VATT-KESKUSTELUALOITTEITA

VATT-DISCUSSION PAPERS

130

STATISTICAL INFERENCE AND LORENZ CURVES: AN EMPIRICAL EXAMPLE USING FINNISH INCOME DATA

Saku Aura

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Helsinki 1996

ISBN 951-561-184-9

ISSN 0788-5016

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Hämeentie 3, 00530 Helsinki, Finland

J-Paino Ky

Helsinki, October 1996

Aura, Saku: Statistical Inference and Lorenz Curves: An empirical example using Finnish Income Data. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, 1996 (C, ISSN 0788-5016; No. 130), ISBN 951-561-184-9.

ABSTRACT: The main theme of this paper is to study statistical inference concerning empirical Lorenz curves and its generalisations. These curves are useful devices for examining welfare and inequality. Asymptotical sampling properties of the Lorenz curve are discussed. An improvement to a standard testing procedure of dominance relations is suggested. It is based on simulating the asymptotic distribution of the test statistic. The empirical part of this study is concerned with the Finnish distribution of disposable income in 1971, 1976, 1981, 1985 and 1990. It is found out that relative inequality has been highest on 1971, but the changes in relative inequality during period 1976-1990 have been almost negligible. Using a generalised Lorenz criterion it is also found out that welfare in income distribution has been rising in chronological order for the whole observation period.

KEY WORDS: Lorenz Curves, Income Distribution, Inequality, Welfare, Statistical Inference, Non-parametric Methods

Aura, Saku: Statistical Inference and Lorenz Curves: An empirical example using Finnish Income Data. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, 1996 (C, ISSN 0788-5016; No. 130), ISBN 951-561-184-9.

TIIVISTELMÄ: Tutkimuksen pääaiheena on eriarvoisuus- ja hyvinvointitarkasteluissa käytettyihin Lorenz-käyriin liittyvä tilastollisen päättelyn teoria. Empiiristen Lorenz-käyrien tärkeimmät tilastolliset ominaisuudet käsitellään sekä kehitetään simulaatioihin perustuva parannus tällä hetkellä standardiaseman saavuttaneeseen testausmenettelyyn. Tutkimuksen empiirisessä osassa menetelmiä sovelletaan Suomen käytettävissä olevien tulojen jakaumaan vuosilta 1971, 1976, 1981, 1985 ja 1990. Tärkein tässä osassa saatu tulos on, että suhteellinen eriarvoisuus on ollut vuonna 1971 suurimmillaan ja muutokset suhteellisessa eriarvoisuudessa vuosien 1976 ja 1990 välisenä aikana ovat olleet pieniä. Niin sanotun yleistetyn Lorenz-kriteerin mielessä on hyvinvointi tulojakauman valossa kasvanut tarkasteluperiodilla kronologisessa järjestyksessä.

ASIASANAT: Lorenz-käyrät, tulonjako, eriarvoisuus, hyvinvointiteoriat, tilastollinen päättely, ei-parametriset menetelmät

Contents

1 Introduction[†]

Lorenz curve and its modifications have been shown to be useful devices (see e.g. Atkinson 1970, Shorrocks 1983) in comparing inequality and welfare in different income distributions. They give intuitively appealing criteria that can be used in ranking distributions without unnecessary parametric assumption about the specific form of the inequality index or the social welfare function. Modifications of the Lorenz curve give various kinds of criteria which give different weights to equity (equal distribution of income) and efficiency (mean income). This paper shortly summarises the most important properties of the Lorenz curve and the generalised Lorenz curve and their relations to various normative ranking criteria. The theorems relating these curves to normative criteria are stated without proof with references to the relevant literature.

The main purpose of this paper is to examine the statistical properties of the Lorenz curve and the generalised Lorenz curve and to examine the problem of dominance testing with respect to these curves. A slight improvement to the standard testing procedure in the literature is suggested in the paper. The modification is based on the standard asymptotic test statistic, but instead of relying on the conservative tabulated values of its asymptotic significance level a Monte Carlo procedure for calculating a better asymptotic significance level is suggested.

The empirical example in this paper is concerned with disposable income of Finnish households in five cross-section data sets covering the period 1971-1990. In the empirical part it is found out that welfare (defined by generalised Lorenz dominance criterion) has risen in this period while the relative inequality (defined by Lorenz criterion) diminished remarkably from 1971 to 1976 and has been fairly constant since then.

[†]This paper is based on my Master's Thesis in Economics at the University of Helsinki. I would like to thank professor Yrjö Vartia of the University of Helsinki and Mr. Ilpo Suoniemi of the Government Institute for Economic Research for supervision, encouragement and advise in doing this research. I remain, however, solely responsible for any remaining errors in this paper. An earlier version of this paper was presented at public economics session of Taloustutkijoiden XIII Kesäseminaari, Jyväskylä.

2 Lorenz curves

The importance of Lorenz curves for the inequality measurement has been known since the seminal article by Atkinson (1970). Atkinson shows that if the means of two distributions[1] are equal, then every symmetric inequality averse social welfare function prefers the distribution that Lorenz dominates[2] the other. If no Lorenz dominance relation can be found (that is when Lorenz curves intersect) one cannot decide the ordering between the two distributions without making further assumptions on the form of the welfare function.[3] The most important property of ordinary Lorenz curves is that every index of relative inequality has a larger value in the distribution that is Lorenz dominated. Therefore, even if the means of the income distributions are different one can use Lorenz dominance criteria to compare relative inequality between distributions.

Definition 1 *The Lorenz Curve of a positive random variable X is*

$$L(p, X) = \int_0^p X(P)dP/\mu_X, \tag{1}$$

where $X(p) = inf\{x|F(x) \geq p\}$ *is the inverse of the distribution function of X,* μ_X *is the mean of X and* $p \in [0, 1]$.

The properties of ordinary Lorenz curve can be found in numerous sources (e.g. Aura 1996 and Lambert 1989). Verbal interpretation of the Lorenz curve is "the value of Lorenz curve in point p tells the proportion of total income received by the poorest p percent of population". An important modification of the ordinary Lorenz curve is the generalised Lorenz curve by Shorrocks (1983).

Definition 2 *The generalised Lorenz GL(p) is the Lorenz curve scaled by the mean of the distribution. The generalised Lorenz curve of a positive random variable X is*

$$GL(p; X) = \int_0^p X(P)dP = \mu_x L(p; X). \tag{2}$$

[1]When comparing discrete populations the size of the population might be considered as a nuisance factor. Usually this problem is overcame by using Sen's (1979) Population Principle which states that the comparisons should be invariant with respect to population replication.

[2]Distribution X is said to Lorenz dominate distribution Y if the Lorenz curve of X is at or above the Lorenz curve of Y on every point in [0, 1].

[3]This does not necessarily mean that one has to specify parametric form of the welfare function. Less restrictive "non-parametric" orderings than the one implied by Lorenz dominance can also be found e.g. by using the aversion to downside inequality (a.k.a. principle of diminishing transfer, third order stochastic dominance) as suggested by Davies and Hoy (1995).

