Empirical Analyses of Inequality and Welfare

By

Esfandiar Maasoumi footnote

Department of Economics, SMU, Texas, 75275-0496

maasoumi@mail.smu.edu

Keywords: *Inequality, welfare, dominance, statistical inference, multidimensional inequality, cluster analysis, PSID, distributions, entropy, Gini, Lorenz, empirical.*

A selective survey of applied research on “income” inequality as well as some related applications involving inequality and welfare indices is presented. Topics such as poverty and mobility are not directly discussed here. But multi-attribute concepts of welfare as extensions of “income” and related developments are presented. Generally speaking, our point of departure is the axiomatically justified measures of inequality centering on Generalized Entropy family of indices, contrasting them with some other measures. Empirical literature is reviewed in this light. Statistical and econometric techniques for use in this general area have developed quite extensively in recent years. But their application in empirical research is not as yet wide-spread. A sampling of available techniques is provided. Cluster analysis and other techniques for aggregation (indexing) of welfare attributes are also briefly discussed. Such associated topics as earnings mobility, human capital investment, growth and development, and the use of inequality indices for the measurement of risk and volatility are not surveyed.

1. Introduction.

The modern theory of income distribution and welfare has experienced radical progress toward a more objective and systematic analysis of welfare and inequality. Value judgements that are implicit in any particular index of inequality or poverty have been made explicit and are better understood. We have a better understanding of the limitations of basing welfare statements on measured income, as well as the challenges of developing more meaningful measures of welfare by means of, for example, equivalence scales, cost of living indices, or multidimensional indices.

On the empirical side, the appropriate statistical tools for drawing inferences about indices and their distributions have been developed. Estimation of inequality indices, their variances, and tests of hypotheses regarding distributional changes, or dominance relations can be conducted according to state-of-art and practical statistical tools. Computer code, even for PCs, is no longer a “problem”.

One does not have to subscribe to either the existence or pre-eminence of Social Welfare Functions (SWFs) in order to appreciate their positive role in providing economy of thought and discipline in this area. For this reason, I have generaly focused on theoretical developments that are based on the SWF approach. Some associated theoretical concepts for ranking distributions are presented in section 2. On the other hand, and relatedly, stochastic dominance concepts and practical statistical distribution theory for ranking distributions have developed and specialized to a point where we might reasonably expect widespread application in routine empirical studies. I have described some of these developments in section 4.

Some generalizations and refinements of the concept of "income" as a measure of welfare have taken place at both the theoretical and empirical levels. For instance, multi-attribute studies have dealt with issues of measurement, aggregation, double counting and scalar index development. Some of these are presented in section 5. The application of these to both international and single country data testify to their practicality and the increasing availability of data. A selective account appears in section 6.

Generally speaking, our point of reference is the axiomatically justified measures of inequality, centering on Generalized Entropy family of measures, contrasting them with some other measures; see section 3. Empirical literature is reviewed in this light.

My aim in this chapter has been to indicate these developments in a single investigation that, while selective, emphasizes the interface between theory and empirical developments and possibilities. In terms of some important details the cited references are more than usually important for readers familiar with theoretical developments but not the statistical techniques, and vice versa. More powerful and rigorous conceptual as well as statistical inference tools now exist in our knowledge base than is apparent from the bulk of this vast literature. Statistical and econometric techniques for use in this general area have developed quite extensively in recent years. But their application in empirical research is not as yet wide-spread. By bringing together this somewhat scattered body of economic concepts and econometric tools and applications, it is hoped that this survey will encourage a statistically sounder development of empirical work in this area. Several recent collected volumes, for example Maasoumi (1991), provide useful supplemental reading. A casual search of related publications in ECONLIT produces a list of several thousand references. The only comfort one may offer a newcomer is that this literature suffers more than its fair share of "high variance" in the quality of investigations. This is particularly so with respect to empirical studies that seem disconnected from the best economic theory, on the one hand, and appropriate statistical analysis (if any), on the other.

Related topics, such as poverty and mobility are not directly discussed here. Such associated topics as earnings mobility, human capital investment, growth and development, and the use of inequality indices for the measurement of risk and volatility are not surveyed.

The table of contents further describes the order of presentation.

2. Comparisons of Income Distributions and Welfare

Since at least Dalton (1920) there has been an awareness of a correspondence between any inequality measure and Social Welfare (SW) evaluations. In the 60s and 70s, the influential contributions of Atkinson and Kolm, among others, made this connection much more precise and the awareness more widespread. See Kolm (1969, 1976$a-b$) and Atkinson (1970). It is true, however, that like competing economic doctrines, at least several poorly motivated or even undesirable (from the viewpoint of their SW functions) inequality and related measures have survived in the important empirical domain which informs public perceptions and policy. The formal SW analysis has nevertheless revolutionized the clarity with which one may readily identify the subjective preferences *and*/*or* prejudices that accompany any welfarist statements made on the basis of indices of inequality and welfare. In this part of the paper we first give a brief outline of the existing theory for ranking distributions (welfare situations). The statistical and empirical counterpart to this theory is described in section 4.

2.1. Ranking Relations

We begin with a SWF-based account of ranking income distributions. One could substitute any other attribute or composite of attributes for "income".

Let $SI = \{1, 2, ...,n\}$ denote the set of "income" units, such as individuals, families or countries, and Y^* to denote the compact set of feasible incomes in R_+^n. A typical n-vector is denoted by $y = (y(i))$, and ordered incomes are $y_1 \geq y_2 \geq .... \geq y_n$, denoted by $Y = (y_1, y_2, ...,y_n)$. Let $SWF = w : Y^* \longrightarrow R$. Throughout this work we assume:

A1. **Anonymity** (Symmetry).

This means that functionals of income, including SWFs, are symmetric and invariant to permutations of y. For instance, $w(y) = w(Y)$.

Take two income vectors y^a and y^b. The first elementary relation between these two is when $y^a(i) \geq y^b(i)$, for all $i \in SI$. That is y^a Pareto dominates y^b, denoted $y^a \geq_p y^b$, iff the element-wise inequality holds. In terms of the ordered incomes, Y, and given anonymity, we have a Rank Dominance criterion, $\geq_R$, which is implied by Pareto dominance. We say that $y^a \geq_R y^b$, iff $Y^a \geq Y^b$. From Saposnik (1981,1983) we have:

Result 1. $y^a \geq_R y^b$ iff $w(y^a) \geq w(y^b)$ for all $w \in W_p$, where W_p denotes the set of symmetric Paretian welfare functions.

It can be shown that $y^a \geq_p y^b$ implies $w(y^a) \geq w(y^b)$ for all $w \in W_p$, which by Result 1 implies rank dominance.

We refered to these relations as "elementary" since they do not reflect any consideration of inequality. Relations that incorporate a preference for equality are usefully represented by W_e, the class of "equality prefering" SWFs. It has become fashionable to require functions in W_e to be Schur concave (S-concave) since this property is consistent with the Pigou-Dalton Principle of transfers. Let B represent a bi-stochastic matrix. Then a function w is S-convave if $w(By) \geq w(y)$; *e.g.*, see Kolm (1976a). There is an important dominance relation, Lorenz Dominance, which provides rankings that are equivalent to ranking by members of W_e. Furthermore, Lorenz dominance can be tested empirically.

Formally, we say y^a Lorenz dominates y^b, denoted $y^a \geq_L y^b$, if its Lorenz curve (see below) is everywhere above that of y^b. Atkinson (1970) and Dasgupta et al.(1973) showed that, when two distributions have equal means, $\mu^a = \mu^b$, Lorenz dominance is equivalent to higher ranking by W_e. Shorrocks (1983) showed a similar equivalence for the empirically more relevant case of unequal means by defining a Generalized Lorenz (GL) curve. The empirical GL is as follows:

$G_y(j/n) = \mu_y L_y(j/n)$, (1)

where $L_y(j/n) = (\sum_{i=1}^{j} y_i)/n\mu_y$, $j 0$ SI, is the $j-th$ coordinate of the empirical Lorenz curve, with $L(0) = 0$ and $L_y(1) = 1$. The GL curve is increasing, continuous, and convex on the unit interval such that $G(0) = 0$ and $G(1) = \mu_y$. When the Generalized Lorenz curve of one distribution is everywhere above that of another we say $y^a \geq_{gl} y^b$, or just y^a Lorenz dominates y^b. From Shorrocks (1983) we find the equivalence result which can form the welfare basis of empirical comparisons that may be made between Lorenz or GL curves:

Result 2. :

$y^a \geq_{gl} y^b$ iff $w(y^a) \geq w(y^b)$ for all $w \, 0 \, W_e$, the class of all increasing and *S*-concave SWFs.

The reader would have noted that, in the hierarchy of dominance and ranking relations, gl dominance is less stringent than the elementary rank dominance defined earlier. Indeed, $y^a \geq_R y^b$ implies $y^a \geq_{gl} y^b$. Rank dominance is equivalent (iff) to the Generalized Lorenz curves being everywhere **steeper**, an unlikely occurance in practice footnote .

Stochastic Dominance:

Dominance relations relative to broad classes of SWFs can be stated in terms of conditions involving the cummulative distribution functions (cdfs) *and*/*or* their quantiles. Let X_1 and X_2 be two income variables (or prospects) with F_1 and F_2 cdfs. Also let $w' \geq 0$ for all $w \in W_p$.

Definition 1. (First Order Stochastic Dominance, FSD) : Relative to the class W_p, $X_1 FSD\, X_2$ iff $F_2(x) \geq F_1(x)$, for all x in the support of X_i.

Now relative to the class of increasing and concave SWFs ($w' \geq 0$, and $w'' \leq 0$), we have,

Definition 2. (Second Order Stochastic Dominance, SSD) : $X_1 SSD\, X_2$ iff

$$\int_0^x F_2(t)dt \geq \int_0^x F_1(t)dt \text{ , for all } x \text{ in the support of } X_i.$$

Equally useful alternative conditions for FSD and SSD are available in terms of quantiles $q_1(p)$ and $q_2(p)$ of X_1 and X_2 which are implicitly defined from $F(X_i \leq q_i(p)) = p, i = 1,2$.

Definition 3. $X_1 FSD\, X_2$ iff $q_1(p) \geq q_2(p)$, for all $0 \leq p \leq 1$.

Definition 4. $X_1 SSD\, X_2$ iff $\int_0^p [q_1(t) - q_2(t)]dt \geq 0$, for all $0 \leq p \leq 1$.

Noting the definition of the continuous Generalized Lorenz curve for X_i as $\int_0^p q_i(t)dt$, the equivalence of SSD and GL dominance conditions becomes evident. Different tests for stochastic dominance are proposed on the basis of the alternative conditions given in definitions 1-4 above. See Maasoumi, Mills, and Zandvakili (1995).

2.2. Equity-efficiency trade offs

For the previously mentioned criteria some would argue that the right balance has not been struck between the **level** of income (efficiency) and equality. For instance, in a Paretian dominance relation, only the levels matter, while in the *S*-concave Paretian SWFs, insufficient weight may be given to equality. In order to remedy this defect some have suggested that, for instance, the SWFs be increasing, *S*-concave and increasing with **equiproportional** increases in all incomes. Such increases, we note, do not increase relative inequality but may be welfare enhancing. Let such a family of SWFs be denoted by W_{ei} where, for any $\lambda \geq 1$, $w(\lambda y) \geq w(y)$, for all $w(.) 0 \, W_{ei}$. Dasgupta et al. (1973) and Shorrocks (1983) have shown that:

Result 3. :

$w(y^a) \geq w(y^b)$ for all $w \, 0 \, W_{ei}$ iff $y^a \geq_{ml} y^b$, where "ml" stands for mean-Lorenz dominance which holds iff $\mu^a \geq \mu^b$ and $y^a \geq_L y^b$.

It is to be noted that "ml" dominance implies, but is not implied by, "gl" dominance defined in relation to GL curves. This is because W_e contains W_{ei}. In a later section we present statistical tests for these dominance relations.

