Journal of Econometrics 42 (1989) 131–144. North-Holland

CONTINUOUSLY DISTRIBUTED ATTRIBUTES AND MEASURES OF MULTIVARIATE INEQUALITY*

Esfandiar MAASOUMI

Southern Methodist University, Dallas, TX 75275, USA

Taking certain aggregates of several continuously distributed welfare attributes as either the *actual* utility or the 'observer's' evaluation of individual welfare, Theil's information measures of inequality in the aggregates are derived and analyzed. The important special case where the joint distribution is the multivariate log-normal is developed as a point of departure. When this joint distribution is not the log-normal, simple approximate formulae for Theil's indices are derived which are readily estimable. An attribute's incremental contribution to multidimensional inequality is derived and analysed.

1. Introduction

In recent contributions the value and the necessity of a multi-dimensional approach to the analysis of welfare and inequality has been argued and assessed by Kolm (1977), Atkinson and Bourguignon (1982), and Maasoumi (1986). In the latter, a number of indices were proposed for the actual measurement of multivariate inequality. When individual utility or welfare is a function of several arguments, the approach suggested in Maasoumi (1986) is to find an aggregate attribute for each individual with a distribution which is in some sense 'closest' to the distributions of all the relevant arguments. Some generalized information criteria of 'divergence' between distributions were employed to assess 'closeness'. Multidimensional inequality was then defined as the inequality in the distribution of the aggregate attribute. The 'ideal' functionals for such aggregates turn out to have the same form as many of the traditional utility functions such as the CES, Cobb–Douglas, etc. Indeed, an alternative interpretation of this approach is that individuals are represented (and ordered) according to their utility functions, and any measure of inequality is equivalent to the specification of a particular Social Welfare Function (SWF) with these utilities as its arguments, see Atkinson (1970) and Kolm (1977).

This paper has several purposes. Following the approach outlined above, we derive and analyse two central inequality measures proposed by Theil (1967)

*Funding support through a number of Indiana University Research Fellowships and grants is gratefully acknowledged. Thanks to Sharon Koga for her skillful typing.

0304-4076/89/$3.50

when several attributes are *continuously* distributed. The special case of a multivariate log-normal distribution is first developed when the attribute aggregator (utility) functions are of the Cobb–Douglas form. This special case is useful in two respects. In the first instance it helps to explicate the important role of the variances and covariances amongst the attributes in determining multivariate inequality. Some conditions of second-order stochastic dominance derived in Atkinson and Bourguignon (1982) are clearly brought out by this central case. Secondly, since many attributes' distributions are generally different from the log-normal, we derive approximate formulae for the inequality indices when the exact distribution is unknown. It is seen that the role played here by the log-normal distribution is similar to that played by the normal distribution in the development of approximate finite sample distribution and moments of econometric estimators. The approximate formulae are derived directly by series expansions of the formulae that define Theil's indices in the continuous case. We find that larger variances, larger positive covariances among the attributes, larger positive skewness, and larger leptokurtosis contribute positively to multivariate inequality. Generalizations of the results in Theil (1967) and Maasoumi and Theil (1979) for the univariate case are incidental.

Section 4 provides some insights concerning the incremental contribution to inequality of each attribute, and sufficient conditions are given under which an *additional* attribute may add to inequality.

The paper is organized as follows: Section 2 briefly describes Theil's indices in the discrete and continuous cases, for any attribute and for their aggregates. Section 3 derives and analyzes the approximate formulae for the same indices without the assumption of log-normality, and an appendix reports some numerical evidence on the measures and the approximate formulae. Section 4 analyzes the incremental value of attributes. Concluding remarks are in section 5.