Shorrocks proves that every strictly increasing (Paretian) and inequality averse[4] social welfare function prefers the distribution that dominates another in the terms of the generalised Lorenz curve (a.k.a. second order stochastic dominance). In addition, another modification of Lorenz curve has been discussed in the literature: the absolute Lorenz curve. However, it will not be discussed here and interested reader may consult other sources for the properties and normative implications of the absolute Lorenz curve (e.g. Aura 1996, Bishop et al. 1989 and Shorrocks 1983).

All the welfare and inequality implications of the Lorenz curves can be expressed in terms of dominance conditions. If a distribution dominates another distribution w.r.t. some Lorenz criterion, then it follows that the corresponding dominance relation is true w.r.t. welfare or equality. In addition to the ordinary Lorenz and generalised Lorenz dominance criterion, other criteria giving alternative welfare or equality conditions can be found. These include joint mean-Lorenz dominance, joint mean-absolute Lorenz dominance and first order (rank) dominance criterion in addition to the absolute Lorenz dominance and the third order dominance relations briefly discussed above. Normative implication of the dominance criteria can be found in Aura (1996), [5] Bishop et al. (1989) and Shorrocks (1983).

Concentration curves are also closely related to the Lorenz curves. Concentration curves are of interest if the problem is to evaluate how different income sources contribute to the total Lorenz curve. Concentration curves come like Lorenz curves in ordinary, generalised and absolute flavours, but the implications of these curves are not discussed here. Good references on concentration curves are Yitzhaki & Slemrod (1991), Lambert (1989) and Bishop et al. (1994).

3 Asymptotic theory for empirical Lorenz curves

Beach & Davidson (1983) were among the first to thoroughly examine the statistical aspects of empirical Lorenz curves without making any parametric assumptions on form of the underlying distributions in the economic literature. They proved that asymptotic inference for Lorenz curves is based on quantities that can easily be estimated non-parametrically. The results by Beach & Davidson cover the case where: 1) underlying distribution is continuous,

[4]Inequality aversion is defined in the discrete case by Schur-concavity of the welfare function (see Shorrocks 1983 and Marshal & Olkin 1979).

[5]Aura (1996) also provides an empirical illustration of the use of these other normative criteria in ranking the same data as in this paper.

2) the data is an i.i.d. sample from the underlying distribution and 3) two first moments of the distribution are finite. Beach & Kalinski (1986) extended the analysis to the case where data is an independently but not identically distributed random sample. This is important for the use of these results to the micro level data where the income data is usually accompanied by sampling weights to account for sampling framework and to correct for non-response.

Let $x_1, x_2, x_3, \ldots, x_n$ be an i.i.d. sample from the underlying income distribution and let $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n)}$ be the ordered sample ($x_{(i)} \leq x_{(i+1)}$). Let G be the the k-dimensional vector of generalised Lorenz ordinates at points p=$[p_1, p_2, p_3, \ldots, p_{k-2}, p_{k-1}, 1]$ so that $G_k = \mu$. The empirical generalised Lorenz ordinate in point p_j is [6]

$$\hat{G}_j = \frac{1}{n} \sum_{i=1}^{[p_j n]} x_{(i)}, \tag{3}$$

where $[p_j n]$ is the greatest integer not larger than $p_j n$. It can be shown (Appendix A) that $\hat{G}_j$ is an asymptotically unbiased and normally distributed estimator for $GL(p_j)$. Expressions for the asymptotic variance/covariance structure of the asymptotically multinormally vector estimator $\hat{G}$ are given in the Appendix A. Furthermore, it can be shown (Appendix A) that

$$\hat{L}_j = \frac{\hat{G}_j}{\hat{\mu}}, \tag{4}$$

where $\hat{\mu}$ is the sample average, is asymptoticaly unbiased and normally distributed estimator for $L(p_j)$ and that vector estimator $\hat{L}$ is asymptotically multinormal.

Naturally the non-parametric approach to statistical inference for Lorenz curves presented above is not the sole possible approach for evaluating random aspects of sample Lorenz ordinates. There exists a vast literature on possible functional forms of the Lorenz curve and because of the one-to-one relation between the generalised Lorenz curve and the distribution function, the literature on parametrically specifying the frequency distributions for income distributions is also relevant here. However, the parametric approach is not discussed here because the author feels that finding a proper parametric form to summarise the income distribution is not an easy task and an erroneous choice of functional form could leave one astray. Nevertheless it should be noted that in a recent article Cowell and Victoria-Feser (1996) show that if there is a possibility of corrupted data (e.g. outliers, recording errors),

[6] This is the truncated version of the empirical generalised Lorenz curve. In practice one uses a interpolated version of the empirical generalised Lorenz curve $\hat{G}_j = \frac{1}{n}\sum_{i=1}^{[p_j n]} x_{(i)} + \frac{p_j n - [p_j n]}{n} x_{([p_j n]+1)}$. These two are asymptotically equivalent so the simpler truncated form is discussed here for convenience. Estimation results presented in this paper are based on the interpolated form. If the data is weighted, then a corresponding formula for the interpolated estimator is needed.

the use of parametric inference with respect to inequality measures can be of advantage. They show that many inequality measures can be estimated using robust parametric estimation methods while the parametric maximum likelihood and non-parametric estimation methods produce non-robust estimators. The lesson to be learnt from these results is that the sensitivity to outliers and the possibility of data corruption are important issues when examining inequality and welfare.[7]

4 Testing dominance relations

The asymptotic statistical theory for Lorenz curves presented above provides only a partial answer to the problem of statistical testing of dominance relations. The problem arises from the nature of Lorenz curves, they are continuous mappings from interval $[0,1]$. The theory presented here is a theory for point-wise inference. Hence before applying the results presented here one must specify a discrete "grid" that is dense enough for being reasonably good approximation for a continuous function. This means specifying a set of points on the interval $[0,1]$ that form a basis of comparison between two Lorenz curves. Typically this means examining the Lorenz curves at decile or vintile points. The author is unaware of any "Kolmogorov-Smirnov type" testing procedures to overcome this limitation.[8]

Testing of the equality of two Lorenz curves can be easily based on the asymptotic results presented above. Ordinary χ^2-test is appropriate for this purpose. The problem with the χ^2-test is that the results do not identify dominance relations. Inequality of the Lorenz ordinates is necessary for the dominance relations but it is not sufficient. In order to get a test that identifies dominance relations one has to give up some power of the test.

Testing of the dominance relations is usually based on the simultaneous inference approach. This methodology was first developed in the context of variance analysis. Let $\hat{G}^1$ and $\hat{G}^2$

[7]The robustness of the results of the empirical part to the possible outliers at very high or low incomes was checked using an ad-hoc procedure: all calculations were done also with bottom and top 1 % of incomes truncated from the sample. Most of the qualitative conclusions remained unchanged with this check: notably all conclusions concerning generalised Lorenz dominance were unchanged and while there were some minor changes between crossing and dominance relations regarding ordinary Lorenz curves the major conclusions of the paper remained unchanged.

