{u(\eta)}{\mu_Y}(q-r)$$

where $t < \eta < a$. We obtain $\dfrac{\Delta L_Y}{q-r} = \dfrac{u(\eta)}{\mu_Y}$. When $q-r \to 0+$ then $\eta \to a-$ and $\dfrac{\Delta L_Y}{q-r} \to \dfrac{u_0}{\mu_Y}$. Hence, $L_Y(p)$ has the left derivative $\left(\dfrac{dL_Y(p)}{dp}\right)_{p=q-} = \dfrac{u_0}{\mu_Y}$.

For every $z > a$ we obtain

$$\Delta L_Y = L_Y(s) - L_Y(q) = \frac{1}{\mu_Y}\int_q^s u(x) f_X(x) dp = \frac{u(\varsigma)}{\mu_X}(s-q),$$

where $a < \varsigma < z$. We obtain $\dfrac{\Delta L_Y}{s-q} = \dfrac{u(\varsigma)}{\mu_Y}$. When $s-q \to 0+$ then $\eta \to a+$ and $\dfrac{\Delta L_Y}{s-q} \to \dfrac{u_0 + d}{\mu_Y}$. Hence, $L_Y(p)$ has the right derivative

$$\left(\frac{dL_Y(p)}{dp}\right)_{p=q+} = \frac{u_0 + d}{\mu_Y} \neq \frac{u_0}{\mu_Y} = \left(\frac{dL_Y(p)}{dp}\right)_{p=q-}. \tag{2.1.12}$$

Consequently, $L_Y(p)$ is continuous in the point $q = F_X(a)$ but it is not differentiable and has a cusp for $p = q$.

**Remark.** If the transformation $u(x)$ is continuous then $d = 0$ and we obtain equality in (2.1.12) and the Lorenz curve is differentiable with the derivative $L'_Y(p) = \frac{y_p}{\mu_Y}$.

For progressive taxations, $u(x)$ is the post-tax income and $\frac{u(x)}{x}$ measures the proportion of post-tax income to the initial income and it is a monotone decreasing function satisfying the condition (i) and the Lorenz curve is increased and $F_Y(y)$ Lorenz dominates $F_X(x)$. If the taxation is a flat tax then (ii) holds and the Lorenz curve and the Gini value remain. The third case in Theorem 2.1.1 indicates that the ratio $\frac{u(x)}{x}$ is increasing and the Gini coefficient increases, but this case has minor practical importance. If transfer policies are studied, then the ratio $\frac{u(x)}{x}$ measures the relative effect of the transfer. If it decreases the relative effect of the transfer decreases with increasing income and the inequality is reduced. If $\frac{u(x)}{x}$ is constant, the transformation $u(x)$ is proportional to the initial income and the Lorenz curve and the Gini value remain.

## 2.2 Additional Properties of Lorenz Curves for Transformed Income Distributions

We follow Fellman (2012b) who considered income *X*, defined on the interval $(a,b)$, where $0 \le a \le x \le b \le \infty$, with the distribution function $F_X(x)$, density function $f_X(x)$, mean $\mu_X$, percentile $x_p$ defined as $F_X(x_p) = p$ and Lorenz curve $L_X(p)$. The general formulae are

$$\mu_X = \int_a^b x\, f_X(x)dx \tag{2.2.1}$$

and

$$L_X(p) = \frac{1}{\mu_X}\int_a^{x_p} x\, f_X(x)dx, \tag{2.2.2}$$

where $a \le x_p \le b$.

We consider the transformation $Y = u(X)$, where $u(\cdot)$ is non-negative, continuous and monotone-increasing. Since the transformation can be considered as a tax ($u(x) \le x$) or a transfer policy ($u(x) \ge x$), the transformed variable is either the post-tax or the post-transfer income.

The mean and the Lorenz curve for variable *Y* are

$$\mu_Y = \int_a^b u(x) f_X(x)dx \tag{2.2.3}$$

and

$$L_X(p) = \frac{1}{\mu_X}\int_a^{x_p} u(x) f_X(x)dx, \tag{2.2.4}$$

In the following, we consider additional properties of the Lorenz curve $L_Y(p)$. If $\frac{u(x)}{x}$ is constant, then according to Theorem 1 (ii), $L_Y(p)=L_X(p)$ and the transformed Lorenz curve is identical with the initial one, a case which will be ignored.

*(a) The ratio $\frac{u(x)}{x}$ is monotonically decreasing.*

According to Theorem 2.1.1, $F_Y(y)$ Lorenz dominates $F_X(x)$. We introduce the values $M$ and $m$ such that $\lim_{x\to a+}\frac{u(x)}{x}=M\le\infty$ and $\lim_{x\to b-}\frac{u(x)}{x}=m\ge 0$. Consequently, $\infty\ge M\ge\frac{u(x)}{x}\ge m\ge 0$.

Let $F_X(x_p)=p$ and $F_X(x_q)=q$. Assume that $p\le q$ and that $a\le x_p\le x\le x_q\le b$ and consequently,

$$M\ge\frac{u(x_p)}{x_p}\ge\frac{u(x)}{x}\ge\frac{u(x_q)}{x_q}\ge m.$$

Note that points $p$ and $q$ are chosen arbitrarily and that the equality signs cannot be ignored because we also include the functions $\frac{u(x)}{x}$, which are not uniformly strict decreasing in the class of transformations. Hence, we have to include members for which equalities hold for almost the whole range and, in addition, sub-intervals in which strict inequalities hold can be chosen arbitrarily short and located arbitrarily within the range $(a,b)$. If one pursues general conditions, the inequalities (2.2.8) and (2.2.9) obtained below cannot be improved. If we assume that $\frac{u(x)}{x}$ is monotonically decreasing, then $u(x)$ must be continuous, otherwise $\frac{u(x)}{x}$ should have positive jumps (Fellman, 2009).

From $\frac{u(x_p)}{x_p} \geq \frac{u(x)}{x}$ it follows that $x_p u(x) \leq x u(x_p)$. The integration over the interval $x_p \leq x \leq x_q$ yields

$$\int_{x_p}^{x_q} x_p u(x) f_X(x)dx \leq \int_{x_p}^{x_q} xu(x_p) f_X(x)dx$$

$$x_p \int_{x_p}^{x_q} u(x) f_X(x)dx \leq u(x_p) \int_{x_p}^{x_q} xf_X(x)dx$$

$$x_p \mu_Y \left(L_Y(q) - L_Y(p)\right) \leq u(x_p)\mu_X \left(L_X(q) - L_X(p)\right)$$

and

$$\left(L_Y(q) - L_Y(p)\right) \leq \frac{u(x_p)\mu_X}{x_p \mu_Y} \left(L_X(q) - L_X(p)\right). \tag{2.2.5}$$

Analogously, it follows from $\frac{u(x)}{x} \geq \frac{u(x_q)}{x_q}$ that $x_q u(x) \geq xu(x_q)$, and we obtain

$$\left(L_Y(q) - L_Y(p)\right) \geq \frac{u(x_q)\mu_X}{x_q \mu_Y} \left(L_X(q) - L_X(p)\right). \tag{2.2.6}$$

Consequently,

$$\frac{u(x_p)\mu_X}{\mu_Y x_p} \left(L_X(q) - L_X(p)\right) \geq \left(L_Y(q) - L_Y(p)\right) \geq \frac{u(x_q)\mu_X}{\mu_Y x_q} \left(L_X(q) - L_X(p)\right). \tag{2.2.7}$$

When $p \to 0$ in (2.2.7), then $L_Y(p) \to 0$, $L_X(p) \to 0$, $\frac{u(x_p)}{x_p} \to M$ and one obtains

$$\frac{M\mu_X}{\mu_Y} L_X(q) \geq L_Y(q) \geq \frac{u(x_q)\mu_X}{\mu_Y x_q} L_X(q). \tag{2.2.8}$$

The lower bound gives an evaluation of how much the Lorenz curve has increased. The upper bound is of minor interest and is commented on later.

When $q \to 1$ in (2.2.7), then $L_Y(q) \to 1$, $L_X(q) \to 1$, $\frac{u(x_q)}{x_q} \to m$ and one obtains

$$1 - \frac{m\mu_X}{\mu_Y}\left(1 - L_X(p)\right) \geq L_Y(p) \geq 1 - \frac{u(x_p)\mu_X}{\mu_Y x_p}\left(1 - L_X(p)\right).$$

In order to compare these inequalities with the inequalities in (2.2.8), we change the argument from *p* to *q*, and the inequalities are

$$\begin{aligned} 1 - \frac{m\mu_X}{\mu_Y}\left(1 - L_X(q)\right) &\geq L_Y(q) \geq \\ &1 - \frac{u(x_q)\mu_X}{\mu_Y x_q}\left(1 - L_X(q)\right). \end{aligned} \tag{2.2.9}$$

The lower bound gives an evaluation of how much the Lorenz curve has increased. The upper bound is of minor interest and is discussed later.

Inequality (2.2.8) is applicable to small values and inequality (2.2.9) to large values of *q*. For small values of *q*, we consider the difference

$$D_1(q) = L_Y(q) - \frac{u(x_q)\mu_X}{\mu_Y x_q} L_X(q) \tag{2.2.10}$$

and for large $q$ we consider the difference

$$D_2(q) = L_Y(q) - 1 + \frac{u(x_q)\mu_X}{\mu_Y x_q}\left(1 - L_X(q)\right). \tag{2.2.11}$$

In general, $\dfrac{dL_Y(q)}{dq} = \dfrac{y_q}{\mu_Y} = \dfrac{u(x_q)}{\mu_Y}$ and $\dfrac{dL_X(q)}{dq} = \dfrac{x_q}{\mu_X}$.

The ratio $\dfrac{u(x)}{x}$ is decreasing and consequently

$$\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) = \frac{d}{dq}\left(\frac{y_q}{x_q}\right) = \frac{d}{dx_q}\left(\frac{y_q}{x_q}\right)\frac{d}{dq}\left(x_q\right) \le 0 .$$

Now we differentiate $D_1(q)$ and obtain

$$\begin{aligned} \frac{d\left(D_1(q)\right)}{dq} &= \frac{u(x_q)}{\mu_Y} - \frac{u(x_q)}{\mu_Y}\frac{\mu_X}{x_q}\frac{x_q}{\mu_X} - L_X(q)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) \\ &= -L_X(q)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) \ge 0 \end{aligned}.$$

Consequently $D_1(q)$ is increasing from zero at $q = 0$ to a maximum $D_1(q_0)$ for $q = q_0$ (say).

Now we differentiate $D_2(q)$ and obtain

$$\begin{aligned} \frac{d\left(D_2(q)\right)}{dq} &= \frac{u(x_q)}{\mu_Y} - \frac{u(x_q)}{\mu_Y} + \left(1 - L_X(q)\right)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) \\ &= \left(1 - L_X(q)\right)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) \le 0 . \end{aligned}$$

Consequently $D_2(q)$ is decreasing from $D_2(q_0)$ to zero when $q \to 1$. The point $q_0$, at which the shift from (2.2.10) to (2.2.11) is performed, is chosen so that $D_1(q_0) = D_2(q_0)$.

Now,

$$L_Y(q_0) - 1 + \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}}\left(1 - L_X(q_0)\right) = L_Y(q_0) - \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}} L_X(q_0);$$

that is,

$$1 - \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}} = 0 \text{ and } \frac{u(x_{q_0})}{x_{q_0}} = \frac{\mu_Y}{\mu_X}.$$

Consequently,

$$D_1(q_0) = D_2(q_0) = L_Y(q_0) - L_X(q_0)$$

Since the ratio $\frac{u(x)\mu_X}{x\mu_Y}$ is decreasing, the difference $\frac{u(x_{q_0})}{x_{q_0}} - \frac{\mu_Y}{\mu_X} = 0$ shifts its sign from plus to minus at point $q_0$. Hemming and Keen (1983) gave the condition for Lorenz dominance that $\frac{u(x)}{x}$ crosses the $\frac{\mu_Y}{\mu_X}$ level once from above. Our results above have shown that the crossing point is $q_0$. The condition obtained can also be otherwise explained. If we write it as $\frac{u(x_{q_0})}{\mu_Y} = \frac{x_{q_0}}{\mu_X}$, we obtain the formula $\frac{dL_Y(q)}{dq}\Big|_{q=q_0} = \frac{dL_X(q)}{dq}\Big|_{q=q_0}$, that is, the Lorenz curves $L_Y(q)$ and $L_X(q)$ have parallel tangents and the distance $L_Y(q_0) - L_X(q_0)$ between the Lorenz curves is maximal for $q = q_0$.

We define the difference function as

$$\tilde{D}(q) = \begin{cases} D_1(q) \text{ for } q \le q_0 \\ D_2(q) \text{ for } q > q_0 \end{cases}, \tag{2.2.12}$$

and the lower bound of $L_Y(p)$ is

$$\tilde{L}(q) = \begin{cases} \dfrac{u(x_q)\mu_X}{\mu_Y x_q} L_X(q) & \text{for} \quad q \le q_0 \\ 1 - \dfrac{u(x_q)\mu_X}{\mu_Y x_q}\left(1 - L_X(q)\right) & \text{for} \quad q > q_0 \end{cases}. \tag{2.2.13}$$

Figure 2.2.1 shows the Lorenz curves $L_Y(q)$, $L_X(q)$, the lower bound $\tilde{L}(q)$ and the difference $\tilde{D}(q)$ between $L_Y(q)$ and the lower bound $\tilde{L}(q)$.

**Remarks.** The variable *Y* Lorenz dominates *X*, and the upper bounds in (2.2.8) and (2.2.9) tells us nothing about the reductions in the inequality. The upper bound contains the maximum value $M$ and one has to take it for granted that it is also inaccurate when M is finite. In addition, there may be situations in which $M = \infty$. The minimum value *m* can be zero, and in this case the upper bound is one and the obvious inequality $L_Y(p) \le 1$ is obtained.

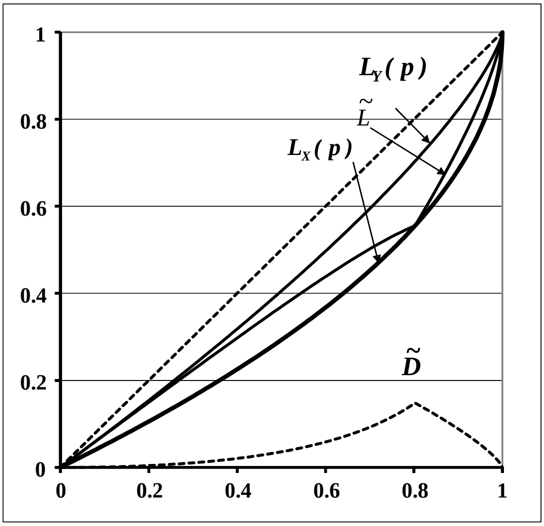



***Figure 2.2.1*** *(Fellman (2012b) A sketch of the Lorenz curves $L_Y(q)$, $L_X(q)$, the lower bound $\tilde{L}(q)$, and the difference $\tilde{D}(q)$ between $L_Y(q)$ and the lower bound $\tilde{L}(q)$ when the transformed variable Lorenz dominates the initial one.*

*(b) The ratio $\frac{u(x)}{x}$ is monotonically increasing.*

The analysis of this case follows similar traces to the earlier study and the results are analogous to our earlier results, but in this case $u(x)$ may be discontinuous. Only the inequality signs have changed their directions. We introduce the values $M$ ($\leq \infty$) and $m$ ($\geq 0$) such that

$$\lim_{x \to a+} \frac{u(x)}{x} = m \text{ and } \lim_{x \to b-} \frac{u(x)}{x} = M$$

and consequently $0 \leq m \leq \frac{u(x)}{x} \leq M \leq \infty$. Note, that in this case the points $p$ and $q$ are also chosen arbitrarily and that the equality signs cannot be ignored because we also include functions $\frac{u(x)}{x}$ which are not uniformly strictly increasing in the class of transformations. Hence, we have to include members for which equalities hold for almost the whole range and, in addition, the subintervals where strict inequalities hold can be arbitrarily short and can be located arbitrarily within the range. If one pursues general conditions, the inequalities (2.2.17) and (2.2.18) obtained below cannot be improved.

If $u(x)$ is discontinuous, the discontinuities can only be a countable number of finite positive jumps. Under such circumstances $u(x)$ is still integrable.

We use the same notations as above and assume that $F_X(x_p)=p$, $F_X(x_q)=q$, that $p \le q$ and consequently that $x_p \le x \le x_q$. Now, $\frac{u(x_p)}{x_p} \le \frac{u(x)}{x} \le \frac{u(x_q)}{x_q}$. Consider $x_p u(x) \ge x u(x_p)$. The integration over the interval $x_p \le x \le x_q$ yields

$$\int_{x_p}^{x_q} x_p u(x) f_X(x)dx \ge \int_{x_p}^{x_q} x u(x_p) f_X(x)dx$$

$$x_p \int_{x_p}^{x_q} u(x) f_X(x)dx \ge u(x_p) \int_{x_p}^{x_q} x f_X(x)dx$$

$$x_p \mu_Y \left(L_Y(q)-L_Y(p)\right) \ge u(x_p)\mu_X \left(L_X(q)-L_X(p)\right)$$

and

$$\left(L_Y(q)-L_Y(p)\right) \ge \frac{u(x_p)\mu_X}{x_p \mu_Y}\left(L_X(q)-L_X(p)\right). \tag{2.2.14}$$

Analogously, if we consider $x_q u(x) \le x u(x_q)$ we obtain

$$x_q \mu_Y \left( L_Y(q) - L_Y(p) \right) \le u(x_q) \mu_X \left( L_X(q) - L_X(p) \right)$$

and

$$\left( L_Y(q) - L_Y(p) \right) \le \frac{u(x_q)\mu_X}{x_q \mu_Y} \left( L_X(q) - L_X(p) \right). \tag{2.2.15}$$

Hence,

$$\frac{u(x_p)\mu_X}{\mu_Y x_p} \left( L_X(q) - L_X(p) \right) \le \left( L_Y(q) - L_Y(p) \right) \le$$
$$\frac{u(x_q)\mu_X}{\mu_Y x_q} \left( L_X(q) - L_X(p) \right). \tag{2.2.16}$$

When $p \to 0$ in (2.2.16), then $L_Y(p) \to 0$, $L_X(p) \to 0$, $\frac{u(x_p)}{x_p} \to m$ and one obtains

$$\frac{m\mu_X}{\mu_Y} L_X(q) \le L_Y(q) \le \frac{u(x_q)\mu_X}{\mu_Y x_q} L_X(q). \tag{2.2.17}$$

Now, the initial variable *X* Lorenz dominates the transformed *Y* and the upper bound is the interesting case.

When $q \to 1$ in (2.2.16), then $L_Y(1) \to 1$, $L_X(q) \to 1$, $\frac{u(x_q)}{x_q} \to M$ one obtains

$$1 - \frac{u(x_p)\mu_X}{\mu_Y x_p} \left( 1 - L_X(p) \right) \ge L_Y(p) \ge 1 - \frac{M\mu_X}{\mu_Y} \left( 1 - L_X(p) \right)$$

After a shift from *p* to *q*, we obtain

$$1-\frac{u(x_q)\mu_X}{\mu_Y x_q}\left(1-L_X(q)\right)\geq L_Y(q)\geq 1-\frac{M\mu_X}{\mu_Y}\left(1-L_X(q)\right). \qquad (2.2.18)$$

Now the upper bound is of interest. Formula (2.2.17) is applicable for small values and formula (2.2.16) for large values of *q*. In the following, we consider the difference between the upper bound in (2.2.17) and the Lorenz curve $L_Y(q)$, that is, for small values of *q*, we obtain

$$D_1(q)=\frac{u(x_q)\mu_X}{\mu_Y x_q}L_X(q)-L_Y(q). \qquad (2.2.19)$$

For large values of *q*, we consider the difference between the lower bound in (2.2.18) and the Lorenz curve $L_Y(q)$, that is, for small values of *q*, we obtain

$$D_2(q)=1-\frac{u(x_q)\mu_X}{\mu_Y x_q}\left(1-L_X(q)\right)-L_Y(q). \qquad (2.2.20)$$

In general, $\dfrac{dL_Y(q)}{dq}=\dfrac{y_q}{\mu_Y}$ and $\dfrac{dL_X(q)}{dq}=\dfrac{x_q}{\mu_X}$.

The ratio $\dfrac{u(x)}{x}$ is increasing and consequently,

$$\frac{d}{dq}\left(\frac{y_q}{x_q}\right)=\frac{d}{dx_q}\left(\frac{y_q}{x_q}\right)\frac{d}{dq}\left(x_q\right)\geq 0.$$

Now we differentiate $D_1(q)$ and note that $\dfrac{u(x_q)}{x_q}$ is increasing and obtain

$$\frac{d\left(D_1(q)\right)}{dq}=\frac{u(x_q)}{\mu_Y}\frac{\mu_X}{x_q}\frac{x_q}{\mu_X}+L_X(q)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right)-\frac{u(x_q)}{\mu_Y}=L_X(q)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right)\geq 0.$$

Consequently $D_1(q)$ is increasing from zero to a maximum for $q_0$.

Now we differentiate $D_2(q)$ and obtain

$$\frac{d\left(D_2(q)\right)}{dq} = +\frac{u(x_q)}{\mu_Y} - \left(1 - L_X(q)\right)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) - \frac{u(x_q)}{\mu_Y}$$
$$= -\left(1 - L_X(q)\right)\frac{\mu_X}{\mu_Y}\frac{d}{dq}\left(\frac{u(x_q)}{x_q}\right) \le 0 .$$

Consequently $D_2(q)$ is decreasing from a maximum to zero. The point denoted $q_0$, at which the shift from $D_1(q)$ to $D_2(q)$ is performed, satisfies $D_1(q) = D_2(q)$.

Now, $1 - \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}}\left(1 - L_X(q_0)\right) - L_Y(q_0) = \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}} L_X(q_0) - L_Y(q_0)$, that is,

$$1 - \frac{u(x_{q_0})\mu_X}{\mu_Y x_{q_0}} = 0 \text{, and } \frac{u(x_{q_0})}{x_{q_0}} = \frac{\mu_Y}{\mu_X} .$$

This condition is identical with the condition in which $\frac{u(x)}{x}$ is decreasing.

Again, the condition $1 - \frac{u(x_p)\mu_X}{\mu_Y x_p} = 0$ can be written $\frac{u(x_{q_0})}{\mu_Y} = \frac{x_{q_0}}{\mu_X}$ and we obtain the formula $\frac{dL_Y(q)}{dq}\Big|_{q=q_0} = \frac{dL_X(q)}{dq}\Big|_{q=q_0}$, that is, the Lorenz curves $L_Y(q)$ and $L_X(q)$ have parallel tangents and the distance between the Lorenz curves is maximal.