2. Multidimensional inequality and the multivariate log-normal attributes

Based on the concepts of entropy and 'information', Theil (1967) proposed two scale-invariant inequality indices, *T1* and *T2*, as follows:

$$T1(x_f) = \sum_{i=1}^{N} x_{if} \log N x_{if}, \tag{1}$$

$$T2(x_f) = -\log N - \frac{1}{N} \sum_{i}^{N} \log x_{if}, \tag{2}$$

where $x_f = (x_{1f}, \ldots, x_{Nf})'$, $x_{if} = X_{if}/\sum_j^N X_{jf}$, and X_{if} is the quantity of an attribute, $f = 1, \ldots, M$, received by the ith individual, $i \in [1, N]$. Both measures are *additively* decomposable by inequality 'within' and 'between' population subgroups. Both belong to the class of decomposable inequality measures (the Generalized Entropy family) which are homogeneous (mean independent), symmetric, and satisfy the Pigou–Dalton Principle of Transfers. Their desirability and unique properties are indicated by the axiomatic studies of Bourguignon (1979), Cowell and Kuga (1981), Foster (1983), and Shorrocks (1980, 1984). Following the approach of Atkinson (1970), Blackorby and Donaldson (1978) analyze a class of welfarist Social Welfare Functions (SWFs) which imply and are implied by these two and other inequality measures. The work of Shorrocks (1980, 1984), in particular, suggests clearly that, if *unambiguous* additive decomposability is required in a study, the choice of Theil's measures, specially the second, is not as arbitrary as it may have seemed.

Let $X_i = (X_{i1}, \ldots, X_{iM})'$ be the ith row of the attribute matrix $X = (X_{if})$ and let the *share matrix* $x = (x_{if})$ be similarly defined. Let $S_i = h_i(X_i)$ denote a suitable aggregator of the elements of X_i and $S^* = (S_1^*, \ldots, S_N^*)$, such that $S_i^* = S_i / \sum_j S_j$, denotes the distribution of the 'aggregate attribute'. The inequality in S^* is referred to here as multivariate inequality. Consequently, the proximity of this distribution to those of its constituent variables is a useful requirement that ensures the preservation of the data 'information' in a way that is relevant for the measurement of distributional inequality. Maasoumi (1985) demonstrates that when $S_i = \prod_{f=1}^{M} X_{if}^{\alpha_f}$, the corresponding S^* distribution is minimally divergent from the M distributions, $x^f = (x_{1f}, \ldots, x_{Nf})$, $f \in [1, M]$. The underlying information criterion is the generalized Kullback–Leibler measure:

$$D_0(S^*, x; \boldsymbol{\alpha}) = \sum_{f=1}^{M} \alpha_f \left\{ \sum_{i=1}^{N} S_i^* \log\left(S_i^* / x_{if}\right) \right\}, \tag{3a}$$

which is minimized when S_i has the Cobb–Douglas form. An alternative is to use the following criterion:

$$D_{-1}(S^*, x; \boldsymbol{\alpha}) = \sum_{f}^{M} \alpha_f \left\{ \sum_{i} x_{if} \log\left(x_{if} / S_i^*\right) \right\}. \tag{3b}$$

According to this criterion, the 'ideal' aggregator is $S_i = \sum_f^M \alpha_f x_{if}$, a linear utility function.[1] One advantage of (3a) over (3b) is that it computes all the divergences from a *common* distribution, the S^*. Another advantage is that the

[1] Based on the notion of ϕ-entropy, further generalizations of information criteria have been considered in Maasoumi (1986). The CES and other utility functions prove 'ideal' with respect to these more generalized criteria.

Cobb–Douglas function allows for some substitutability between the attributes, whereas the linear function associated with (3b) is too restrictive in this regard. In the multivariate context substitution effects are clearly important and should be taken into account. However the function S_i is defined, interpreted, and (possibly) estimated, it is clearly seen that its form introduces a new degree of arbitrariness in any multivariate approach which allows for interaction between the attributes.

When the constituent attributes are treated as *continuous* variables, we denote their continuous aggregate by the variable S. Then, Theil's measures are defined as follows:

$$T1(S) = \mathrm{E}[(S/\mathrm{E}S)\log(S/\mathrm{E}S)], \tag{4}$$

$$T2(S) = \mathrm{E}\log(\mathrm{E}S/S) = \log \mathrm{E}S - \mathrm{E}\log S. \tag{5}$$