2.3. Relative versus Absolute inequality

Kolm (1976*a*,*b*) and Shorrocks (1983), among others, have studied the consequences of the subjective, absolutist, requirement that inequality should not change when **equal increments** are added to all incomes. In order to accommodate this absolute notion of inequality, one may define an empirically implementable Absolute Lorenz curve to play the same role for absolute inequality analysis as the classic Lorenz curve does for relative inequality. Let the empirical absolute Lorenz curve be defined as follows, see Moyes (1987), :

$A_j(j/n) = \sum_{i=1}^{j} (y_i - \mu_y)/n$, for $j \, 0$ SI. (2)

$= G_y(j/n) - (j/n)G_y(1)$ (3)

We define Absolute Lorenz dominance as $y^a \geq_A y^b$ iff the absolute Lorenz curve for y^a is everywhere above that of y^b. When the means are equal (or taken to be equal because of a statistically insignificant difference), one does not need to construct the absolute Lorenz curve since, in that case, $y^a \geq_A y^b$ iff $y^a \geq_L y^b$, thus the ordinary Lorenz curve is all that must be fitted

Once again, if one wishes to require that absolute "equality preserving" redistributions that increase mean income (efficiency) be welfare increasing, one would need to define a Mean Absolute Lorenz dominance relation $\geq_{ma}$ as follows:

Mean Absolute Lorenz Dominance: $y^a \geq_{ma} y^b$ iff $\mu^a \geq \mu^b$, and $y^a \geq_A y^b$.

Let W_a denote the class of all SWFs such that for any $\lambda \geq 0$, $w(y + \lambda . l_n) \geq w(y)$, where l_n is the summation vector of ones. Shorrocks (1983) provides the welfare basis for empirical rankings done according to the mean absolute Lorenz criterion. He showed that $y^a \geq_{ma} y^b$ iff $w(y^a) \geq w(y^b)$ for all $w \in W_a$. Again, the hierarchy among the above dominance relations is revealed by noting that $\geq_{ma}$ implies, but is not implied by, $\geq_{ml}$. Intuitively, this last relation follows because equal increment increases to all incomes, while **absolute** equality preserving, increase **relative** equality.

3. Inequality Measures and SWFs

It is of general interest to compare distributions across countries, or over time for the same region, perhaps in order to assess any impact of economic policy or events. This requires that we estimate the various curves defined above in order to evaluate any dominance relations statistically. This involves two sets of operations. The first is to obtain parametric or non-parametric estimates of the Lorenz curve and its derivative forms discussed above. The second is to statistically test for the existence of the types of inequality relations, or differences, between the curves found in the first stage.

It is also an independently interesting and common practice to compare distributions on the basis of specific inequality indices. Indeed, it is very common to explicitly or implicitly rank income distributions on the basis of these estimated inequality measures without proceeding to a comparison of Lorenz-type curves. As we shall see below, this practice may be inconclusive. This is because these index comparisons are based on specific cardinalizations of SWFs, and "direct" relation between inequality measures and SWFs is merely monotonic, not one-to-one. Welfare relations are best investigated in terms of Lorenz-type curves and by testing for stochastic dominance which provide welfare rankings that are valid over whole classes of welfare functions (inequality measures).

3.1. Measures of Inequality.

Kolm (1969) and Atkinson (1970) have provided clear and influential formalizations of the relationship between SWFs and inequality measures footnote . Since then there has been much progress in both expanding on this important relationship and in utilizing it for more informed analyses of inequality measures. Kolm and Atkinson considered a utilitarian and individualistic welfare function which was increasing in incomes and equality prefering. Let x be the income variable, μ_x its mean, and y_e the "equal equivalent income"; *i.e.*, the level of income which if received by everyone would leave social welfare at the same level as for a given income vector. Thus $y_e < \mu_x$ so long as there is any inequality, and a measure of divergence between these two would indicate the degree of welfare loss due to inequality. Atkinson (1970) and Kolm (1969) argued that, see also Blackorby and Donaldson (1978),

$$I(x) = 1 - y_e/\mu_x \quad (4)$$

may be a good measure of "relative" inequality. Indeed, one could just as well take:

$$I(x) = 1 - SWF(x)/\mu_x \quad (5)$$

where $SWF(x)$ is the "average" or mean SWF. This should make clear that the SWF approach does not by itself identify a unique inequality index even when a particular SWF is agreed upon. Nevertheless, the last definition has been used to define "abbreviated" SWFs as the basis for empirical implementation, as follows:

$$SWF(x) = \mu_x[1 - I(x)] \quad (6)$$

An important example of the measures generically defined in (5) is the Atkinson family of inequality measures :

$$A_r = 1 - [\int_0^\infty x^r dF]^{\frac{1}{r}}/\mu_x, \; r \leq 1, \quad (7)$$

where F is the *c.d.f* of income. Similarly, the Generalized Entropy (GE) family of indices is given by:

$$I_\gamma = \frac{1}{\gamma(\gamma+1)} \int_0^\infty (x/\mu_x)[(x/\mu_x)^\gamma - 1]dF, \; \gamma \text{ real}, \quad (8)$$

This family includes Herfindahl's, variance of logarithms, square of the coefficient of variation, and Theil's first and second measures, I_0 and I_{-1} , respectively. Also, up to a monotonic transformation, there is a unique member of GE corresponding to each member of the Atkinson family. γ is the degree of aversion to relative inequality; the higher its absolute value the greater is the sensitivity of the measure to inequality (transfers) in the tail areas of the distribution.

The axiomatic derivation technique that identifies GE is constructive and is to be appreciated as an important breakthrough in organizing learning and knowledge in this area. By borrowing from functional theory first developed in "information theory", see Maasoumi (1993), one is forced to put down an **explicit** set of properties (axioms) which may or may not be defensible from the point of view of SWFs. Using these axioms as explicit constraints on the function space one then obtains the appropriate inequality index. To exemplify, let us follow Bourguignon (1979) or Shorrocks (1980, 1984) in their discussion of the "fundamental welfare axioms" of symmetry, continuity, Principle of Transfers, and additive decomposability which identify GE as the desirable scale invariant family of inequality measures.

Axiom 1. The inequality index (function) is symmetric in incomes.

This is equivalent to anonymity which requires that only income matters not the identity of its recepient.

Axiom 2. Principle of transfers holds.

This requires that inequality decrease if we redistribute from a single richer individual to a poorer one, leaving their respective ranking and all the other individuals' incomes unchanged.

Axiom 3. Continuity.

This is relatively inocuous, helping in the mathematical derivations and in comparing different populations. None of the well known inequality indices violates this requirement. Note that it does allow for the practically non-sensical zero inequality.

Axiom 4. Invariance to scalar multiplication.

This is a serious limitation as it restricts attention to "relative" inequality. This is so since this requirement implies mean invariance; doubling everyone's income would leave inequality unchanged. Questions of "efficiency" can only be taken up by absolute inequality measures. None of the popular measures violates this requirement!

The class of functionals satisfying Axioms 1-4 is still too large. Also, any further axioms are less likely to command consensus. In fact, any further requirements must be justified by plausible considerations of such things as policy, empirical necessity, and practical interest. The most commonly invoked of such requirements is:

Axiom 5. Additive decomposability (aggregation consistency).

This requirement, later strengthened as an "aggregation consistency" axiom by Shorrocks (1984), says that total inequality must be the sum of a "between group" component, obtained over group **means**, and an additive component which is a weighted sum of "within group" inequalities. As we shall argue in later sections, this kind of decomposability is very useful for controlling and dealing with heterogeneity of populations, and as a means of unambiguously identifying the sources of inequality and those that are affected by it.

In their various incarnations, Axioms 1-5 together identify the GE family as the "ideal" family of indices. But other axiom sets have been given which "justify" other inequality indices, see Blackorby and Donaldson (1978), and Dagum (1990). One of the most popular and enduring inequality indices is the Gini ratio given by:

$$G = (2/\mu_x)\int_0^\infty x[F - 1/2]dF \qquad (9)$$

$$= 2\int_0^\infty [F - L]dF \qquad (10)$$

where L is the continuous version of the Lorenz curve defined earlier.

Using the above formulae, the SWF corresponding to each measure is readily derived. The SWF for Gini depends on F, the distribution of income, and hence involves interpersonal comparison of utilities. Some regard such interpersonal considerations as desirable; see Dagum (1993). Indeed such comparisons may be inevitable, especially in developing socially acceptable lower bounds for inequality; zero inequality is merely a mathematical point of departure that is seldom "equitable". But some researchers, "while recognizing the extreme difficulty of any meaningful specification of interpersonal comparisons of utility", are content to leave these specifications implicit and ambiguously "derived from a well specified and accepted income inequality measure such as the Gini ratio"; Dagum (1993, page 7)! This begs the question of what is an "acceptable and well specified" income inequality measure in the first place. Furthermore, if the useful **additive** decomposability requirement of axiom 5 is imposed, such measures as Gini and variance of logarithms must be excluded. The latter two measures provide ambiguous decompositions of overall inequality by population subgroups; see Shorrocks (1984). For the GE family, for instance, a discretized (estimation) formula that helps to demonstrate its decomposability is as follows:

$$I_\gamma = \sum_{r=1}^{R}[Y_r/\sum_j^n y_j]^{\gamma+1}(n_r/n)^\gamma I_\gamma^r + I_\gamma^b \qquad (11)$$

where $Y_r/\sum_j y_j$ is the share of total income to group r, $r = 1, 2, ..., R$, and n_r is the number of units in that group. I^r is the "within" group GE inequality which is defined over the income shares within the $r-th$ group, and I^b is the "between" group GE inequality defined over the R group means. Shorrocks (1984) has convincingly argued that Theil's second measure ($\gamma = -1$) provides the most unambiguous answer to such fundamental questions as : How much of the overall inequality is due to the inequality in the $r-th$ group (for example age, gender, race, educational groups)? Having a good idea about the incidence of inequality, or poverty, is an essential pre-requisite for devising well-directed and appropriate remedial action. It is also essential in establishing lower bounds for inequality that reflect acceptable differences on the basis of experience, education and skills, or other social norms. Such additive decomposability and "aggregation consistency" criterion, requiring that inequality increases if one or more I^r increase (I^b constant), are violated by Gini! We shall see further supporting arguments in favor of requiring additive decomposability in empirically relevant applications. In particular, a property of GE measures provides a partial but important degree of robustness with respect to the thorny problem of defining individual "income" when we admit both the realism of heterogeneity, on the one hand, and the difficulty of correcting for it, on the other.

Theil's two measures were further studied in Maasoumi and Theil (1979) and Maasoumi (1989b) with a view to determine their characteristics in terms of the moments of distributions. Nagar-type approximations were developed around the lognormal distribution in much the same way as Nagar and Edgeworth expansions are developed in econometrics around the normal distribution. This development is helpful and may be considered when estimation and inference issues are presented in section 4 below. Briefly, the two measures may be defined as follows:

$$I_0 = E[\frac{x}{E(x)} \log_e \frac{x}{E(x)}], \qquad I_{-1} = E(\log_e \frac{E(x)}{x}) \tag{12}$$

Let $\log x = z$, $Ez = \mu$, $\text{var } z = \sigma^2$, $\gamma_1 = E(z-\mu)^3/\sigma^3$ as the skewness, and $\gamma_2 = E(z-\mu)^4/\sigma^4 - 3$ as the kurtosis of the log income distribution. Assuming the existence of the first four moments, the following small-σ expansions were given by Maasoumi and Theil (1979):

$$I_0 = \frac{1}{2}\sigma^2[1 + \frac{2}{3}\sigma\gamma_1 + \frac{1}{4}\sigma^2\gamma_2 + o(\sigma^2)] \tag{13}$$

$$I_{-1} = \frac{1}{2}\sigma^2[1 + \frac{1}{3}\sigma\gamma_1 + \frac{1}{12}\sigma^2\gamma_2 + o(\sigma^2)] \tag{14}$$

Note that, when z has a lognormal distribution, both indices equal $\sigma^2/2$. These formulae can be used when the underlying distribution is not known. They allow us to see that positive skewness and leptokurtosis increase inequality, and that I_0 is more sensitive to positive skewness (high income groups) and fat tails (large extreme income groups) than I_{-1}. Maasoumi (1989b) extended these formulae to the multivariable case. He reported some empirical examples indicating that, (i) these expansions underestimated inequality somewhat for household income and net equity in housing in the PSID panels, and (ii) these expansions are not suitable for very small γ_1 and γ_2 distributions, such as the almost flat distribution of years of schooling in the same data set. As we shall discuss in section 4, however, these inequality measures may be (and often are) directly and nonparametrically estimated by the Method of Moments techniques.