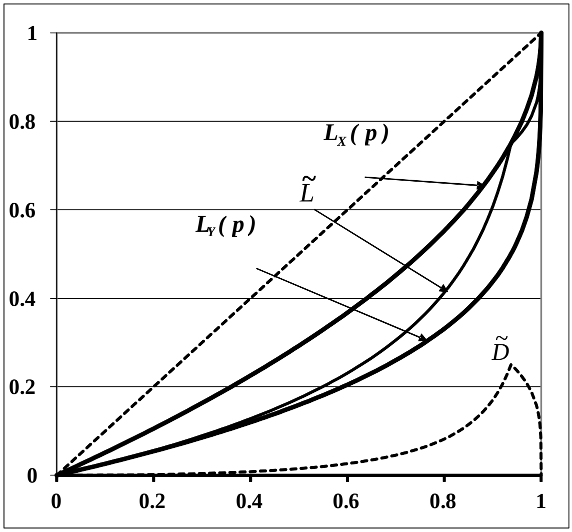



***Figure 2.2.2*** *(Fellman 2012b) A sketch of the Lorenz curves $L_Y(q)$, $L_X(q)$, the upper bound $\tilde{L}(q)$, and the difference $\tilde{D}(q)$ between the upper bound $\tilde{L}(q)$ and $L_Y(q)$ when the transformed variable is Lorenz dominated by the initial one.*

We define the difference function as

$$\tilde{D}(q) = \begin{cases} D_1(q) \text{ for } q \le q_0 \\ D_2(q) \text{ for } q > q_0 \end{cases}, \qquad (2.2.21)$$

and the upper bound of $L_Y(q)$ is

$$\tilde{L}(q) = \begin{cases} \dfrac{u(x_q)\mu_X}{\mu_Y x_q} L_X(q) & \text{for} \quad q \le q_0 \\ 1 - \dfrac{u(x_q)\mu_X}{\mu_Y x_q}\left(1 - L_X(q)\right) & \text{for} \quad q > q_0 \end{cases}. \qquad (2.2.22)$$

In Figure 2.2.2, we sketch the Lorenz curves $L_Y(q)$, $L_X(q)$, the upper bound $\tilde{L}(q)$ and the difference $\tilde{D}(q)$ between the upper bound $\tilde{L}(q)$ and $L_Y(q)$.

Now the lower bounds are of minor interest because the initial variable $X$ Lorenz dominates $Y$. Note that $m=0$ is possible in some situations and the lower bound in (2.2.17) can be zero. Note that $M$ can be great and even $M=\infty$ is possible in some situations and the lower bound in (2.2.18) can be even negative.

**Example 2.2.1 The Pareto distribution.** Consider income $X$ with the Pareto distribution $F_X(x)=1-x^{-\alpha}$ and $f_X(x)=\alpha x^{-\alpha-1}$, where $\alpha>1$ and $x\geq 1$.

Now, $\mu_X=\dfrac{\alpha}{\alpha-1}$ and the Lorenz curve $L_X(p)=1-(1-p)^{\frac{\alpha-1}{\alpha}}$.

From $F_X(x_p)=1-x_p^{-\alpha}=p$ we obtain $x_p=(1-p)^{-\frac{1}{\alpha}}$.

Let the transformation be $Y=u(x)=x^{\beta}$ ($0<\beta<1$) so that the function $\dfrac{u(x)}{x}=\dfrac{x^{\beta}}{x}=x^{\beta-1}=\dfrac{1}{x^{1-\beta}}$ is decreasing. We obtain $\mu_Y=\dfrac{\alpha}{\alpha-\beta}$, the Lorenz curve $L_Y(p)=1-(1-p)^{\frac{\alpha-\beta}{\alpha}}$ and

$$D_1(q)=1-\frac{1-\beta}{\alpha-1}(1-q)^{\frac{\alpha-\beta}{\alpha}}-\frac{(\alpha-\beta)}{(\alpha-1)}(1-p)^{\frac{1-\beta}{\alpha}}$$

and

$$D_2(q)=\frac{(1-\beta)}{(\alpha-1)}(1-q)^{\frac{\alpha-\beta}{\alpha}}$$

$$\tilde{D}(q)=\begin{cases} D_1(q)=1-\dfrac{1-\beta}{\alpha-1}(1-q)^{\frac{\alpha-\beta}{\alpha}}-\dfrac{(\alpha-\beta)}{(\alpha-1)}(1-p)^{\frac{1-\beta}{\alpha}} & \text{for} \quad q\leq q_0 \\ D_2(q)=\dfrac{(1-\beta)}{(\alpha-1)}(1-q)^{\frac{\alpha-\beta}{\alpha}} & \text{for} \quad q>q_0 \end{cases}$$

$$\tilde{L}(q)=\begin{cases}\dfrac{(\alpha-\beta)}{(\alpha-1)}\left((1-q)^{\frac{1-\beta}{\alpha}}-(1-q)^{\frac{\alpha-\beta}{\alpha}}\right) & \text{for} \quad q \le q_0 \\ 1-\dfrac{(\alpha-\beta)}{(\alpha-1)}\left((1-q)^{\frac{\alpha-\beta}{\alpha}}\right) & \text{for} \quad q > q_0\end{cases}$$

For $\beta < 1$, the ratio $\dfrac{u(x)}{x}$ is decreasing, this case being sketched in Figure 2.2.1, and if $\beta > 1$ the ratio $\dfrac{u(x)}{x}$ is increasing, this case being sketched in Figure 2.2.2.

## 2.3 Regional and Temporal Variation in the Income Inequality

We start with an example from Finland.

**Example 2.3.1. Finland 1971-1990.** We illustrate our methods using data from Finland from 1971 to 1990 (Fellman et al., 1996). The theoretical analyses of the Finnish data are presented more in detail in Chapter 5. The base $x$ for taxes includes all taxable income. From this we subtract direct taxes $t$ to get the base for all non-taxable benefits $b$. These are child allowances and housing subsidies. We have standardized the income variables to be comparable across households of different sizes using the OECD equivalence scale, which assigns the weight of 1.0, 0.7 and 0.5 equivalent adults to the first and additional adults and children, respectively. We show in Table 2.3.1 the estimated generalized Gini coefficients for different values of the parameter ν and the relevant income concepts. Under the actual column we see the inequality indices for original income, $x$, post-tax pre-transfer income $y = x - t$ and final income

$y+b=x-t+b$. Household disposable income per equivalent adult is equal to $x-t+b$.

We observe in Figure 2.3.1 that the Gini coefficients of original income for all v′s decrease monotonically over the period, 1971-1990, indicating decreasing income inequality.

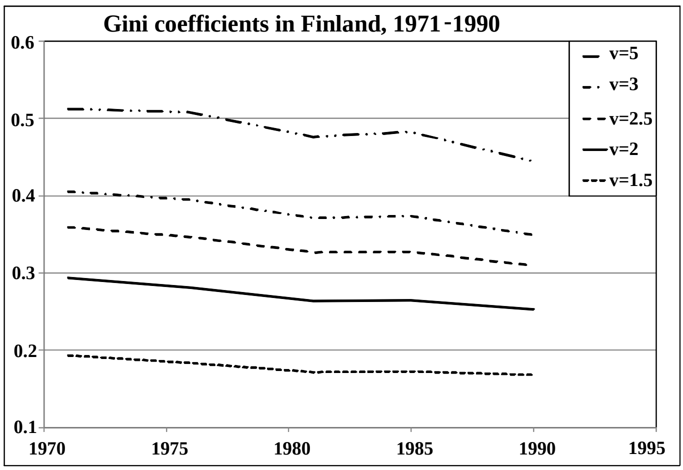



***Figure 2.3.1*** *Generalized Gini coefficients in Finland, 1971-1990, for different v′s.*

Eriksson and Jäntti (1997) showed that, in Finland, earnings inequality dropped dramatically between 1971 and 1975 and continued to decrease until 1985. From 1985 to 1990 there was a substantial increase in the inequality of earnings, comparable in magnitude to that found in the UK and US. Furthermore, they showed that the rise in inequality increased in Finland between 1985 and 1990 but this followed a sharp decline during the 1970s and early 1980s. The Figure 2.3.1 indicates that the conclusions given by Fellman et al. (1996) and Eriksson and Jäntti (1997) are similar for the period up to 1985 but after that they differ.

***Table 2.3.1*** *Inequality of income in Finland 1971-1990. Generalized Gini coefficients for pre-tax, post-tax and post-transfer income for actual incomes.*

| | | | | **Actual** |
|---|---|---|---|---|
| | **Year** | $x$ | $x-t$ | $x-t+b$ |
| | 1971 | 0.193 | 0.176 | 0.173 |
| | 1976 | 0.183 | 0.168 | 0.165 |
| 1.5 | 1981 | 0.171 | 0.159 | 0.154 |
| | 1985 | 0.172 | 0.155 | 0.151 |
| | 1990 | 0.168 | 0.147 | 0.145 |
| | 1971 | 0.294 | 0.271 | 0.267 |
| | 1976 | 0.281 | 0.259 | 0.255 |
| 2.0 | 1981 | 0.264 | 0.246 | 0.239 |
| | 1985 | 0.265 | 0.241 | 0.235 |
| | 1990 | 0.253 | 0.224 | 0.222 |
| | 1971 | 0.359 | 0.333 | 0.328 |
| | 1976 | 0.346 | 0.320 | 0.315 |
| 2.5 | 1981 | 0.326 | 0.303 | 0.295 |
| | 1985 | 0.327 | 0.299 | 0.291 |
| | 1990 | 0.309 | 0.275 | 0.273 |
| | 1971 | 0.405 | 0.378 | 0.372 |
| | 1976 | 0.394 | 0.365 | 0.359 |
| 3.0 | 1981 | 0.371 | 0.346 | 0.335 |
| | 1985 | 0.373 | 0.342 | 0.332 |
| | 1990 | 0.349 | 0.313 | 0.311 |
| | 1971 | 0.512 | 0.483 | 0.476 |
| | 1976 | 0.507 | 0.472 | 0.465 |
| 5.0 | 1981 | 0.476 | 0.447 | 0.433 |
| | 1985 | 0.482 | 0.448 | 0.434 |
| | 1990 | 0.444 | 0.405 | 0.402 |

Source: Fellman et al. (1996).
Note: $x$ is actual pre-tax income, $t$ denotes taxes and $b$ benefits.

Gottschalk & Smeeding (1997) compared the trends in inequality in different countries during the last decades in the 20th century. They noted marked differences. The first group consists of countries that experienced at least as large an increase in inequality as in the United States. This group includes only United

Kingdom. A second group which experienced substantial increases in inequality, but less than the United States and United Kingdom includes Canada, Australia and Israel. France, Japan, The Netherlands, Sweden and Finland form a third group with positive, but quite small changes in earnings inequality over the 1980s. Figure 2.3.1 agrees with the findings by Gottschalk and Smeeding. While even the Nordic countries experienced some increase in earnings during the 1980s, they started from very low levels, resulting from a long secular decline in inequality. Finally, Italy and Germany form a small group that experienced no measurable increase in earnings inequality during the 1980s.

Bach et al. (2009) analyzed income distributions in Germany (1932-2003) using several indicators of income inequality. They found a modest increase of the Gini coefficient, a substantial drop of median income and a remarkable growth of the income share accruing the economic elite that is the 0.001 percent of persons in the population. Their findings are supported by a relative difference between mean and median income that measures the skewness of the distribution: a rise in this measure of inequality indicates that incomes in the upper half of the distribution have increased more than the lower half.

In contrast to the findings for Finland (Fellman et al., 1996), income inequality in the United States has increased dramatically over the past 30 years. For instance, for households headed by working-age individuals, market incomes in the upper part of the distribution show an upwards trend in almost all periods since 1978, while they increased remarkably little in the middle and show large and sustained declines at the bottom during and after recessions. This is particularly true for the recent economic crisis.

Levy and Murnane (1992) presented a thorough study of the income distribution in US and discussed the variation in the inequality. For males they found that the inequality moved from stability or gradual increases in the 1970s to rapid increases in 1980s. For females they noted that annual earnings inequality moved from modest decline in the 1970s to increases in 1980s. They gave detailed interpretation of these general findings based on the variations in the composition of the labour force.

Yun (2006) studied the earnings inequality in US, 1969-1999 using different inequality measures; the ninetieth-tenth percentile log wage differential, the coefficient of variation, the Gini coefficient, the Theil index and the variance of log earnings. All measures identify an increase in the inequality. The increasing trends varied. The inequality was stable until 1980, steadily increased from 1980 to 1986, was stable again from 1987 to 1992 an increased thereafter. Autor et al. (2008) considered income inequality in US, 1963-2005. They found increasing trends and stressed that this trend was not an “episodic” one, but a continuing increase reflecting the mechanical confounding effects of changes in labour force composition. They provided an overview of the literature on U.S. wage inequality and discussed if the substantial increase since the 1980s can be considered as an episodic event or a continuous development.

Heathcote et al. (2010) conducted a systematic empirical study of cross-sectional inequality in the United States. They found a large and steady increase in wage inequality between 1967 and 2006. Taxes and transfers compress the level of income inequality, especially at the bottom of the distribution, but have little effect on the overall trend. Meyer and Sullivan (2011) found that post-tax income inequality started to increase later (in the late 1970s) than that of pre-tax income and that its increase in the 1980s occurred at a slower rate.

Analysing earlier results for US, Gottschalk and Danziger (2005) found that the development of male wage and family income inequality were largely comparable over the period 1975 to 2002. Bargain et al. (2011) noted increasing income inequality during the late 1970s and early 1980s. Furthermore, they stated that the usual approach for evaluating the role of taxation as a driver of overall inequality trends is to compare income inequality measured before and after taxes (see e.g. Gottschalk & Smeeding 1997). However, tax burdens and their impact on the income distribution are determined by both tax schedule and tax base. For instance, a given progressive income tax schedule redistributes more when the distribution of taxable incomes becomes more dispersed, and very little if everybody earns about the same (Musgrave & Thin 1948; Dardanoni & Lambert 2002). Bargain et al. (2011) concluded that main findings are as follows. The increase in post-tax income inequality was slower than that of pre-tax inequality indicating that the redistributive role of the tax system has increased over time. However, their decomposition reveals that most of this increase in redistribution was not due to the policy effect but a mechanical consequence of the rising inequality in pre-tax income.

## 2.4 Estimation of Gini Coefficients

Fellman (2012a) analysed the estimation of Gini coefficients using Lorenz curves. Primary income data yields the most accurate estimates of the Gini coefficient. However, the estimation must often be based on tables with grouped data or on Lorenz curves. The Lorenz curves are usually defined for five quintiles or for 10 deciles**.** As explained above in Section 1.1 the Gini coefficient is defined as the ratio of the area between the diagonal and the Lorenz curve and the area of the whole triangle under the diagonal. For five quintiles, the trapezium rule is the most commonly used method. However, this rule yields for every trapezium positive bias for the estimate of the area under the Lorenz curve and, consequently, the rule causes negative bias for the Gini coefficient. Simpson ́s rule is better fitted to the Lorenz curve, but demands an even number of subintervals of the same length. That is, Gini coefficients can be based on Lorenz curves given in deciles.

Various attempts have been made to produce more exact estimates. Gastwirth (1972) introduced interval estimates of the Gini coefficient in order to measure the accuracy of the estimates. Needleman ́s study (1978) starts from the trapezium estimate of the Gini coefficient $G_L$. He then introduces an improved upper estimate $G_U$. His final estimate follows the "two-thirds rule" that is $G = \frac{G_L}{3} + \frac{2G_U}{3}$ . McDonald and Ransom (1981) considered the $\Gamma$ density, applied Monte Carlo methods and introduced lower and upper bounds of the Gini estimates.

Golden (2008) showed how a quick approximation of the Gini coefficient can be calculated empirically, using numerical data in cumulative income quintiles. Fellman (2012a) compared different methods. He applied Simpson ́s rule and considered Lorenz curves with deciles. In addition, he used Lagrange polynomials and generalizations of Golden ́s method.

There are several different situations and, consequently, alternative analyses of Gini coefficients have to be performed. When Lorenz curves are considered, the simplest situation is that they are defined for five quintiles or for 10 deciles. In the first case, the most commonly used method is the trapezium rule. For Simpson ́s rule, the number of subintervals should be even and the intervals should have the same length. This means, for example, that Lorenz curves with 10 deciles are suitable. One has three $L$ values for each doubled subinterval. The area under this part of the Lorenz curve is estimated so that the Lorenz curve is approximated by a parabola obtaining the same $L$ values. Consequently, the comparison of different rules can be performed for Lorenz curves with deciles.

Following Fellman (2012a) we assume a Lorenz curve $L(p)$ with deciles. Let the observed values of the cumulative Lorenz curve be $p_i$ and $L_i$ for

$i = 0, 1, \ldots, 10$. Note that $p_i = i/10$, ($i = 0, 1, \ldots, 10$), that $L_0 = 0$ and that $L_{10} = 1$. According to the trapezium rule, the estimated area under the Lorenz curve is

$$\tilde{I} = ½\sum_{i=0}^{9}(L_{i+1} + L_i)(p_{i+1} - p_i) \tag{2.4.1}$$

and the estimated Gini coefficient, $G_T$ is $1 - 2\tilde{I}$. Every trapezium yields a positive bias to the estimated area, as can be seen in Figure 2.4.1. Since the biases obtained add and no elimination of biases can be performed, the estimated Gini coefficient always has a negative bias.

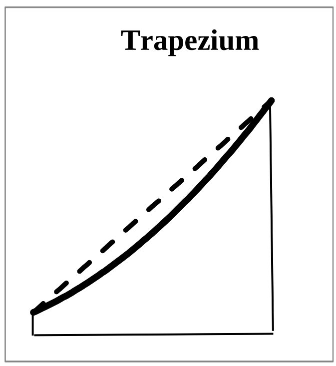



***Figure 2.4.1*** *A sketch showing the bias in the trapezium rule.*

Compared to the trapezium rule, Simpson΄s rule gives more accurate approximations. As stressed above, Simpson΄s rule demands two restrictions: the number of subintervals has to be even and the subintervals have to be of equal length. In order to obtain Simpson΄s rule, the subintervals should be grouped two by two. Each doubled subinterval has three $L$ values. The area under this part of the Lorenz curve is estimated such that a parabola obtaining the same $L$ values approximates the Lorenz curve. Simpson΄s rule obviously yields exact results for quadratic curves but, in general, this also holds for cubic curves. Assuming *2n* subintervals, the approximate area formula for a doubled interval is

$$\tilde{I}_i = \frac{1}{3n}\left(L_i + 4L_{i+1} + L_{i+2}\right),$$

the total sum is

$$\tilde{I} = \frac{1}{3n}\sum_{i=0}^{4}\left(L_{2i} + 4L_{2i+1} + L_{2i+2}\right) \tag{2.4.2}$$

and $G_S = 1 - 2\tilde{I}$ .

Golden (2008) gave a detailed account of an alternative method based on Lorenz curves with quintiles. He considered $p$ and $L$ in percentages. The layout of the method is presented in Table 2.4.1. First he determined where the cumulative income shortfall is greatest and defined Z as the largest quintile point of the cumulative income shortfall from perfect equality divided by 100. In order to obtain the largest cumulative income shortfall he defined the transformed variable $\tilde{L}_i = L_{i-1} + 20$. This transformation, $\tilde{L}_i = L_{i-1} + 20$, indicates a search for an interval at which $L_i$ shifts from increases faster than $p_i$ to slower increases. For low $i$ ́s, the transformed value $\tilde{L}_i > L_i$. Later, there is a first $i$ value such that $\tilde{L}_i < L_i$. For this value, one finds an interval for which $L$ is closely parallel with the diagonal, the greatest shortfall is obtained, and one defines $q = (20i - \tilde{L}_i)/100$. The estimated Gini coefficient in percentages, $G_G$, is $G_G = 50q(3-q)$. When this method was applied to 621 income observations, Golden (2008) noted that his approach performed better than the trapezium rule, also stressing that his method could be applied to Lorenz curves with deciles.

Fellman (2012a) generalized Golden ́s method in the following way. If the Lorenz curves are given in deciles, then Golden ́s transformation should be $\tilde{L}_i = L_{i-1} + 10$ and if the $p_i$ ́s are not equidistant, then one has to define

$\tilde{L}_i = L_{i-1} + p_i - p_{i-1}$. Following Golden´s rule, these processes have to continue until $\tilde{L}_i < L_i$. Then introduce $q = (p_i - \tilde{L}_i)/100$ and $G_G = 50q(3-q)$.

***Table 2.4.1*** *A layout of a Lorenz curve with deciles. Following Golden (2008), the data is given in percentages. The transformed* $\tilde{L}_{20i} = L_{20i-20} + 20$ *values appear in the text.*

| ***i*** | **0** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|---|
| $p_i$ | 0 | 20 | 40 | 60 | 80 | 100 |
| $L_i$ | $L_0 = 0$ | $L\,L_{20}$ | $L_{40}$ | $L_{60}$ | $L_{80}$ | $L_{100} = 100$ |
| $\tilde{L}_i$ | $\tilde{L}_0 = 0$ | $\tilde{L}_{20}$ | $\tilde{L}_{40}$ | $\tilde{L}_{60}$ | $\tilde{L}_{80}$ | $\tilde{L}_{100}$ |

In many empirical situations, the income distribution $F(x)$ is given in grouped tables. If the mean of or total incomes in the groups are known, the cumulative distribution can be considered as a Lorenz curve, but the subintervals are usually not of constant length. The trapezium rule holds, but it still yields a positive bias for the area and negative bias for the Gini coefficient.

An obviously better alternative is to approximate the Lorenz curve with Lagrange´s interpolation (Berrut & Trefethen, 2004). Lagrange polynomials of the second degree can be considered as a generalisation of Simpson´s rule and do not demand subintervals of equal length, but the number of subintervals should still be even. The polynomials obtained have to be integrated in order to yield approximate areas and Gini coefficients. If the subintervals are of the same length, the Lagrange polynomial method is identical with Simpson´s rule.

Fellman (2012a) applied the Lagrange interpolation of second degree. However, he had to assume an even number of subintervals. Now the Lagrange polynomial is

$$\begin{aligned} L(p) = \sum_{i=0}^{n-1} \Bigg( & L_{2i} \frac{(p-p_{2i+1})(p-p_{2i+2})}{(p_{2i}-p_{2i+1})(p_{2i}-p_{2i+2})} \\ & + L_{2i+1} \frac{(p-p_{2i+2})(p-p_{2i})}{(p_{2i+1}-p_{2i+2})(p_{2i+1}-p_{2i})} + L_{2i+2} \frac{(p-p_{2i+1})(p-p_{2i})}{(p_{2i+2}-p_{2i+1})(p_{2i+2}-p_{2i})} \Bigg) \end{aligned} \tag{2.4.3}$$

This approximate polynomial must be integrated in order to obtain an estimate of the area under the Lorenz curve.

The comparison between different estimation methods is in general difficult to perform. These difficulties are mainly caused by the fact that the true Gini coefficient is unknown, but sometimes, where more detailed studies have already resulted in very accurate estimates, the comparisons are possible. Some authors (e.g., Gastwirth, 1972; Mehran, 1975; McDonald & Ransom, 1981; Rigo, 1985; Giorgi & Pallini, 1987) have introduced interval estimates, but these are often rather broad and it is still difficult to identify the best method. Such comparison problems are eliminated if the numerical estimations are applied to theoretical distributions.

Needleman (1978) stated that as the Lorenz curve is convex, the trapezium approximation is always greater than the actual area under the curve, so that the estimate based on this approximation is always less than the actual value of the coefficient. Furthermore, he noted that most authors using the trapezium approximation indicate that they are aware of the bias involved, but either assume the error so small as to be insignificant, or else use a large number of intervals in the belief, usually justified, that the bias will then be negligible. McDonald and Ransom (1981) introduced lower and upper bounds of the Gini estimates. In order to estimate the bounds of the Gini coefficient estimates, they considered the income to have a $\Gamma$ density, that is, $g(y)=\dfrac{\beta^{\alpha} y^{\alpha-1} e^{-y\beta}}{\Gamma(\alpha)}$ with corresponding $G = \frac{\Gamma(\alpha + ?}{\Gamma(\alpha+1)\sqrt{\pi}}$ and $\mu = \alpha / \beta$ and applied Monte Carlo methods.