[8]Empirical results of this paper are quite robust to the choice of the grid. All of the calculations in the empirical part were also done with evaluating the dominance at every percentage point (total 100 points), but all the conclusions remained the same. It can be easily seen from the formula for asymptotic covariance of the estimators that nearby points on empirical generalised Lorenz curve are highly correlated so that adding extra points to analysis beyond a certain degree does not add information into analysis.

be the vector estimators of generalised Lorenz ordinates in two different populations.[9] Let $\hat{d} = \hat{G}^1 - \hat{G}^2$ be the vector of estimated differences in generalised Lorenz ordinates. Under null hypothesis $d = 0$ the two distributions are equivalent. Using the asymptotic theory one can obtain expressions for the asymptotic variance/covariance structure of the elements of $\hat{d}$ and find that $\hat{d}$ is asymptotically unbiased and multinormal estimator for d. For distribution 1 to strongly dominate distribution 2 it must be that $d_i \geq 0$ for all i and $d_i > 0$ at least for one i. The opposite case would mean that distribution 2 dominates. If there exists a pair i, j so that $d_i > 0$ and $d_j < 0$ then the generalised Lorenz curves intersect and the distributions are non-comparable in the generalised Lorenz sense.

The problem of dominance testing is to separate those differences that are not significantly different from zero from those that are. This a problem of simultaneous inference. Usually it is solved by using the Studentized Maximum Modulus (SMM)-distribution. It gives slightly smaller critical values for the test statistic than the use of more familiar Bonferroni inequality. SMM(k, df) is the distribution of maximum of the absolute standardised independent t_{df}-variates in a k-dimensional random sample. It can be shown that regardless of the covariance matrix between test statistics, the SMM-distribution gives an upper bound to the critical value of test statistics (Stoline & Ury, 1979).

In order to test the null hypothesis that two generalised Lorenz curves are equal against three other possible outcomes (the first one dominates the second; the second dominates the first or intersection) one first calculates standardised test statistics [10]

$$t_i = \frac{d_i}{\sqrt{\hat{\sigma}^1_{ii}/n_1 + \hat{\sigma}^2_{ii}/n_2}} \tag{5}$$

in every point of the grid (e.g. in every decile point). Test of stochastic dominance can be obtained by basing the inference on two test statistic:

$$t_+ = \max\{0, t_i\} \tag{6}$$

and

$$t_- = \min\{0, t_i\}. \tag{7}$$

If only t_+ is statistically significant then the first distribution dominates. On the other hand, if only t_- is statistically significant then second distribution dominates and finally if both

[9] The same theory applies to the problem of testing Lorenz-ordinates so that in testing of the ordinary Lorenz dominance one may substitute L for G.

[10] The author apologises for a slight abuse of standard notation. $\hat{\sigma}^1_{ii}$ means the i:th diagonal element of the estimated asymptotic covariance matrix for the first sample.

t_+ and t_- are statistically significant then the generalised Lorenz curves intersect. Standard procedure to evaluate the significance of t_+ and t_- is to compare them with SMM(k, ∞) critical values. This approach does not use the information of the covariance between the test statistics and therefore gives an overly conservative asymptotic test.

In the paper a Monte Carlo procedure for calculating the significance level of observed t_+ and t_- is suggested. Test statistics for Lorenz curves tend to be highly correlated and this can be used to "reduce the dimensionality" of the test statistic vector. Simulation of the significance level is done by generating random variables under the null hypothesis. The consistent estimator of the covariance matrix is used in the simulation. The simulation procedure is repeated a large number (call this n) of times and the number of observations when maximum standardised absolute values are greater than t_+ and t_- are counted (call these i_+ and i_- respectively). Simulated significance levels are then given by i_+/n and i_-/n.

The adequate number of simulations was determined to be 100,000. This was done by considering the simulation as random sampling from binomial distribution where the parameter p (=true level of significance) is unknown. It can be easily determined by Bayesian approach using a non-informative prior (beta(0.5,0.5)) that the observed significance level is precise enough even near $p = 0.01$ when the number of simulations is 100,000. For details see Aura (1996).

Table 1 Critical values of SMM-distribution when $k = 19, 20$ and the range of simulated critical values when $\alpha = 0.10, 0.05, 0.01, 0.001$.

	Significance level 0.10	Significance level 0.05	Significance level 0.01	Significance level 0.001
SMM k=19	2.774	3.004	3.466	4.044
SMM k=20	2.791	3.016	3.479	4.056
Simulations	[2.173,2.224]	[2.466,2.511]	[3.037,3.069]	[3.700,3.800]

N.B. The simulations were not done when the absolute value of the test statistic exceeded four (approximately the 0.001 critical level of SMM-distribution). The figures in brackets give the range of simulated critical values in the comparisons across Finnish income data.

From Table 1 one can see that taking into account the correlation between test statistics enhances the power of the test. The difference between critical values from SMM-distribution and simulations are sometimes as high as 0.5 in our data. This means that the outcome of the test can be quite different when one uses simulated significance levels instead of the overly conservative tabulated significance levels.

5 Empirical example using Finnish Income Data

This section illustrates how this methodology can be used to examine two questions about Finnish income distribution. The first one is 'how has inequality developed between 1971 and 1990' and the second is 'how has welfare developed between these years'. Inequality is examined by Lorenz dominance criterion and welfare by generalised Lorenz dominance criterion.

The data consist of Finnish Household Survey data in 1971, 1976, 1981, 1985 and 1990. Disposable income is studied here. This includes factor income and net transfers. The household disposable income is adjusted by the OECD-equivalence scale,[11] which gives weight of one for the first adult, 0.7 for each subsequent adults and 0.5 for each child. This process gives the 'household member equivalence scale adjusted income'. This is weighted by the number of members in household so that e.g. a household of 2 adults and one child with disposable income of FIM 22,000 would be treated as three separate individuals with incomes of FIM 10,000. The justification for this procedure is to take into account the economies of scale of people living in larger households and still be able to study welfare between individuals and not between households. From the economic theory point of view this procedure would be correct only if the OECD-scale is the true equivalence scale and that the division of income is done within each household so that every member of the household is on the same cardinal utility level (whatever that means).

Table 2 The summary statistics of the data. The figures are given in terms of 1990 currency.