In the remainder of this section we analyze the important special case in which the S_i takes the Cobb–Douglas form and the joint distribution of the M attributes is the multivariate log-normal. Thus, let $z = (z_1, \ldots, z_M)'$ represent the continuously distributed attributes and $\lambda_f = \mathrm{E}z_f$, $f = 1, \ldots, M$. Following Theil (1967), we note that $x_{if} = z_f/N\lambda_f$ and $S_i^* = S/N\mathrm{E}(S)$. If z has a joint log-normal distribution with mean $\mu = (\mu_1, \ldots, \mu_M)$, $\mu_f = \mathrm{E}\log z_f$, and covariance matrix $\Sigma = (\sigma_{fk})$, then $S = \prod_{f=1}^{M} z_f^{\alpha_f}$ will also have a log-normal distribution; see Aitchison and Brown (1957). Formally, it may be verified that

$$S \sim \Lambda(\alpha'\mu, \alpha'\Sigma\alpha). \tag{6}$$

Using the results in Kendall and Stuart (1969, p. 169), we find that

$$\log \mathrm{E}S = \tfrac{1}{2}\alpha'\Sigma\alpha + \alpha'\mu. \tag{7}$$

From (5)–(7), Theil's second measure is obtained as follows:

$$T2(S) = \tfrac{1}{2}\alpha'\Sigma\alpha \geq 0 \tag{8}$$

$$= \sum_{f=1}^{M} \alpha_f^2\left(\tfrac{1}{2}\sigma_{ff}\right) + \sum_{f=1}^{M} \alpha_f \sum_{f<k}^{M} \sigma_{fk}\alpha_k. \tag{9}$$

Either by direct calculations or by analogy with the univariate case, we find

$$T1(S) = \tfrac{1}{2}\alpha'\Sigma\alpha = T2(S). \tag{10}$$

Expressions (8) and (10) may be recognized as generalizations of the univariate result in Theil (1967). For any log-normal variate z_f, both of Theil's measures are given by $\frac{1}{2}\sigma_{ff}$, half the variance of $\log z_f$. From (9) we may deduce that multivariate inequality is made up of two parts: a weighted sum of attribute inequalities and, in the second term of (9), adjustments due to covariation in the attributes (see section 4 for a similar decomposition). In the terminology of Atkinson and Bourguignon (1982), a multivariate distribution (second-order) dominates another if it has *lower* variances and covariances. In particular, when their marginal distributions have identical variances (σ_{ff}), lower (or negative) covariances indicate greater equality. These observations follow from the homogeneity of Theil's measures *and* the fact that their implicit SWF's belong to the class u^{--} defined in Atkinson and Bourguignon (1982). $T1(S)$ does not belong to the class of inequality measures proposed in Atkinson (1970). $T2(S)$ is a member of the latter class, however, with a constant 'degree of relative inequality aversion' equal to one and a constant elasticity of substitution in S of the same amount. For the log-normal attributes, however, these two measures provide the same ranking of distributions.

3. More general distributions and aggregate functions

Despite its empirical plausibility in some cases, the log-normal distribution does not generally fit the 'true' distribution of many attributes of interest. Hence we provide certain approximate formulae for $T1$ and $T2$ which can be conditionally used for any aggregator function, S, having a leptokurtic distribution. Our results here follow the developments in Maasoumi and Theil (1979) who were concerned with univariate (income) inequality.

For the aggregate attribute S, let $p = \log_e S - \mathrm{E}(\log_e S)$ and $\sigma_S^2 = \mathrm{E}p^2$. Theil' measures are expressible as follows:

$$T2(S) = \log_e \mathrm{E}e^p, \tag{11}$$

$$T1(S) = \mathrm{E}(pe^p)/\mathrm{E}e^p - T2(S). \tag{12}$$

Using Taylor's expansions of such expressions as e^p, $\log \mathrm{E}e^p$, pe^p, and $1/\mathrm{E}e^p$, the following approximations may be verified to the order given:

$$T1(S) = \tfrac{1}{2}\sigma_S^2\left[1 + \tfrac{2}{3}\gamma_{1S}\sigma_S + \tfrac{1}{4}\gamma_{2S}\sigma_S^2 + \mathrm{O}(\sigma_S^2)\right], \tag{13}$$

$$T2(S) = \tfrac{1}{2}\sigma_S^2\left[1 + \tfrac{1}{3}\gamma_{1S}\sigma_S + \tfrac{1}{12}\gamma_{2S}\sigma_S^2 + \mathrm{O}(\sigma_S^2)\right], \tag{14}$$

where the skewness coefficient, γ_l, and kurtosis, γ_2, are given by

$$\gamma_{1S} = \mathrm{E}p^3/\sigma_S^3 \quad \text{and} \quad \gamma_{2S} = \mathrm{E}p^4/\sigma_S^4 - 3. \tag{15}$$