3.2 Income, income units, equivalence scales, and related issues.

Income data are often aggregated by groups. Frequently, the least aggregated data on disposable incomes *and/or* expenditures are for household units. The occasional individual level data that is available may be regarded as inherently unrepresentative since many individuals, such as children and non-working adults, do consume but have no incomes. This state of the data has lead to a proliferation of methods for adjusting household incomes for their size *and/or* composition. An accessible recent survey of the various approaches is given in Coulter et al.(1992). The authors present an extensive account of the various "equivalence scales" measures that have been developed in order to adjust money income (or expenditure) by a metric that acts as an exchange rate that obtains a needs-adjusted "income" for a household unit. The simplest such adjusted income, of course, is the per capita income in which all adult members of the unit have the same weight. It is one thing to object to this admittedly simplistic adjustment, but quite another to find a satisfactory alternative. There have been two strands of work on measuring equivalence scales. One is *statistical/econometric* and has largely depended on the use of household survey and expenditure data in order to estimate systems of demand equations. Lewbel (1994, this handbook), and Slesnick (1991) provide very informed surveys of these techniques. Pollak and Wales (1979) and Pollak (1991) made clear the by now classic identification problem of this line of inquiry. As Pollak (1991) argues, using household demographic characteristics, prices, quantities and expenditures may allow "situation comparisons" but not "welfare comparisons" which, albeit, is what is really needed when one wishes to develop a common welfare metric that would serve as a unit by which to measure household equivalent "income". Blundell and Lewbel (1991) clarify this point further and make a gallant effort to isolate the types of utilities and cost functions that are partially identified. They show that **relative** equivalence scales are recoverable from the usual empirical demand analyses, and are useful to the extent that they allow identification of changes in response to price changes. Alternatively, one may make the empirically unsupported assumption that scales are independent of the reference utility level. This can be assumed if preferences are homothetic, often rejected empirically, or if certain other testable restrictions on otherwise rather flexible and popular quasi-homothetic or translog cost functions are validated. Blundell and Lewbel (1991) demonstrate the testing strategy. Lewbel's later work with tests of restrictions in non-parametric regressions which avoid specific functional specifications for demand functions is very relevant here. Unfortunately, however, Blundell and Lewbel (1991) do not find support for the independence assumption using UK Family Expenditure Survey data.

This conundrum is not unique to the econometric attempt at finding more sophisticated scales. There are other methods for devising equivalence scales, some purely subjective, some partially based on expenditure data, or assistance programs' computations, and partially on subjective weightings of observers. Other scales have been computed on the basis of welfare theory by using direct questionaire techniques of Van Praag and his associates; see van Praag (1991). As Coulter et al. (1992) show, the distributional measures, such as inequality or poverty indices, which are computed on the basis of different adjustments to money incomes exhibit great sensitivity toward which equivalence scale is used. It is generally true, however, that inequality measures tend to be smaller when incomes are adjusted than with money incomes. This finding depends, however, on the degree of sensitivity to transfers of income at various parts of the income distribution, among other things. See Coulter et al. (1992) for a sensitivity analysis based on UK data and the members of the GE family. The analyses of redistributions and distributional change is particularly sensitive to the definition of "needs" and the corresponding scales used to translate money incomes.

A further question concerns the components of income such as wages, interests, gains, transfers, etc. The distributions of such components are seldom similar. Also, sometimes "expenditure" is used as a proxy for income requiring additional caution in inferential statements and interpretations.

4. Statistical Inference

While not as yet fully evident in the typical empirical report on inequality, there are few, if any, estimates *and/or* models that cannot be accompanied by standard errors and test statistics. There exists sufficient inference theory specialized to deal with the types of statistics we have in this area. There are also a growing number of computer programs that are accessible to most researchers, but relatively straightforward code can be written by anyone familiar with computation. In this section we provide a sample of existing techniques and theory to encourage further reliance on statistical inference in this area. There is a legitimate argument in some cases against an "overkill", however. This argument flows from the truism that in this area we sometimes deal with large samples which do not justify too much concern for precision (sampling variance)! This argument is occasionally contradicted by large standard errors, and it may be turned around in order to justify the reporting of even more statistical measures of precision and tests. This is because almost all of the useful statistical theory in this area is based on asymptotic approximations which are supposed to do well for large samples! A remaining practical problem is that of the ease with which large samples may be handled by desktop machines. For most practitioners, mainframe computing will remain relevant for a few more years.

Statistical results described here may be conveniently put in three groups. The first is simply a collection of estimators of various inequality and related measures with their variances. The second is tests for such things as significance of distributional change, either directly or through measured changes in summary indices such as inequality indices. I include here tests for dominance and other tests for comparing or ranking distributions. The third group will be concerned with the related issues such as estimating statistically adequate distributions for, say, income data, or evaluating distribution generating models that are inspired by a combination of stylized facts of this literature and a priori plausible diffusion processes. We exemplify the existing techniques in each of these areas in turn.

4.1. Estimation and variances

Almost all indices of interest in this area are functions of the moments of underlying distribution of whatever attribute (*s*) we wish to investigate. Speaking in general terms therefore, the Method of Moments (MM) and the associated statistical inference theory has always been available and known for use in this area. Much of the work refered to below is then seen to be useful specializations and detailed derivations of the MM technique. Taken a step further, one might say that tests of various hypotheses in this area are also specific examples of tests associated with the MM technique. In particular, the reader is encouraged to see these tests as specific instances of the Generalized Method of Moments (GMM) techniques now dominating inference in non-linear dynamic models of modern econometrics.

There are two ways of approaching estimation of inequality and other similar indices. In an **indirect approach**, the first step is to fit a parametric distribution to the data, and second step is to derive the index as a function of (the moments of) the fitted distribution or its normalized incomplete moments. This approach is most developed for the Gini coefficient and other measures which are directly related to the Lorenz curves.

The second approach is direct since we can avoid parametric, or even non-parametric estimation of a distribution function, and directly estimate the index by the Method of Moments (MM) technique which is well suited to most applications in this area. This second approach has many well known advantages amongst which we merely emphasize the avoidance of a priori theorizing about distribution forms. This point ought not be overemphasized in this area, however. As will be discussed in the next section, there are now well established flexible functional forms, such as the generalized beta of the second kind, which fit the income data quite well; see McDonald and Mantrala (1993) for a recent discussion. Maximum Likelihood estimation of the parameters of such distributions forms the basis of MLE estimates of desired indices. The latter ML estimators generally inherit the desirable asymptotic properties of consistency and efficiency, leading to a well developed theory of asymptotic inference. They are, however, vulnerable to potential misspecification inconsistency, especially if the all-important tail areas are poorly fitted by the parametric distribution function chosen.

A third approach based on obtaining the Maximum Entropy (ME) distribution of income data is not yet implemented, see Maasoumi (1993), and may be regarded as an "optimal" method of moments implementation of the indirect approach above. In general, however, since all index estimators are functions of sample moments, the estimation of income distributions as the first step of the indirect method is often superfluous and cumbersome. One may avoid it if the sole objective is to estimate inequality and similar other indices. See also the small$-\sigma$ approximation technique of Maasoumi and Theil (1979) described earlier in this chapter.

We first consider the direct MM estimation of inequality measures and their standard errors which permit construction of asymptotic confidence intervals. Of course, the most general treatment will have to be in terms of equivalence scale-adjusted incomes for individuals. Thus, following Cowell (1989), if the total household income is $y(i)$, its characteristics vector is c_i, and the "adult equivalent" function is $\zeta_i = \zeta(y(i), c_i)$, the "income" variable is $x_i = y(i)/\zeta_i$ with corresponding weights ζ_i. We deal with the simplest example of ζ described earlier, *i.e.*, household size h_i. Then we may write the GE family of inequality indices as follows :

$$I_\beta = [\mu_{1\beta}\mu_{11}^{-\beta}\mu_{10}^{\beta-1} - 1]/\beta(\beta-1)\ , \beta \neq 1,0 \qquad (15)$$

where $\beta = \gamma + 1$ indicates the relation between the normalization here and the one in (8)-(11) of section 3.1, and μ_{ij} are the raw moments of the joint distribution of household size H and "income" X; for instance

$\mu_{i\beta} = \iint h^i x^\beta dF(h,x)$, $i = 1,2, -\infty < \beta < \infty$ (16)

Theil's two measures are given by:

$I_0 = \log \mu_{11} - \log \mu_{10} - \tau_{10}/\mu_{10}$ (17)

$I_{-1} = \tau_{11}/\mu_{11} - \log \mu_{11} + \log \mu_{10}$ (18)

where,

$\tau_{ij} = \iint h^i x^j \log x \, dF(h,x)$ $i,j = 0,1,2$ 19)

MM estimators of I_β are obtained by replacing the population moments with their sample counterparts. For example, $\mu_{i\beta}$ is estimated by:

$m_{i\beta} = \sum_l^n h_l^i x_l^\beta / n$, for n households, (20)

Evidently, inequality measures are functions of population moments, hence the usual techniques for derivation of their approximate variances and asymptotic distributions apply. For example, if the vectors of the population and sample moments are denoted by, respectively μ and m, we let $I_\beta = g(\mu)$ and its estimator, $\hat{I}_\beta = g(m)$. Under regularity conditions (certainly with random sampling or certain forms of stratified sampling), $\sqrt{n}(m - \mu)$ is asymptotically normal with zero mean and covariance matrix $\sum$. And if $g(.)$ is differentiable, we may base inferences on the following asymptotic distribution [*E.g.*, see Amemiya (1985)]:

Theorem 1:

$\sqrt{n}(\hat{I}_\beta - I_\beta)\ N(0, V)$ (21)

where the asymptotic variance matrix is computed from:

$V = G^\vee \sum G / n$ (22)

and the elements of $G = \partial g/\partial\mu$ are obtained as follows:

$\partial g/\partial\mu_{1\beta} = \varphi_\beta/\mu_{1\beta}$, $\beta \neq 0,1$, (23)

$\partial g/\partial\mu_{11} = -\beta.\varphi_\beta/\mu_{11}$, $\beta \neq 1$ (24)

$= -\varphi_1/\mu_{11}$, $\beta = 1$ (25)

$\partial g/\partial\mu_{10} = (\beta - 1)\varphi_\beta/\mu_{10}$, $\beta \neq 0$ (26)

$= \varphi_0/\mu_{10}$, $\beta = 0$ (27)

where,

$\varphi_\beta = \mu_{1\beta}\mu_{11}^{-\beta}\mu_{10}^{\beta-1}/\beta(\beta - 1)$, $\beta \neq 0,1$ and, $\varphi_i = \tau_{1i}/\mu_{1i} + (2i - 1)$, $i = 0,1$. (28)

Cowell (1989) has derived the following expressions for an estimator of V, $\hat{V}$:

$\hat{V} = \hat{V}_0 + \hat{V}_1 + \hat{V}_2$ (29)

where " ^" indicates that the sample moments replace the population moments in the following expressions:

$V_0 = \mu_{11}^{-2\beta}\mu_{10}^{2\beta-2}(\mu_{2,2\beta} - \mu_{1\beta}^2)/n\,\beta^2(\beta - 1)^2$ (30)

$V_1 = [(\frac{\mu_{22}}{\mu_{11^2}} - 1) - \frac{2}{\beta}(\frac{\mu_{2,\beta+1}}{\mu_{1\beta,\beta+1}} - 1)]\varphi_\beta^2/\ n(\beta - 1)^2$ (31)

$V_2 = \varphi_\beta^2[(\frac{\mu_{20}}{\mu_{10^2}} - 1)/\beta^2 + 2(\frac{\mu_{2\beta}}{\mu_{1\beta}\mu_{10}} - 1)/\beta^2(\beta - 1)$

$- 2\,(\frac{\mu_{21}}{\mu_{11}\mu_{00}} - 1)/\beta(\beta - 1)]/n$, $\beta \neq 0,1$ (32)

For the special cases, for example Theil's second measure at $\beta = 1$, we would estimate the following variance using sample moments in place of the population moments:

$V = (\tau^* - \tau_{11}^2)/n\mu_{11}^2 + \varphi_1^2(\frac{\mu_{22}}{\mu_{11^2}} + 1) - 2\varphi_1(\frac{\tau_{22}}{\mu_{11^2}} + 1)/n$

$+(\frac{\mu_{20}}{\mu_{10^2}} + 1)/n + 2(\tau_{21} - \varphi_1\mu_{21})/n\mu_{10}\mu_{11}$ (33)

where $\tau^* = \iint h^2 x^2 (\log x)^2 dF(h,x)$. See Cowell (1989) for the case $\beta = 0$ and further details as well as the corresponding formulae after group decomposition is conducted. If μ_{10} is a priori fixed we may drop V_2. If μ_{11} is also known both V_1 and V_2 become unnecessary.