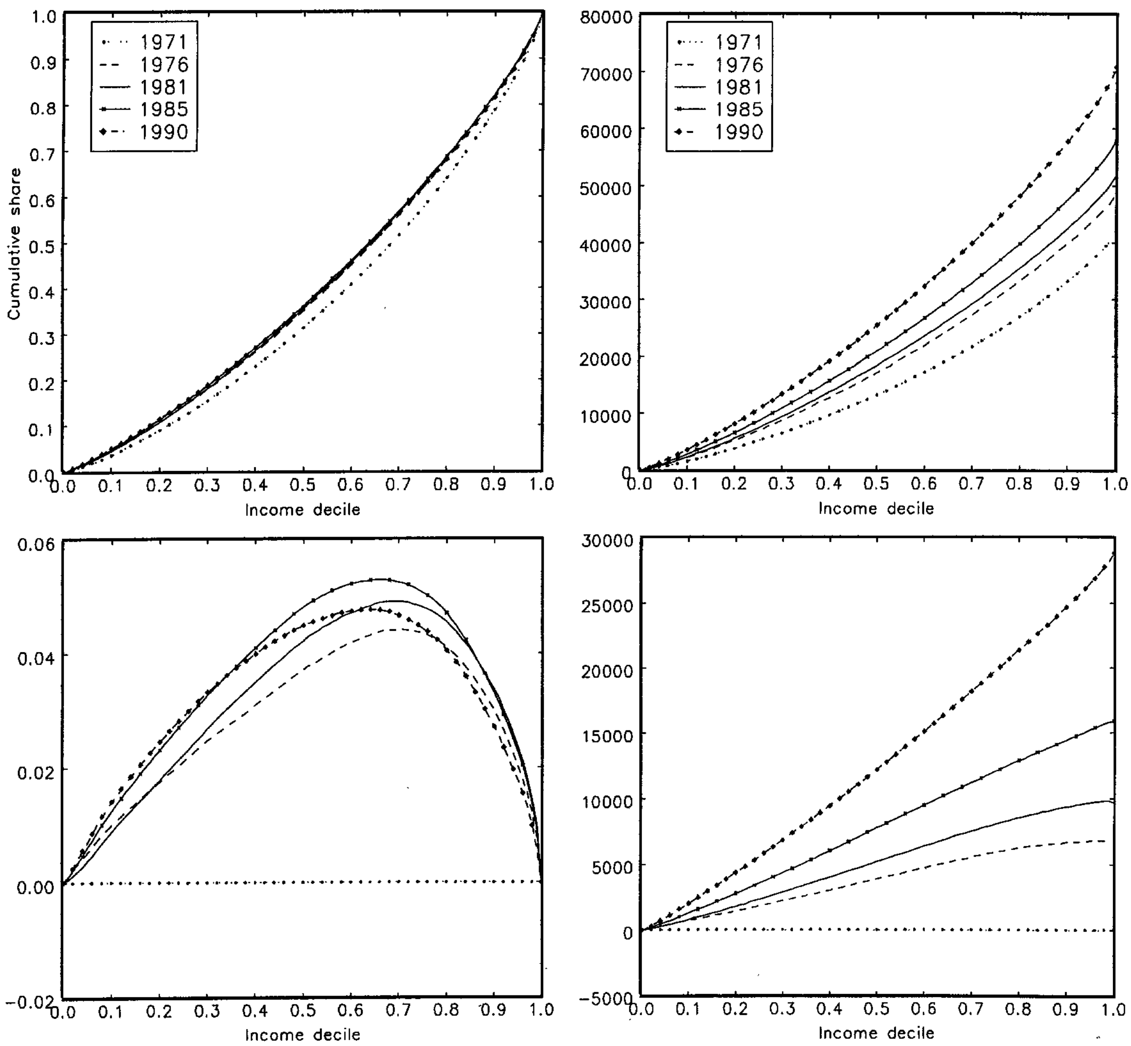
Year	N	Mean Income	Standard Deviation	Minimum Income	Maximum Income	Gini Coefficient
1971	2986	42060.14	22306.01	151.10	290112.44	0.2698
1976	3348	48672.23	19681.97	511.34	247686.77	0.2130
1981	7386	51717.42	20002.49	416.31	282010.34	0.2071
1985	8200	57928.45	21941.42	115.36	229472.82	0.1998
1990	8253	70872.11	31042.99	1671.18	1246873.70	0.2050

Five households from 1990 were excluded from the sample because their disposable income was negative. All the calculations in this paper were done using sampling weights in order to account for sampling framework and to correct for non-response. The weighting scheme is described in Laaksonen (1988).

[11]The choice of equivalence scale can be critical for welfare analysis. The above scale is widely used but its relative merits with respect to other equivalence scales are not clear. The use of equivalence scales in welfare analysis is thoroughly discussed in Coulter et al. (1992).

5.1 Results from Lorenz curves

Figure 1 Top panel: i) Lorenz curves and ii) Generalised Lorenz curves for 1971, 1976, 1981, 1985 and 1990. Bottom panel: iii) difference of Lorenz curves from the Lorenz curve of 1971 and iv) difference of generalised Lorenz curves from the generalised Lorenz curve of 1971.



From the upper left panel of Figure 1 one can see that Lorenz curves from 1976 to 1990 are almost indistinguishable. The Lorenz curve of 1971 is below every other curve, so leaving aside the statistical aspect one could conclude that the relative inequality has been highest in 1971 and that the differences in the period 1976 to 1990 are small.

The bottom left panel of Figure 1 displays the differences of Lorenz curves from the Lorenz curve 1971. It can be seen that the curve of 1985 is above the curve of 1976 in every point. No other pure dominance relation can be seen after 1976 without using statistical tests.

The statistical analysis changes the outcome drastically. If one uses significance level of 0.01

Table 3 Pairwise statistical comparisons of Lorenz curves.

Year	Year	t_+	p-value	t_-	p-value	χ^2	p-value
1990	1985	2.369	0.0670	-2.585	0.0384	44.01	< 0.001
1990	1981	7.036	< 0.01	-2.516	0.0459	96.82	< 0.001
1990	1976	5.427	< 0.01	-0.847	0.8155	86.86	< 0.001
1990	1971	16.448	< 0.01	-	-	351.11	< 0.001
1985	1981	4.928	< 0.01	-0.500	0.9718	47.38	< 0.001
1985	1976	4.947	< 0.01	-	-	60.86	< 0.001
1985	1971	16.891	< 0.01	-	-	367.07	< 0.001
1981	1976	2.162	0.1142	-2.527	0.0479	41.35	0.0022
1981	1971	14.576	< 0.01	-	-	271.17	< 0.001
1976	1971	11.410	< 0.01	-	-	201.17	< 0.001

N.B. The t_+ statistic tests the hypothesis that the Lorenz curve of the year in the first column is above the Lorenz curve of the year in the second column in at least one point. When t_+ or t_- is equal to zero, it is marked as - in the table. When the calculated p-value is less than 0.01 it is marked by < 0.01. The χ^2-statistic is ordinary χ^2-test for the equality of two Lorenz curves. Its significance levels can be compared with significance levels of the dominance test if one wishes to determine how much of the power of the test has to be given up in order to be able to identify dominance hypothesis.

as a criterion of statistical significance then the dominance conditions is satisfied in 8 of the 10 comparisons. The Lorenz curves of 1990 and 1985 are equivalent using this criterion and so are also the Lorenz curves of 1981 and 1976. All other comparisons with this significance level indicate that relative inequality has been diminishing in chronological order. Note also that even if the dominance test with significance level of 0.01 does not reject the null hypothesis in the two comparisons, the χ^2-statistic does. Therefore the conclusion is that we know that there is some evidence that Lorenz curves are not equivalent, but we do not know how to interpret the evidence.[12]

If one employs significance levels higher than 0.01 (either 0.05 or 0.10) the results in the comparisons change in three cases (Table 4). Hence with 10 pairs to compare, we have 7 unambiguous dominance relations in contrast with only 5 found using non-statistical criteria. The three pairs that are left unranked are worth further examination.