In order to perform comparisons between the estimated and theoretical Gini coefficients Fellman (2012a) analysed classes of theoretical Lorenz curves with varying Gini coefficients. He compared Gini estimates for the Pareto distributions. If one defines the Pareto distribution as $F(x) = 1 - x^{-\alpha}$, where $x \geq 1$ and $\alpha > 1$. Then the frequency function is $f(x) = \alpha x^{\alpha-1}$, the mean is $\mu = \frac{\alpha}{\alpha - 1}$, the quantiles are $x_p = \left(\frac{1}{1-p}\right)^{\frac{\alpha-1}{\alpha}}$, the Lorenz curve $L(p) = 1 - (1-p)^{\frac{\alpha-1}{\alpha}}$ and the Gini coefficient $G = \frac{1}{2\alpha - 1}$. Fellman considered $1.5 \leq \alpha \leq 5.0$, then the Gini coefficient satisfies the inequalities $0.111 \leq G \leq 0.500$. This $G$ interval corresponds to the most common Gini coefficients. Fellman’s results appear in Table 2.4.2 and Figure 2.4.2. Note that Simpson ́s and Golden ́s rules yield similar accuracy, but the trapezium rule shows the largest errors for all levels of Gini coefficients. This theoretical study indicates that Golden ́s rule is not uniformly better than the trapezium rule.

Gastwirth (1972) presents interval estimations of the Gini coefficient. The exact Gini estimate on Current Population Surveys (CPS) income data for 1968 was computed by Tepping, his result being 0.4014. Gastwirth ́s Table 2 shows Tepping ́s data grouped into a 10 subgroup Lorenz curve. He compares his Gini interval estimates with Tepping ́s finding. Gastwirth (1972) considers a minimum of restrictive conditions, obtaining the interval $0.3883 < G < 0.4083$. Mehran (1975) suggests an alternative estimation method, obtaining the interval estimate $0.3883 < G < 0.4087$. The grouping limits are not equidistant and one cannot apply Simpson ́s rule. Applying the trapezium rule yields *0.3883* and the negative bias is apparent. The Lagrange rule yields *0.4033* and the modification of the Golden rule yields the rather inaccurate estimate *0.3740*.

***Table 2.4.2*** *(Fellman, 2012a). The estimation of the Gini coefficient applied to the Lorenz curve for the Pareto distributions. Note that the estimated Gini coefficients according to the trapezium rule are inaccurate and show negative biases. Simpson ́s and Golden ́s rules yield similar accuracy, but Golden is best for large Gini values.*

| | | Estimates | | | Error | |
|---|---|---|---|---|---|---|
| **G** | **Trapezium** | **Simpson** | **Golden** | **Trapezium** | **Simpson** | **Golden** |
| 11.11 | 10.858 | 11.044 | 11.104 | -0.253 | -0.067 | -0.008 |
| 12.50 | 12.206 | 12.419 | 12.529 | -0.294 | -0.081 | 0.029 |
| 14.29 | 13.935 | 14.185 | 14.370 | -0.350 | -0.101 | 0.084 |
| 16.67 | 16.235 | 16.535 | 16.833 | -0.431 | -0.132 | 0.166 |
| 20.00 | 19.442 | 19.816 | 20.291 | -0.558 | -0.184 | 0.291 |
| 25.00 | 24.223 | 24.717 | 25.476 | -0.777 | -0.283 | 0.476 |
| 33.33 | 32.102 | 32.820 | 34.026 | -1.232 | -0.513 | 0.693 |
| 50.00 | 47.481 | 48.730 | 50.317 | -2.519 | -1.270 | 0.317 |

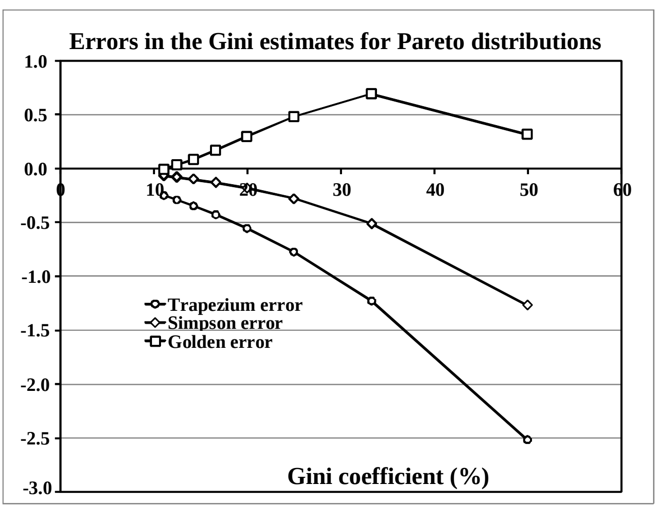



***Figure 2.4.2*** *Estimation errors in the Gini coefficients estimated by the trapezium, Simpson, and Golden rules. Note that Simpson ́s and Golden ́s rules yield similar accuracy, but the trapezium rule shows the largest errors (Fellman, 2012a).*

Lorenzen (1980) presents information about the total distribution of income for households in Germany in 1973 in his *Tabelle 2*. The Gini coefficient calculated by Lorenzen is based on data pooled in his *Tabelle 3*, which yielded *0.30*. Using Lorenzen´s *Tabelle 3*, Fellman performed a comparison of the estimates obtained based on the trapezium rule and the Lagrange rule. The available empirical data cannot yield a comparison of the accuracy of the two methods. The estimated Gini coefficient according to the trapezium rule shows negative biases compared to Lorenzen´s result, being *0.2920*. The Lagrange interpolation yields the estimate *0.3486* and the modified Golden method 0.3002.

This study indicates that the biased trapezium rule is almost always inferior and shows negative biases. No method however is uniformly optimal. Note that Simpson´s and Golden´s rules yield similar accuracy. Golden´s method is usually of medium quality, but its accuracy fluctuates.

## References

[1] Autor, D., Katz, L., Kearney, M. (2008). Trends in U.S. wage inequality: Revising the revisionists, *Review of Economics and Statistics 90 (2)*:300-323.

[2] Bach, S., Corneo, G., Steiner, V. (2009). From bottom to top: the entire income distribution in Germany, 1992-2003. *Review of Income and Wealth 55*:303-330.

[3] Bargain, O., Dolls, M., Immerwoll, H., Peichl, A., Pestel, N., Sieglioch, S. (2011). *Tax policy and income inequality in the U.S., 1978—2009: A decomposition approach*. ECINEQ WP 2011:215.

[4] Berrut, J. -P. & Trefethen, L. N. (2004). Barycentric Lagrange interpolation. *SIAM Review 46 (3)*:501-517.

[5] Dardanoni, V., Lambert, P. J. (2002). Progressivity comparisons, *J Publ Econ*, *86:* 99-122.

[6] Eriksson, T., Jäntti, M. (1997). The distribution of earnings in Finland 1971–1990. *Europ. Econ. Rev. 41*:1763-1779.

[7] Fellman, J. (1976). The effect of transformations on Lorenz curves. *Econometrica 44*:823-824.

[8] Fellman, J. (2009). Discontinuous transformations, Lorenz curves and transfer policies. *Social Choice and Welfare 33(2)*:335-342. DOI 10.1007/s00355-008-0362-4.

[9] Fellman, J. (2011). Discontinuous transfer policies with given Lorenz curve. *Advances and Applications in Statistics Volume 20 No. 2 February Issue of the ADAS*, 2011:133-141.

[10] Fellman. J. (2012a). Estimation of Gini coefficients using Lorenz curves. *Journal of Statistical and Econometric Methods vol.1, no. 2*:31-38

[11] Fellman, J. (2012b). Properties of Lorenz curves for transformed income distributions. *Theoretical Economics Letters*, *2*, 487-493 doi:10.4236/tel.2012.25091 Published Online December 2012, (http://www.SciRP.org/journal/tel)

[12] Fellman, J., Jäntti, M., Lambert, P. (1996). *Optimal Tax-transfer Systems and Redistributive Policy: The Finnish Experience*. Swedish School of Economics and Business Administration Working Papers 324.

[13] Gastwirth, J. L. (1971). A General Definition of the Lorenz Curve. *Econometrica, 39(6)*:1037-1039.

[14] Gastwirth, J. L. (1972). The estimation of the Lorenz curve and Gini coefficient. *Rev. Economics and Statistics 54*:306-316.

[15] Giorgi, G. M. & Pallini, A. (1987). About a general method for the lower and upper distribution-free bounds on Gini´s concentration ratio from grouped data. *Statistica 47*:171-184.

[16] Golden, J. (2008). A simple geometric approach to approximating the Gini coefficient. *J. Economic Education 39(1)*:68-77.

[17] Gottschalk, P., Danziger, S. (2005). Inequality of wage rates, earnings and family income in the United States, 1975-2002. *Review of Income and Wealth 51* (2): 231-254.

[18] Gottschalk, P. and Smeeding, T. (1997). Cross-national comparisons of earnings and income inequality. *Journal of Economic Literature 35*: 633-87.

[19] Heathcote, J., Perri, F., Violante, G. L. (2010). *Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967—2006*. Federal Reserve Bank of Minneapolis, Research Department Staff Report 436, October 2009. 60 pp.

[20] Hemming, R. and Keen, M. J. (1983). Single crossing conditions in comparisons of tax progressivity. *Journal of Public Economics 20*, 373-380.

[21] Jakobsson, U. (1976). On the measurement of the degree of progression. *Journal of Public Economics 5*:161-169.

[22] Kakwani, N. C. (1977). Applications of Lorenz curves in economic analysis. *Econometrica 45*:719-727.

[23] Lambert, P. J. (2001). *The Distribution and Redistribution of Income: A mathematical analysis*. (3rd edition) Manchester: Manchester University Press. xiv+313 pp.

[24] Levy, F., Murnane, R. J. (1992). U.S. earnings levels and earnings inequality: A review of revcent trends and proposed explanations. *Journal of Economic Literature, XXX*: 1333-1381.

[25] Lorenz, M. O. (1905). Methods for measuring concentration of wealth. *J. Amer. Statist. Assoc. New Series, No. 70*: 209-219.

[26] Lorenzen, G. (1980). Was ist ein “echtes” Konzentrationsmaß? *Allgemeines Statistisches Archiv 4*:390-400.

[27] Lydall, H. F. (1968). *The Structure of Earnings*. (London: Oxford University press).

[28] McDonald, J. B & Ransom, M. R. (1981). An analysis of the bounds for the Gini coefficient. *Journal of Econometrics 17*:177–188

[29] Mehran, F. (1975). Bounds on the Gini index based on observed points of the Lorenz curve. *J Amer Statist. Assoc. JASA 70*:64-66.

[30] Meyer, B. and Sullivan, J. (2011). Consumption and income poverty over the business cycle. *Research in Labor Economics 32*: 51-82.

[31] Musgrave, R. and Thin, T. (1948). Income tax progression, 1929-1948, *Journal of Political Economy 56*: 498—514.

[32] Needleman, L. (1978). On the approximation of the Gini coefficient of concentration. *The Manchester School 46*:105-122.

[33] Rigo, P. (1985). Lower and upper distribution free bounds for Gini ́s concentration ratio. *Proceedings International Statistical Institute, 45$^{th}$ Session, Amsterdam, Contributed Papers, Book 2*:629-630.

[34] Yun, M-S. (2006). Earnings inequality in USA, 1969-99: Comparing inequality using earnings equations. *Review of Income and Wealth 52(1)*:127-144.

# 3

# Taxation

In Chapter one and Chapter two, we have introduced the central properties of income distributions and the methods how to analyse income distributions and redistributions. We have also given example how to estimate distributions and concentration measures in empirical data. In this chapter we apply the methods on the effects of taxation policies.

## 3.1 A Class of Tax Policies

Following Fellman (2001) we consider a pre-tax income $X$, assumed given, with the distribution function $F_X(x)$, density function $f_X(x)$, mean $\mu_X$, Lorenz curve $L_X(p)$ and the Gini coefficient $G_X$. Now, we consider a class of tax policies characterized by the transformation $Y = u(X)$ where $u(\cdot)$ is non-negative, monotone increasing and continuous with the properties

$$\mathbf{U}: \begin{cases} u(x) \leq x \\ u'(x) \leq 1 \\ E(u(X)) = \mu_X - \tau \end{cases} . \tag{3.1.1}$$

The function $u(x)$ is the post-tax income associated with the pre-tax income $x$ and $\tau$ is the mean tax. The monotonicity of $u(x)$ indicates that the internal order of the incomes remains the same after taxation. The taxation reduces the income and consequently, the first condition in (3.1.1) is obvious. The second guarantees that also the taxes increase monotonically with increasing initial income $x$, and the third indicates that the different tax policies yield the same total amount of taxes when applied to the given pre-tax incomes. In order to give a more realistic definition of the class of tax policies, Fellman (2001) introduced the restriction $u'(x) \leq 1$. Earlier in Fellman (1995) and in Fellman et al. (1996, 1999), this restriction was not assumed. Therefore, some of the results in those studies differ slightly from the results in later papers and in this study.

The class $\boldsymbol{U}$ of tax policies contains both progressive and non-progressive policies and is therefore an adaptive tool for inequality and welfare studies. The policies in $\boldsymbol{U}$ do not have a Lorenz ordering. Accordingly, the Lorenz curves corresponding to post-tax distributions generated by members of $\boldsymbol{U}$ may intersect.

Assume set of arbitrary policies $u_i(x)$, ( $i=1,\dots,k$ ), belonging to $\boldsymbol{U}$. Consider their linear combination

$$u_\theta(x)=\sum_{i=1}^{k}\theta_i u_i(x) \quad \theta_i \geq 0\,(\,i=1,\dots,k\,,\,)\ \sum_{i=1}^{K}\theta_i=1. \tag{3.1.2}$$

We obtain

$$u_\theta(x)=\sum_{i=1}^{k}\theta_i u_i(x)\leq\sum_{i=1}^{k}\theta_i x=x\sum_{i=1}^{k}\theta_i=x\,, \tag{3.1.3}$$

$$u'_\theta(x)=\sum_{i=1}^{k}\theta_i u'_i(x)\leq\sum_{i=1}^{k}\theta_i=1 \tag{3.1.4}$$

and

$$E\left(\sum_{i=1}^{k}\theta_i u_i(X)\right)=\sum_{i=1}^{k}\theta_i E\big(u_i(X)\big)=\sum_{i=1}^{k}\theta_i\big(\mu_X-\tau\big)=\mu_X-\tau\,. \tag{3.1.5}$$

Hence, $u_\theta(x)$ belongs to $\boldsymbol{U}$ and $\boldsymbol{U}$ is a convex class of policies.

Denote by $L_i(p)$ the Lorenz curves, $G_i$ the Gini coefficients corresponding to the policies $u_i(x)$ ( $i=1,\dots,k$ ). From the fact that integration is a linear operator we obtain the Lorenz curve $L_\theta(p)$ and the Gini coefficient $G_\theta$

$$L_\theta(p)=\frac{1}{\mu-\tau}\int_0^{x_p}\left(\sum_{i=1}^{k}\theta_i u_i(x)\right)f_X(x)dx=$$

$$\sum_{i=1}^{k}\theta_i \frac{1}{\mu-\tau}\int_0^{x_p} u_i(x) f_X(x)dx = \sum_{i=1}^{k}\theta_i L_i(p) \tag{3.1.6}$$

and

$$G_\theta = 1 - 2\int_0^1 L_\theta(p)dp = \sum_{i=1}^{k}\theta_i \int_0^1 L_i(p)dp = \sum_{i=1}^{k}\theta_i G_i \,. \tag{3.1.7}$$

Conversely, if we consider a Lorenz curve satisfying (3.1.6) and (3.1.7) it corresponds to a policy of the form (3.1.2) and belongs to ***U***. Hence, the classes of Lorenz curves and of Gini coefficients are also convex and we can summarize all results in:

**Theorem 3.1.1.** The class **U** and the classes of Lorenz curves and of Gini coefficients corresponding to the policies in **U** are convex.

Now we study the class (3.1.1) of policies in more detail. First we analyse a policy which serves as a benchmark for the members of policies. Consider

$$u_0(x) = \begin{cases} x & x \le a_0 \\ \\ a_0 & x > a_0 \end{cases}, \tag{3.1.8}$$

that means that for incomes $x \le a_0$ there is no tax and for $x > a_0$ the tax is $x - a_0$ so that the post-tax is constantly equal to $a_0$.

We prove that there is a unique value $a_0$ such that $E(u_0(X)) = \mu_X - \tau$ and consequently, the corresponding policy belongs to ***U***. For an arbitrary *a* we obtain

$$E(u_0(X)) = \int_0^a x f_X(x)dx + \int_a^\infty a f_X(x)dx =$$

$$\mu_X L_X\left(F_X(a)\right)+a\left(1-F_X(a)\right) \tag{3.1.9}$$

The function

$$e(a)=\mu_X L_X\left(F_X(a)\right)+a\left(1-F_X(a)\right) \tag{3.1.10}$$

starts from the value $e(0)=0$ and has the derivative

$$e'(a)=\mu_X \frac{a}{\mu_X} f_X(a)+1-F_X(a)-af_X(a)=1-F_X(a)\geq 0\,. \tag{3.1.11}$$

From the fact that the mean $\mu_X$ exists, it follows that

$$\lim_{a\to\infty} e(a)=\lim_{a\to\infty}\mu_X L_X\left(F_X(a)\right)+\lim_{a\to\infty} a\left(1-F_X(a)\right)=\mu_X$$

because

$$\lim_{a\to\infty}\mu_X L_X\left(F_X(a)\right)=\mu_X$$

and

$$0\leq \lim_{a\to\infty} a\left(1-F_X(a)\right)=\lim_{a\to\infty} a\int_a^\infty f_X(x)dx=\lim_{a\to\infty}\int_a^\infty a f_X(x)dx\leq \lim_{a\to\infty}\int_a^\infty x f_X(x)dx=0\,.$$

Hence, the function $e(a)$ is continuous and monotone increasing from $e(0)=0$ to $e(\infty)=\mu_X$ and consequently, there exists a unique $a_0$ such that $E\left(u_0(X)\right)=e(a_0)=\mu_X-\tau$. This value $a_0$ satisfies the inequality $a_0\geq\mu_X-\tau$ (with equality if and only if $F_X\left(a_0\right)=0$). For this value of $a_0$ the tax policy $u_0(x)$ belongs to **U**.

Define $p_0=F_X(a_0)$. For $p\leq p_0$,

$$L_0(p) = \frac{1}{\mu_X - \tau} \int_0^{x_p} x f_X(x) dx = \frac{\mu_X}{\mu_X - \tau} L_X(p)$$

and for $p > p_0$

$$L_0(p) = \frac{\mu_X}{\mu_X - \tau} L_X(p_0) + \frac{1}{\mu_X - \tau} \int_{x_{p0}}^{x_p} a f_X(x) dx = \frac{\mu_X}{\mu_X - \tau} L_X(p_0) + \frac{a_0}{\mu_X - \tau}(p - p_0).$$

Hence, the corresponding Lorenz curve is

$$\mathrm{L}_0(p) = \begin{cases} \dfrac{\mu_X}{\mu_X - \tau} L_X(p) & \mathrm{p} \le p_0 \\ \\ \dfrac{\mu_X}{\mu_X - \tau} L_X(p_0) + \dfrac{a_0}{\mu_X - \tau}(p - p_0) & \mathrm{p} > p_0 \end{cases} . \qquad (3.1.12)$$

By definition given above, $F_X(a_0) = p_0$ and $x_{p_0} = a_0$. In the point $p = p_0$ the derivative to the left is

$$\frac{\mu_X}{\mu_X - \tau} L'_X(p_0) = \frac{x_{p_0}}{\mu_X - \tau} = \frac{a_0}{\mu_X - \tau}$$

and to the right is

$$\frac{a_0}{\mu_X - \tau}.$$

Therefore the derivative exists also in the point $p = p_0$ and the Lorenz curve (3.1.12) has a continuous derivative within the interval (0, 1).

Consider an arbitrary transformation $u(x)$ with the properties (3.1.1). Then according to Figure 3.1.1, $u(x) \le u_0(x) = x$ for $x \le a_0$.

From the fact that the function $u(x)$ is an increasing function it follows that there exists a unique $x* \geq a_0$ such that $u(x) < a_0$ for $x < x*$ and $u(x) \geq a_0$ for $x \geq x*$. Hence,

$$u(x) < u_0(x) \text{ for } x < x* \text{ and } u(x) \geq u_0(x) \text{ for } x \geq x*.$$

The difference

$$D(p) = \frac{1}{\mu_X - \tau} \int_0^{x_p} \left(u_0(x) - u(x)\right) f_X(x) dx, \tag{3.1.13}$$

where $F_X(x_p) = p$, increases monotonically from zero to a maximum for $p* = F_X(x*)$, where after it decreases monotonically to zero. Hence, $u_0(x)$ generates a post-tax income distribution that Lorenz dominates all tax policies of the given class $\boldsymbol{U}$ (Fellman, 1995, 2001; Fellman et al., 1996, 1999). Furthermore, it also Lorenz dominate the flat tax policy $\hat{u}(x) = \dfrac{\mu_X - \tau}{\mu_X} x$, whose mean is $\mu_X - \tau$ and Lorenz curve $L_X(p)$. Consequently, $L_0(p) \geq L_X(p)$ and $u_0(x)$ Lorenz dominates the initial income variable *X*.

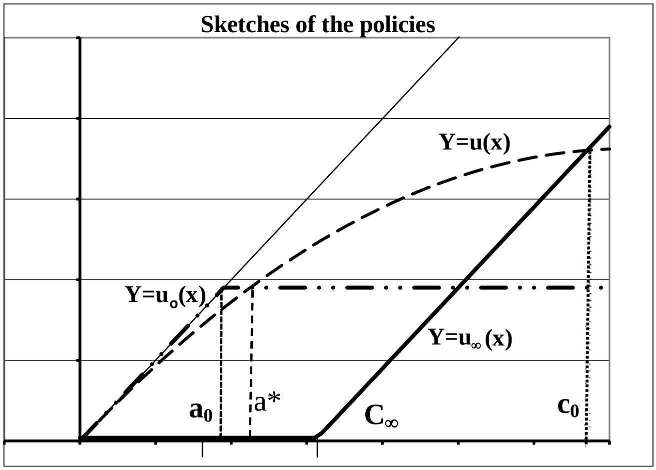



***Figure 3.1.1*** *Sketches of the two extreme tax policies:* $Y = u_o(X)$ *and* $Y = u_\infty(X)$*, and an arbitrary policy* $Y = u(x)$ *(after Fellman, 2001, 2002, 2014).*

Let $G_0$ be the Gini coefficient corresponding to $u_0(x)$. We obtain

$$\begin{aligned} G_0 &= 1 - 2\int_0^1 L_0(p)dp \geq 1 - 2\int_0^1 L_X(p)dp \\ &= 1 - 2\frac{\mu_X}{\mu_X - \tau}(1 - G_X) = G_X - \frac{\tau}{\mu_X - \tau}(1 - G_X) \end{aligned}. \tag{3.1.14}$$

The policy (3.1.8) Lorenz dominates the class ***U*** and therefore we obtain that the lower bound $G_X - \dfrac{\tau}{\mu_X - \tau}(1 - G_X)$ in (3.1.14) is a lower bound of the Gini coefficients of all policies in ***U***.