The Gini Coefficient:

Gini is a commonly used measure which does not provide the full additive decomposability. Less desirable but useful decompositions are possible for Gini. Its asymptotic distribution may be derived as above. However, Gini is a function of both the moments and *U*-statistics. That is, Gini (GI) may be estimated by:

$GI = d/2m_{10}m_{11}$ (34)

$d =_{i,}\ \sum_{i,j0S} K_2(h_i, x_i, h_j, x_j)/n(n - 1)$ (35)

where d is the Gini mean difference, S denotes the set of all permutations of integer pairs in n, and K_2 is the symmetric kernel defined by,

$K_2(h_i, x_i, h_j, x_j) = h_i h_j |x_i - x_j|$ (36)

By (35)-(36) d is a *U*-statistic; see Cowell (1989). Therefore,

$\text{var}\,(d) = 2[2(n - 2)\eta_1 + \eta_2]/n(n - 1)$ (37)

where $\eta_2 = var(K_2)$, $\eta_1 = var(K_1)$, and $K_1(h_i,x_i) = E(K_2|h_i,x_i)$.

For inferences the following may be used for variance of GI :

$\text{var}\,(GI) = (1/2m_{10}m_{11})^2\{var(d) + d^2 var(m_{10})/m_{10}^2 + d^2 var(m_{11})/m_{11}^2$

- 2d $cov(d,m_{10})/m_{10} - 2d\ cov(d,m_{11})/m_{11}$

$+2d^2$ cov $(m_{10},m_{11})/m_{10}m_{11}\}$ (38)

Following Cowell (1989) Appendix I gives the estimators of each term above which involve simple sample moment computations with a few computations based on order statistics. It is recommend that the data be ordered only once before any computations since computations that do not require ordered observations are unaffected by this operation.

From the above discussion, see also Gastwirth et al. (1986, 1989), confidence intervals may be computed on the basis of an asymptotic normal distribution and the above variance formulae. But it is often the case that, because of large samples, these inequality indices are estimated quite precisely. For example, Cowell (1989) reports the computations in Table 1 below based on a White-Nonwhite decomposition of a sample of about six thousand families in the tenth wave of the Michigan Panel Study of Income Dynamics.

Table 5.1 Inequality indices for white-nonwhite decomposition of Michigan PSID sample

β	Whites	Nonwhites	Between	All
	$I_\beta, \sqrt{V_2}$	$I_\beta, \sqrt{V_2}$	$I_\beta, \sqrt{V_2}$	$I_\beta, \sqrt{V_2}$
-1.0	0.340, 0.0157	0.506, 0.0397	0.023, 0.0035	0.424, 0.0167
-0.5	0.279, 0.0110	0.388, 0.0257	0.021, 0.0031	0.327, 0.0109
0	0.251, 0.0099	0.336, 0.0221	0.019, 0.0028	0.285, 0.0094
0.5	0.245, 0.0113	0.322, 0.0236	0.018, 0.0025	0.272, 0.0106
1.0	0.259, 0.0161	0.336, 0.0310	0.017, 0.0022	0.284, 0.0152
2.0	0.392, 0.0590	0.492, 0.0913	0.015, 0.0018	0.425, 0.0580
GI	0.381, 0.0161	0.442, 0.0228		

$n_1 = 3636$, $n_2 = 2356$, $m_{10} = 68.98$ (sic), $m_{11} = 410046.19$;

4.2. Tests for Distributional Differences.

Since our welfare criteria are often functions of ordered incomes, the relevant statistical inference techniques here routinely rely upon sample quantiles and related statistics. For instance, the Generalized Lorenz curves are functions of quantiles. Thus an asymptotic distribution theory similar to the previous section is available which relies upon the following well known distribution of sample quantiles (*E.g.*, see Cramer (1946)):

Theorem 2:

$\sqrt{n}(\hat{q} - q)$ is asymptotically $N(0, \Omega_q)$ where the typical covariance element is,

$\omega_{ij} = p_i(1-p_j)/f(q_{p_i})f(q_{p_j})\ i \le j = 1,2,....,K$ (39)

In the above theorem $p = F(q_p)$, and $q_p = X(p) = \inf[x : F(x) \ge p)$ is the usual inverse function defining the quantile q_p. $\hat{q} = (\hat{q}_{p_1},......,\hat{q}_{p_K})^\vee$ are the corresponding sample quantiles obtained from the sample distribution $F_n(\hat{q}_{p_i}) = p_i$ and are consistent estimates of $q = (q_{p_1},....,q_{p_K})^\vee$. It is assumed that $F(x)$ is strictly monotonic with continuously differentiable density $f(.)$; see Siddiqui (1960), and Sen (1972).

For instance, for the income variable X the Generalized Lorenz curve $GL_x(p) = \int_0^p q_p(t)dt$. Thus tests of GL dominance are equivalent to tests of second order stochastic dominance defined over quantiles. In what follows we first obtain the distribution of interval means which forms the basis of testing for rank dominance. It is then shown that the asymptotic distribution of the Generalized Lorenz function is derived from the distribution of these interval means. This derived distribution forms the basis of some tests of Generalized Lorenz and other forms of dominance described earlier.

To implement any of these tests we require estimates of ω_{ij} which depend on the unknown population density. Here again there is a choice to be made between parametric, nonparametric, and computer intensive methods of estimation. In the indirect approach alluded to earlier, a parametric estimate of the income distribution and, often, a closed form expression for the index to be estimated is obtained. But to be consistent with a nonparametric approach, however, one may use nonparametric methods of estimating the density and the covariances. For instance, $X_n^\vee(p) = 1/f(q_p)$ may be estimated consistently by:

$[\hat{q}_{p+b_n} - \hat{q}_{p-b_n}]/2b_n$ (40)

where the band $b_n \to 0$ as $n \to \infty$ such that $nb_n \to \infty$.

In order to test for rank dominance income coordinates need to be computed. To do so one can first approximate the inverse function $X(p)$ by a step function at p_k, $k = 1,2,....,K$. The conditional (interval) mean is $\mu_k = E(x|q_{k-1} \le x \le q_k)$, and the conditional variance is similarly defined. These two moments are then estimated by:

$\hat{\mu}_k = (\sum_{i=t\ k}\hat{q}_{p_i})/(r_k - r_{k-1})$ (41)

$\hat{\sigma}_k^2 = [\sum_{i=t\ k}(\hat{q}_{p_i} - \hat{\mu}_k)^2]/(r_k - r_{k-1})$ (42)

where $r_k = [np_k]$ indicates the integer part of the argument, and $t_k = r_k - 1$.

Define $\mu = (\mu_1,\mu_2,......,\mu_K)^\vee$, and $y^* = (y_1^*,y_2^*,......,y_K^*)^\vee$, where $y_k^* = \mu_k\Delta p_k$ represents the $k-th$ group's income, and $s = (s_1,s_2,....,s_K)^\vee$ as the vector of income shares $s_k = y_k^*/\mu_0$, with μ_0 as the overall mean. The following theorem gives the desired asymptotic distribution; for example see Bishop et al. (1989):

Theorem 3:

$\sqrt{n}(y^* - y^*)$ is asymptotically distributed as $N(0, \sum_y)$ where,

$$\sigma_{yjk} = \beta_j(q_k - q_{k-1} - \beta_k) \,, j \leq k, \quad (43)$$

$$\sigma_{ykk} = \Delta p_k(\sigma_k^2 + \mu_k^2) - p_k q_k^2 - p_{k-1} q_{k-1}^2 + (q_k - q_{k-1})\beta_k - \beta_k^2 \quad (44)$$

and

$$\beta_k = p_k q_k - p_{k-1} q_{k-1} - \Delta p_k \mu_k \quad (45)$$

If two independent samples of n_a and n_b observations are used to obtain the estimates $\hat{\mu}^a$ and $\hat{\mu}^b$, and we denote the asymptotic variance of these estimates by $\sum_m = (\sigma_{ik})$, $\sigma_{ik} = \sigma_{yik}/\Delta p_i \Delta p_k$, we may invoke Cramer's linear transformation theorem in conjunction with Wald's theorem on the distribution of quadratic forms in Normal variates to obtain the following statistic to test $H_0 : \mu^a = \mu^b$:

$$\eta_n = (\hat{\mu}^a - \hat{\mu}^b)' \sum^{-1} (\hat{\mu}^a - \hat{\mu}^b) \; \chi_K^2 \quad (46)$$

where $\sum = \sum_m^a / n_a + \sum_m^b / n_b$.

Similarly, equality of Lorenz curves may be tested as a hypothesis of equality of sums of conditional means, $l'\mu$, using the following statistic with an asymptotic (standard) Normal distribution:

$$S_l = l'(\hat{\mu}^a - \hat{\mu}^b)/(l' \sum l)^{1/2} \quad (47)$$

This statistic is, however, useful for testing $l'(\mu^a - \mu^b) \leq 0$ against the one sided alternative of $l'(\mu^a - \mu^b) > 0$, but not for the null of $\mu^a \geq \mu^b$. Thus it will have poor power against alternatives with a mixture of lower and higher group conditional means (intersecting curves).

Testing for Generalized Lorenz dominance requires cumulative means and variances, *i.e.*,:

$$\gamma_k = E[x|x \leq q_k], \text{ and } \omega_k^2 = E[(x - \gamma_k)^2 | x \leq q_k] \quad (48)$$

See Beach et al (1995) for the joint asymptotic distribution of the sample estimators of γ_k and ω_k^2 .

The vector of K Generalized Lorenz ordinates is $G = (p_1\gamma_1, ..., p_K\gamma_K)'$, where $p_l\gamma_l = \sum_{i=1}^{l} \mu_i \Delta p_i$ demonstrates the relation with the conditional means. Thus Beach and Davidson (1983) — see also Butler and McDonald (1989) — utilized the asymptotic distributions given earlier to demonstrate the asymptotic normality of $\sqrt{n}(G - G)$, where G is a consistent estimate of G with variance matrix $\Omega_g = (w_{ij})$ such that:

$$w_{ij} = p_i[\omega_i^2 + (1 - p_j)(q_i - \gamma_i)(q_j - \gamma_j) + (q_i - \gamma_i)(\gamma_j - \gamma_j)] \quad (49)$$

A test of **equality** of Generalized Lorenz curves may then be based on:

$$\eta_g = (G^a - G^b)' \hat{\Omega}^{-1} (G^a - G^b) \; \chi_K^2$$

(50)

where

$$\Omega = \Omega_g^a / n_a + \Omega_g^b / n_b$$

(51)

Beach et al (1995) are concerned with dominance relations when Lorenz curves cross at one or more points. Focusing on "transfer sensitive" inequality criteria, they propose tests that are in the spirit of the tests described above.

Similar procedures for testing mean-Lorenz and absolute Lorenz dominance are given in, for instance, Bishop et al. (1989).

Since the hypothesis of stochastic dominance is a set of inequality restrictions that is not well represented by the implicit nulls of the above tests, one may follow Bishop et al (1989) and Beach and Richmond (1985) by using the union intersection techniques in a multiple comparison of the individual elements of the moment vectors above, tested one at a time and sequentially. This requires size adjustments in order for **multiple comparison** inferences to be at the desired test size. Richmond's Studentized Maximum Modulus test is one alternative, see Beach and Richmond (1985), and Bishop et al. (1988*a,b*) and Bishop et al.(1989). Another is to use Bonferroni inequalities to make the size ajustments.

Thus η tests are like the Kullback-Leibler and Kolmogorov-Smirnov tests, see Maasoumi (1993) and McFadden (1989), respectively, in being useful to determine the "equality" of two distributions. But these tests are not all able to identify dominance or unrankable intersections. On the otherhand, tests such as S_1 above are appropriate for identifying dominance relations, but not for equivalent or intersecting distributions. Only the rather cumbersome union intersection tests allow a partial ordering of distributions into, dominated, dominant, equivalent, and non-comparable. That is, if the tests at all ordinates do not reject equality the two distributions are indifferent; If the tests at one or more ordinates reject equality in the same direction, we have dominated ranking. In other outcomes we may infer intersection (non-comparability).