The pair 1990 and 1985 is left unranked if unanimous ranking by the three different significance level is used as a dominance criterion. It is difficult to arrive in any clear-cut conclusion on the change in relative inequality between this pair. Using 0.01 as the significance level these two years are Lorenz equivalent, but with 0.05 as significance level 1985 dominates and with 0.10 as significance level the Lorenz curves intersect. The problem how to make some

[12]The power of the simulation test would merit closer attention. The analysis could be extended in a general decision theoretic framework where one would use some natural loss functions for type I and type II errors.

Table 4 Outcomes of the statistical analysis for Lorenz curves using different significance levels α.

Year	Year	Eye ball comparison	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
1990	1985	X_L	X_L	$<_L$	$\sim_L$
1990	1981	X_L	X_L	X_L	$>_L$
1990	1976	X_L	$>_L$	$>_L$	$>_L$
1990	1971	$>_L$	$>_L$	$>_L$	$>_L$
1985	1981	X_L	$>_L$	$>_L$	$>_L$
1985	1976	$>_L$	$>_L$	$>_L$	$>_L$
1985	1971	$>_L$	$>_L$	$>_L$	$>_L$
1981	1976	X_L	$<_L$	$<_L$	$\sim_L$
1981	1971	$>_L$	$>_L$	$>_L$	$>_L$
1976	1971	$>_L$	$>_L$	$>_L$	$>_L$

N.B. Symbol $>_L$ means that the first year Lorenz dominates the second, $<_L$ means that second year dominates, $\sim_L$ means that there is no statistically significant difference and X_L means that the curves intersect. The eye ball comparison column refers to the non-statistical Lorenz comparison, where the sample Lorenz curves are assumed to correspond exactly to the underlying population Lorenz curves.

conclusion based on this kind of evidence is left unsolved. The author would argue that no gross error is made if it is stated that these two years have Lorenz curves that are, if not completely equivalent, very close to each other. The author feels that not too much weight should be given to the fact that the Gini coefficient for 1985 is smaller than for 1990, since the difference is so small that it is probably statistically insignificant. Furthermore, it seems to be that 1990 dominates 1985 in the poorest part of the income distribution. So the jury for these years is still out and near-equivalence is the conclusion to be reached with this pair.

With 1990 and 1981 one can reach a more decisive conclusion than with 1990 and 1985. It is statistically beyond reasonable doubt that 1990 dominates 1981 until the 0.35 fractile of the income distribution and if 1981 dominates 1990 in some part of the distribution it occurs after the 0.65 fractile. Therefore if one is more interested in the relative position of the poor and lower middle class as opposed to the upper middle class and near-rich then one would conclude that 1990 has less relative inequality than 1981.

The pair 1981 and 1976 is also troublesome. The result of near-equivalence of these two Lorenz curves can be reached most easily if one considers the maximum distance between these two curves. The maximum of difference between 1981 Lorenz ordinates to 1976 Lorenz ordinates is 0.544 percentage points and the minimum is -0.157 percentage points. Even though the negative difference is significant at 0.05 level the possible difference is so small that one might reasonably argue that it should be left unnoticed.

To compare the statistical analysis with common sense a non-statistical comparison using a rule of thumb was also done. The rule was adapted from Atkinson et al. (1995, 87), which was published by OECD, and it states that[13] "if maximum absolute difference between two Lorenz curves is less than one percentage point, then there is no difference." Using this rule the 1971 is dominated by every other year in this study and all the other years are equivalent. This leads us to the conclusion on the changes in relative inequality. If we accept the income definitions and equivalence scale[14] as the relevant ones for welfare and inequality analysis we can conclude that the difference between 1971 and other years in relative inequality are large and the differences in period 1976 to 1990 are small in absolute value and in comparison to the differences with 1971.

5.2 Results from generalised Lorenz curves

Generalised Lorenz curves give stronger results than ordinary Lorenz curves with the present data. Even without considering the statistical significance of possible crossings one can conclude that 9 out of 10 comparisons give a clear dominance results. The only possible crossing is between 1981 and 1976. In all other comparisons one would conclude that a later year dominates the previous one in generalised Lorenz sense.

Year 1976 would seem to dominate 1981 in the first vintile. However, the difference is not statistically significant if any reasonable significance level is employed. The conclusion of the test is that 1981 dominates 1976. This would mean that welfare (as defined by preference of every Paretian inequality averse social welfare function) has increased in chronological order in Finnish income distribution during the observation period. This is natural if one examines the change in the mean income and the change in relative inequality. The relative inequality (as defined by Lorenz criterion) has been fairly constant in period 1976-1990 and the mean income has steadily increased and has almost doubled between 1971-1990. It can be concluded that the changes in the mean income have dominated the change in relative inequality in overall contributions to the change in generalised Lorenz curve.

[13] One possible justification for using this rule is that the systematic (non-random) error (see e.g. Cowell and Victoria-Feser (1996)), which is hard to take into account, can be large in comparison to the (random) sampling error in real world income data.

[14] All the calculations were also done in per capita and per household terms. Per capita calculations give qualitatively the same conclusions and per household results are very different in many cases. Even though "true equivalence scale" is an unknown concept and is likely to remain so the author argues that per capita and per OECD-unit calculations are more relevant than per household calculations. Per household calculations treat a family of two adults and five children equivalent as a single adult, which is quite absurd in welfare terms.

Table 5 Pairwise statistical analysis for the generalised Lorenz curves.

Year	Year	t_+	p-value	t_-	p-value	χ^2	p-value
1990	1985	34.399	< 0.01	-	-	1441.04	< 0.001
1990	1981	51.677	< 0.01	-	-	3074.17	< 0.001
1990	1976	51.422	< 0.01	-	-	3745.40	< 0.001
1990	1971	68.556	< 0.01	-	-	7841.10	< 0.001
1985	1981	19.180	< 0.01	-	-	536.75	< 0.001
1985	1976	23.968	< 0.01	-	-	1013.52	< 0.001
1985	1971	45.429	< 0.01	-	-	4313.03	< 0.001
1981	1976	8.328	< 0.01	-0.6721	0.8898	188.12	< 0.001
1981	1971	30.303	< 0.01	-	-	1877.49	< 0.001
1976	1971	19.622	< 0.01	-	-	711.17	< 0.001

Table 6 Outcomes of the statistical analysis for generalised Lorenz curves using different significance levels α.

Year	Year	Eye ball comparison	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
1990	1985	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1990	1981	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1990	1976	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1990	1971	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1985	1981	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1985	1976	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1985	1971	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1981	1976	X_{GL}	$>_{GL}$	$>_{GL}$	$>_{GL}$
1981	1971	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$
1976	1971	$>_{GL}$	$>_{GL}$	$>_{GL}$	$>_{GL}$

In Aura (1996) it is found out that the results presented here for welfare (generalised Lorenz) dominance are not restricted to the class of Paretian inequality averse social welfare functions. The dominance with respect to all symmetric Paretian welfare functions requires that the distribution rank dominates (that is $x_{(i)} \geq y_{(i)}$ for all i) the other distribution. Using the dominance of vintile group means (see Bishop et al. 1989) as an empirical counterpart of rank dominance it is found in Aura (1996) that the statistical chronological dominance also holds for this criterion.