Consider another extreme policy

$$u_\infty(x) = \begin{cases} 0 & x < c_\infty \\ x - c_\infty & x \geq c_\infty \end{cases}. \tag{3.1.15}$$

It takes everything from the poorest whose income is below $c_\infty$ and a constant amount $c_\infty$ from the riches whose income is greater than $c_\infty$. A sketch of $u_\infty(x)$ is also presented in Figure 3.1.1. Below we prove that there exists a value $c_\infty$ such that $u_\infty(x)$ satisfies the condition (3.1.1) and belongs to ***U***. For an arbitrary $c$ we obtain

$$\begin{aligned}
E(u_\infty(X)) &= \int_0^\infty u_\infty(x) f_X(x)dx = \int_c^\infty (x-c) f_X(x)dx \\
&= \int_c^\infty x f_X(x)dx - \int_c^\infty c f_X(x)dx \\
&= \int_0^c x f_X(x)dx + \int_c^\infty x f_X(x)dx - \int_0^c x f_X(x)dx - \int_c^\infty c f_X(x)dx \\
&= \mu_X(1 - L_X(F_X(c)) - c(1 - F_X(c))
\end{aligned} \qquad (3.1.16)$$

Consider the function $e(c) = \mu_X(1 - L_X(F_X(c)) - c(1 - F_X(c))$. From the fact that $\mu_X$ exists then $e(0) = \mu_X$ and

$$\lim_{c\to\infty}(e(c)) = \lim_{c\to\infty}(\mu_X(1 - L_X(F_X(c))) - \lim_{c\to\infty}(c(1 - F_X(c))) =$$
$$-\lim_{c\to\infty} c(1 - F_X(c)) = \lim_{c\to\infty} c\int_c^\infty f_X(x)dx \le \lim_{c\to\infty}\int_c^\infty x f_X(x)dx = 0.$$

Consider the derivative $e'(c)$. Now

$$e'(c) = -\mu_X \frac{c}{\mu_X} f_X(c) - (1 - F_X(c)) + c f_X(c) = -(1 - F_X(c)) < 0$$

and $e(c)$ is monotone decreasing from $\mu_X$ to zero. Hence, there exists a unique value $c_\infty$ such that

$$e(c_\infty) = \mu_X(1 - L_X(F_X(c_\infty)) - c_\infty(1 - F_X(c_\infty)) = \mu_X - \tau$$

and the policy (3.1.15) belongs to U.

Let $F_X(c_\infty) = r_\infty$ and we get the condition

$$e(c_\infty) = \mu_X\left(1 - L_X(r_\infty)\right) - c_\infty\left(1 - r_\infty\right) = \mu_X - \tau$$

or equivalently

$$\mu_X L_X(r_\infty) + c_\infty\left(1 - r_\infty\right) = \tau\,. \tag{3.1.17}$$

The corresponding Lorenz curve is

$$L_\infty(p) = \begin{cases} 0 & p < r_\infty \\ \dfrac{\mu_X}{\mu_X - \tau}\left(L_X(p) - L_X(r_\infty)\right) - \dfrac{c_\infty(p - r_\infty)}{\mu_X - \tau} & p \geq r_\infty \end{cases}. \qquad 3.1.18)$$

This Lorenz curve is continuous and has a derivative in the whole interval $(0,1)$ because in the point $p = r_\infty$ the derivative to the left is zero and to the right is $\dfrac{\mu_X}{\mu_X - \tau}\dfrac{c_\infty}{\mu_X} - \dfrac{c_\infty}{\mu_X - \tau} = 0$. Figure 3.1.1 gives examples of the extreme policies $Y = u_0(X)$ and $Y = u_\infty(X)$, and an arbitrary policy $Y = u(x)$.

A sketch of the Lorenz curves $L_X(p)$, $L_0(p)$, and $L_\infty(p)$ is given in Figure 3.1.2.

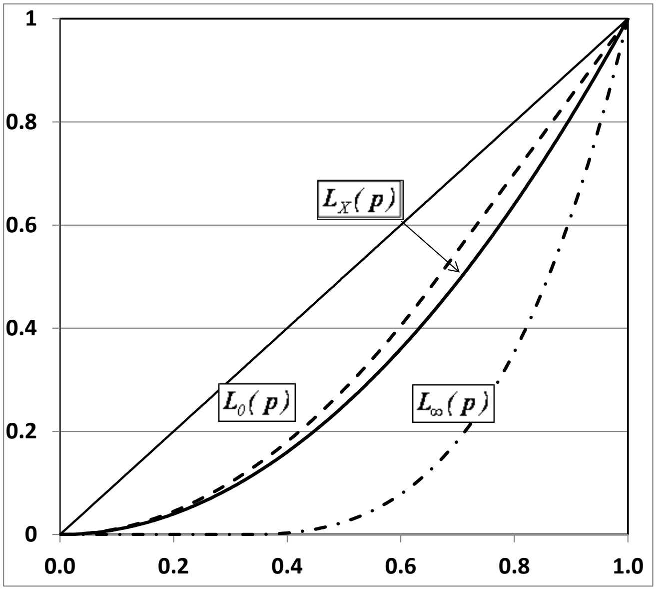



***Figure 3.1.2*** *The region between the extreme Lorenz curves is the region of attainable Lorenz curves (Fellman, 2001, 2014).*

We can prove (Fellman, 2001).

**Theorem 3.1.2.** The Lorenz curve $L_\infty(p)$ is inferior to all Lorenz curves corresponding to the class **U**.

**Proof.** Consider an arbitrary policy $u(x)$ in the class ***U***. For $x < c_\infty$, we get $u(x) \geq u_\infty(x)$. As a consequence of the condition $u'(x) \leq 1$ the curve $u(x)$ crosses $u_\infty(x)$ in only one point (say) $c_0 > c_\infty$ and for large $x$ values $u(x) \leq u_\infty(x)$. Hence, $u(x) \geq u_\infty(x)$ for $x < c_0$ and $u(x) \leq u_\infty(x)$ for $x \geq c_0$. Furthermore,

$$\int_0^\infty u_\infty(x) f_X(x)dx = \int_0^\infty u(x) f_X(x)dx = \mu_X - \tau .$$

The difference

$$D(p)=\int_0^{x_p}\left(u(x)-u_\infty(x)\right)f_X(x)dx$$

increases monotonically from 0 to a maximum $\int_0^{c_0}\left(u(x)-u_\infty(x)\right)f_X(x)dx$ for $x_p=c_0$ whereupon it decreases monotonically to 0 for $x_p\to\infty$. This behaviour proves the theorem.

The extreme Lorenz curves $\mathrm{L}_0(p)$ and $L_\infty(p)$ define a closed region of attainable Lorenz curves (c.f. Figure 3.1.2).

Now we evaluate the corresponding Gini coefficient $\mathrm{G}_\infty$. Consider the function

$$L_m(p)=\frac{\mu_X}{\mu_X-\tau}\left(L_X(p)-L_X(r_\infty)\right)-\frac{c_\infty}{\mu_X-\tau}\left(p-r_\infty\right) \qquad (3.1.19)$$

For $p<r_\infty$, $L_m(p)<0$ and for $p\geq r_\infty$ $L_m(p)=L_\infty(p)$. Hence $L_\infty(p)\geq L_m(p)$ for all $p\in[0,1]$. If we use (3.1.17), we get

$$=1-\frac{\mu_X}{\mu_X-\tau}(1-G_X)+2L_X(r_\infty)-\frac{c_\infty}{\mu_X-\tau}(1-r_\infty)^2=$$

$$<G_X+\frac{\tau(1+G_X)}{\mu_X-\tau}-\frac{c_\infty}{\mu_X-\tau}\leq G_X+\frac{\tau(1+G_X)}{\mu_X-\tau}.$$

In fact, this upper bound $G_X+\frac{\tau(1+G_X)}{\mu_X-\tau}$ is the same as the bound given in Fellman (1995). There the bound was stricter, since it was obtained without the derivative restriction in (3.1.1).

As a consequence of the formula (3.1.14) and Theorem 3.1.2, the Gini coefficient $G_X + \frac{\tau(1+G_X)}{\mu_X - \tau}$ is the maximum and $G_X - \frac{\tau}{\mu_X - \tau}(1-G_X)$ is the minimum of the Gini coefficients for the class ***U***. Hence we obtain for every policy $u(x)$ the inequalities

$$G_X - \frac{\tau(1-G_X)}{\mu_X - \tau} \le G_u \le G_X + \frac{\tau(1+G_X)}{\mu_X - \tau}. \qquad (3.1.20)$$

For the generalized Gini coefficient,

$$G(\nu) = 1 - \nu(1-\nu)\int_0^1 (1-p)^{\nu-2} L(p)dp$$

proposed by Yitzhaki (1983), Fellman (2001) obtained a similar formula

$$G_X(\nu) - \frac{\tau(1-G_X(\nu))}{\mu_X - \tau} \le G_u(\nu) \le G_X(\nu) + \frac{\tau(\nu - 1 + G_X(\nu))}{\mu_X - \tau}. \qquad (3.1.21)$$

For $\nu = 2$ the formula (3.1.21) is identical with (3.1.20).

Consider the welfare index $W = \mu(1-G)$ developed by Sen (1973) and later discussed by Lambert (2001, Chapter 5). For this index, Fellman (2001) obtained the simple inequality formula

$$W_X - 2\tau \le W_u \le W_X. \qquad (3.1.22)$$

From the deduction of the bounds in (3.1.20), (3.1.21) and (3.1.22), it follows that the formulae hold for arbitrary pre-tax income distributions. For a specific pre-tax income distribution, these bounds can be sharpened. This can be explained in the following way. Let us consider the lower bound in (3.1.20). For all pre-tax income distributions, the Lorenz curve $L_{u_0}(p)$ has a linear part, which starts from $p = p_0$ and which corresponds to the tax-paying part of the population (c.f. formula (3.1.8) and Figure 3.1.2). The accuracy of the lower bound depends on this linear part. The value of $1 - p_0$ indicates the proportion of taxpayers in the population and the accuracy of the bound increases as $1 - p_0$ decreases. Hence, the lower bound is accurate when there are very few but very high-income taxpayers.

Now we consider the upper bound in (3.1.20). The Lorenz curve $L_\infty(p) \equiv 0$ for $p < r_\infty$ and this part of the Lorenz curve influences the accuracy (c.f. formula (3.1.18) and Figure 3.1.2). For small values of $c_\infty$ and $r_\infty$ we obtain good accuracy. This is the case when the tax-paying ability of the low-income individuals is good, i.e. they are not extremely poor.

The strength of the obtained bounds is that they are independent of the distribution $F_X(x)$ and depend only on the basic quantities $G_X$, $\mu_X$ and $\rho$. In addition, the formulae obtained, are simple functions of these quantities. Furthermore, we observe that if $\tau \to 0$ then both the upper and lower bounds in (3.1.19), (3.1.20) and (3.1.21) converge towards $G_X$ $G_X(\nu)$ and $W_X$, respectively, indicating that the approximations presented have not introduced any "bias".

We have observed that ***U*** contains policies that increase and decrease inequality. Therefore, the intervals given for the indices are wide and the obtained bounds cannot be used as approximations of the indices of a specific policy in ***U***. The central role of these intervals is that they define limits for attainable index values and consequently give indications of the redistributive power of the class ***U***.

Let $G_0$ be the Gini coefficient of the policy (3.1.8) and $G_\infty$ be the Gini coefficient of the policy (3.1.15). Obviously, $\min_{\boldsymbol{U}} G = G_0 < G_\infty = \max_{\boldsymbol{U}} G$. Now we prove that the set of Gini coefficients which corresponds to the class (3.1.1) is compact, that is:

**Theorem 3.1.3.** There is a member of the class **U** with a prescribed Gini coefficient $\tilde{G} \in [G_0, G_\infty]$.

**Proof.** Let the prescribed Gini coefficient be $\tilde{G} \in [G_0, G_\infty]$. Construct a member of the class U as a linear combination of (3.1.8) and (3.1.15). We get $\tilde{G} = \theta G_0 + (1-\theta) G_\infty$ and the prescribed value of the Gini coefficient is obtained for

$$\tilde{\theta} = \frac{\tilde{G} - G_0}{G_\infty - G_0}. \tag{3.1.23}$$

**Remark.** Theorem 3.1.3 says that there exists at least one member of the class $\boldsymbol{U}$ that results in a post-tax income distribution with a prescribed Gini coefficient within the closed interval $[G_0, G_\infty]$. In general, this policy is not unique, but the extreme coefficients $G_0$ and $G_\infty$ are attainable only by the extreme policies.

One can also prove the analogous theorem:

**Theorem 3.1.4.** There is a member of the class **U** whose Lorenz curve satisfies the condition, $L_u(\tilde{p}) = \tilde{l}$ where $\tilde{l} \in [L_\infty(\tilde{p}), L_0(\tilde{p})]$.

**Proof.** The solution can be constructed by a linear combination of the policies (3.1.8) and (3.1.15). The prescribed condition is obtained for $\tilde{l} = \theta L_\infty(\tilde{p}) + (1-\theta) L_0(\tilde{p})$ and

$$\tilde{\theta} = \frac{\tilde{l} - L_0(\tilde{p})}{L_\infty(\tilde{p}) - L_0(\tilde{p})}. \tag{3.1.24}$$

Every point within the closed region, limited by the Lorenz curves $L_0(p)$ and $L_\infty(p)$, is attainable by a Lorenz curve corresponding to a member of the class ***U***. This means that there exists a policy that gives a post-tax income distribution such that the lowest proportion $\tilde{p}$ of income receivers receives exactly the proportion $\tilde{l}$ of the total amount of post-tax income. Within the class ***U***, the solution is not necessarily unique.

Consider the Lorenz curve $L_X(p)$ and the Lorenz curve $L_u(p)$, for an arbitrary member of the class ***U***. According to the general theory, we have $L'_X(p) = \frac{x_p}{\mu_X}$ and $L'_u(p) = \frac{y_p}{\mu_X - \tau}$. Now, $y_p = u(x_p)$ and, hence, $y_p \le x_p$ and we obtain

$$\frac{L'_u(p)}{L'_X(p)} \le \frac{\mu_X}{\mu_X - \tau}. \tag{3.1.25}$$

This is a necessary restriction on feasible Lorenz curves for members of the class ***U***. In general, there may be Lorenz curves between the extreme ones that do not correspond to policies in the class ***U***. The inequality (3.1.25) indicates that the Lorenz curve for the transformed variables cannot differ markedly from the Lorenz curve of *X*. This is especially notable for small values of $\tau\left(\tau / \mu_X\right)$. For the extreme policies (3.1.8) and (3.1.15) equality in (3.1.25) is obtained for the sub intervals $(0 \le p \le p_0)$ and $(r_\infty \le p \le 1)$, respectively. These properties stress the optimality of the extreme policies.

## 3.2 Attainable Lorenz Curves

In this Section we present necessary and sufficient conditions under which a given Lorenz curve can be obtained by a member of the class ***U***. These conditions are related to stochastic dominance of first order. Maasoumi and Heshmati (2000) presented stochastic dominance of first, second and third order and how they can be defined by alternative equivalent conditions.

Let $V$ and $W$ be non-negative stochastic variables having the distributions $F_V(v)$ and $F_W(w)$, the means $\mu_V$ and $\mu_W$ and the Lorenz curves $L_V(p)$ and $L_W(p)$, respectively. Using our notations the Maasoumi and Heshmati definition of stochastic dominance of first order is:

**Definition 3.2.1**. *The variable V First Order Stochastic Dominates W if and only if any one of the following equivalent conditions holds:*

*i.* $E[g(V)] \ge E[g(W)]$ *for all increasing functions g.*

*ii.* $F_V(v) \le F_W(v)$ *for all v.*

*iii.* $v_p \ge w_p$ *for all* $0 \le p \le 1$.

In this study of income distributions we restrict our investigations on non-negative continuous stochastic variables. For these the Lorenz curves are differentiable and we can prove the following lemma.

**Lemma 3.2.1.** *Let V and W be continuous non-negative stochastic variables having the distributions* $F_V(v)$ *and* $F_W(w)$, *the means* $\mu_V$ *and* $\mu_W$ *and the Lorenz curves* $L_V(p)$ *and* $L_W(p)$, *respectively, then the conditions:*

*i. V first order stochastic dominates W.*

*ii.* $F_V(v) \le F_W(v)$ *for all v.*

*iii.* $v_p \ge w_p$ *for all p* $(0 < p < 1)$.

*iv.* $\frac{L'_W(p)}{L'_V(p)} \le \frac{\mu_V}{\mu_W}$ *for all p* $(0 < p < 1)$.

*are equivalent.*

**Proof.** The equivalence between (i), (ii) and (iii) is given in Definition 3.2.1. Now, we only have to prove the equivalence between (iv) and (iii) (say). The connection between (iii) and (iv) are the formulae

$$L'_V(p) = \frac{v_p}{\mu_V} \text{ and } L'_W(p) = \frac{w_p}{\mu_W}$$

a) Assume that (iii) holds

Now,

$$1 \ge \frac{w_p}{v_p} = \frac{\mu_W L'_W(p)}{\mu_V L'_V(p)},$$

$$\frac{L'_W(p)}{L'_V(p)} \le \frac{\mu_V}{\mu_W}$$

and (iv) is obtained.

b) Assume that (iv) holds. Now

$$\frac{\mu_V}{\mu_W} \geq \frac{L'_W(p)}{L'_V(p)} = \left(\frac{w_p}{\mu_W}\right)\left(\frac{v_p}{\mu_V}\right)^{-1},$$

$$1 \geq \frac{w_p}{v_p}, \; v_p \geq w_p$$

and the proof is completed.

**Remark.** The condition (iv) in Lemma 3.2.1, being equivalent with (i), (ii) and (iii), indicates that we have obtained a new criterion for stochastic dominance of first order between two non-negative stochastic variables.

In Section 3.1 formula (3.1.25) we have noted that stochastic dominance of first order is a necessary condition that the transformed distribution is a post-tax income distribution corresponding to a policy of the class ***U***. In the following we obtain sufficient conditions.

At first we consider the class

$$\mathbf{U^*}: \begin{cases} u(x) \leq x \\ E(u(X)) = \mu_X - \tau \end{cases} \tag{3.2.1}$$

This class, presented in Fellman (1995) and in Fellman et al. (1996, 1999), is defined as the initial class ***U*** without the restriction

$$u'(x) \leq 1 \tag{3.2.2}$$

and consequently, $\boldsymbol{U} \subseteq \boldsymbol{U^*}$. Now we prove

**Theorem 3.2.1.** (Fellman, 2002, 2014) Consider a differentiable Lorenz curve $\overline{L}(p)$ and a stochastic variable Y with the corresponding distribution $\overline{F}_Y(y)$ with the mean $(\mu_X - \tau)$. Then the necessary and sufficient conditions that the Lorenz curve $\bar{L}(p)$ is an attainable Lorenz curve of a member of **U***, $\bar{F}_Y(y)$ being the corresponding distribution and $\bar{u}(x) = (\mu_X - \tau)\bar{L}'(F_X(x))$ being the corresponding transformation, is that one of the following equivalent conditions holds:

i. X first order stochastic dominates Y.

ii. $F_X(x) \le \bar{F}_Y(x)$ for all x.

iii. $y_p \le x_p$ for all p ($0 < p < 1$) or.

iv. $\dfrac{\bar{L}'(p)}{L'_X(p)} \le \dfrac{\mu_X}{\mu_X - \tau}$ for all p ($0 < p < 1$).

**Proof.** Assume that the presumptive post tax income distribution is $\bar{F}_Y(y)$ ($\bar{f}_Y(y)$) with the mean $\mu_X - \tau$. We introduce the quantiles $x_p$ and $y_p$, where $F_X(x_p) = p$ and $\bar{F}_Y(y_p) = p$. These quantiles can also be defined as $x_p = F_X^{-1}(p)$ and $y_p = \bar{F}_Y^{-1}(p)$. In Section 3.1 we noted that

$$y_p \le x_p \text{ for all p } (0 < p < 1) \tag{3.2.3}$$

and this condition still holds for the class ***U****. Consequently it is a necessary condition for $\bar{F}_Y(y)$ to be an attainable post-tax income distribution. From (3.2.3) it follows that

$$F_X(x_p) = p = \bar{F}_Y(y_p) \le \bar{F}_Y(x_p) \text{ for all p } (0 < p < 1).$$

The condition

$$F_X(x) \le \bar{F}_Y(x) \text{ for all x} \tag{3.2.4}$$

being equivalent with (3.2.3) is also a necessary condition that the post-tax income distribution $\overline{F}_Y(y)$ corresponds to a tax policy belonging to ***U****. From formula (3.2.4) we obtain

$$\overline{F}_Y^{-1}(F_X(x)) \le \overline{F}_Y^{-1}(\overline{F}_Y(x)) = x \text{ for all x.} \tag{3.2.5}$$

In the following we prove that the condition that the distribution $\overline{F}_Y(y)$ satisfies (3.2.5) is sufficient, that is, $\overline{F}_Y(y)$ is a post-tax income distribution for a member of the class ***U****. Consequently, the condition (3.2.4), being equivalent, is also sufficient. Consider a distribution $\overline{F}_Y(y)$ with mean $\mu_X - \tau$ satisfying (3.2.3). According to the definition of a distribution function we have

$$P(Y \le y) = \overline{F}_Y(y)\,. \tag{3.2.6}$$

The cumulative distribution function $\overline{F}_Y(y)$ is monotone increasing and

$$P(\overline{F}_Y(Y) \le \overline{F}_Y(y)) = \overline{F}_Y(y)\,. \tag{3.2.7}$$

If $Z = \overline{F}_Y(Y)$ and $z = \overline{F}_Y(y)$, then $Y = \overline{F}_Y^{-1}(Z)$, $y = \overline{F}_Y^{-1}(z)$ and

$$P(Z \le z) = z\,. \tag{3.2.8}$$

Consider the initial distribution $F_X(x)$. Then

$$z = P(Z \le z) = P(F_X^{-1}(Z) \le F_X^{-1}(z))\,. \tag{3.2.9}$$

Let $X = F_X^{-1}(Z)$ and $x = F_X^{-1}(z)$ then $Z = F_X(X)$ and $z = F_X(x)$.

Now,

$$y = \bar{F}_Y^{-1}(z) = \bar{F}_Y^{-1}(F_X(x)) = \bar{u}(x) \text{ (say).} \tag{3.2.10}$$

Hence $\bar{u}(x)$ is continuous and monotone increasing. In addition, from (3.2.5) follows that $\bar{u}(x)$ satisfies the condition

$$\bar{u}(x) \le x \tag{3.2.11}$$

and $\bar{u}(x)$ belongs to ***U**** and the distribution $\bar{F}_Y(y)$ corresponds to a policy belonging to the class ***U**** and the sufficiency is obtained.

Let us now consider Lorenz curves. First we give the conditions that a specific Lorenz curve (and the corresponding distribution $\bar{F}_Y(y)$ ) can be attained by a member of the class ***U****. Let us consider an arbitrary Lorenz curve $\bar{L}(p)$ with the conditions

i. $\bar{L}(p)$ has a continuous derivative of the first order ($\bar{L}'(p)$).

ii. $\lim_{p \to 1}(1-p)\bar{L}'(p) = 0$.