The following summary of an empirical study in Bishop et al (1989) is indicative of the characteristics and the advantages of conducting **statistical** comparisons relative to the traditional numerical comparisons. Bishop et al. (1989) conduct a comparison of cost of living adjusted income distributions of the 50 states in the US with the corresponding distribution for the US as a whole. As they note, purely numerical comparison precludes the inference of "indifference". Secondly, only 18 states can be ordered by rank dominance, compared with 46 states using the statistical multiple comparison tests. In terms of Generalized Lorenz curves, numerical comparison ranks 32 states versus 47 statistically. And comparisons of ordinary Lorenz curves ranks 35 states numerically versus 50 (all) statistically. Three other findings are worth emphasizing. (i) Statistical inferences regarding rank dominance (Pareto principle) are quite decisive and render Generalized Lorenz comparisons redundant. The latter criterion, while equality prefering, appears to be heavily weighted toward efficiency (higher mean incomes dominate) for these data. (ii) Rank and ordinary Lorenz comparisons yield inferences that are frequently different from those obtained from joint mean absolute Lorenz comparisons (only 13 cases agree). (iii) Consistent with (ii), Kuznets' hypothesis of lower inequality for higher mean (richer) populations generally does (not) find statistical support when relative (absolute) measures of inequality are employed. It is worth noting, however, that the mean absolute Lorenz tests fail to rank very many of the states.

In a recent paper Bishop et al (1994) use CPS data to re-examine comparisons of U.S. income inequality over time. In particular, they find that adjustment for the practice of "top coding" can uncover an increase in inequality in the 1980s based on Lorenz dominance tests. Top-coding can cause omissions of some very high incomes with a bias effect that has been estimated at about the same order of magnitude as in-kind payments.

4.2.1. Recent Tests for Stochastic Dominance:

The above inference techniques have been extended to test for "stochastic dominance" relations of desired orders. There is some independent progress in this respect in D. McFadden (1989), A. Kaur, B.L.S. Prakasa Rao, and H. Singh (1994), K.Xu, G. Fisher, and D. Wilson (1995), Xu (1995), C. Beach, R. Davidson, and G. Slotsve (1995), and L. Klecan, R. McFadden, and D. McFadden (1991). McFadden (1989) contains a simple Fortran code for computing Smirnov tests of first and second order stochastic dominance in the case of independent "portfolios". For the test of first order dominance Smirnov statistic is computed which has a known distribution. For second order dominance a similar statistic is computed but without a known distribution. McFadden suggests Monte Carlo methods for computing its significance levels. Klecan et al (1991) extend this work by studying the asymptotic distribution of these tests when the assumption of independence is relaxed. Beach et al (1995) derive the joint asymptotic distribution of cumulative means and coefficients of variation to test for Lorenz dominance of "transfer sensitive" inequality measures when Lorenz curves cross. Their work, like the ones described above, see also G. Anderson (1994), tests "equality" hypotheses. For this reason they are also forced to apply Studentized Max Modulus (SMM) techniques for practical inferences. Xu (1995) and Xu et al (1995) are distinct, however, in that the weak inequality nature of the hypothesized dominance relation is recognized and appropriately tested. They adapt the test of inequality constraints due to Kodde and Palm (1986) and Wolak (1989) to test for first and second order (Generalized Lorenz) stochastic dominance on the basis of the sample quantiles. Theorem 2 above provides the central result that guides both the Beach et al (1995) and the Xu et al (1995) contributions. Maasoumi, Mills, and Zandvakili (1995) have bootstrap applications of the McFadden tests for income distributions. Xu et al (1995) report a financial application. Porter (1978) is a survey of earlier empirical applications which, however, assumed the distribution function is known rather than estimated.

4.3. Fitting Income Distributions

As was mentioned earlier, the indirect method of statistical inference on indices and welfare comparisons requires the determination and estimation of a parametric distribution function for the income data. Some generative theories have been put forward to justify certain families of distributions, but modern work has generally found greater success with statistical specification searches conducted over very general (flexible) functional families that subsume all the previously proposed cases as special cases.

Gibrat (1931), Rutherford (1955), Champernowne (1953), Sargan (1957), Mandelbrot (1960), and Dagum and Lemmi (1989) exemplify attempts at describing a causal model for the generation of income (or wealth) distributions. We refer the reader to the last citation above for a general account on the necessary properties of an income distribution, including some "stylized facts", and useful diffusion processes that underly many of the popular as well as statistically successful distribution families.

The most successful among the latter is the four parameter family of generalized beta of the second kind (GB2) studied by McDonald (1984) and others. It appears that the model generated and studied by Majumdar and Chakravarty (1990) is a reparametrization of GB2. Either of these two includes such special cases as the Singh-Maddala (1976) and Dagum (1977) models. A graph given in McDonald (1984) provides the relationship between the various distribution families and some old standbys as Lognormal, Pareto, Weibull, Exponential and Gamma; See Appendix II.

McDonald (1984) describes both the GB2 and a GB1. GB2 is also known as Feller-Pareto and Generalized *F*. Special and limiting cases of GB2 include Burr type 3 (Dagum), Burr type 12 (Singh-Maddala), Generalized Gamma (GG), Gamma, Lognormal, Weibull, Fisk, Rayleigh, chi-square, exponential and the Beta of the second kind (B2). The densities for GB2 and GB1 are given as follows:

$$\text{gb2}\,(x;a,b,p,q) = ax^{ap-1}/b^{ap}B(p,q)[1+(x/b)^a]^{p+q}\ , x \geq 0 \qquad (52)$$

$$\text{gb1}\,(x;a,b,p,q) = ax^{ap-1}[1-(x/b)^a]^{q-1}/b^{ap}B(p,q)\ 0 \leq x \leq b \qquad (53)$$

where $B(.,.)$ is the beta function. Many of the well known inequality indices and criteria may be shown to be functions of the moments of these two distributions, *i.e.*, their four parameters. For instance, for the Gamma distribution, Gini = $1/p.B(p, .5)$; see McDonald (1984).

Given an estimation criterion such as Minimum χ^2 or Maximum Likelihood, fitting GB1 and GB2 will provide better results than fitting any of their respective special cases. This general approach minimizes the risk of misspecification commonly associated with parametric methods. Indeed GB2 is shown by several authors to perform very well in a number of studies; see

McDonald (1984), and Majumdar and Chakravarty (1990). A comparison of Method of Moments estimates of the GE measures with the parametric estimates from GB2 has not been conducted yet.

To estimate the unknown parameters, denote n_r as the number of observations in $r - th$ of R income groups, $I_r = [x_{r-1}, x_r]$, $I_1 = [0, x_1]$, $I_R = [x_{R-1}, \infty]$. Efficient estimates using minimum χ^2 or MLE may be obtained based on the multinomial likelihood function. Briefly, if $p_r(\alpha) = \int_{I_r} f(x, \alpha)dx$ represents the probability of I_r according to the pdf $f(.)$, the min χ^2 estimators of α are obtained by minimizing:

$$\chi^2 = n \sum_{r=1}^{R} \frac{[p(\alpha) - n_r/n]^2}{p_r(\alpha)} \quad (54)$$

See Majumdar and Chakravarty (1990). MLE of α is obtained from the log-likelihood function:

$$l(\alpha) = \ln(n!) + \sum_{r=1}^{R} [n_r \ln(p_r) - \ln(n_r!)] \quad (55)$$

see McDonald (1984). These methods produce asymptotically equivalent estimators. In general there is no reason to expect these estimators to be optimal relative to such fit criteria as Sum of Squared Error (SSE) and Sum of Absolute Errors (SAE), commonly employed to judge how well a distribution fits the data. See McDonald and Mantrala (1993) for a comparative analysis that demostrates the role of groupings and the fit criteria in infering the "optimal" distribution of incomes. They use different estimation methods for constant 1991 dollar CPS income data for all families in the years 1970, 1980, and 1990. Likelihood ratio tests are available for the nested models such as the special cases of GB2 and GG. χ^2 tests tend to always reject due to the large sample sizes!. GB2 indeed does extremly well, followed by the Generalized Gamma and the Gamma. Lognormal is consistently the worst, by any of the fit criteria, including SSE and SAE.

Recently Wilfling (1992) and Wilfling and Kramer (1993) have given conditions, respectively, for the Lorenz dominance rankings amongst the members of the GB2 family and the Singh-Madala distributions. Wilfling gives a necessary and sufficient condition for the Burr distribution, and then a necessary and a sufficient condition for the GB2 to be dominant. Using the parameter estimates obtained by Singh-Maddala and McDonald, Wilfling and Kramer are able to conclude that only 1970 and 1965 US income distributions are Lorenz ranked. Thus rankings other than that achieved by Gini coefficient will be obtained with alternative inequality indices.

There is of course a non-parametric alternative to this first step, that of kernel and similar estimators. It would seem that the nonparametric alternative would be useful for "eyeballing" exercises that may indicate suitable functional families for income and other attribute distributions. The parametric method, or the direct MM method of estimating indices has advanced too far and too well, however, for further explorations in the non-parametric arena to be very informative. Finally we note the approximate small$-\sigma$ formulae of Maasoumi and Theil (1979) provide a method of estimating indices on the basis of the first four sample moments of the data. After all, this is all the information that is used to obtain Pearson-type approximations to unknown distributions.

Measurement Error:

The issue of measurement error in data does not admit universal statements or solutions. Elsewhere I have discussed some remedies, see Maasoumi and Nickelsburg (1988), and Maasoumi and Zandvakili (1990). Time aggregation to reduce transitory components, and categorization of even continuous variable observations, are examples of partial remedies. Sometimes a little summary information, such as group means and population sizes, captures much of the relevant information. Thus if within-group variation is insignificant grouping would be a useful means of avoiding measurement error without incurring a heavy cost in underestimating inequality indices; see Cowell and Mehta (1982). It should be pointed out, however, that the relative error in non-income data is likely exaggerated in comparison to estimated errors in such routine data as in the national income accounts. An example is the size of the "underground" economy as the systematically unreported component of national incomes. The latter error is possibly greater than in categorizing individuals in several age, educational, or quality of life groups. Of course, systematic errors arising out of poor survey designs admit some correction techniques that deserve more frequent application.

The difficulty in practice is to know when, if, and where measurement errors occur! An "extreme observation" is not necessarily an error, and unlike statistical averaging operations, inequality and poverty measures may be intended to be sensitive to precisely such outlying values. There is some recent attempt to develop "robustness" properties of estimates of inequality measures. This strikes me as somewhat odd since we know much about the theoretical robustness or sensitivity of such measures to extreme (tail area) values. For instance, if measurement errors occur in tail area observations, inequality indices with high degree of inequality aversion will be more affected than the Gini, say, which is relatively insensitive to tail areas. Evidently, this lack of "statistical" robustness does not provide a meaningful argument in favor of poverty insensitive welfare analysis.

Grouped data:

Much of income data made available by governments and other agencies is in the form of grouped data. This entails loss of information to various degrees. Determining the degree of loss is a rather complicated question which depends, among other things, on the "sufficient statistics" of the underlying distribution as well as the functional that defines the desired index to be estimated. A simple example may suffice. In the linear regression model, if it is desired to obtain the OLS estimates of the coefficients, the only statistics of consequence are the second moments of the data. Therefore, any grouping (aggregation) of the original data which would not (substantially) alter these moments or their estimates, entails no **consequential** loss of information. As we have seen different inequality indices are different functions of the moments of data and, in particular, reflect different degrees of sensitivity toward the tail areas of distributions. Understanding their sensitivity to grouping is essential for both data collectors and investigators.

From the existing literature, for instance see Cowell and Mehta (1982), Gastwirth (1972, 1975), Gastwirth and Glauberman (1976), Kakwani and Podder (1976), and Davies and Shorrocks (1989), several important facts emerge. One is that knowledge of group means as well as the group sizes is very important. Between-group inequality can be computed readily in order to establish a lower bound for total inequality. Finer lower and upper bounds have been obtained for Gini and the decomposable family of measures in the work cited above. When data are grouped by income levels, empirical studies have shown that "true" Gini is about 2/3 *rd* of the way from its lower bound, whereas members of the GE family are about 1/3 of the way from their lower bound. I find such statements to be in need of serious qualification. They can imply that "within group" inequality is generally small. This is not so. Firstly, the grouping in the above studies are by income, leaving potentially little income variation within income intervals. Grouping by other criteria, such as education, age, gender and race, can and often does leave a great deal of **income** variation within groups. Secondly, the above reported assessments of "1/3, 2/3" rules are empirical and based on samples which may not be representative across countries, regions, and time. The need for ungrouped micro-data that shed light on within group inequality is real and unmitigated by general approximation rules and interpolation methods.

To assess the performance of interpolation methods, for instance, Copwell and Mehta (1989) present a very good analysis of the relative efficacy and characteristics of such techniques as "split histogram", polynomial approximations, and Pareto and other density fits to within group intervals. When only group information is available all of the above techniques are perhaps uncharitably characterized as "information manufacturing", they all endow group intervals with within group information that is not data based. Genralizations of empirical experiences with such techniques are rather dangerous, especially across populations and data types.