Furthermore, it may be noted that even if generalised Lorenz dominance can rank distributions with respect to welfare it does not render inequality comparisons unnecessary. Ordinary Lorenz comparison are well suited for a time series of inequality data and cross country comparisons because the problem of converting all income distributions to a common currency does not affect them. Furthermore, the joint mean-Lorenz dominance criterion can be used as a welfare criterion that gives more relative weight to equality considerations than the generalised Lorenz criterion.

6 Conclusions

In the empirical part of the paper it is found that changes in inequality in the distribution of Finnish disposable income have been small between 1976-1990 in comparison to the changes between 1971 and 1976. Welfare as defined by preference for a large class of social welfare functions has been increasing during the whole period. Methodologically the paper is in line with other studies (e.g. Bishop et al. 1989) which indicate that using statistical dominance criteria (that is, statistically analyse the sampling variability of observed Lorenz curves) allows the researcher to draw more powerful conclusions about welfare and inequality.

A slight modification of the standard testing procedure has been suggested in this paper. It is found that using simulations to evaluate the significance level of observed test statistic produces a slightly more powerful asymptotic test than basing test on the overly conservative tabulated asymptotic critical values. The small sample comparison of the two testing procedures is a possible direction for further study. Analysis of asymptotical sampling properties of concentration curves, which are closely related to Lorenz curves, present another possible line for further study.

For empirical research the results of this paper are encouraging. Using various Lorenz criteria and statistical inference in analysing micro level income data makes it possible to draw strong conclusions about welfare and inequality in income distribution. Possible directions for future empirical studies present extension of the analysis to deal with different income concepts and analysis of the effects of taxes and transfers to the final income distribution. This analysis could be done with methodology close to the one discussed here, using Lorenz and concentration curves.

7 References

ATKINSON, A. B. (1971): On the Measurement of Inequality. Journal of Economic Theory 2, 244-263.

ATKINSON, A. B. - RAINWATER, L. - SMEEDING, T. M. (1995): Income Distribution in OECD Countries. OECD Social Policy Studies No. 18, Paris.

AURA, S. P. (1996): Lorenz Curves, Welfare and Statistical Inference (in Finnish). Government Institute for Economic Research, Research report 35, Helsinki.

BEACH, C. M. - DAVIDSON, R. (1983): Distribution-Free Statistical Inference with Lorenz Curves and Income Shares. Review of Economic Studies 4:163, 723-735.

BEACH, C. M. - KALINSKI, S. F. (1986): Lorenz Curve Inference with Sample Weights: an Application to the Distribution of Unemployment Experience. Applied Statistics 35:1, 38-45.

BEACH, C. M. - RICHMOND, J. (1985): Joint Confidence Intervals for Income Shares and Lorenz Curves. International Economic Review 26:2, 439-450.

BISHOP, J. A. - CHOW, K. V. - FORMBY, J. P. (1994): Testing for Marginal Changes in Income Distribution with Lorenz and Concentration Curves. International Economic Review 35:2, 479-488.

BISHOP, J. A. - FORMBY, J. P. - THISTLE, P. D. (1989): Statistical Inference, Income Distributions, and Social Welfare. S. 49-81. In SLOTTJE, D. J. (ed.): Research on Economic Inequality. JAI Press, London.

COULTER, F. A. E. - COWELL, F. A. - JENKINS, S. P. (1992): Differences in Needs and Assessment of Income Distributions. Bulletin of Economic Research 44:2, 77-124.

COWELL, F. A. - VICTORIA-FESER, M.-P. (1996) Robustness Properties of Inequality Measures. Econometrica 64:1, 77-101.

DAVIES, J. - HOY, M. (1995): Making Inequality Comparison When Lorenz Curves Intersect. American Economic Review 85:5, 980-986.

LAAKSONEN, S. (1988): Correcting for Nonresponse in Household Data (in Finnish). Central Statistical Office of Finland, Studies 147. Helsinki.

LAMBERT, P. J. (1989): The Distribution and Redistribution of Income: a Mathemamatical analysis. Basil Blackwell, Oxford.

MARSHALL, A. W. - OLKIN, I. (1979): Inequalities: Theory of Majorization and Its Applications. Academic Press, New York.

RAO, C. R. (1973): Linear Statistical Inference and Its Applications. Second Edition. John Wiley & Sons, New York.

SEN, A. (1979): Survey on Welfare and Real Income. Journal of Economic Literature 17, 1-45.

SHORROCKS, A. F. (1983): Ranking Income Distributions. Economica 50, 3-17.

STIGLER, S. M. (1974): Linear Functions of Order Statistics with Smooth Weight Functions. Annals of Statistics 2:4, 676-693.

STOLINE, M. R. - URY, H. K. (1979): Tables of the Studentized Maximum Modulus Distribution. Technometrics 21:1, 87-93.

YITZHAKI, S. - SLEMROD, J. (1991): Welfare dominance: An Application to Commodity Taxation. American Economic Review 81:3, 480-496.

Appendix A: Asymptotic results

Below some proofs of asymptotic properties of the estimators are given. The results given here were first obtained by Beach & Davidson (1983) the proofs presented here are somewhat different from theirs. Both methods utilise the results by Stigler[1] (1974) concerning linear functions of order statistics given in the following lemma.

Let $x_1, x_2, x_3 \ldots, x_n$ be an i.i.d. sample and let $x_{(1)}, x_{(2)}, x_{(3)} \ldots, x_{(n)}$ be the corresponding ordered sample. Let

(1) $$S_n = \frac{1}{n}\sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) x_{(i)},$$

be the class of statistics of interest and assume that:

i) The Kernel function J is bounded and is continuous a.e. in F^{-1}.

ii) X has finite variance.

Lemma 1 Let the above assumptions hold. Then $\sqrt{n}(S_n - E(S))$ is asymptotically normal, where

(2) $$E(S) = \lim_{n\to\infty} E(S_n) = \int_0^1 J(P)X(P)dP$$

is the asymptotic expectation of S_n and

(3) $$\sigma^2(S) = n\mathrm{Var}(S) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(F(x))J(F(y))[F(\min(x,y)) - F(x)F(y)]dxdy.$$

is the asymptotic variance[2] of S_n. (Stigler 1974, 677-684.)

Furthermore, it is shown in Stigler (688) that statistics in the form of

(4) $$S'_n = \frac{1}{n}\sum_{i=1}^{n} J_n\left(\frac{i}{n+1}\right) x_{(i)},$$

where the functions J_n are bounded and for every continuation point p of J it holds that $J_n \to J$ uniformly in some open neighbourhood of p, have same limiting distribution as S_n.

[1]Stigler's results cover a wider spectrum of situations where the asymptotic normality of linear functions of order statistics holds. Here only the results relevant for the situation at hand are discussed.