These conditions imply that the corresponding distribution $\bar{F}_Y(y) = M(\frac{y}{\mu})$, where $M(\cdot)$ is the inverse function to $\bar{L}'(p)$, is continuous and has a finite mean μ. When the Lorenz curve $\bar{L}(p)$ and the mean μ are given then the corresponding income distribution is unique (Fellman, 1976, 1980).

Consider a Lorenz curve $\bar{L}(p)$ and the corresponding distribution $\bar{F}_Y(y)$ with the mean $\mu_X - \tau$. We have

$$\bar{L}'(p) = \frac{y_p}{\mu_X - \tau}$$

and

$$y_p = \bar{F}_Y^{-1}(p) = (\mu_X - \tau)\bar{L}'(p) = (\mu_X - \tau)\bar{L}'(F_X(x_p)).$$

From these formulae it follows that $\bar{u}(x) = (\mu_X - \tau)\bar{L}'(F_X(x))$. Hence, the condition

$$\bar{u}(x) = (\mu_X - \tau)\bar{L}'(F_X(x)) \le x \qquad (3.2.12)$$

is a necessary condition for attainability. On the other hand let us assume that (3.2.12) holds. Let $\bar{F}_Y(y)$ be the distribution, which corresponds to $\bar{L}(p)$ and has the mean $\mu_X - \tau$. Then

$$x_p \ge (\mu_X - \tau)\bar{L}'(p) = (\mu_X - \tau)\bar{L}'(\bar{F}_Y(y_p)) = y_p,$$

and

$$y_p \le x_p \text{ for all p } (0 < p < 1). \qquad (3.2.13)$$

Consequently, the condition (3.2.12) is also sufficient and the theorem is proved.

Now we add the restriction $u'(x) \le 1$ and consider the initial class ***U*** of policies. For this class the necessary and sufficient condition is given in

**Theorem 3.2.2.** Consider a twice differentiable Lorenz curve $\bar{L}(p)$ and a stochastic variable Y with the corresponding distribution $\bar{F}_Y(y)$ with the mean $(\mu_X - \tau)$ and define $\bar{u}(x) = \bar{F}_Y^{-1}(F_X(x))$. Then necessary and sufficient condition that the Lorenz curve $\bar{L}(p)$ is an attainable Lorenz curve of a member of **U**, $\bar{F}_Y(y)$ being the corresponding distribution and $\bar{u}(x)$ being the corresponding transformation, is that one of the following equivalent conditions holds:

i. $\dfrac{\bar{L}''(p)}{L''_X(p)} \leq \dfrac{\mu_X}{\mu_X - \tau}$ for all p $(0 < p < 1)$ or equivalently,

ii. $f_X(x) \leq \bar{f}_Y(y)$ where $y = \bar{F}_Y^{-1}(F_X(x)) = \bar{u}(x)$.

**Proof.** If we assume that $\bar{L}(p)$ has the second derivative $\bar{L}''(p)$ we can add the restriction $\bar{u}'(x) \leq 1$ into (3.2.1) in order to obtain the class 3.1.1. We have the derivatives of first order $\bar{L}'(p) = \dfrac{y_p}{\mu_X - \tau}$ and $L'_X(p) = \dfrac{x_p}{\mu_X}$. According to the formula (1.3.2) the derivatives of the second order are $\bar{L}''(p) = \dfrac{1}{(\mu_X - \tau)\bar{f}_Y(y_p)}$ and $L''_X(p) = \dfrac{1}{\mu_X f_X(x_p)}$.

Note that if we, according to (3.2.10), define $y = \bar{F}_Y^{-1}(F_X(x)) = \bar{u}(x)$ then we obtain

$$\bar{u}'(x) = \frac{f_X(x)}{\bar{f}_Y(\bar{F}_Y^{-1}(F_X(x)))} \leq 1. \qquad (3.2.14)$$

and

$$f_X(x) \leq \bar{f}_Y(y) \text{ where } y = \bar{F}_Y^{-1}(F_X(x)) = \bar{u}(x) \qquad (3.2.15)$$

Hence, for every $p$ $(0 < p < 1)$ we have

$$f_X(x_p) \leq \bar{f}_Y(y_p). \qquad (3.2.16)$$

Consequently, (3.2.16) can be written

$$\frac{\bar{L}''(p)}{L''_X(p)} \le \frac{\mu_X}{\mu_X - \tau}. \tag{3.2.17}$$

This is a necessary condition that the transformation $y = \bar{F}_Y^{-1}(F_X(x)) = \bar{u}(x)$ in (3.2.14) belongs to the class ***U***. We can reverse the steps from (3.2.17) to (3.2.14) and consequently, (3.2.17) is also sufficient and the proof is completed.

If (3.2.17) is integrated we obtain

$$\mu_X L'_X(p) \ge (\mu_X - \tau)\bar{L}'(p), \tag{3.2.18}$$

or alternatively

$$\frac{\bar{L}'(p)}{L'_X(p)} \le \frac{\mu_X}{(\mu_X - \tau)} \tag{3.2.19}$$

which is identical with the condition (3.1.25). The integration step from (3.2.17) to (3.2.19) is not reversible so the condition (3.2.19) is only necessary for the class ***U*** given in (3.1.1) but, as proved above, necessary and sufficient for the class *U** given in (3.2.1). This difference can be explained so that there can exist policies belonging to the class *U** but not belonging to ***U***. Explicitly, such policies do not satisfy the condition $\bar{u}'(x) \le 1$.

The condition (iv) implies after integrations that

$$(\mu_X - \tau)\bar{L}(p) \le \mu_X L_X(p). \tag{3.2.20}$$

indicating *Generalized Lorenz Dominance* (GLD). The integration step from $\bar{L}'(p) \le \frac{\mu_X}{(\mu_X - \tau)} L'_X(p)$ given in (iv) in Theorem 3.2.1 to the condition (3.2.20) is not reversible. Consequently, GLD is only a necessary condition, or otherwise expressed, stochastic dominance implies GLD (cf. Lambert 2001 p. 49).

## 3.3 Classes of Non-differentiable Tax Policies

The transformed variable $Y = u(X)$ is the income after the taxation (Fellman, 2001, 2002; Fellman et al., 1996, 1999). In order to obtain a realistic class of policies we included in Fellman (2001, 2002) the additional restriction $u'(x) \le 1$. This condition indicates that the tax paid is an increasing function of the income *x*. In order to allow that the function $u(x)$ is not differentiable everywhere, we replace in this study the derivative restriction by the more general condition $\Delta u(x) \le \Delta x$ (Fellman, 2013). According to this restriction the function $u(x)$ is continuous and the tax is an increasing function of the income *x*. In fact, the increment in the tax is $\Delta x - \Delta u(x) \ge 0$. If $u'(x) \le 1$ holds then it follows that

$$\Delta u(x) = u(x + \Delta x) - u(x) = u'(\xi)\Delta x \le \Delta x ,$$

but the condition $\Delta u(x) \le \Delta x$ is more general and does not imply differentiability. We intend to show that the assumption $\Delta u(x) \le \Delta x$ is sufficient for the whole theory.

Now, the class of tax policies is

$$\mathbf{U:} \begin{cases} u(x) \le x \\ \Delta u(x) \le \Delta x \\ E(u(X)) = \mu_X - \tau \end{cases} . \tag{3.3.1}$$

We consider the extreme policies

$$u_0(x) = \begin{cases} x & x \le a_0 \\ a_0 & x > a_0 \end{cases} \tag{3.3.2}$$

and

$$u_{\infty}(x)=\begin{cases}0 & x\leq c_{\infty}\\ x-c_{\infty} & x>c_{\infty}\end{cases}. \tag{3.3.3}$$

The function $u_0(x)$ in (3.3.2) is not differentiable in the point $a_0$ and $u_{\infty}(x)$ in (3.3.3) in the point $c_{\infty}$, but the condition $\Delta u(x)\leq \Delta x$ holds for all *x*. Already in (3.1.12) we obtained that the Lorenz curve corresponding to (3.3.2) is

$$L_0(p)=\begin{cases}\dfrac{\mu_X}{\mu_X-\tau}L_X(p) & p\leq p_0\\ \dfrac{\mu_X}{\mu_X-\tau}L_X(p_0)+\dfrac{a_0}{\mu_X-\tau}\left(p-p_0\right) & p>p_0\end{cases}, \tag{3.3.4}$$

where $p_0=F_X(a_0)$ and according to (3.1.18) the Lorenz curve corresponding to (3.3.3) is

$$L_{\infty}(p)=\begin{cases}0 & p<r_{\infty}\\ \dfrac{\mu_X}{\mu_X-\tau}\left(L_X(p)-L_X(r_{\infty})\right)-\dfrac{c_{\infty}(p-r_{\infty})}{\mu_X-\tau} & p\geq r_{\infty}\end{cases}, \tag{3.3.5}$$

where $p_{\infty}=F_X(c_{\infty})$.

The policy (3.3.2) is optimal, that is, it Lorenz dominates all the policies in the class **U**, and the policy (3.3.3) is Lorenz dominated by all policies in **U** (Fellman, 2001, 2002).

In the following we show how the main result in Fellman (2002) can be obtained when we replace the restriction $\bar{u}?\, x)\leq 1$ by the more general restriction $\Delta\bar{u}(x)\leq\Delta x$. The function $\bar{u}(x)$ may be piecewise differentiable as the transformations (3.3.2) and (3.3.3). We consider post-tax income distributions with the mean $\mu_X-\tau$. Without the restriction $\Delta\bar{u}(x)\leq\Delta x$, the necessary and sufficient condition that a given Lorenz curve $\overline{L}(p)$ ($\overline{F}_Y(y)$) corresponds to a member of the class **U** is that the initial distribution $F_X(x)$ stochastically dominates $\overline{F}_Y(y)$. The inclusion of the restriction $\Delta\overline{u}(x) \le \Delta x$ results that the stochastic dominance is only necessary, that is the transformed distribution $\overline{F}_Y(y)$ must satisfy additional conditions.

Assume a given differentiable Lorenz curve $\overline{L}(p)$ with a continuous derivative. These conditions can be assumed because the corresponding transformation $\overline{u}(x)$ has to be continuous satisfying the condition $\Delta\overline{u}(x) \le \Delta x$. Starting from $\overline{L}(p)$, the connection between $\overline{L}(p)$ and the post-tax distribution $\overline{F}_Y(y)$ with the mean $\mu_X - \tau$ is that $\overline{F}_Y(y) = M\left(\frac{y}{\mu_X - \tau}\right)$, where $M(\cdot)$ is the inverse function of $\overline{L}'(p)$. The corresponding transformation is $\overline{u}(x) = y = (\mu_X - \tau)\overline{L}'(F_X(x))$. The condition $\Delta\overline{u}(x) \le \Delta x$ can be written

$$\begin{aligned}\Delta\overline{u}(x) &= (\mu_X - \tau)\left(\overline{L}'(F_X(x+\Delta x)) - \overline{L}'(F_X(x))\right)\\ &= (\mu_X - \tau)\left(\overline{L}'(p+\Delta p) - \overline{L}'(p)\right)\end{aligned}$$

where $p = F_X(x)$ and $p + \Delta p = F_X(x+\Delta x)$. On the other hand, we can write

$$\Delta\overline{u}(x) = (\mu_X - \tau)\left(\overline{L}'(p+\Delta p) - \overline{L}'(p)\right) = \left(y_{p+\Delta p} - y_p\right),$$

where $y_p$ and $y_{p+\Delta p}$ are defined by $p = \overline{F}_Y(y_p)$, $p + \Delta p = \overline{F}_Y(y_{p+\Delta p})$.

If we assume that $\overline{u}(x)$ is piecewise differentiable, then $\overline{L}'(p)$ and $\overline{F}_Y(y)$ are piecewise differentiable.

If we assume that the density functions $f_X(x)$ and $\overline{f}_Y(y)$ exist, we obtain

$$\Delta p = F_X(x+\Delta x) - F_X(x) = f_X(\xi)\Delta x,$$

where $x < \xi < x + \Delta x$ and

$$\begin{aligned}\Delta p &= \bar{F}_Y(y_{p+\Delta p}) - \bar{F}_Y(y_p) = \bar{f}_Y(\eta)(y_{p+\Delta p} - y_p)\\ &= \bar{f}_Y(\eta)\big(\bar{u}(x_p + \Delta x) - \bar{u}(x_p)\big)\end{aligned}$$

where $\bar{f}_Y(y) = \bar{F}_Y'(y)$ and $y_p < \eta < y_{p+\Delta p}$.

Consequently,

$$p = \bar{F}_Y(y_p) = F_X(x_p)$$

and

$$y_p = \bar{F}_Y^{-1}\big(F_X(x)\big).$$

From $f_X(\xi)\Delta x = \Delta p = \bar{f}_Y(\eta)\Delta\bar{u}(x)$ and from the condition $\Delta\bar{u}(x) \le \Delta x$ it follows that

$$f_X(\xi)\Delta x = \bar{f}_Y(\eta)\Delta\bar{u}(x) \le \bar{f}_Y(\eta)\Delta x$$

and consequently, $\dfrac{f_X(\xi)}{\bar{f}_Y(\eta)} \le 1$. If we let $\Delta x \to 0$, then $\Delta p \to 0$, $\xi \to x$ and $\eta \to y_p$ and we obtain $\dfrac{f_X(x)}{\bar{f}_Y(y_p)} \le 1$. This condition can also be written $h(x)$ or $\dfrac{f_X(x)}{\bar{f}_Y(y)} \le 1$ when $h(x)$. Hence, all the results in Fellman (2002) still hold, but the proof had to be slightly modified.

## 3.4 Discussion

In this chapter we reconsidered the effect of variable transformations on the redistribution of income. The aim was to generalise the conditions considered in earlier papers. Particularly we were interested if we can drop the assumptions of continuity and differentiability of the transformations. The main result is that with a slight modification of the proof the additional condition $\frac{f_X(x)}{\bar{f}_Y(y)} \leq 1$ is obtained.

We have obtained that, if we demand sufficient and necessary conditions, theorems earlier obtained, still hold and the continuity assumption can be included in the general conditions. The main result is that continuity is a necessary condition if one pursues that the income inequality should remain or be reduced.

The study of the class of tax policies indicated that the differentiability, earlier assumed, can be dropped but if one wants to retain the realism of the class the transformations should still be continuous and satisfy the restriction $\Delta\bar{u}(x) \leq \Delta x$. The earlier results obtained and presented in Fellman (2001, 2002) still hold.

Empirical applications of the optimal policies of a class of tax policies and the class of transfer policies considered here have been discussed in Fellman et al. (1996, 1999). There we developed "optimal yardsticks" to gauge the effectiveness of given real tax and transfer policies in reducing inequality.

## References

[1] Fellman, J. (1976). The effect of transformations on Lorenz curves. *Econometrica 44*:823-824.

[2] Fellman, J. (1980). *Transformations and Lorenz curves.* Swedish School of Economics and Business Administration Working Papers 48, 18 pp.

[3] Fellman, J. (1995). Intrinsic mathematical properties of classes of income redistributive policies. Swedish School of Economics and Business Administration Working Papers, 306, 26 pp.

[4] Fellman, J. (2001). Mathematical properties of classes of income redistributive policies. *European Journal of Political Economy 17*:195-209.

[5] Fellman, J. (2002). The redistributive effect of tax policies. *Sankhya Ser. B 64*:1-11.

[6] Fellman, J. (2013). Properties of Non-Differentiable Tax Policies. *Theoretical Economics Letters* 2013, 3, 142-145. doi:10.4236/tel.2013.33022. Published Online June 2013. (http://www.scirp.org/journal/tel).

[7] Fellman, J. (2014). The properties of a class of tax policies. Advances and *Applications in Statistics, ADAS. 39(2)*:125-148.

[8] Fellman, J., Jäntti, M., Lambert, P. (1996). *Optimal Tax-transfer Systems and Redistributive Policy: The Finnish Experience*. Swedish School of Economics and Business Administration Working Papers 324.

[9] Fellman, J., Jäntti, M., Lambert, P. (1999). Optimal tax-transfer systems and redistributive policy. *Scandinavian Journal of Economics 101*:115-126.

[10] Lambert, P. J. (2001). *The Distribution and Redistribution of Income: A Mathematical Analysis*. (3rd edition) Manchester: Manchester University Press. xiv+313 pp.

[11] Maasoumi, E. & Heshmati, A. (2000). Stochastic dominance amongst Swedish income distributions. *Econometric Reviews 19*:287-320.

[12] Sen, A. (1973). *On Economic Inequality*. Clarendon Press, Oxford.

[13] Yitzhaki, S. (1983). On an extension of the Gini index. *International Economic Review 24*:617-628.

# 4

# Transferring

In Chapter one and two, we have introduced the central properties of income distributions and the methods how to analyse income distributions and redistributions. We have also given example how to estimate distributions and concentration measures in empirical data. In Chapter 3 we have presented the effect of taxation on the income distribution and inequality. In this chapter we apply the theory in order to analyse the effects of transfer policies.

## 4.1 The Class of Transfer Policies

In this section we present the results of a study of a class of transfer (benefit) policies. Below we compare some results concerning transfer policies with our earlier results concerning tax policies. Consider an initial income distribution, defined in Chapter 3, with the distribution function $F_X(x)$, density function $f_X(x)$, mean $\mu_X$, Lorenz curve $L_X(p)$, the Gini coefficient $G_X$, generalized Gini coefficient $G_X(\nu)$ (Yitzhaki, 1983) and welfare index $W_X$ (Sen, 1973). Following Fellman (1995, 2001) and Fellman et al. (1996, 1999), we consider the class of transfer policies characterized by the transformation $Y = h(X)$, where $h(\cdot)$ is non-negative, monotone-increasing and continuous with the properties

$$\text{H:} \begin{cases} h(x) \geq x \\ \\ E(Y) = \mu_X + \rho \end{cases}. \tag{4.1.1}$$

The function $h(x)$ is income including government cash transfers associated with the original income $x$ and ρ is the mean transfer. The scenario pursued here can apply as well to an income policy; in that case $h(x)$ is income after an increase according to the policy. The monotonity of $h(x)$ indicates that the initial income order remains fixed. The first formula in (4.1.1) is obvious and the second indicates that the class ***H*** of transfer policies is constrained to distribute a given amount of benefit (ρ).

The class ***H*** contains both progressive and non-progressive policies and is therefore an adaptive tool for inequality and welfare studies. It is necessary already at this stage to point out that the transformed incomes corresponding to the policies in ***H*** do not have a Lorenz ordering.

We present some general results analogous to the results holding for the class of tax policies in Chapter 3. For details, see Fellman (1995 and 2001). Consider a set of arbitrary policies $h_i(x), (i=1,\ldots,k)$, belonging to ***H***. Following the analyses in Section 3.1 we obtain that the transformation

$$h_\theta(x)=\sum_{i=1}^{k}\theta_i h_i(x) \quad \theta_i \geq 0 \quad (i=1,\ldots,k) \quad \sum_{i=1}^{k}\theta_i=1, \tag{4.1.2}$$

also belongs to ***H*** because

$$h_\theta(x)=\sum_{i=1}^{k}\theta_i h_i(x) \geq \sum_{i=1}^{k}\theta_i x = x\sum_{i=1}^{k}\theta_i = x, \tag{4.1.3}$$

and

$$E\left(\sum_{i=1}^{k}\theta_i h_i(x)\right)=\sum_{i=1}^{k}\theta_i E\left(h_i(X)\right)=\sum_{i=1}^{k}\theta_i\left(\mu_X+\rho\right)=\mu_X+\rho\,. \tag{4.1.4}$$

Denote the corresponding Lorenz curves by $L_i(p), (i=1,\ldots,k)$ and $L_\theta(p)$ and the corresponding Gini coefficients by $G_i, (i=1,\ldots,k)$ and $G_\theta$, then $h_\theta(x)$ has the Lorenz curve

$$L_\theta(p)=\sum_{i=1}^{k}\theta_i L_i(p) \tag{4.1.5}$$

and the Gini coefficient

$$G_\theta = \sum_{i=1}^{k} \theta_i G_i \,. \tag{4.1.6}$$

Consequently, we obtain a theorem which is analogous to Theorem 3.1.1.

**Theorem 4.1.1.** (Fellman, 1995 and 2001) The class **H** and the classes of Lorenz curves and Gini coefficients corresponding to the policies in **H** are convex.

In order to obtain the range of the policies, we consider the member

$$h_0(x) = \begin{cases} b_0 & x \le b_0 \\ x & x > b_0 \end{cases}, \tag{4.1.7}$$

i.e. all incomes below the level $b_0$ are raised up to $b_0$ and all incomes above this level remain as they were. The policy (4.1.7) is an example of the minimum salary policy.

For an arbitrary value of *b*,

$$\begin{aligned} E(h_0(x)) &= \int_0^b b f_X(x)dx + \int_b^\infty x f_X(x)dx \\ &= bF_X(b) + \mu_X - \mu_X L_X(F_X(b))e(b) \end{aligned} .$$

Now, $e(0) = \mu_X$, and we obtain

$$\lim_{b\to\infty}(e(b)) = \lim_{b\to\infty} bF_X(b) + \mu_X - \mu_X \lim_{b\to\infty} L_X(F_X(b)) = \infty \,.$$

The derivative $e'(b) = F_X(b) > 0$ and $e(b)$ is monotone increasing. Consequently, there is a unique $b_0$ such that $E(h_0(X)) = \mu_X + \rho$. The policy (4.1.7) yields an income distribution that Lorenz dominates all the income distributions for the class ***H*** and has the Lorenz curve (Fellman, 1995, 2001; Fellman et al., 1996, 1999)

$$L_0(p) = \begin{cases} \dfrac{b_0}{\mu_X + \rho} p & p \le q_0 \\ \dfrac{b_0}{\mu_X + \rho} q_0 + \dfrac{\mu_X}{\mu_X + \rho}(L_X(p) - L_X(q_0)) & p > q_0 \end{cases} \tag{4.1.8}$$

where $F_X(b_0) = q_0$.

The Lorenz curve $L_0(p)$ has continuous derivative because for $p \le q_0$, $L_0'(p) = \dfrac{b_0}{\mu_X + \rho}$ and for $p > q_0$, $L_0'(p) = \dfrac{\mu_X}{\mu_X + \rho} L'_X(p) = \dfrac{\mu_X}{\mu_X + \rho} \dfrac{x_p}{\mu_X} = \dfrac{x_p}{\mu_X + \rho}$ which converges towards $\dfrac{b_0}{\mu_X + \rho}$ when $p \to q_0$.