There is a somewhat constructive line of attack on this problem in Davies and Shorrocks (1989). Here, while grouping by incomes is again the basis of the excercise, a potentially useful guideline is developed for the collector of grouped data. Davies and Shorrocks search for optimal grouping of data by varying group sizes in order to maximize inequality. They find that (i) it is better to have groups that almost equally share in incomes (not equal income intervals), and (ii) it is important to have a finer grouping of the extreme tail area than is typically done. Once again, the degree of inequality aversion of an index influences the results.

Many useful grouping computations are incorporated in a menu driven PC program written by F. Cowell at the London School of Economics.

A final comment is worthwhile on the performance of Method of Moments and approximation methods, such as the Maasoumi-Theil expansions alluded to earlier, when only grouped data is available. It is known that estimators of moments of order higher than the second can be quite unstable and imprecise. The loss of efficiency in these estimators due to grouping can be large, especially since they are asymptotically justified and grouping results to small sample sizes. Further, grouping may generally result in underestimation. Some may argue that, faced with this difficulty, fitting a distribution to either the whole data or to each group is called for. But reliable fits require sample information which is not available with grouped data.

5. Multivariable Welfare and Inequality

As may be surmized from the discussion in section 3.1, once the reality of heterogeneity amongst the members of a household and between households is admitted, the notion of "income inequality" itself becomes ambiguous. What types of heterogeneities are to be adjusted for, and how, before one may measure a "pure" income inequality?

There are at present three approaches that are at least partially motivated by the difficulties of distributional comparisons briefly alluded to earlier. The **first** is to search for measures that are in some sense less sensitive to incorrect methods of scaling incomes. Decomposable measures, whether of inequality or poverty, provide some protection. Additive decomposability offers an opportunity to "control" for heterogeneity sources that are classifiable when data are collected. Gender, age, education, income category, marital status, family size, race, ethnicity, geographic location, employment status, and many other attributes, are examples of very useful and observable characteristics which explain some sources of heterogeneity. As Coulter et al (1992) rightly argue, see also Maasoumi and Nickelsburg (1988), and Maasoumi and Zandvakili (1986, 1989, 1990), the between group component of the GE family of inequality measures is, inevitably and perhaps appropriately, the only component that is not free of how heterogeneity is defined *and*/*or* adjusted for. The within group components and their weights are free of such "contaminations". But decomposability has its limits, both practical and because it requires comparisons of possibly many conditional inferences.

A **second approach** in recent years has moved away from distributional comparisons on the basis of indices such as inequality measures. This requires an evaluation of rankings of the type discussed in section 2. In particular, the work of Atkinson, Bourguignon and others focuses on deriving conditions for the stochastic dominance of one distribution over another on the basis of welfare functions which are, in a sense, decomposed or separable for different population groups. This separation also opens the way for allowing different welfare evaluations for different groups. Welfare comparisons of this type do not require equivalence scales and quantitative separation of incomes and needs. Empirical implementation requires a sequential formation and comparisons of the Generalized Lorenz curves, defined earlier, starting with the "neediest group", then combining the two neediest groups if the first comparison test is passed, and so on. See Atkinson and Bourguignon (1987, 1989).

A **third** line of inquiry is the multidimensional approach to welfare analysis which could be conducted both in terms of distributional indices, and in terms of dominance relations between welfare functions. When attributes other than money incomes are taken into account and allowed to explicitly enter the SWFs, a powerful aspect of the axioms of "fundamentality" and "anonymity" may be invoked which partially justifies common representations in preferences, in functional forms and for "representative agent" formalisms. Simply put, if we enter into the preference functions all the attributes that we think would matter in distinguishing differences amongst individuals, then the need for ex post and often arbitrary distinctions in functional representations and other adjustments would be reduced. The difficulty is data availability, but this has been greatly remedied in recent years and seen to be far from an exclusive difficulty with the multidimensional approach.

Following this third line has led to an advocacy of welfare comparisons that go beyond money incomes. See Sen (1977), Atkinson and Bourguignon (1982), and Maasoumi (1986, 1989a). As we will note below, this approach can incorporate elements of the first approach as well by identifying measures of mulivariate inequality, say, which are decomposable and thus more robust to incorrect scale adjustment techniques that may still be employed.

Note that in the second approach mentioned above one still needs to define separable welfare functions for a priori distinct groups. This may be seen as a relaxation of the total "symmetry" requirement without which one could not claim any dominance relations free of needs considerations. This multidimensional approach can then be seen as also complementing this second approach by reducing the need for consideration of large numbers of groups which make empirical implementation difficult, and clearcut outcomes next to impossible. Further, under certain conditions, decomposability is possible both by population groups and by attributes; see Maasoumi (1986). Axiomatic conditions for this latter type of decomposability bear some further investigation.

In the way of a brief account of mutidimensional or multiattribute analysis, let x_{ij} denote the measurement of attribute $j = 1,2,...,m$, associated with individual (unit, household, country) $i = 1,2,....,n$. Define the welfare matrix $X = (x_{ij})$ and consider any scalar measure as a function of X and the joint distribution of the m attributes. Unlike the univariate case, it has proven difficult to develop "consensus" axioms which may characterize an ideal measure of multivariable inequality; see Tsui (1992a) for an excellent attempt. One of the main difficulties here is that, whatever the axiom sets, there is an inevitable aggregation of the m attributes that results in any scalar measure. In view of this truism, Maasoumi (1986) proposed a two step procedure whereby this aggregation issue is dealt with directly and explicitly. Once an "ideal" aggregation function is determined, the choice of an ideal measure of inequality may be guided by the extensive analysis of that issue in the univariate literature. The latter has consistently identified the GE family of inequality measures.

The aggregation of attributes in the first step has been addressed by several authors. Two broad approaches may be identified. The first is based on measures of closeness and affinity which may identify either attributes that are similar in some sense, *and/or* determine a "meanvalue", or aggregate, which most closely represents the constituent attributes. The second approach which is axiomatic lays down properties that we may agree an aggregate function should possess. This second approach, recently developed by Tsui (1992b), inherits the difficulties of arriving at consensus properties which parallel the difficulty of adopting a criterion of "closeness" in the first approach. But the latter difficulty has seen some resolution in information theory which also seems to justify members of the Generalized Entropy family as ideal "closeness" criteria. This topic is beyond the scope of the present chapter. The interested reader may see Maasoumi (1993).

Let us define the following GE measure of closeness or diversity between the m densities of m attributes and S_i, the aggregate "well-being" function for the $i - th$ unit:

$$D_\gamma(S, X; \alpha) = \sum_{j=1}^m \alpha_j \{\sum_{i=1}^n S_i[(S_i/x_{ij})^\gamma - 1]/\gamma(\gamma + 1)\} \qquad (56)$$

where $\alpha_j s$ are the weights attached to each attribute. Minimizing D_γ with respect to S_i such that $\sum S_i = 1$, produces the following "optimal" aggregation functions:

$$S_i \alpha (\sum_j^m \alpha_j x_{ij}^{-\gamma})^{-1/\gamma} \quad \gamma \neq 0, -1 \qquad (57)$$

$$S_i \alpha \Pi_j x_{ij}^{\alpha_j} \quad \gamma = 0 \qquad (58)$$

$$S_i \alpha \sum_j \alpha_j x_{ij} \quad \gamma = -1 \qquad (59)$$

These are, respectively, the hyperbolic, the generalized geometric, and the weighted means of the attributes, see Maasoumi (1986). These solutions include many of the more arbitrarily proposed aggregates in empirical applications. For instance, the weighted mean is equivalent to the popular principal components when $\alpha_j s$ are the elements of the first eigen vector of the $X'X$ matrix, see Ram (1982) and Maasoumi (1989a).

Elsewhere the author has proposed a measure of multidimensional inequality which is the GE index of the S_i distribution just obtained; see Maasoumi (1986) where the properties of this measure and its decomposability by groups as well as by attributes is discussed. There have been several applications of this measure to US data, to international data, and in "country studies". We will briefly present some of these application here. But it is worthwhile to first note a possible difficulty with potential "double counting" of the same latent welfare components by inclusion of measurements on apparently distinct attributes. This issue is studied by Hirschberg, Maasoumi, and Slottje (1991) for international data. The basic idea is to detect "clusters" of attributes which are statistically similar. Once this is accomplished, a "representative" aggregator attribute is chosen for each cluster. These representative or composite attributes are then included in the desired but lower dimensional multivariate welfare analysis. The approach of Hirschberg et al (1991) is based on statistical clustering techniques as well as a new entropy based criterion. In Hirschberg et al. (1991) 24 attributes of well being were analysed for 120 countries. These included such attributes as GNP and related concepts, literacy and mortality rates, labor force participation rates, basic amenities (*e.g.*, radios and roads), militarization indices, health status, infrastructure indicators, political freedom and civil liberty measurements; see Table 2 for a listing and definitions. Interesting and quite plausible "clusters" were identified based on several criteria of similarity. Table 2 is one example in which the "dots" indicate clustering in five categories on the basis of the minmax distance criterion. The authors then proceeded to compute aggregate measures of well being on the basis of the "representative" attributes for the five clusters, and computed Maasoumi's measures of multidimensional inequality. This type of study also allows an investigation of the robustness of inferences, for example, with respect to levels of aggregation (clustering), weighting factors, and aversion to inequality (parameter β of GE).

6. Further empirical applications :

The applications discussed here include measurement of world inequality in several attributes including incomes, dynamic analysis of inequality and mobility in the US based on panel data, and indices of US "well-being" based on several attributes. For details of data sources , variable definitions, and technical details, the original sources should be referenced.

*Table*5.2, $D_{kl}(4)$–five clusters*

N T R E D S D C F M M P H M R M U W P C H I C P
W E A N P O P C T L L G S – D F N L R L P N L H
S L D C L G M E T F S L I F B F F Y

NWS 6 6 6 6 6 6 6 6 6
TEL 6 6 6 6 6 6 6 6 6
RAD 6 6 6 6 6 6 6 6 6
ENC 6 6 6 6 6 6 6 6 6
DP 6 6 6 6 6 6 6 6 6
SOL 6 6
DPG 6 6
CC 6 6 6 6 6 6 6 6 6
FTM 6 6 6 6 6 6 6 6 6
MLE 6 6 6 6 6 6 6 6 6
MLT 6 6 6 6 6 6 6 6 6
PG 6 6 6 6 6 6 6 6 6 6
HS 6 6 6 6 6 6 6 6 6 6
$M-F$ 6 6 6 6 6 6 6 6 6 6
RDS 6 6 6 6 6 6 6 6 6 6
MFL 6 6
UNI 6 6
WLF 6
PR 6 6 6 6 6 6 6 6 6 6
CL 6 6 6 6 6 6 6 6 6 6
HPB 6 6 6 6 6 6 6 6 6 6
INF 6 6 6 6 6 6 6 6 6 6
CLF 6 6 6 6 6 6 6 6 6 6
PHY 6 6 6 6 6 6 6 6 6 6

* $D_{kl}(4) = \max_{i \in C_l, j \in C_k} ‖X_i - X_j‖$ is a complete linkage method of clustering between all clusters C_l and C_k. This is a MinMax selection rule.

NWS: number of daily newspapers HS: average household size
TEL: telephones per capita $M-F$: male — female literacy rates
RAD: radio receivers per 1000 people *RDS* : *nat'l* territory per km road
ENC: rate of energy consumption *MFL* : *male*/*female* life expectancy
per capita rate
DP: per capita real GDP UNI: uniform distribution
SOL: soldier to the civilian ratio WLF: % women in the labor force
DPG: growth rate in real GDP PR: political rights variable

CC: caloric consumption per capita CL: civil liberty
FTM : *female*/*male* literacy rates HPB: people per hospital beds
MLE: male life expectancy INF: infant mortality rate per 1000
MLT: male literacy rate CLF: % children in the labor force
PG: population growth rate PHY: people per physician

6.1 Multivariable International Inequality

In two recent papers, Maasoumi (1989a) and Maasoumi and Jeong (1985), information theoretic aggregate measures described above were utilized in a comparative analysis of well-being in the world. Traditional analysis in this area has hithertofore been based almost exclusively on GNP or GDP. Typically, the size distribution of one of these attributes would be characterized by a Gini measure of inequality. This approach suffers from two principal shortcomings. The first is the use of Gini index, and the second is a more widely understood and accepted problem

[Tables 5.3-5.4 go here].

with the exclusive identification of well-being with simple income measures. The latter is particularly troubling in international comparisons where many significant other factors cannot be controlled amongst diverse attitudes, cultures, and socio-economic arrangements for individual and social attainment of welfare. Other relevant attributes have now been measured and made available by international agencies. Among the problems that must be faced, measurement accuracy, "independence" of attributes (the double counting problem), how to combine the chosen attributes in composite indices, and the choice of an appropriate index of (*e.g*) inequality, may be mentioned. In previous sections we have indicated some methods for dealing with these problems.