[2]The expression asymptotic variance/covariance is used rather loosely in this appendix. The scaling factor n^{-1} is omitted in the discussion.

Appendix A: (continued)

Estimators of Generalised Lorenz Ordinates

Assume that Income Distribution has a absolutely continuous frequency distribution that is positive over its support and has finite variance. The estimator of generalised Lorenz ordinate (equation 3 in page 4) can be written in the form

$$\hat{G}_j = \frac{1}{n}\sum_{i=1}^{n} I^{p_j}(i/n)x_{(i)}, \tag{5}$$

where the function I^{p_j} is Heaviside function, which has the value 1 if its argument is less than p_j and 0 otherwise. It can easily be shown that $\hat{G}_j$ belongs to the class of statistics in the form J_n', where the function

$$J_n(u) = I^{p_j}\left(\frac{n+1}{n}u\right). \tag{6}$$

Proposition 1 Under the above assumptions for the following are true:
i) $\hat{G}_j$ is assymptotically normal.
ii) $\lim_{n\to\infty} E(\hat{G}_j) = G_j$ (the order of bias is $o(n^{-\frac{1}{2}})$).
iii) Asymptotic variance of $\hat{G}_j$ is

$$\sigma_{jj} = n\mathrm{Var}(\hat{G}_j) = p_j[\lambda_j^2 + (1-p_j)(X(p_j)-\mu_j)^2], \tag{7}$$

where $X(p_j)$ is the value of the fractile function of X at p_j, $\mu_j = E[X|F(X) \le p_j]$ defines a conditional expectation of X and $\lambda_j^2 = E[X - E[X|F(X) \le p_j]|F(X) \le p_j]^2$ defines a conditional variance.
iv) Estimator vector $\hat{G}$ is asymptotically multinormal.
v) Let $p_k \ge p_j$. Asymptotical covariance between $\hat{G}_j$ and $\hat{G}_k$ is

$$\sigma_{jk} = p_j[\lambda_j^2 + (1-p_k)(X(p_j)-\mu_j)(X(p_k)-\mu_k) + (X(p_j)-\mu_j)(\mu_k-\mu_j)]. \tag{8}$$

Proof of Proposition 1
i) Follows directly from Lemma 1.
ii) Follows from the equation 2 of Lemma 1.
iii) From equation 3 of Lemma 1

$$\begin{aligned}\sigma_{jj} &= \int_0^\infty\int_0^\infty I^{p_j}(x)I^{p_j}(y)[F(\min(x,y)) - F(x)F(y)]dxdy \\ &= \int_0^{X(p_j)}\int_0^{X(p_j)}[F(\min(x,y)) - F(x)F(y)]dxdy\end{aligned}$$

Appendix A: (continued)

$$
\begin{aligned}
&= \int_0^{X(p_j)} \left(\int_0^y F(x)dx + \int_y^{X(p_j)} F(y)dx - \int_0^{X(p_j)} F(x)F(y)dx \right) dy \\
&= \int_0^{X(p_j)} \left(yF(y) - \int_0^{X(p_j)} xf(x)dx + \int_y^{X(p_j)} xF(y) - F(y)(p_jX(p_j) - p_j\mu_j) \right) dy \\
&= \int_0^{X(p_j)} \left(X(p_j) - p_jX(p_j) + p_j\mu_j)F(y) - \int_0^y xf(x)dx \right) dy \\
&= (X(p_j) - p_jX(p_j) + p_j\mu_j)(p_jX(p_j) - p_j\mu_j) - \int_0^{X(p_j)} \int_0^y xf(x)dxdy \\
&= p_j(1-p_j)(X(p_j) - \mu_j)^2 + p_j\mu_jX(p_j) - p_j\mu_j^2 - \int_0^{X(p_j)} \int_x^{X(p_j)} xf(x)dydx \\
&= p_j(1-p_j)(X(p_j) - \mu_j)^2 + p_j\mu_jX(p_j) - p_j\mu_j^2 - p_jX(p_j)\mu_j + p_j(\lambda_j^2 + \mu_j^2) \\
&= p_j[\lambda^2 + (1-p_j)(X(p_j) - \mu_j)^2]
\end{aligned}
$$

iv) Let $p_j \neq p_k$ and consider a linear combination of $\hat{G}_j$ and $\hat{G}_k$. All linear combinations of these two statistics belong to the class of statistics S_n' so they are asymptotically normal. Using Cramér-Wold device (Rao 1972, 128) it follows that the asymptotic distribution is multinormal. The generalisation to dimensions above two is straightforward.

v) Let $p_k \geq p_j$. Considering the sum $G_k + G_j$ and using the indentity $\sigma_{jk} = n\mathrm{Cov}(\hat{G}_j, \hat{G}_k) = \frac{1}{2}n(\mathrm{Var}(\hat{G}_j + \hat{G}_k) - \mathrm{Var}(\hat{G}_j) - \mathrm{Var}(\hat{G}_k))$ it follows from iv) and equation 3 of Lemma 1 that

$$
\begin{aligned}
\sigma_{jk} &= \frac{1}{2}(\int_0^\infty \int_0^\infty (I^{p_j}(x) + I^{p_k}(x))(I^{p_j}(y) + I^{p_k}(y))[F(\min(x,y)) - F(x)F(y)]dxdy \\
&\quad -n\mathrm{Var}(\hat{G}_j) - n\mathrm{Var}(\hat{G}_k)) \\
&= \frac{1}{2}(\int_0^{X(p_k)} \int_0^{X(p_k)} (I^{p_j}(x) + 1)(I^{p_j}(y) + 1)[F(\min(x,y)) - F(x)F(y)]dxdy \\
&\quad -n\mathrm{Var}(\hat{G}_j) - n\mathrm{Var}(\hat{G}_k)) \\
&= \frac{1}{2}(\int_0^{X(p_k)} \int_0^{X(p_k)} (I^{p_j}(x)I^{p_j}(y) + I^{p_j}(x) + I^{p_j}(y) + 1)[F(\min(x,y)) \\
&\quad -F(x)F(y)]dxdy - n\mathrm{Var}(\hat{G}_j) - n\mathrm{Var}(\hat{G}_k)) \\
&\overset{\text{(iii)}}{=} \frac{1}{2}\int_0^{X(p_k)} \int_0^{X(p_k)} (I^{p_j}(x) + I^{p_j}(y))[F(\min(x,y)) - F(x)F(y)]dxdy \\
&= \int_0^{X(p_k)} \int_0^{X(p_k)} I^{p_j}(x)[F(\min(x,y)) - F(x)F(y)]dxdy \\
&= n\mathrm{Var}(\hat{G}_j) + \int_0^{X(p_j)} \int_{X(p_j)}^{X(p_k)} [F(y) - F(x)F(y)]dxdy \\
&= n\mathrm{Var}(\hat{G}_j) + \int_0^{X(p_j)} F(Y)dy \int_{X(p_j)}^{X(p_k)} [1 - F(x)]dx \\
&= n\mathrm{Var}(\hat{G}_j) + pj(X(p_j) - \mu_j)[(1-p_k)X(p_k) - (1-p_j)X(p_j) + p_k\mu_k - p_j\mu_j] \\
&= p_j[\lambda^2 + (1-p_k)(X(p_j) - \mu_j)(X(p_k) - \mu_k) + (X(p_j) - \mu_j)(\mu_k - \mu_j)]
\end{aligned}
$$

Appendix A: (continued)

The above expression shows that, all the elements of covariance matrix can be estimated without any parametric assumption about the form of distribution function.