The function $L_C(p) = \dfrac{b_0}{\mu_X + \rho} q_0 + \dfrac{\mu_X}{\mu_X + \rho}(L_X(p) - L_X(q_0))$ is not a Lorenz curve, but it is convex and has $L_T(p) = \dfrac{b_0}{\mu_X + \rho} p$ as tangent in the point $q_0$. Hence, $L_C(p) \ge L_0(p)$ for all $p \in [0, 1]$. From $L_0(1) = 1$, we obtain $b_0 q_0 - \mu_X L_X(q_0) = \rho$ and the corresponding Gini coefficient is

$$G_0 \ge 1 - 2\int_0^1 L_C(p)dp = 1 - 2\frac{\mu_X}{\mu_X + \rho}\int_0^1 L_X(p)dp - 2\frac{b_0 q_0}{\mu_X + \rho} + 2\frac{\mu_X}{\mu_X + \rho} L_X(q_0)$$

$$G_X + \frac{\rho}{\mu_X + \rho}(1 - G_X) - 2\frac{\rho}{\mu_X + \rho} =$$

$$G_X - \frac{\rho}{\mu_X + \rho}(1 + G_X).$$

Consequently, $G_0$ satisfies the inequality

$$G_0 \geq G_X - \frac{\rho}{\mu_X + \rho}(1 + G_X). \tag{4.1.9}$$

This bound is obviously a lower bound of the Gini coefficient of all policies in ***H***.

For the transfer policies, a lowest Lorenz curve cannot be found, but we can attain arbitrarily closely an inferior Lorenz curve (Fellman, 2001). Consider the sequence of transfer policies

$$\boldsymbol{H}_S: \; h_i(x) = \begin{cases} x & x < x_i \\ x + k_i(x - x_i) & x \geq x_i \end{cases} \quad i = 1, 2, \ldots \tag{4.1.10}$$

These policies give no benefits to the poorest part of the population ($x < x_i$), but positive benefits to the richest part ($x \geq x_i$). We construct the sequence so that $\boldsymbol{H}_S \subseteq \boldsymbol{H}$ and that their Lorenz curves converge towards an inferior Lorenz curve. If we define $k_i (> 0)$ so that $\int_{x_i}^{\infty} k_i(x - x_i) f_X(x)dx = \rho$, then every $h_i(x)$ is continuous and monotone increasing, $h_i(x) \geq x$ and $E(h_i(X)) = \mu_X + \rho$. Hence, $\boldsymbol{H}_S \subseteq \boldsymbol{H}$ and the corresponding Lorenz curve is (Fellman, 1995 and 2001):

$$L_i(p) = \begin{cases} \dfrac{\mu_X}{\mu_X + \rho} L_X(p) & p < q_i \\ \dfrac{\mu_X}{\mu_X + \rho} L_X(q_i) + \\ \dfrac{\rho}{\mu_X + \rho} \dfrac{\mu_X\left(L_X(p) - L_X(q_i)\right) - x_i(p - q_i)}{\mu_X\left(1 - L_X(q_i)\right) - x_i(1 - q_i)} & p \geq q_i \end{cases}, \tag{4.1.11}$$

where $F_X(x_i) = q_i$ .

If we choose the sequence $i = 1,2,\ldots$ so that $x_i \to \infty, q_i \to 1$ and hence, $k_i \to \infty$ in (4.1.10) we obtain a limit Lorenz curve

$$L_\infty(p) = \begin{cases} \dfrac{\mu_X}{\mu_X + \rho} L_X(p) & p < 1 \\ 1 & p = 1 \end{cases} . \qquad (4.1.12)$$

This Lorenz curve has no well-defined income distribution.[4] It does not correspond to a member of the class ***H*** but we can come arbitrarily close to it by choosing $q_i$ in (4.1.11) arbitrarily close to *1*, that is, the proportion of benefit-receivers tends towards zero. We can prove:

**Theorem 4.1.2.** (Fellman, 1995). The Lorenz curve $L_\infty(p)$ is inferior to the Lorenz curves for the class **H**.

**Proof.** Choose an arbitrary policy $h(x)$ in ***H***. We can evaluate its Lorenz curve in the following way:

$$L_h(p) = \frac{1}{\mu_X + \rho} \int_0^{x_p} h(x) f_X(x) dx \geq \frac{1}{\mu_X + \rho} \int_0^{x_p} x f_X(x) dx = \frac{\mu_X}{\mu_X + \rho} L_X(p) = L_\infty(p)$$

This inequality holds for all $p \leq 1$.

Figure 4.1.1 gives examples of the Lorenz curves $L_X(p)$, $L_0(p)$ and $L_\infty(p)$. The Lorenz curves (4.1.8) and (4.1.12) define the semi-closed region of attainable Lorenz curves (Figure 4.1.1).

[4]Vaguely speaking the limit inferior Lorenz curve in (4.1.12) corresponds to a policy, which gives all benefits to the richest income-receiver.

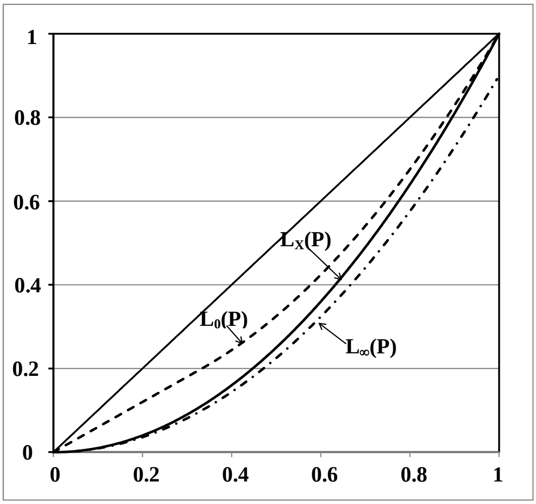



***Figure 4.1.1*** *The Lorenz curves* $L_X(p)$, $L_0(p)$ *and* $L_\infty(p)$. *The region between the extreme Lorenz curves is the region for attainable Lorenz curves (Fellman, 2001).*

The Gini coefficient $G_\infty$ corresponding to the Lorenz curve (4.1.12) is

$$G_\infty = 1 - 2\int_0^1 L_\infty(p)dp = 1 - 2\frac{\mu_X}{\mu_X + \rho}\int_0^1 L_X(p)dp =$$

$$1 - \frac{\mu_X}{\mu_X + \rho}(1 - G_X) = G_X + \frac{\rho}{\mu_X + \rho}(1 - G_X). \tag{4.1.13}$$

The final formula (4.1.13) is an exact equality and is an upper bound for all Gini coefficients for the policies in ***H***. However, it does not correspond to any member of the class ***H***. From Theorem 4.1.2 and the convergence of the policies in $\boldsymbol{H}_S$ it follows that (4.1.13) is a supremum of the Gini coefficients belonging to ***H***. For the Gini coefficient, the generalized Gini coefficient and the welfare index we obtain the bounds (Fellman, 1995):

$$G_X - \frac{\rho}{\mu_X + \rho}(1 + G_X) \le G_h \le G_X + \frac{\rho}{\mu_X + \rho}(1 - G_X), \tag{4.1.14}$$

$$G_X(\nu)-\frac{\rho}{\mu_X+\rho}(\nu-1+G_X(\nu))\le G_h(\nu)\le G_X(\nu)+\frac{\rho}{\mu_X+\rho}(1-G_X(\nu)) \quad (4.1.15)$$

and

$$W_X \le W_h \le W_X + 2\rho\,. \quad (4.1.16)$$

From the deduction of the bounds in (4.1.14), (4.1.15) and (4.1.16) it follows that the formulae hold for arbitrary income distributions. For a specific income distribution the lower bounds in (4.1.14) and (4.1.15) and the upper bound in (4.1.16) can be sharpened and the accuracy of the bounds was discussed in detail in Fellman (1995, 2001). The strength of the bounds is that they are independent of the distribution $F_X(x)$ and depend only on the basic quantities $G_X$, $\mu_X$ and $\rho$. In addition, the formulae obtained, are simple functions of these quantities. If $\rho\to 0$ then both the upper and lower bounds in (4.1.14), (4.1.15) and (4.1.16) converge towards $G_X$ , $G_X(\nu)$ and $W_X$ , respectively, indicating that the approximations performed do not cause any "bias".

The bounds presented cannot be used as approximate formulae for a specific policy in ***H***. The central role of these intervals is that they define limits for the attainable index values and hence give indications of the redistributive effect of the class of transfer policies. C.f. with the analysis of the formulae (3.1.20), (3.1.21) and (3.1.22) for the tax policies in Section 3.1.

Every Gini coefficient belonging to the semi-closed interval $\tilde{G}\in[G_0,G_\infty)$ and every point within the semi-closed region limited by the Lorenz curves $L_\infty(p)$ (excluded) and $L_0(p)$ (included) is attainable by a member of the class ***H***. These results can be given in following theorems:

**Theorem 4.1.3.** (Fellman, 1995)**.** There is a member of the class **H** with a prescribed Gini coefficient $\tilde{G} \in [G_0, G_\infty)$.

**Proof.** Choose $q_i$ in (4.1.11) so that the corresponding member of the sequence (4.1.10) has a Gini coefficient $G_i$ which exceeds $L_h(p)$. Construct a member of the class $H$ as a linear combination of (4.1.7) and this member of (4.1.11). We get $\tilde{G} = \theta G_0 + (1-\theta) G_i$ and the prescribed value of the Gini coefficient is obtained for

$$\tilde{\theta} = \frac{\tilde{G} - G_0}{G_i - G_0}. \tag{4.1.17}$$

This means that there exists a policy that gives a post-transfer income distribution with the Gini coefficient $\tilde{G}$.

**Remark.** Theorem 4.1.3 says that there exists at least one member of the class ***H*** that results in a post-transfer income distribution with a prescribed Gini coefficient within the closed interval $[G_0, G_\infty)$. In general, this policy is not unique, but the extreme coefficient $G_0$ is attainable only by the extreme policy.

We can also prove:

**Theorem 4.1.4.** (Fellman, 1995). There is a member of the class **H** that satisfies the condition $L_h(\tilde{p}) = \tilde{l}$, where $\tilde{l} \in (L_\infty(\tilde{p}), L_0(\tilde{p})]$.

**Proof.** Choose a member from the set (4.1.10) such that its corresponding Lorenz curve $L_i(p)$ satisfies the inequality $L_i(\tilde{p}) < \tilde{l}$. The solution to Theorem 4.1.4 can be constructed by a linear combination of the policy (4.1.7) and the chosen member of (4.1.10). The prescribed condition is obtained for $\tilde{l} = \theta L_i(\tilde{p}) + (1-\theta) L_0(\tilde{p})$ and

$$\tilde{\theta} = \frac{\tilde{l} - L_0(\tilde{p})}{L_1(\tilde{p}) - L_0(\tilde{p})}. \qquad (4.1.18)$$

Every point within the closed region, limited by the Lorenz curves $L_0(p)$ and $L_\infty(p)$, is attainable by a Lorenz curve corresponding to a member of the class $\boldsymbol{H}$. This means that there exists a policy that gives a post-transfer income distribution such that the lowest proportion $\tilde{p}$ of income receivers receives exactly the proportion $\tilde{l}$ of the total amount of post-transfer income. Within the class $\boldsymbol{H}$, the solution discussed in Theorem 4.1.4 is in general not unique.

**Remark.** In fact, Theorem 4.1.4 can be generalised to the whole region $\tilde{l} \in \left[L_\infty(\tilde{p}), L_0(\tilde{p})\right]$. If the given point is located on the lower border, that is $L_h(\tilde{p}) = L_\infty(\tilde{p})$, then the solution is a member of the sub-set $\boldsymbol{H}_S$ under the restriction $q_i \geq \tilde{p}$.

Consider the Lorenz curve $L_X(p)$ and a Lorenz curve $L_h(p)$ for an arbitrary member $Y = h(X)$. According to the general theory, we have

$$L'_X(p) = \frac{x_p}{\mu_X}$$

and

$$L'_h(p) = \frac{y_p}{\mu_X + \rho}.$$

Now, $y_p = h(x_p)$ and $y_p \geq x_p$, indicating that $Y$ stochastically dominates $X$. Furthermore, we obtain

$$\frac{L'_h(p)}{L'_X(p)} \geq \frac{\mu_X}{\mu_X + \rho}. \tag{4.1.19}$$

This is a necessary restriction on feasible Lorenz curves for members of the class ***H***. In general, there may be Lorenz curves between the extreme ones that do not correspond to policies in the class ***H***. The inequality (4.1.19) indicates also that the Lorenz curve for the transformed variables cannot differ markedly from the Lorenz curve of $X$. This is especially notable for small values of $\rho$ $(\rho / \mu_X)$. For the extreme policy (4.1.7) and for the sequence of policies in (4.1.10), equality in the formula (4.1.19) is obtained for whole subintervals; $q_0 \leq p \leq 1$ and $0 \leq p \leq q_i$, respectively. For the inferior Lorenz curve (4.1.12), equality holds within the semi-closed interval $0 \leq p < 1$. These properties stress the optimality of the extreme policies. In the next section we give necessary and sufficient conditions that a given Lorenz curve corresponds to a transfer policy belonging to the class *H*.

## 4.2 Attainable Lorenz Curves

Above we noted that among post-transfer income distributions there exist distributions with given coefficients and distributions whose Lorenz curves have given, prescribed co-ordinates $(p, l)$. However, we also stressed that every Lorenz curve within the admissible region is not necessarily attainable. Now, we derive necessary and sufficient conditions that a given Lorenz curve corresponds to a transfer policy belonging to the given class ***H***. A similar study has been performed for tax policies in Chapter 3 and in Fellman (2001, 2002).

In general, let $U$ and $V$ be non-negative stochastic variables having the distributions $F_U(u)$ and $F_V(v)$, the means $\mu_U$ and $\mu_V$ and the Lorenz curves $L_U(p)$ and $L_V(p)$, respectively. Stochastic dominance of first, second and third order can be defined by alternative equivalent–conditions. Some of these are given by Maasoumi and Heshmati (2000). (cf. also Davidson and Duclos 2000, Klonner, 2000 and Zheng, 2000). Using our notations, the Maasoumi-Heshmati (2000) definition of stochastic dominance of first order is (c.f. Definition 3.2.1).

**Definition 4.2.1.** *The variable U First Order Stochastic Dominates V if and only if any one of the following equivalent conditions holds:*

*i.* $E[g(U)] \geq E[g(V)]$ *for all increasing functions g.*

*ii.* $F_U(u) \leq F_V(u)$ *for all u.*

*iii.* $u_p \geq v_p$ *for all p* $(0 < p < 1)$.

We can prove the following theorem (c.f. Lemma 3.2.1).

**Theorem 4.2.1.** Let U and V be non-negative stochastic variables having the distributions $F_U(u)$ and $F_V(v)$, the means $\mu_U$ and $\mu_V$ and the Lorenz curves $L_U(p)$ and $L_V(p)$, respectively, then the conditions:

(i) U stochastically dominates V.

(ii) $F_U(v) \leq F_V(v)$ for all v.

(iii) $u_p \geq v_p$ for all p $(0 < p < 1)$.

(iv) $\dfrac{L'_V(p)}{L'_U(p)} \leq \dfrac{\mu_U}{\mu_V}$ for all p $(0 < p < 1)$.

are equivalent.

**Proof.** The equivalence between (i), (ii) and (iii) is given in Definition 4.2.1. Now, we only have to prove the equivalence between (iii) and (iv) (say). The connection between (iii) and (iv) are the general formulae

$$L'_U(p) = \frac{u_p}{\mu_U} \text{ and } L'_V(p) = \frac{v_p}{\mu_V}. \tag{4.2.1}$$

a) Assume that (iii) holds. Now, using (4.2.1) we obtain

$$1 \geq \frac{v_p}{u_p} = \frac{\mu_V L'_V(p)}{\mu_U L'_U(p)}$$

and the condition (iv) is obtained.

b) Assume that (iv) holds. Now

$$\frac{\mu_U}{\mu_V} \geq \frac{L'_V(p)}{L'_U(p)} = \left(\frac{v_p}{\mu_V}\right)\left(\frac{u_p}{\mu_U}\right)^{-1}$$

$$1 \geq \frac{v_p}{u_p},$$

$$u_p \geq v_p$$

and the proof is completed.

**Remark.** The condition (iv) in Theorem 4.2.1, being equivalent with (i), (ii) and (iii), indicates that (iv) is a criterion for stochastic dominance of first order between two positive stochastic variables. This criterion was also presented in Fellman (2003) and in Chapter 3.

In Section 4.1 we have noted that stochastic dominance of first order is a necessary condition that a transformed distribution is a post-transfer income distribution corresponding to a policy of the class ***H***. In the following we obtain sufficient conditions. Our results can be given in the following theorem, which is analogous to Theorem 3.2.1.

**Theorem 4.2.2.** Consider a Lorenz curve $\bar{L}_Y(p)$ and a corresponding stochastic variable Y with the distribution $\bar{F}_Y(y)$ and the mean $(\mu_X + \rho)$. Then the necessary and sufficient condition that the Lorenz curve $\bar{L}_Y(p)$ is an attainable Lorenz curve of a member of **H**, $\bar{h}(x) = (\mu_X + \rho)\bar{L}'(F_X(x))$ being the member and $\bar{F}_Y(y)$ being the corresponding post-transfer distribution, is that one of the following equivalent conditions hold:

i. Y stochastically dominates X.

ii. $F_X(x) \geq \bar{F}_Y(x)$ for all x.

iii. $y_p \geq x_p$ for all p $(0 < p < 1)$ or.

iv. $\dfrac{\bar{L}'_Y(p)}{L'_X(p)} \geq \dfrac{\mu_X}{\mu_X + \rho}$.

**Proof.** Assume that the presumptive post-transfer income distribution is $\bar{F}_Y(y)$ with the mean $(\mu_X + \rho)$. We introduce the quantiles $x_p$ and $y_p$, where $F_X(x_p) = p$ and $\bar{F}_Y(y_p) = p$. These quantiles can also be defined as $F_X^{-1}(p) = x_p$ and $\bar{F}_Y^{-1}(p) = y_p$. In Section 4.1 we noted that

$$y_p \geq x_p \text{ for all p } (0 < p < 1) \tag{4.2.2}$$

is a necessary condition for $\bar{F}_Y(y)$ to be an attainable post-transfer income distribution. From (4.2.2) it follows that

$$\overline{F}_Y(x_p) \le \overline{F}_Y(y_p) = p = F_X(x_p) \text{ for all p } (0 < p < 1).$$

The condition

$$F_X(x) \ge \overline{F}_Y(x) \text{ for all x} \tag{4.2.3}$$

being equivalent with (4.2.2), is also a necessary condition that the post-transfer income distribution corresponds to a transfer policy belonging to ***H***. From formula (4.2.3) we obtain

$$\overline{F}_Y^{-1}(F_X(x)) \ge \overline{F}_Y^{-1}(\overline{F}_Y(x)) = x \text{ for all x.} \tag{4.2.4}$$

In the following we prove that the condition that the distribution satisfies (4.2.3) is sufficient, that is, the distribution is a post-transfer income distribution for a member of the class ***H***. Consequently, the equivalent conditions (i) and (iii) are also sufficient.

Consider a distribution with mean $(\mu_X + \rho)$ satisfying (4.2.3). According to the definition of a distribution function we have

$$P(Y \le y) = \overline{F}_Y(y).$$

The cumulative distribution function is monotone increasing and

$$P(\overline{F}_Y(Y) \le \overline{F}_Y(y)) = \overline{F}_Y(y). \tag{4.2.5}$$

If $Z = \overline{F}_Y(Y)$ and $z = \overline{F}_Y(y)$, then $Y = \overline{F}_Y^{-1}(Z)$, $y = \overline{F}_Y^{-1}(z)$ and

$$P(Z \le z) = z. \tag{4.2.6}$$

The transformed variable $Z = \overline{F}_Y(Y)$ is uniformly distributed over the interval (*0*, *1*) and (4.2.6) is an old well-known result. Consider the initial distribution $F_X(x)$. Then

$$z = P(Z \le z) = P(F_X^{-1}(Z) \le F_X^{-1}(z)). \qquad (4.2.7)$$

Let $X = F_X^{-1}(Z)$ and $x = F_X^{-1}(z)$ then $Z = F_X(X)$ and $z = F_X(x)$.

Now,

$$y = \overline{F}_Y^{-1}(z) = \overline{F}_Y^{-1}(F_X(x)) = \overline{h}(x) \text{ (say).} \qquad (4.2.8)$$

Hence $\overline{h}(x)$ is continuous and monotone increasing. In addition, from (4.2.4) follows that $\overline{h}(x)$ satisfies the condition

$$\overline{h}(x) \ge x \qquad (4.2.9)$$

and $\overline{h}(x)$ belongs to ***H*** and the distribution $\overline{F}_Y(y)$, having the mean $(\mu_X + \rho)$, corresponds to a policy belonging to the class ***H*** and the sufficiency is obtained.

Let us now introduce Lorenz curves and obtain the conditions that a specific Lorenz curve (and the corresponding distribution $\overline{F}_Y(y)$) can be attained by a member of the class ***H***. Let us consider an arbitrary Lorenz curve $\overline{L}(p)$ with the conditions:

i. $\overline{L}(p)$ has a continuous derivative of the first order ($\overline{L}'(p)$).

ii. $\lim_{p \to 1}(1-p)\overline{L}'(p)$ =*0*.

These conditions imply that the corresponding distribution $\overline{F}_Y(y) = M(\frac{y}{\mu})$, where $M(\cdot)$ is the inverse function to $\overline{L}'(p)$ is continuous and has a finite mean (Fellman, 1976, 1980). In those papers, it was assumed that the second derivative exists but this condition is not necessary in this context. In general, when the Lorenz curve $\overline{L}(p)$ and the mean are given, the corresponding income distribution is unique.

Consider the distribution $\overline{F}_Y(y)$ with the mean $\mu_X + \rho$. We have $\overline{L}'(p) = \dfrac{y_p}{\mu_X + \rho}$ and $y_p = \overline{F}_Y^{-1}(p) = (\mu_X + \rho)\overline{L}'(p)$. From these formulae it follows that $\overline{h}(x) = (\mu_X + \rho)\overline{L}'(F_X(x))$. Hence, the condition

$$\overline{h}(x) = (\mu_X + \rho)\overline{L}'(F_X(x)) \geq x \tag{4.2.10}$$

is a necessary condition for attainability. On the other hand let us assume that (4.2.10) holds. Let $\overline{F}_Y(y)$ be the distribution, which corresponds to $\overline{L}(p)$ and has the mean $\mu_X + \rho$. Then

$$x_p \leq (\mu_X + \rho)\overline{L}'(F_X(x_p)) = (\mu_X + \rho)\overline{L}'(p) = (\mu_X + \rho)\overline{L}'(\overline{F}_Y(y_p)) = y_p,$$

and $y_p \geq x_p$ for all $p$ ($0 < p < 1$). Consequently, (4.2.10) is also sufficient and the proof is completed.

The content in Theorem 4.2.2 indicates that every stochastic variable *Y*, which stochastically dominates *X* and whose mean is $\mu_X + \rho$ and whose Lorenz curve belongs to the semi-closed Lorenz region, is attainable by a policy belonging to ***H***. The condition (iv) implies after integrations that

$$(\mu_X + \rho)\overline{L}(p) \geq \mu_X L_X(p)\,. \tag{4.2.11}$$

indicating *Generalized Lorenz Dominance* (GLD). The integration step from $\overline{L}'(p) \geq \frac{\mu_X}{(\mu_X + \rho)} L_X'(p)$ given in (iv) in Theorem 4.2.2 to the condition (4.2.11) is not reversible. Consequently, GLD is only a necessary condition, or otherwise expressed, stochastic dominance implies GLD (cf. Lambert 2001 p. 49).