The second difficulty, the almost exclusive reliance on the Gini measure in applied studies, is also resolved by using the family of Generalized Entropy (GE) measures. Here economic theory and information theory together help to clarify a revealing correspondence between degrees of relative aversion to inequality (tail area transformations which are so important for policy analysis) and the choice of summary indices of inequality, poverty, etc. Gini, being more sensitive to changes in the central area of distributions, is seen to be less appropriate than some entropy based measures for the study of dynamic distributional changes that may be expected as a result of typical development or tax-based policies.

In Maasoumi (1989a) per capita GNP was considered as well as two composite ("representative") indices of Basic Needs (BN) indicators and Physical Quality of Life Indicators (PQLI). Each of the latter is made up of other components reflecting well-being. The two entropy-based Theil measures were computed for each of these three attributes for a distribution of countries in the middle to late 70s. In addition, the first entropy based composite indices of GNP-BN and GNP-PQLI were proposed providing the first multi-dimensional measures of well-being based on information theory. As was stated earlier, this approach also provided an interpretation for some of the previously proposed composite measures, such as the Principal Components (PC) used by Ram (1982). Tables 3-4 show world inequality thus measured on the basis of the author's multidimensional versions of Theil's two measures (M_1 and M_2). For the country-specific values of the well-being indices the reader is referred to Maasoumi (1989a). The aggregation method described in section 5 above was used to compute the weighted arithmatic (S_1^*) and weighted geometric (S_2^*) means as composite indices of well-being in Tables 3-4.

The results in Tables 3-4 show that: (i) multiattribute inequality is lower than income inequality, but much higher than in "quality of life" indicators; (ii) within group inequality is lowest amongst the "industrialized" nations, and generally increases with less development or more oil(!); (iii) centrally planned economies of this era are the most unequal and varied group in income, but relatively equal in qualitative indicators; finally, between group inequality among these groups is about two or more times higher than average within group inequality.

Several choices of attribute weights were tried in the above study, but further work must be done on this question, on other indexing (aggregation) functions, and other members of the GE family. Indeed, the results on Maximum Entropy regression functions discussed recently in Maasoumi (1993) would suggest that even richer aggregate functions than the "CES" forms of section 5 can be derived if further side conditions are available in specific applications.

6.2 US Inequality

A similar application of these techniques to US panel data on (i) household total real income, (ii) net equity in housing, and (iii) education (years of schooling of household head) was reported in Maasoumi and Nickelsburg (1988). Michigan Panel Study of Income Dynamics (PSID) provides a rich source of data on several

thousand households, their demographic characteristics, their economic status in terms of income, education, housing equity, transfer payments, etc, starting in 1968. The above mentioned study focused on the two information measures of inequality proposed by Theil, denoted by I(.) in Table 5, for each of the three attributes, and for an aggregate function, *S*. *S* was computed as the generalized geometric mean of the three attributes. Table 5 is an example of the findings for four age groups in a random sample of 267 households, using Principal Components (PC) weights for the "α" coefficients in the aggregation functions. The

Gini coefficient of “income” is also given for comparison.

In Table 5 we first note that Gini is rather insensitive both accross groups and over time. This may be contrasted with Theil’s measure of income inequality. Secondly and as expected, between group inequality is generally higher than within age group inequality, especially in education where their ratio does not decline over time. This between group inequality is picked up by the multivariate measure which moderates the corresponding income inequality values. A moderate decline in “inequality” is indicated over the decade of 70s because of a moderate decline in inequality between age groups. Interpreting the “within group” component as an age corrected (independent) inequality value, we see that it is both relatively small and unchanging.

In the above two applications, the important decomposition properties of the GE measures prove crucial in successfully controlling for such group characteristics as economic system, geographical location, level of industrialization, and GNP level in the case of countries, and for level of education, age, race, gender and income level, in the case of individuals and households in the second study.

[Table 5.5 goes here].

6.3 Time Aggregated Measures of Income

A third set of applications is the subject of Maasoumi and Zandvakili (1986,1989 and 1990). Whereas aggregation over different attributes was the focus of the previous studies above, **aggregation over time** for a single attribute, household

income, is studied here. Earlier, Shorrocks (1978, 1981) had considered simple sum of incomes over time and proposed a measure of “mobility”. In contrast, this author’s more general CES aggregates can allow for general income substitution and weighting schemes between different time periods. This can provide for more plausible means of controlling for transitory changes in attribute distributions, and provide mobility profiles as the aggregation interval is lengthened. Tables 6-7 depict a sample of the findings in our cited studies.

In Table 6 the computed inequalities are for real annual household incomes. There are four panels, one for each level of aversion to inequality (β). In each panel there are thirteen years of inequality decomposed between and within four (4) groupings done by the years of schooling of the household head. Education is clearly an “equalizer” of incomes; the higher the educational attainment of a group the lesser is the inequality within that group. This finding is remarkably robust to values of β. Changing parameters like β in search of robust inferences is strongly recommended. We note that most of the total inequality is within educational groups, not between them. The ratio of within to between group inequality does depend on the aversion parameter, but this is hardly surprising since large β indices are very sensitive to very small or very large incomes within some groups. Finally, we note that, whatever β, inequality has generally increased among the PSID households.

Table 7 is similarly arranged except that in each panel the thirteen entries represent successive levels of income aggregation. Annual, snap-shot volatility in inequality observed in Table 6 is absent. The profiles show that 13 year incomes are about as unequal as they were in the first year of 1968.

Once again the decomposition properties of the GE measures allow us to control for such characteristics as race, education, age, sex and income level. In the absence of control for such “co-factors”, the previous literature in this area has been unconvincing in its attribution of distributional changes and shifts to policy decisions and other socio-economic conditions.

Similar applications to the definition and measurement of industrial *and*/*or* trade “concentration” is possible and is of considerable value in economics and political science. Several recent collected volumes (see references) contain modern empirical studies in this field.

7. Concluding Remarks

This survey has discussed a selection of theoretical and statistical techniques that have been developed in recent years and which promise to come to widespread use in empirical studies. Some omissions from this coverage have been unavoidable. Among the more consequential of these in the empirical literature one should mention the sound work on earnings mobility and wage differentials, see the recent book by Atkinson et al. (1992) for a survey. We have also neglected the literature on measuring wealth and its distribution, as well as numerous country studies, especially those dealing with issues of development policy, growth, poverty and taxes. Much of this vast literature is difficult to interpret, however, since its findings generally depend on the Gini coefficient, coefficient of variation, and the variance of logarithms. These are inequality (dispersion) indices which either violate some near-consensus and elementary welfare properties, or some very useful additive decomposability characteristics which this author regards as practically indispensable.

[Tables 5.6-5.7 go here].

APPENDIX

In order to estimate the variance of *GI*, the Gini coefficient, we note that:

$d = \frac{1}{n}\sum_{i=1}^{n} d_i$, $d_i = \frac{1}{n-1}\sum_{j=1}^{n-1} d_{ij}$, and $d_{ij} = h_i h_j |x_i - x_j|$. For sample variances:

$var(d_{ij}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n-1} d_{ij}^2 - d^2 = \frac{2n}{n-1}(m_{20}m_{21} - m_{21}^2) - d^2$ (A1)

$var(d_i) = \frac{1}{n}\sum_{i=1}^{n} d_i^2 - d^2 = \frac{1}{n}\sum_{i=1}^{n} q_i^2 - d^2$ (A2)

where q_i is computed iteratively for each order statistic using the sample (h_i^*, y_i), $i = 1,2,....,n$, of observations ordered by incomes:

$q_i = \frac{h_i^*}{n-1}\{\sum_{j=1}^{i} 2h_j^*(y_i - y_j) + nm - ny_i m_{10}\}$ (A3)

From the definition of $var(d)$ in the text and (A1)-(A2), a consistent estimate of $var(d)$ is given by:

$2\,\{2(n-1)var(d_i) - var(d_{ij})\}/(n-2)(n-3)$ (A4)

See Cowell (1989), Gastwirth (1974), and Glasser (1962) for more details.

The rest of the terms in var (*GI*) are estimated as follows:

$v\hat{a}r(m_{10}) = (m_{20} - m_{10}^2)/(n-1)$ (A5)

$v\hat{a}r(m_{11}) = (m_{22} - m_{11}^2)/(n-1)$ (A6)

$c\hat{o}v(m_{10}, m_{11}) = (m_{21} - m_{10}m_{11})/(n-1)$ (A7)

$c\hat{o}v(d, m_{10}) = 2 \text{ cov } (d_i, h_i)/(n-2)$

$= 2\{\frac{1}{n}\sum_{i=1} h_i d_i - m_{10}d\}/(n-2)$ (A8)

$c\hat{o}v(d, m_{11}) = 2 \text{ cov } (d_i, h_i x_i)/(n-2)$

$= 2\{\frac{1}{n}\sum_{i=1}^{n} h_i x_i d_i - m_{11}d\}/(n-2)$ (A9)

References

Aigner, D.J., and A.J. Heins (1967),"A social welfare view of the measurement of income inequality," **Review of Income and Wealth**, 13, 12-35.

Amemiya, T. (1985). **Advanced Econometrics**, Harvard University Press, Cambridge, MA.

Anderson, G.J. (1994),"Nonparametric tests of stochastic dominance in income distributions," Working paper (mimeo.), University of Toronto.

Atkinson, A. (1970),"On the Measurement of Inequality," **Journal of Economic Theory**, II, 244-63.

Atkinson, A., and F. Bourguignon (1982),"The comparison of multi-dimensioned distributions of economic status,"**Review of Economic Studies**, 49, 183-201.

______________________________ (1987),"Income Distribution and Differences in Needs," in G.R. Feiwel (ed.), **Arrow and the Foundations of the Theory of Economic Policy**, Macmillan: London.

_____________________________ (1989),"The design of Direct Taxation and Family Benefits," **Journal of Public Economics**, 41, 3-29.

Atkinson, A., F., Bourguignon, F., and C. Morrison (1992). **Empirical studies of earnings mobility**, Fundamentals of Pure and Applied Economics, vol. 52, Distribution Section, Harwood Academic.

Beach, C.M. and R. Davidson (1983),"Distribution-free statistical inference with Lorenz curves and income shares," **Review of Economic Studies**, 50, 723-735.

_________and J. Richmond (1985),"Joint confidence intervals for income shares and Lorenz curves," **International Economic Review**, 26, 439-450.

Beach, C.M., R. Davidson, and G.A. Slotsve (1995),"Distribution-free statistical inference for inequality dominance with crossing Lorenz curves," Unpublished paper N95A03, GREQAM, Universites $d^{\vee}$*Aix*-Marseille .

Bishop, J. A., J. P. Formby, and P. D. Thistle, (1989),"Statistical inference, income distributions, and social welfare," **Research on Economic Inequality**, 1, 49-82.

Bishop, J. A., S. Chakraborti, and P.D. Thistle (1988a),"Asymptotically distribution free inference for generalized Lorenz curves," **Review of Economics and Statistics**.???

______________________________(1988b),"Relative inequality, absolute inequality, and welfare: Some statistical tests," **Economics Letters**, 26, 291-294.

Bishop, J.A., *J – R.* Chiou, and J. Formby (1994),"Truncation bias and the ordinal evaluation of income inequality," **Journal of Business & Economic Statistics**, 12, 1, 123-128.

Blackorby, C. and D. Donaldson (1978),"Measures of relative equality and their meaning in terms of social welfare," **Journal of Economic Theory**, 18, 59-80.

Blundell, R. and A. Lewbel (1991),"The information content of equivalence scales," **Journal of Econometrics**, Annals issue, 50, *No.*1/2, 49-68.

Bourguignon, F. (1979),"Decomposable Income Inequality Measures," **Econometrica**, 47, 901-20.

Butler, R.J., and J.B. McDonald (1989),"Using incomplete moments to measure inequality," **Journal of Econometrics**, 42, 1, 109-120.

Champernowne, D.G. (1953),"A model for inome distribution," **Economic Journal**, 53, 318-351.