Estimators of Ordinary Lorenz Ordinates

The derivation of the asymptotic distribution of ordinary Lorenz ordinates from the asymptotic distribution of generalised Lorenz ordinates is straightforward. Let $\hat{G}$ be the k-dimensional estimator of generalised Lorenz ordinates and let $\hat{g}_k = \hat{\mu}$. Consider the function $h : \mathcal{R}^k \rightarrow \mathcal{R}^{k-1} : h_i = x_i/x_k$. Note that $h(G) = L$, where L is the $k-1$ dimensional vector (the k:th element of Lorenz ordinate vector is uninteresting because it is by definition equal to 1) of Lorenz ordinates and consider the asymptotic distribution of $h(\hat{G})$. Using (see Rao 1972, 385-386) delta-method for differentiable functions of asymptotically normal quantities one easily obtains that $h(\hat{G})$ is asymptotically normal with expectation L and covariance matrix $H'\Sigma H$, where Σ is the covariance matrix of generalised Lorenz ordinates and

$$H = \begin{bmatrix} 1/\mu & 0 & 0 & \dots & 0 & -G_1/\mu^2 \\ 0 & 1/\mu & 0 & \dots & 0 & -G_2/\mu^2 \\ 0 & 0 & 1/\mu & & 0 & -G_3/\mu^2 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & 1/\mu & -G_{k-1}/\mu^2 \end{bmatrix}$$

is the Jacobian matrix of h evaluated at G. All the elements of matrix $H'\Sigma H$ can be estimated with no parametric assumptions on the form of the distribution function. (Beach & Davidson 1983.)

Appendix B: Tabulation of the Curves

Table 1 Lorenz ordinates L(p) and generalised Lorenz ordinates GL(p) of disposable income for years 1971, 1976, 1981, 1985 and 1990 at vintile points. The number below the estimated ordinate is its estimated asymptotic standard error. Note that all generalised Lorenz ordinates are expressed in 1990 currency.

%	L(p) 1971	L(p) 1976	L(p) 1981	L(p) 1985	L(p) 1990	GL(p) 1971	GL(p) 1976	GL(p) 1981	GL(p) 1985	GL(p) 1990
5.0	1.42	1.95	1.79	2.02	2.13	597.4	947.3	925.5	1172.0	1510.1
	0.052	0.048	0.039	0.035	0.036	22.42	24.46	21.18	21.36	25.95
10.0	3.56	4.56	4.45	4.81	4.97	1497.4	2221.4	2299.6	2787.2	3525.6
	0.083	0.068	0.057	0.049	0.048	37.23	35.33	32.42	30.68	35.72
15.0	6.14	7.53	7.48	7.94	8.11	2582.2	3662.6	3870.2	4599.7	5744.4
	0.109	0.090	0.070	0.061	0.060	49.59	48.73	40.94	39.73	44.65
20.0	9.02	10.77	10.81	11.33	11.48	3792.6	5241.7	5590.5	6564.7	8134.0
	0.133	0.108	0.082	0.072	0.072	61.98	60.80	50.07	47.99	53.46
25.0	12.15	14.26	14.38	14.95	15.05	5109.0	6940.1	7437.8	8661.2	10668.9
	0.155	0.126	0.092	0.081	0.085	73.61	73.17	58.46	56.34	63.34
30.0	15.50	17.98	18.19	18.78	18.83	6519.9	8751.1	9408.9	10878.5	13345.6
	0.178	0.142	0.103	0.091	0.097	86.91	84.99	67.85	65.40	72.67
35.0	19.12	21.90	22.23	22.82	22.80	8040.0	10659.8	11495.7	13219.7	16160.9
	0.200	0.158	0.111	0.100	0.109	100.25	98.27	76.38	74.19	82.94
40.0	22.97	26.06	26.47	27.07	26.96	9660.2	12682.4	13689.8	15680.9	19108.4
	0.219	0.174	0.119	0.108	0.121	111.06	113.05	85.26	83.32	93.16
45.0	27.03	30.45	30.92	31.52	31.31	11370.9	14819.2	15989.8	18259.4	22187.5
	0.239	0.187	0.126	0.116	0.133	124.59	126.93	94.07	92.65	103.53
50.0	31.35	35.07	35.58	36.17	35.84	13185.8	17070.7	18400.8	20953.4	25402.5
	0.258	0.200	0.133	0.123	0.145	138.90	141.09	103.81	101.65	114.76
55.0	35.95	39.92	40.46	41.01	40.59	15121.6	19428.1	20925.2	23758.8	28768.6
	0.277	0.211	0.138	0.129	0.157	155.26	155.65	112.13	110.39	128.15
60.0	40.84	45.03	45.56	46.06	45.58	17177.5	21916.4	23560.9	26683.3	32305.7
	0.294	0.221	0.142	0.134	0.168	174.36	171.14	122.17	119.98	140.68
65.0	46.03	50.38	50.90	51.33	50.80	19361.2	24522.9	26326.6	29733.6	36002.2
	0.309	0.229	0.146	0.138	0.179	188.30	185.33	132.17	129.14	153.18
70.0	51.59	56.01	56.51	56.84	56.26	21700.2	27261.0	29223.3	32928.9	39874.4
	0.322	0.235	0.148	0.141	0.190	209.17	201.57	142.36	140.56	166.07
75.0	57.57	61.92	62.39	62.64	62.01	24214.5	30138.7	32265.0	36289.3	43949.2
	0.333	0.239	0.148	0.143	0.200	229.12	215.02	152.44	151.29	181.15
80.0	64.04	68.16	68.59	68.75	68.09	26936.5	33173.1	35474.3	39825.4	48258.8
	0.339	0.241	0.146	0.144	0.209	254.89	231.38	164.62	161.89	196.03
85.0	71.11	74.76	75.17	75.20	74.57	29910.2	36387.7	38873.8	43562.9	52848.3
	0.341	0.240	0.144	0.141	0.218	279.13	246.62	174.59	173.89	211.74
90.0	78.84	81.82	82.20	82.14	81.54	33158.4	39821.9	42512.9	47582.9	57790.8
	0.337	0.233	0.136	0.134	0.224	302.78	263.77	186.67	187.82	229.74
95.0	87.59	89.60	89.95	89.87	89.35	36840.0	43612.5	46521.3	52060.7	63325.0
	0.315	0.209	0.118	0.114	0.224	331.02	286.31	201.87	206.55	254.92
100.0	100.00	100.00	100.00	100.00	100.00	42060.1	48672.2	51717.4	57928.5	70872.1
	-	-	-	-	-	408.20	340.15	233.03	242.30	341.71