## 4.3 Discontinuous Transfer Policies with a Given Lorenz Curve

In earlier papers we have studied classes ***H*** of continuous transfer policies, defined in (4.1.1) and below in (4.3.1). In this section we consider an expanded class ***H****, containing discontinuous policies, defined below in (4.3.2)**,** and generalize the results holding for class ***H*** to class ***H****. A realistic transformation describing a general transfer policy must be continuous, but we also are prepared to consider situations where discontinuous policies are plausible.

For class ***H*** we have obtained supreme and inferior Lorenz curves $L = L_0(p)$ in (4.1.8) and $L = L_\infty(p)$ in (4.1.12). In addition, we have proved that there are policies belonging to ***H*** with given Gini coefficients or Lorenz curves passing through given points in the $(p, L)$ plane. The necessary and sufficient conditions under which a given Lorenz curve $\overline{L}(p)$ corresponds to a member of class ***H*** of transfer policies are equivalent to the condition that the transformed variable $\overline{Y} = h(X)$ stochastically dominates the initial variable *X*.

The notations in this section will be similar to those in our earlier sections. Let the income be $X$ with the distribution function $F_X(x)$, density function $f_X(x)$, mean $\mu_X$ and Lorenz curve $L_X(p)$. The basic formulae are

$$\mu_X = \int_0^\infty x\, f_X(x)dx$$

and

$$L_X(p) = \frac{1}{\mu_X}\int_0^{x_p} x\, f_X(x)dx,$$

where $F_X(x_p) = p$.

We introduce the transformation $Y = h(X)$, where $h(\cdot)$ is non-negative and monotone increasing. Since the transformation can be considered as a tax ($h(x) \le x$) or a transfer ($h(x) \ge x$) policy, the transformed variable $Y$ is either the post-tax or post-transfer income. The mean and the Lorenz curve for variable $Y$ are

$$\mu_Y = \int_0^\infty h(x) f_X(x)dx$$

and

$$L_Y(p) = \frac{1}{\mu_Y}\int_0^{x_p} h(x) f_X(x)dx.$$

A general theorem concerning Lorenz dominance is (Fellman, 1976; Jakobsson, 1976; Kakwani, 1977 and also given in Theorem 1.4.1):

**Theorem 4.3.1.** Let $X$ be a non-negative, random variable with distribution $F_X(x)$, mean $\mu_X$ and Lorenz curve $L_X(p)$. Let $h(x)$ be a non-negative, monotone-increasing function; let $Y = h(X)$ and let $E(Y) = \mu_Y$ exist. The Lorenz curve $L_Y(p)$ of Y exists, and the following results hold:

i. $L_Y(p) \geq L_X(p)$ if $\dfrac{h(x)}{x}$ is monotone decreasing.

ii. $L_Y(p) = L_X(p)$ if $\dfrac{h(x)}{x}$ is constant.

iii. $L_Y(p) \leq L_X(p)$ if $\dfrac{h(x)}{x}$ is monotone increasing.

Recently, Egghe (2009) returned to Theorem 4.3.1 and gave a new proof. In addition, he showed that the theorem is not true for the dual transformation.

Fellman (1980, 2003) introduced the class of transfer policies

$$\text{H: } \begin{cases} h(x) \geq x \\ E(h(X)) = \mu_X + \rho \end{cases}, \tag{4.3.1}$$

where $h(x)$ is non-negative, monotone increasing and continuous.

This class was considered in Section 4.1. Now we modify this class allowing $h(x)$ to be discontinuous and define

$$\boldsymbol{H^*}\text{: } \begin{cases} h(x) \geq x \\ E(h(X)) = \mu_X + \rho \end{cases}, \tag{4.3.2}$$

where $h(x)$ is non-negative and monotone increasing.

If $h(x)$ is discontinuous and satisfies Theorem 4.1.1 and the transformation should result in an increasing transformed variable with finite mean then the discontinuities can only consist of finite positive jumps and the number of jumps can be assumed to be finite or countable. (Fellman, 2009).

Consider an optimal policy which Lorenz dominates all policies in ***H****. According to Theorem 4.1.1, $\frac{h(x)}{x}$ must be monotonically decreasing. Consequently, it must be continuous because if $h(x)$ has a discontinuity point, then the ratio $\frac{h(x)}{x}$ has a positive jump and cannot be monotonously decreasing. Consequently, although class ***H****, in comparison with the initial class ***H***, also contains discontinuous policies, the policy

$$h_0(x) = \begin{cases} b_0 & x \le b_0 \\ x & x > b_0 \end{cases}, \tag{4.3.3}$$

being optimal among all continuous policies, is still optimal for the class ***H****. It has the Lorenz curve

$$L_0(p) = \begin{cases} \dfrac{b_0}{\mu_Y + \rho} p & p \le q_0 \\ \dfrac{b_0}{\mu_Y + \rho} q_0 + \dfrac{\mu_X}{\mu_Y + \rho}\left(L_X(p) - L_X(q_0)\right) & p > q_0 \end{cases}, \tag{4.3.4}$$

where $q_0 = F_X(b_0)$. The inferior Lorenz curve presented in Section 4.1 can be obtained from the sequence (Fellman, 2001)

$$\boldsymbol{H_s}\!: \; h_i(x) = \begin{cases} x & x < x_i \\ x + k_i(x - x_i) & x \ge x_i \end{cases} \quad (i = 1, 2, \ldots). \tag{4.3.5}$$

Define $k_i$ (>0) so that

$$\int_{x_i}^{\infty} k_i(x - x_i) f_X(x) dx = \rho ,$$

then $h_i(x)$ is continuous and monotone increasing, $h_i(x) \geq x$ and $E(h_i(X)) = \mu_X + \rho$. Hence, $\boldsymbol{H_s} \subseteq \boldsymbol{H} \subseteq \boldsymbol{H^*}$, and the corresponding Lorenz curve is (Fellman, 2001)

$$L_{hi}(p) = \begin{cases} \dfrac{\mu_X}{\mu_X + \rho} L_X(p) & p < q_i \\ \dfrac{\mu_X}{\mu_X + \rho} L_X(q_i) & \\ + \dfrac{\rho}{\mu_X + \rho} \dfrac{\mu_X(L_X(p) - L_X(q_i)) - x_i(p - q_i)}{\mu_X(1 - L_X(q_i)) - x_i(1 - q_i)} & p \geq q_i \end{cases}. \tag{4.3.6}$$

If we choose the sequence $i = 1,2,\ldots$ so that $x_i \to \infty$, $q_i \to 1$, and hence, $k_i \to \infty$ in (6), we obtain a limit Lorenz curve

$$L_\infty(p) = \begin{cases} \dfrac{\mu_X}{\mu_X + \rho} L_X(p) & p < 1 \\ 1 & p = 1 \end{cases}. \tag{4.3.7}$$

Independently of the existence of discontinuity points we can still prove (Fellman, 2009).

**Theorem 4.3.2.** The Lorenz curve $L_\infty(p)$ is inferior to the Lorenz curves for class **H***.

**Proof**. Consider an arbitrary, continuous or discontinuous policy $h(x)$ in $\boldsymbol{H^*}$ with Lorenz curve $L_h(p)$. Using the condition $h(x) \geq x$, we can evaluate $L_\infty(p)$ in the following way:

$$L_h(p) = \frac{1}{\mu_X + \rho} \int_0^{x_p} h(x) f_X(x) dx \geq \frac{1}{\mu_X + \rho} \int_0^{x_p} x f_X(x) dx = \frac{\mu_X}{\mu_X + \rho} L_X(p) = L_\infty(p) . \quad (4.3.8)$$

This inequality holds for all $0 \leq p \leq 1$.

In addition, ***H*** $\subseteq$ ***H****, and consequently, there exist policies belonging to ***H**** with given Gini coefficients or Lorenz curves passing through given points in the $(p, L)$ plane. Hence, class ***H**** of transfer policies containing discontinuous policies satisfies the same properties as the initial class ***H*** discussed in Fellman (1980) and Fellman (2001). Figure 4.3.1 includes the Lorenz curves $L_X(p)$, $L_0(p)$ and $L_\infty(p)$. The figure also shows a Lorenz curve, $L_Y(p)$ with a cusp corresponding to a discontinuous transformation.

In Fellman (2003) and in Section 4.2 we obtained necessary and sufficient conditions under which a given differentiable Lorenz curve $\bar{L}(p)$ corresponds to a member of a given class of transfer policies. These conditions are equivalent to the condition that the transformed variable $Y = h(X)$ stochastically dominates the initial variable *X*.

Now we generalize the results, including the classes with discontinuous transformations. A discontinuous transformation $h(x)$ can only have a countable number of positive finite steps, and every jump in the transformation $h(x)$ results in a cusp in the Lorenz curve.

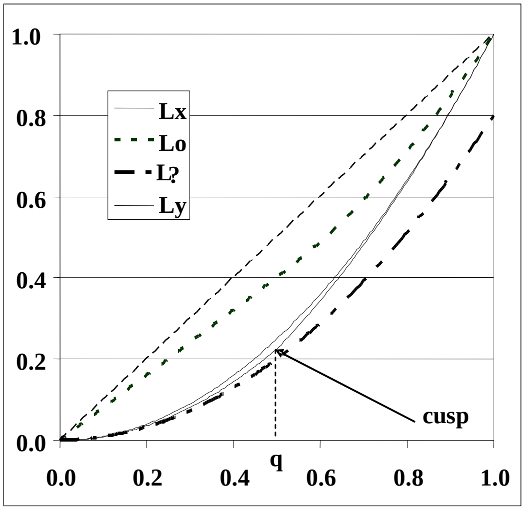



***Figure 4.3.1*** *A sketch of the Lorenz curves $L_X(p)$ and $L_Y(p)$, when $h(x)$ is discontinuous for $x = a$ and $q = F_X(a)$. Note the cusp of $L_Y(p)$ at the point $p = q$. The figure also includes the maximum and minimum Lorenz curves $L_0(p)$ and $L_\infty(p)$ for the transfer policies in* ***H**** *(Fellman, 2003).*

Consider a Lorenz curve $\overline{L}(p)$ which is convex and differentiable everywhere with the exception of a countable number of cusps. More general Lorenz curves cannot be considered. The corresponding distribution is $\overline{F}_Y(y) = M(\frac{y}{\mu})$, in which $M(\cdot)$ is the inverse function to $\overline{L}'(p)$ and $Y$ is assumed to have the mean $\mu = \mu_X + \rho$ (Fellman, 1980). If $\overline{L}(p)$ has a cusp, then the derivative $\overline{L}'(p)$ and the function $M(\cdot)$ have positive jumps.

In general, when Lorenz curve $\overline{L}(p)$ and the mean of the corresponding distribution are given, the income distribution is unique (see Chapter 1). Now, we prove that the conditions obtained earlier still hold for class ***H****, that is, we will characterize attainable Lorenz curves even though they are not universally differentiable.

Fellman (2003) has noted that stochastic dominance of first order is a necessary condition for a transformed distribution to be a post-transfer income distribution corresponding to a policy of class ***H*** in (4.3.1). The result is given in:

**Theorem 4.3.3.** Consider Lorenz curve $\overline{L}_Y(p)$ and a corresponding stochastic variable Y with the distribution $\overline{F}_Y(y)$ and the mean $(\mu_X + \rho)$. Then the necessary and sufficient condition that the Lorenz curve $\overline{L}_Y(p)$ is an attainable Lorenz curve of a member of **H***,

$$\overline{h}(x) = (\mu_X + \rho)\overline{L}'(F_X(x)),$$

being the member and $\overline{F}_Y(y)$ being the corresponding post-transfer distribution, is that one of the following equivalent conditions hold:

i. Y stochastically dominates X.

ii. $F_X(x) \geq \overline{F}_Y(x)$ for all x.

iii. $y_p \geq x_p$ for all p ($0 < p < 1$) or.

iv. $\dfrac{\overline{L}'(p)}{L_X'(p)} \geq \dfrac{\mu_X}{\mu_X + \rho}$.

When we prove this theorem for class ***H****, we have to show that $y_p \geq x_p$ holds for distribution $\overline{F}_Y(y)$. The proof of Theorem 4.2.2 given earlier in Fellman (2003) can be applied as such for $\overline{h}(x)$ wherever it is continuous, but the discontinuity points need special attention. Class ***H**** of transfer policies containing discontinuous policies satisfies the same properties as the initial class discussed in Fellman (1980) and Fellman (2003), and we obtain the transformation

$$y = \bar{h}(x) = (\mu_X + \rho)\bar{L}'(F_X(x)).$$

If $\bar{L}(p)$ has a cusp for $p = q$, then $\bar{h}(x)$ has a jump for $x_q$. Consider a neighbourhood $x_q - h < x < x_q + h$, where $x_q$ is the only discontinuity point of $\bar{h}(x)$, and choose a $\delta > 0$ so small that $x_q - h < x_{q-\delta} < x_{q+\delta} < x_q + h$.

Let $\lim_{\delta \to 0+} h(x_{q-\delta}) = y_{q-}$ and $\lim_{\delta \to 0+} h(x_{q+\delta}) = y_{q+} > y_{q-}$.

Now, the transformation $\bar{h}(x)$ is continuous for all $\delta > 0$, and

$$y_{q-\delta} = (\mu_X + \rho)\bar{L}'(F_X(x_{q-\delta})) \geq x_{q-\delta}.$$

When $\delta \to 0_+$, the inequality holds for the limits, and we obtain $y_{q-} \geq x_q$. Similarly, we obtain

$$y_{q+\delta} = (\mu_X + \rho)\bar{L}'(F_X(x_{q+\delta})) \geq x_{q+\delta},$$

and when $\delta \to 0_+$, the inequality holds for the limits, and we obtain $y_{q+} \geq x_q$. Hence, $y_p \geq x_p$ for all *p*, and $Y = \bar{h}(X)$ stochastically dominates the initial variable *X*.

## 4.4 Discussion

In this chapter we have studied the effects of transfer policies. A realistic transformation describing a general transfer policy must be continuous. The generalized class ***H**** of transfer policies containing discontinuous policies satisfies the same properties as the initial class discussed in Fellman (1980) and Fellman (2003). The theory presented here is obviously applicable in connection with other income redistributive studies such that the discontinuity can be assumed to be realistic. If the problem is reductions in taxation, then the tax reduction for a taxpayer can be considered as a new benefit (Fellman, 2001). Consequently, the class of transfer policies ***H**** can be used for comparisons between different tax-reducing policies. If transfers are increased, the effect of increases on a receiver can also be studied through transfer policies ***H****. In general, such changes may be mixtures of several different components and discontinuity cannot be excluded, and the continuity assumption can be dropped. One general result is still that continuity is a necessary condition if one expects that income inequality should remain or be reduced. Analogously, tax increases and transfer reductions can be considered as new tax policies (Fellman, 2001).

Empirical applications of the optimal policies among a class of tax policies and the class of transfer policies have been discussed in Fellman et al. (1999), where we developed "optimal yardsticks" to gauge the effectiveness of given real tax and transfer policies in reducing inequality. We return to these problems in the next chapter.

## References

[1] Davidson, R., Duclos, J. -Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica 68*:1435-1464.

[2] Egghe, L. (2009). The theorem of Fellman and Jakobsson: a new proof and dual theory. *Mathematical and Computer Modelling 50*:1595-1605.

[3] Fellman, J. (1976). The effect of transformations on Lorenz curves. *Econometrica 44*:823-824.

[4] Fellman, J. (1980). *Transformations and Lorenz curves.* Swedish School of Economics and Business Administration Working Papers: 48, 18 pp.

[5] Fellman, J. (1995). *Intrinsic Mathematical Properties of Classes of Income Redistributive Policies*. Swedish School of Economics and Business Administration Working Papers 306.

[6] Fellman, J. (2001). Mathematical properties of classes of income redistributive policies. *European Journal of Political Economy 17*:195-209.

[7] Fellman, J. (2002). The redistributive effect of tax policies. *Sankhya Ser. B 64*:1-11.

[8] Fellman, J. (2003). On Lorenz curves generated by a given class of transfer policies. In R. Höglund, M. Jäntti, G. Rosenqvist eds., *Statistics, Econometrics and Society*: Essays in honour of Leif Nordberg. Statistics Finland, Research Reports Number 239, 27-40.

[9] Fellman, J. (2009). Discontinuous transformations, Lorenz curves and transfer policies. *Social Choice and Welfare 33(2)*:335-342. DOI 10.1007/s00355-008-0362-4.

[10] Fellman, J., Jäntti, M., Lambert, P. (1996). *Optimal Tax-transfer Systems and Redistributive Policy: The Finnish Experience*. Swedish School of Economics and Business Administration Working Papers 324.

[11] Fellman, J., Jäntti, M., Lambert, P. (1999). Optimal tax-transfer systems and redistributive policy. *Scandinavian Journal of Economics 101*:115-126.

[12] Jakobsson, U. (1976). On the measurement of the degree of progression. *Journal of Public Economics 5*:301:316.

[13] Kakwani N. C. (1977). Applications of Lorenz curves in economic analysis. *Econometrica 45*:719-727.

[14] Klonner, S. (2000). The first-order stochastic dominance ordering of the Singh Maddala distribution. *Economics Letters 69*:123-128.

[15] Lambert, P. J. (2001). *The Distribution and Redistribution of Income: A mathematical analysis*. (3rd edition) Manchester: Manchester University Press. xiv+313 pp.

[16] Maasoumi, E., Heshmati, A. (2000). Stochastic dominance amongst Swedish income distributions. *Econometric Reviews 19*:287-320.

[17] Sen, A. (1973). *On Economic Inequality*. Clarendon Press, Oxford.

[18] Yitzhaki, S. (1983). On an extension of the Gini index. *International Economic Review 24*:617-628.

[19] Zheng, B. (2000). Poverty orderings. *J. of Economic Surveys 14*:427-440.

# 5

# Optimal Redistributive Tax Transfer Policy

Fellman et al. (1996, 1999) developed an approach to evaluating the design of the tax system and social transfer system scheme by relating the redistributive properties that would have occurred under an optimal design of taxes and transfers. They presented the findings which generalise the findings presented in Fellman (1976) and Fei (1981) and identified the optimal tax and transfer policies in a wide class of possible policies, constrained only to raise a given amount in tax revenue and/or distribute a given amount of cash benefits. These optimal policies for the given budget size would maximize welfare in the distribution of disposable money income in the absence of distinct effects. The extent to which observed policies fall short of this ideal, in reducing inequality, was measured by new indices, in which a distributional judgement parameter can be set to reflect alternative degrees of inequality aversion and to carry out sensitivity analysis.

In the Finnish case for the period 1971-1990, transfers were found not to be very efficient in redistribution income across households, whereas tax policies came much closer to the inequality reducing effect of an optimal pattern (Fellman et al., 1999).

## 5.1 Optimal Tax Policy

Consider, as above, the before tax income distribution, assumed given with the distribution function $F_X(x)$, density function $f_X(x)$, mean $\mu_X$, Lorenz curve $L_X(p)$ and the Gini coefficient $G_X$. Following Chapter 3, we consider a class of tax policies characterized by the transformation $Y = u(X)$ where $u(\cdot)$ is non-negative, monotone increasing and continuous with the properties

$$\text{U: } \begin{cases} u(x) \le x \\ u'(x) \le 1 \\ E(u(X)) = \mu_X - \tau \end{cases}, \tag{5.1.1}$$

where $u(x)$ is the post tax income associated with pre-tax income $x$ and the mean tax, assumed given. Consequently, we consider the same class of policies as in (3.1.1). Fellman et al. (1999) considered a slightly different class because the derivative condition was not assumed. Following Fellman et al (1999) we consider here impact effects only, not allowing individual agents for example to adjust their labour supplies in anticipation of the particular tax policy in the class which may be applied.

The polar case presented in (3.1.8):

$$\mathbf{u_0}\text{: } u_0(x) = \begin{cases} x & x \le a_0 \\ a_0 & x > a_0 \end{cases}, \tag{5.1.2}$$

serves as a reference or benchmark for what follows. Here for incomes $x \le a_0$ there is no tax, but for incomes $x > a_0$ the tax is $x - a_0$. In Section 3.1 we have shown that there exists a unique value $a_0$ such that $E(u_0(X)) = \mu_X - \tau \le a_0$ (with equality if and only if $F_X(a_0) = 0$), and that the after-tax Lorenz curve given in (3.1.12) is

$$\mathrm{L}_0(p) = \begin{cases} \dfrac{\mu_X}{\mu_X - \tau} L_X(p) & \mathrm{p} \le p_0 \\ \dfrac{\mu_X}{\mu_X - \tau} L_X(p_0) + \dfrac{a_0}{\mu_X - \tau}(p - p_0) & \mathrm{p} > p_0 \end{cases}, \tag{5.1.3}$$

where $p_0 = F_X(a_0)$. We have shown that the Lorenz curve (5.1.3) is the highest for the whole class of transformations (5.1.1). In Chapter 3 we stated that it

Lorenz dominates the initial income distribution and that irrespectively of the inclusion of the derivative restriction in (5.1.1) or not, the optimal policy is the same. For proofs of these and all subsequent mathematical assertions, see Chapter 3, Fellman (1995) and Fellman et al. (1996).

Although not all members of the class of policies under consideration are progressive, i.e. inequality reducing the policy $u_0(x)$ generates a post-tax income distribution that Lorenz dominates all tax policies of the given class ***U*** (Fellman, 1995, 2001; Fellman et al., 1996, 1999). Consequently, it also Lorenz dominate the flat tax policy $\hat{u}(x) = \frac{\mu_X - \tau}{\mu_X} x$, whose Lorenz curve is $L_X(p)$. Consequently, $L_0(p) \geq L_X(p)$ and $u_0(x)$ Lorenz dominates the initial income variable *X*.

Following the Atkinson (1970) theorem, (5.1.3) therefore implies maximal social welfare in this class, and $u_0(x)$ is in this sense optimal.

The generalized Gini coefficient of Yitzhaki (1983) for income after this optimal tax policy is

$$G(\nu) = 1 - \nu(1-\nu)\int_0^1 (1-p)^{\nu-2} L_0(p)dp, \tag{5.1.4}$$

which may be expressed in terms of the original Lorenz curve $L_X(p)$ using (5.1.3). According to the formula (3.1.21) a lower limit of this generalized Gini coefficient is $G_X(\nu) - \frac{\tau(1 - G_X(\nu))}{\mu_X - \tau}$. Here $\nu$ is a distributional judgement parameter, increases in which connote a more inequality-averse stance on the part of the social-decision maker. The case $\nu = 2$ is that of the ordinary Gini coefficient.

Now consider any actual (non-optimal) tax policy with mean tax $\tau$ and let $G_X(\nu)$ and $G_{X-T}(\nu)$ be the generalized Gini coefficients for pre- and post-tax income, respectively. Let furthermore, $G_0(\nu)$ be the generalized Gini coefficient for the optimal policy $u_0(x)$ in (5.1.2). Fellman et al. (1999) proposed to measure the effectiveness of this actual policy by the index:

$$I_T(\nu) = \frac{G_X(\nu) - G_{X-T}(\nu)}{G_X(\nu) - G_0(\nu)} \times 100 \tag{5.1.5}$$

which records its inequality-reducing performance as a percentage $I_T(\nu)$ of the maximum reduction that could have been achieved with the same tax yield τ. This is in contrast with some existing approaches to the measurement of redistributive effect, namely those of Musgrave and Thin (1948), Pechman and Okner (1974) and Blackorby and Donaldson (1984), which express actual inequality reduction as a percentage of pre-tax inequality and equality. In the first two cases cited, the Gini coefficient is used, and in the last the Atkinson index, to measure inequality. See Lambert (2001, Section 8.4) for more of this. Our index thus uses the optimal tax policy as a yardstick, whereas the others use the pre-tax distribution. In fact, the Pechman and Okner construction, if not the other two, use an implicit “optimal” yardstick, in essence comparing actual redistribution with that occurring if all income units were given the same post-tax income (i.e., it uses perfect redistribution, with zero net budget, as a reference). By confining attention to the class of tax policies which satisfy the government budget constraint to assess the effectiveness of an actual tax, or index has, as Fellman et al. (1999) stressed, more realism and direct appeal. It also incorporates the distributional judgement parameter ν which can be varied to carry out sensitivity analysis.