Coulter, F.A.E., F.A. Cowell, and S.P. Jenkins (1992),"Differences in Needs and Assessments of Income Distributions," **Bulletin of Economic Research**, 44, 2, 77-124, University of Essex.

Cowell, F.A. (1989),"Sampling variance and decomposable inequality measures," **Journal of Econometrics**, 42, No. 1, 27-42.

Cowell, F.A., and K. Kuga (1981),"Inequality Measurement: An Axiomatic Approach," **European Economic Review**, 15, 287-305

Cowell, F.A. and F. Mehta (1982),"The estimation and interpolation of inequality measures," **Review of Economic Studies**, 49, 273-290.

Cramer, H. (1946), **Mathematical Methods of Statistics**, Princeton University Press.

Dagum, C. (1977),"A new model of personal income distribution : Specification and estimation," **Economie Appliquee**, 30, 413-437.

_________ (1990),"On the relationship between income inequality measures and social welfare functions," **Journal of Econometrics**, Annals Issue, 43, 91-102.

_________ (1993),"The social welfare bases of Gini and other income inequality measures," **Statistica**, anno LIII, no. 1, 1-30.

Dagum, C., and A. Lemmi (1989),"A contribution to the analysis of income distribution and income inequality, and a case study: Italy," in D.J. Slottje (ed.) **Research on Economic Inequality**, vol. 1., 123- 157, JAI Press.

Dalton, H. (1920),"The measurement of the inequality of incomes," **Economic Journal**, 30, 348-361.

Dasgupta, P., A. Sen, and D. Starrett (1973),"Notes on the measurement of inequality," **Journal of Economic Theory**, 6, 180-187.

Davies, J.B. and A.F., Shorrocks (1989),"Optimal grouping of income and wealth data," **Journal of Econometrics**, 42, 97-108.

Gastwirth, J.L. (1972),"The estimation of the Lorenz curve and Gini index," **Rev. Econ. and Statis.**, 63, 3, 306-316.

_______________ (1974),"Large sample theory of some measures of inequality," **Econometrica**, 42, 191-196.

______________ (1975),"The estimation of a family of measures of economic inequality," **Journal of Econometrics**, 61-70.

Gastwirth, J.L., and M. Glauberman (1976),"The interpolation of the Lorenz curve and Gini index from grouped data," **Econometrica**, 44, 3, 479-483.

Gastwirth, J.L., T.K. Nayak, and A.N. Krieger (1986),"Large sample theory for the bounds on the Gini and related indices from grouped data," **Journal of Business and Economic Statistics**, 4, 269-273.

Gastwirth, J.L., T.K. Nayak, and J.-L. Wang (1989),"Statistical properties of measures of between-group income differentials," **Journal of Econometrics**, 42, No.1, 5-19.

Gibrat, R. (1931). **Les Inegalites Economiques**, Sirey, Paris.

Glasser, G.J. (1962),"Variance formulas for the mean difference and the coefficient of concentration," **Journal of the American Statistical Association**, 57, 648-654.

Hirschberg, J.G., E. Maasoumi, and D.J. Slottje (1991),"Cluster analysis for measuring welfare and quality of life across countries," **Journal of Econometrics**,

Annals issue, 50, *No*.1/2, 131-150.

Hong, Y. and H. White (1993),"Consistent nonparametric entropy-based testing," Mimeo., UCSD.

Juhn, C. , K.M. Murphy, and B. Pierce (1991),"Accounting for the slowdown in Black-White wage convergence," In **Workers and Their wages**: **Changing Patterns in the United States**., Marvin H. Kosters (Ed.) Washington: Americam enterprise Institute Press.

__________________________ (1993),"Wage inequality and the rise in returns to skill," **Journal of Political Economy**, 101, no.3., 410-442.

Kakwani, N.C., and N. Podder (1976),"Efficient estimation of the Lorenz curve and the associted inequality measures from grouped observations," **Econometrica**, 44, 137-149.

Kaur, A., B.L.S. Prakasa Rao, and H. Singh (1994),"Testing for second-order stochastic dominance of two distributions," **Econometric Theory**, 10, 849-866.

Klecan, L., R. McFadden, and D. McFadden (1991),"A robust test for stochastic dominance," Working paper, Economics Dept., MIT.

Kodde, D.A., and F.C. Palm (1986),"Wald criteria for jointly testing equality and inequality restrictions," **Econometrica**, 50, 1243-1248.

Kolm, S. C. (1969),"The optimal production of social justice," in **Public Economics**, J. Margolis and H. Guitton (eds.), MacMillan: London, 145-200; Also presented at IEA conference on Public Economics, Biaritz, 1966.

Kolm, S.C. (1976a),"Unequal inequalities : I," **Journal of Economic Theory**, 12, 416-442.

___________(1976b),"Unequal inequalities : II," **Journal of Economic Theory**, 13, 82-111.

Lillard, L., and Willis, R.(1978),"Dynamic Aspects of Earnings Mobility," **Econometrica**, 46, 985-1012.

Maasoumi, E.(1986),"The Measurement and Decomposition of Multi-Dimensional Inequality," **Econometrica**, 54, 991-97.

Maasoumi, E.(1989a),"Composite Indices of Income and Other Developmental Indicators : A General Approach," **Research in Income Inequality**, forthcoming.

_____________(1989b),"Continuously ditributed attributes and measures of mutivariate inequality," **Journal of Econometrics**, 42, 1, 131-144.

___________. (1991), **Measurement of Welfare and Inequality**, (Ed.), Annals of Econometrics, **Journal of Econometrics**, Vol. 50, No. 1/2.

___________. (1993),"A Compendium to Information Theory in Economics and Econometrics," **Econometric Reviews**, Vol 12, 3, 1-49.

Maasoumi, E., and $J-H$., Jeong (1985),"The trend and the measurement of world inequality over extended periods of accounting," **Economics Letters**, 19, 295-301.

Maasoumi, E., and G. Nickelsburg (1988),"Multivariate Measures of Well Being and an Analysis of Inequality in the Michigan Data," **Journal of Business and Economic Statistics**, 6, 3, 327-334.

Maasoumi, E. and H. Theil (1979),"The effect of the shape of the income distribution on two inequality measures," **Economics Letters**, 4, 289-291.

Maasoumi, E., and S. Zandvakili (1986),"A Class of Generalized Measures of Mobility with Applications," Economics Letters, 22, 97-102.

_________________,(1989),"Mobility Profiles and Time Aggregates of Individual Incomes," **Research on Income Inequality**, 195-218..

__________________,(1990),"Generalized entropy measures of mobility for different sexes and income levels," **Journal of Econometrics**, 121-133.

Maasoumi, E., J. Mills, and S. Zandvakili (1995),"Bootstrapped tests of dominance for income distributions in the US," Mimeo, Department of Economics, SMU, Dallas, TX.

Majumdar, A.,and S.T., Chakravarty (1990),"Distribution of personal income: Development of a new model and its application to U.S. income data," **Journal of Applied Economics**, 5, 189-196.

Mandlebrot, B. (1960),"The Pareto-Levy law and the distribution of income," **International Economic Review**, 1, 79-

McDonald, J. (1984),"Some generalized functions for the size distribution of income," **Econometrica**, 52, 647-663.

McDonald, J.B., and R.J. Butler (1990),"Regression models for positive random variables," **Journal of Econometrics**, 43, No. 1/2, 227-251.

McDonald, J.B., and A. Mantrala (1993),"Apples, Oranges and Distribution Trees," Working Paper, Bringham young University.

McFadden, D. (1989),"Testing for stochastic dominance," in Part II of T. Fomby and T.K. Seo (eds.) **Studies in the Economics of Uncertainty** (in honor of J. Hadar), Springer-Verlag.

Moyes, P. (1987),"A new concept of Lorenz dominance," **Economics Letters**, 23, 203-207.

Pollak, R.(1991),"Welfare comparisons and situation comparisons," **Journal of Econometrics**, Annals issue, 50, *No*.1/2, 31-48.

Pollak, R., and T. Wales (1979),"Welfare comparisons and equivalence scales," American Economic review, 69, 216-221.

Porter, R. (1978),"Portfolio applications: Empirical studies," in G. Whitmore and M. Findley (eds.) **Stochastic Dominance**, Lexington: Lexington Books, pp. 117-161.

Ram, R.(1982),"Composite Indices of Physical Quality of Life, Basic needs Fulfillment, and Income : A 6Principal Component6 representation," **Journal of Development Economics**, 11, 227-47.

Rutherford, R.S.G.(1955),"Income distributions: A new model," **Econometrica**, 23, 277-294.

Saposnik, R. (1981),"Rank dominance in income distribution," **Public Choice**, 36, 147-151.

____________(1983),"On evaluating income distributions: Rank dominance, the Suppes-sen grading principle of justice, and Pareto optimality," **Public Choice**, 40, 329-336.

Sargan ,J.D. (1957),"The distribution of wealth," **Econometrica**, 25, 568-90, reprinted as Chapter 3, in E. Maasoumi (ed.) Contributions to Econometrics : J.D. Sargan, Vol. 1, (1988), Cambridge University Press.

Sen, A. (1970) **Collective choice and social welfare**, Holden Day: San Francisco, (reprinted, North-Holland, 1979).

________(1977) ,"On weights and measures: Informational constraints in social welfare analysis," **Econometrica**, 45, 7, 1539-1572.

Sen, P. K. (1972),"On the Bahadur representation of sample quantiles for sequences of φ-mixing random variables," **Journal of Multivariate Analysis**, 2, 77-95.

Shorrocks, A.F. (1978),"Income Inequality and Income Mobility," **Journal of Economic Theory**, 19, 376-93.

__________(1980),"The class of additively decomposable inequality measures," **Econometrica**, 48, 613-625.

__________(1981),"Income Stability in the United States," Ch. 9 in **The Statics and Dynamics of Income**, Klevmarken, N.A., and J.A. Lybeck (eds.), Tieto Ltd, Bristol.

___________(1983),"Ranking income distributions," **Economica**, 50, 3-17.

__________(1984), "Inequality Decomposition by population Subgroups," **Econometrica**, 52, 1369-85.

Siddiqui, M.M.(1960),"Distribution of quantiles in samples from a bivariate population," **Journal of Research of the National Bureau of Standards**, Mathematics and Mathematical Physics, 64B, 3, 145-150.

Singh, S.K., and G.S. Maddala (1976),"A function for the size distribution of incomes," **Econometrica**, 44, 963-970.

Slesnick, D.T. (1991),"Normative index numbers," in E. Maasoumi (1991), cited above, **Journal of Econometrics**, pp. 107-130.

Slottje, D.J.,(1989), **Econometric Issues in Analysing Inequality**, **Part I**: **Estimation**
Special Annals Issue of **Journal of Econometrics**.

______________(1990), **Economic issues in analysing inequality**. **Part II**: **Modelling methods**; Special Annals Issue of **Journal of Econometrics**.

Theil, H. (1967), **Economics and Information Theory**, North-Holland, Amsterdam.

__________(1989),"The development of international inequality : 1960-1985," **Journal of Econometrics**, Annals issue, 42, No.1, 145-155.

Tsui, $K-Y$. (1992a),"Multidimensional generalizations of the relative and absolute inequality indices: The Atkinson-Kolm-Sen approach," Chinese University of Hong Kong, Mimeo. November.

Tsui, $K-Y$. (1992b),"Composite indices of well-being: Axiomatic foundations," Chinese Univ. of Hong Kong, Dept. of Economics, Mimeo., September.

U.S. Bureau of the Census, Current Population Reports : Consumer Income, Series $P-60$ (various years).

van Praag, B.M.S. (1991),"Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept," **Journal of Econometrics**, special Annals edited by Maasoumi cited above.

Wilfling, B. (1992),"A sufficient condition for Lorenz domination of generalized beta distribution of the second kind," mimeo, University of Dortmund, Germany.

Wilfling, B., and W. Kramer (1993),"The Lorenz-ordering of Singh-Maddala income distributions," **Economics Letters**, 43, 53-57.

Wolak, F. (1989),"Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models," **Econometric Theory**, 5, 1-35.

Xu, K., G. Fisher, and D. Wilson (1995),"New distribution-free tests for stochastic dominance," Working paper No. 95-02, February, Dept. of Economics, Dalhousie University, Halifax, Nova Scotia.

Xu, K. (1995),"Asymptotically distribution-free statistical test for generalized Lorenz curves:An alternative approach," Dept. of Econ., Dalhousie University, Halifax, Nova Scotia, presented at the 7th World Congress of Econometric Society, Japan.