## 5.2 Optimal Transfer Policy

Consider the income $Y$ with distribution and density functions $F_Y(y)$, $f_Y(y)$, mean $\mu_Y$ and Lorenz curve $L_Y(p)$. In conformity with Chapter 4, we study a whole class of transfer policies characterized by a transformation $h(Y)$, where $h(\cdot)$ is non-negative, monotone increasing and continuous with the properties

$$\begin{cases} h(y) \geq y \\ E(h(Y)) = \mu_Y + \rho \end{cases}, \tag{5.2.1}$$

where $h(y)$ is the income, including cash transfer from government, associated with original income $y$. These properties indicate that no income decreases, that the internal order of the incomes remains the same and that all the policies raise the mean income to $\mu_Y + \rho$, where $\rho$ is the mean benefit, taken as given. The scenario pursued here can apply as well to an income policy: in that case $h(y)$ is the income after a policy-induced increase.

The polar case which serves as a reference or benchmark for what follows is:

$$h_0(y) = \begin{cases} b & y \leq b \\ y & y > b \end{cases} \tag{5.2.2}$$

i.e., all incomes below the level $b$ are raised to $b$ and all incomes above this level remain. It is shown in Section 4.1 that there exist a unique level $b_0$ such that

$$E(h_0(Y)) = \mu_Y + \rho$$

and for which the Lorenz curve for income including benefits according to (4.1.8):

$$L_0(p) = \begin{cases} \dfrac{b_0}{\mu_Y + \rho} p & p \le q_0 \\ \dfrac{b_0}{\mu_Y + \rho} q_0 + \dfrac{\mu_Y}{\mu_Y + \rho} (L_0(p) - L_0(q_0)) & p > q_0 \end{cases} \tag{5.2.3}$$

(where $q_0 = F_Y(b_0)$). Fellman et al. (1999) gave a slightly different layout of the Lorenz curve, but the relation $b_0\, q_0 - \mu_Y L_0(q_0) = \rho$ between the variables given in Section 4.1 proves the mathematical identity between the proposed formulae. The Lorenz curve $L_0(p)$ is the highest for the whole class of transformations defined by (5.2.1) and higher than $L_Y(p)$, thus engendering highest social welfare in the class (again not all policies in the class (5.2.1) are inequality reducing).

The generalized Gini coefficient for income after this optimal transfer policy is

$$G_0(\nu) = 1 - \nu(1-\nu)\int_0^1 (1-p)^{\nu-2} L_0(p)dp, \tag{5.2.4}$$

which may be expressed in terms of the original Lorenz curve by using (5.2.3).

As with taxes, the effectiveness of any actual (non-optimal) transfer policy with mean benefit ρ and pre- and post-benefit generalized Gini coefficients $G_Y(\nu)$ and $G_{Y+B}(\nu)$, respectively, may be measured in index form by

$$I_B(\nu) = \frac{G_Y(\nu) - G_{Y+B}(\nu)}{G_Y(\nu) - G_1(\nu)} \times 100. \tag{5.2.5}$$

Expressing the performance as a percentage of the maximum inequality reduction achievable for a given budget ρ. The index in (5.2.5) can also be used to assess the inequality-reducing performance of an incomes policy $h(Y)$ , measured against the optimal income policy $h_0(Y)$ for the same average increase $\rho$ in peoples incomes.

## 5.3 The Optimal Redistributive Tax-transfer Policy

We characterize each tax and transfer policy by the mean tax τ and the mean transfer $\rho$ where, we assume that $\rho \le \tau$. The transformation of the original incomes can be performed in two steps, first the taxation which reduces mean income from $\mu_X$ by an amount τ, and then the distribution of cash benefits so that the mean increases to $\mu_X - \tau + \rho$.

In this situation, the optimal tax and the optimal transfer policy of Sections 5.1 and 5.2 can be joined to given tax and transfer strategy. Under the assumption that both τ and ρ are taken as given, the joint strategy can be proved optimal, and actual combined tax and transfer programs can be gauged against it for their welfare. The rigorous assumption that both $\tau$ and $\rho$ are taken as given is necessary for the optimality. Under the weaker assumption that only the difference $\tau - \rho$ is taken as given, perfect redistribution will be attainable (Fellman et al., 1999).

Following Fellman et al. (1999), we start with the taxation. The optimal tax policy Lorenz dominates any other tax policy. Let $Y_0$ be post-tax income under the optimal tax policy and let $Y_u$ be post-tax income under an arbitrary tax policy. Assume that $E(Y_0) = E(Y_u)$ and denote as above the corresponding Lorenz curves $L_0(p)$ and $L_u(p)$. Under arbitrary taxation the poorest part of the population (after taxes) is poorer than under the optimal taxation (no taxes paid).

If, after taxation, we consider the benefit policy, then for an optimal income distribution, the optimal benefit policy must be performed. This means all benefits must go to the poor. Then the minimum income under the optimal taxation, $b_0$ (say), is greater than the minimum income under the arbitrary taxation, $b_u$. Consider the Lorenz curve after the benefit. Let the breaking points in (5.2.3) be $q_0$ and $q_u$, respectively. Obviously $q_0 \geq q_u$ . Hence, $L_0(p) \geq L_u(p)$ for $p < q_u$ and for $p > q_0$. For $q_u \leq p \leq q_0$ the curved part in $L_u(p)$ is convex and monotone and cannot intersect twice the linear part in $L_0(p)$ . Hence, $L_0(p) \geq L_u(p)$ for all $0 \leq p \leq 1$. Consequently, if we join the optimal tax policy and the optimal benefit policy, the joint policy is optimal.

As shown in Fellman et al. (1999) if $b_0 \geq a_0$ then $q_0 = 1$ and $b_0 = \mu_X - \tau + \rho$ in which case the optimal policy creates perfect equality. If, on the other hand, $b_0 < a_0$ then $q_0 \leq p_0$ and the final Lorenz curve $L_D$ is defined by:

$$L_D(p) = \begin{cases} \dfrac{b_0}{\mu_X - \tau + \rho} p & p \leq q_0 \\[2ex] \dfrac{b_0}{\mu_X - \tau + \rho} q_0 + \\ \dfrac{\mu_X}{\mu_X - \tau + \rho}\left(L_X(p) - L_X(q_0)\right) & q_0 < p \leq p_0 \\[2ex] \dfrac{b_0}{\mu_X - \tau + \rho} q_0 + \dfrac{\mu_X}{\mu_X - \tau + \rho}\left(L_X(p_0) - L_X(q_0)\right) + \\ \dfrac{a_0}{\mu_X - \tau + \rho}(p - p_0) & p > p_0 \end{cases} \qquad (5.3.1)$$

They derived, in a short and straightforward manner the result of Fei (1981). The class of combined tax-transfer policies in which (5.3.1) is optimal is Fei ́s class of "equity-oriented fiscal programs"; moreover (5.3.1) is Fei´s "two-valued program" shown to be optimal in his Theorem 7 (whose proof is complex and combinatorial). The analysis of Fellman et al. (1999) thus extends Fei ́s insight to the more general case of fiscal programs with a non-balanced budget, in which the mean excess tax revenue $\tau - \rho > 0$ can be devoted to publicly provided goods and services repayment of debt, etc. Fellman et al. (1999) showed that in the case of these more general fiscal programs, where the tax yield $\tau$ and benefit budget $\rho$ are both specified, the two-valued program with "floor value" $b_0$ and "ceiling value" $a_0$ (in Fei ́s terminology) is also optimal. In particular, the analysis extends Fei ́s Theorem 4, in which he shows (for $\tau = \rho$ ) that either $a_0 = b_0$ (the "maximal rational budget" engendering perfect equality) or $a_0 < b_0$. Fei also proves, in his Theorem 5, that $a_0$ is decreasing and $b_0$ is increasing, in the common value $\tau = \rho$; our own analysis proves that more generally, $a_0$ is decreasing in $\tau$ and $b_0$ is increasing in $\rho$ (by construction).

Finally using the generalised Gini coefficient $G_D(\nu)$ for income after the optimal tax and benefit, system namely

$$G_D(\nu) = 1 - \nu(\nu - 1)\int_0^1 (1 - p)^{\nu - 2} L_D(p)dp \tag{5.3.2}$$

which is determined by the original distribution $L_X(p)$ according to (5.3.1), the inequality-reducing performance of any actual (non-optimal) combined tax and benefit policy with mean tax $\tau$ and mean benefit $\rho$ can be assessed. Let

$$I_{T,B}(\nu) = \frac{G_X(\nu) - G_{X-T+B}(\nu)}{G_X(\nu) - G_D(\nu)} x100 \tag{5.3.3}$$

be the index, where $G_{X-T+B}(\nu)$ is the generalized Gini coefficient for disposable income after application of the actual tax and benefit policy and $G_D(\nu)$ is the generalized Gini coefficient for the optimal tax and benefit system**.**

In next section we use the index (5.3.3) and present the analysis of the combined effect of taxation and benefit rules in Finland, 1971-1990.

## 5.4 Empirical Illustration: Finland 1971-1990

Fellman et al. (1999) illustrated their methods using data from Finland from 1971 to 1990. The data used were drawn from the Household Budget Surveys (HBS) in Finland 1971, 1976, 1981, 1986 and 1990, a series of cross-sectional studies which are comparable over time. The income data in these surveys stem from tax and other administrative registers and can be considered to be of high quality. The sample size varies from 1296 in 1971 to 2897 in 1990. The taxation and benefit rules are the rules valid for the period 1971-1990. The sample is restricted to those households with positive disposable income. These data are also used in Example 2.3.1 in Section 2.3.

The base $x$ for taxes includes all taxable income, such as earnings self-employment income, capital income, work-related and taxable transfers and private transfers. From this we subtract direct taxes $t$ to get the base for all non-taxable benefits $b$. These was taken in this application to be the two major benefit schemes that have remained non-taxable throughout time period covered, namely child allowances and housing subsidies. During the period, 1971-1990, child allowances are paid to the households at a flat rate per each child under the age of 16 (17 in 1990). From the third child onwards, the sum per child increases. Housing subsidies have been means-tested throughout the time period and are therefore negatively correlated with the tax base.

The income variables were standardized to be comparable across households of different sizes using the OECD equivalence scale, which assigns the weight of 1.0, 0.7 and 0.5 equivalent adults to the first and additional adults and children, respectively. Household disposable income per equivalent adult is equal to $x - t + b$ (Fellman et al., 1999).

In Table 5.4.1 we show *inter alia* the effectiveness indices $I_T(\nu)$, $I_B(\nu)$ and $I_{T,B}(\nu)$ estimated by Fellman et al. (1999) from the data (along with some other statistics discussed below). Following Fellman et al. (1999), the threshold for the optimal tax was calculated by the following simple procedure. They fixed the threshold to be equal to the *i*th income unit ́s pre-tax income, $x(i)$ say, and collected all income above $x(i)$ of the income units that have higher income. If the total tax thus collected was higher than the actually collected amount, the threshold was set at $x(i-1)$. This procedure was then repeated until the tax threshold led to less taxes being collected when the threshold was set to $x(k)$. The optimal post-tax income is then $x(i)$ for $i < k$ and $x(k)$ for $i \geq k$. The benefit threshold and post-benefit income distribution were analogously estimated. The effectiveness of the actual tax system measured by our index, i.e., the inequality reduction of actual taxes relative to the optimal policy, declines from 1971 to 1981 and rises thereafter, thus having a slight U-shaped pattern over time. The inequality effectiveness of benefits declined between 1971 and 1990 – with exception of 1981. The combined effectiveness of taxes and transfers followed the same U-shaped pattern as that of taxes alone. For instance, using $\nu = 2$, in 1990 taxes achieved a 17.7% reduction in the Gini coefficient on moving from pre-tax to post-tax (but pre-benefit) income relative to the optimal tax policy. On moving from actual post-tax income to post-benefit income is reduced by 4.3% relative to the optimal benefits. On the other hand moving from pre-tax and pre-transfer income toi disposable income would achieve a 15.2 % reduction in equality relative to the optimal combined tax and transfer policy.

***Table 5.4.1*** *Redistributive effectiveness of taxes and benefits in Finland, 1971-1990, measured using generalized Gini coefficients.*

| | | **Taxes** | | | **Benefits** | | | **Taxes and benefits** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **Actual** | **Optimal** | **Maxi-mum** | **Actual** | **Optimal** | **Maxi-mum** | **Actual** | **Optimal** | **Maxi-mum** |
| ν | Year | $D_T(\nu)$ | $I_T(\nu)$ | $P_T(\nu)$ | $D_B(\nu)$ | $I_B(\nu)$ | $P_B(\nu)$ | $D_{T,B}(\nu)$ | $I_{T,B}(\nu)$ | $P_{T,B}(\nu)$ |
| | 1971 | 8.8 | 17.3 | 0.51 | 1.7 | 14.3 | 0.12 | 10.4 | 17.1 | 0.61 |
| | 1976 | 8.2 | 12.9 | 0.63 | 1.8 | 13.0 | 0.14 | 9.8 | 13.4 | 0.73 |
| 1.5 | 1981 | 7.0 | 11.2 | 0.63 | 3.1 | 17.9 | 0.18 | 9.9 | 13.1 | 0.76 |
| | 1985 | 9.9 | 15.0 | 0.66 | 2.6 | 11.8 | 0.22 | 12.2 | 14.8 | 0.83 |
| | 1990 | 12.5 | 17.5 | 0.71 | 1.4 | 8.0 | 0.17 | 13.7 | 16.3 | 0.84 |
| | 1971 | 7.8 | 18.3 | 0.43 | 1.5 | 10.5 | 0.14 | 9.2 | 16.8 | 0.55 |
| | 1976 | 7.8 | 14.0 | 0.56 | 1.5 | 10.0 | 0.15 | 9.3 | 13.5 | 0.69 |
| 2.0 | 1981 | 6.8 | 12.4 | 0.55 | 2.8 | 13.7 | 0.21 | 9.5 | 13.2 | 0.72 |
| | 1985 | 9.1 | 15.5 | 0.58 | 2.5 | 9.7 | 0.26 | 11.3 | 14.3 | 0.79 |
| | 1990 | 11.5 | 17.7 | 0.65 | 0.9 | 4.3 | 0.21 | 12.3 | 15.2 | 0.81 |
| | 1971 | 7.2 | 19.7 | 0.37 | 1.5 | 9.3 | 0.16 | 8.6 | 16.9 | 0.51 |
| | 1976 | 7.5 | 15.1 | 0.50 | 1.6 | 8.8 | 0.18 | 9.0 | 13.8 | 0.65 |
| 2.5 | 1981 | 7.1 | 14.3 | 0.49 | 2.6 | 11.3 | 0.23 | 9.5 | 13.8 | 0.69 |
| | 1985 | 8.6 | 16.2 | 0.53 | 2.7 | 9.2 | 0.29 | 11.0 | 14.3 | 0.77 |
| | 1990 | 11.0 | 18.6 | 0.59 | 0.7 | 3.1 | 0.23 | 11.7 | 14.9 | 0.78 |

*Source:* Fellman et al. (1999).
*Notes.* The reduction in equality $D$ is measured as the percentage decline in the generalized Gini coefficient due to actual taxes, benefits, or both. The optimal inequality reduction $I$ is measured as the actual decline in pre-tax, (transfer or tax and transfer) income inequality as a percentage of the optimal decline. See text, especially equations (5.1.5), (5.2.5) and (5.3.3) for exact definitions. The maximal decline $P$ is measured as a proportionate reduction that would occur if the optimal policy were implemented. These are related as $D = I \times P$ . Note that $D$ and $I$ are expressed as percentages whilst $P$ is a fraction. Differences between $D$ and $I \times P$ in the reported figures are due to rounding errors.

The inequality effectiveness of benefits is always smaller than that of taxes. This is unsurprising as the actual tax schedule in Finland is progressive during the period covered by the data. However, the main benefit studied, the child allowance, depends only on the number of children in the household. The optimal tax schedule thus only increases, rather than introduces, progressivity, whereas the optimal benefit policy would redistribute child allowances heavily to the lower tail, thus greatly increasing the inequality reduction of the actual benefits. The central argument for this is that the tax rate is related to the individual money incomes and not to the equivalent income calculated for the whole household.

The indices $I_T(\nu)$, $I_B(\nu)$ and $I_{T,B}(\nu)$ presented by Fellman et al. (1999) measure the effectiveness of tax and benefit policies relative to optimal yardsticks which are conditional on the budget sizes $\rho$ and $\tau$. Consider the Pechman and Okner (1974) indices

$$D_T = \frac{G_X - G_{X-T}}{G_X} x100 \qquad (5.4.1)$$

of inequality impact ( here for taxes). There is a simple relationship between our indices and those of Pechman and Okner (suitably generalized for $\nu \neq 2$ ). It is as follows

$$D_T(\nu) = I_T(\nu) P_T(\nu) \qquad (5.4.2)$$

$$D_B(\nu) = I_B(\nu) P_B(\nu) \qquad (5.4.3)$$

$$D_{T,B}(\nu) = I_{T,B}(\nu) P_{T,B}(\nu) . \qquad (5.4.4)$$

Where the terms

$$P_T(\nu) = \left(G_X(\nu) - G_0(\nu)\right) / G_X(\nu) ,$$

$$P_B(\nu) = \left(G_X(\nu) - G_1(\nu)\right) / G_X(\nu)$$

and

$$P_{T,B}(\nu) = \left(G_X(\nu) - G_D(\nu)\right) / G_X(\nu),$$

express in proportionate terms the maximum inequality reduction that could have been achieved with the given budget sizes (Fellman et al., 1999).

## 5.5 Concluding Remarks

Following Fellman et al. (1999) we have demonstrated the properties of optimal tax and benefit policies and shown how to gauge the effectiveness of actual (non-optimal) tax and benefit policies, as well as combined tax-benefit-systems, using the inequality impact of optimal policy as a yardstick. This has resulted in new indices for income taxes which contrast markedly with some existing indices of redistributive effect (progressivity), which either involve no optimal yardstick or at best a very unrealistic one. The new optimal yardstick is, of course, not fully realistic. It serves as a benchmark, just as, for example, the 45° line of perfect equality, though unattainable, is taken routinely as the yardstick against which to measure inequality using the Gini coefficient.

In the case of benefit systems, the indices lend themselves directly to another use: to measure the inequality performance of an incomes policy. Furthermore, all of the indices incorporate an inequality aversion parameter, and can be used to assess the contribution of “targeting” to observed inequality trends, along with that of budget size. Fellman et al. (1999) illustrated this by an application to Finnish data (and showed, incidentally, that the findings were quite robust to changes in the assumed inequality aversion of the evaluator).

Fellman et al. (1999) stressed that all of the constructed indices are impact measures, which take the pre-tax income distribution as exogenous to the choice of tax and benefit policies from classes which would have the given mean budget size ($\tau$ or $\rho$). With more sophisticated modelling, for example of people΄s preferences over consumption and leisure or, more ambitiously, in a computable general equilibrium environment, one could in principle devise indices of policy effectiveness with superior welfare properties – but these would not be measurable from published income data.

Another restrictive assumption of the mathematical modelling is that taxes and government transfers do not disturb the ranking of income units from poorest to richest by their living standards (equivalent incomes). Some lump-sum elements in the tax code (e.g. child allowances) can cause reranking in equivalent income terms, as can benefits going to people on the basis of factors outwith the equivalence scale (e.g. single mothers, the handicapped etc.). By using the Lorenz dominance criterion, Fellman et al. (1999) neglected any wider consideration of social needs.

## References

[1] Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*:244-263.

[2] Blackorby, C., Donaldson, D. (1984). Ethical social index numbers and the measurement of effective tax/benefit progressivity. *Canadian journal of Economics 17*:693-694.

[3] Fei, J. C. H. (1981). Equity oriented fiscal programs. *Econometrica 49*:869-881.

[4] Fellman, J. (1976). The effect of transformations on Lorenz curves. *Econometrica 44*, 823-824.

[5] Fellman, J. (1995). *Intrinsic Mathematical Properties of Classes of Income Redistributive Policies*. Swedish School of Economics and Business Administration Working Papers 306.

[6] Fellman, J. (2001). Mathematical properties of classes of income redistributive policies. *European Journal of Political Economy 17*:195-209.

[7] Fellman, J., Jäntti, M., Lambert, P. (1996). *Optimal Tax-transfer Systems and Redistributive Policy: The Finnish Experience*. Swedish School of Economics and Business Administration Working Papers 324.

[8] Fellman, J., Jäntti, M., Lambert, P. (1999). Optimal tax-transfer systems and redistributive policy. *Scandinavian Journal of Economics 101*:115-126.

[9] Lambert, P. J. (2001). *The Distribution and Redistribution of Income: A mathematical analysis*. (3rd edition) Manchester: Manchester University Press. xiv+313 pp.

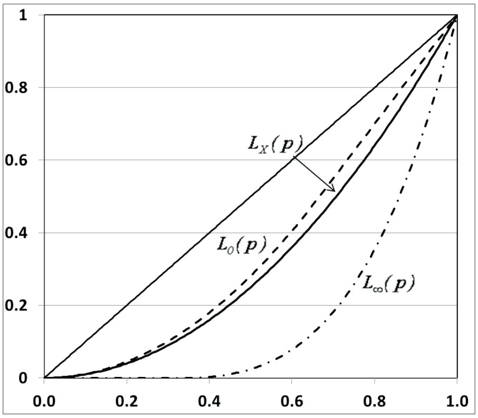
[10] Musgrave, R. A., Thin, T. (1948). Income tax progression, 1929-48. *Journal of Political Economy 56*:498-514.

[11] Pechman, J. A., Okner, B. (1974). Who Bears the Tax Burden? Brookings Institution, Washington DC.

[12] Yitzhaki, S. (1983). On an extension of the Gini index. *International Economic Review 24*:617-628.

Different skew models such as the lognormal and the Pareto have been proposed as suitable income distributions, but such specific distributions are usually applied in empirical investigations. For general studies more wide-ranging tools have been considered. The central and most commonly applied theory is connected to the Lorenz curve and the Gini coefficient. Without any assumptions concerning specific distributions, this theory enables analyses of temporal and regional variations in the income inequalities. Particularly, it is a valuable tool for studies of the effect of taxes and transfers to the redistribution of income. Taxation and transferring will be given individual presentations. In this study I have collected the central parts of my contributions to the theory of income distributions and furthermore, I have tried to locate my results within the framework of the general literature.

The region between the extreme Lorenz curves is the region of attainable Lorenz curves (Fellman, 2001, 2014).




To order additional copies of this book, please contact:
Science Publishing Group
book@sciencepublishinggroup.com
www.sciencepublishinggroup.com

ISBN 978-1-940366-25-8
9 781940 366258 >

Price: US $90