In the case of the Cobb–Douglas S, $\sigma_S^2 = \alpha'\Sigma\alpha$ can be used in (13)–(15). Of course, these formulae are also applicable for any single attribute.

Several properties of (13)–(14) may be noted:

(i) For positively skewed (with long tails on the right) and leptokurtic distributions, both indices exceed their value for the log-normal distribution ($\frac{1}{2}\sigma_S^2$).
(ii) In terms of skewness alone, the excess of *T1* over $\frac{1}{2}\sigma_S^2$ is twice that of *T2*.
(iii) In terms of kurtosis alone, the excess of *T1* over $\frac{1}{2}\sigma_S^2$ is three times that of *T2*.

All of these properties [particularly (ii) and (iii)] reflect the relative sensitivity of these two measures to transfers at different parts of the distributions. When the logarithms of the attributes are nearly normally distributed, both *T1* and *T2* are close to $\frac{1}{2}\sigma_S^2$. Clearly, the accuracy of these approximations [$O(\sigma_S^4)$] is quite good so long as σ_S^2 is small (less than one). Another condition for the validity of the approximations is the existence of at least the first four moments of the underlying distributions.

Estimation and inference based on formulae (13) and (14), for any attribute or for their aggregate, is rather straightforward. Under random sampling, the first four *sample* moments are all that we need to *consistently* estimate σ_S^2, γ_{1S} and γ_{2S}. For the Cobb–Doublas aggregate, a consistent estimate of Σ is the second sample moment matrix. Note that in this case $\log S$ is a *linear* function of $\log Z_f$, and the sample-based estimates of σ_S^2, γ_{1S}, and γ_{2S} are evidently consistent under *random sampling.*

In the appendix to this paper we present some numerical evidence based on the data of the Michigan Panel Study of Income Dynamics. Except in the case of extremely flat distributions (near-zero inequalities), the approximate formulae perform well. The difficulty with approximation of near-zero quantities is a common feature of mathematical approximations.

4. Incremental contribution to multivariate inequality

In welfare theory, 'fundamentalism' or the assumptions of 'anonymity' and 'impartiality' suggest that the greater the number of attributes considered the fuller is the representation and the ordering of the individuals in a population [see Kolm (1977) for a formal statement of the same]. A single attribute such as income provides a poor representation. In practice, however, one cannot or need not consider a very large set of attributes. One reason is that the attendant measurement problems and errors tend to mount rapidly, particularly in the case of qualitative indicators and factor components of variables such as 'wealth'. Another reason is that the incremental contribution to inequality of additional attributes may be too small to justify their inclusion, particularly for a priori less important or questionable variables. It is thus useful to be able to identify this incremental contribution in a relatively simple manner. This is the primary aim of this section.

We pursue this question first in the special case which led to (8)–(10), and then in the general case of discrete variates whose distributions are *unknown*. From eq. (9), the separate contribution to inequality of the fth attribute, I_f, say, is given by

$$I_f = \tfrac{1}{2}\alpha_f^2\sigma_{ff} + \alpha_f \sum_{k \neq f}^{M} \alpha_k \sigma_{fk}, \tag{16}$$

where the symmetry of Σ has been utilized. Since $\alpha_f \geq 0$,[2] we observe that

$$I_f > 0 \quad \text{if} \quad \sum_{k \neq f}^{M} \alpha_k \sigma_{fk} \geq 0. \tag{17}$$

This is a sufficient condition, however, and when $\sum_{k \neq f} \alpha_k \sigma_{fk} < 0$ a general condition is

$$I_f \geq 0 \quad \text{iff} \quad \left| \sum_{k \neq f}^{M} (\alpha_k/\alpha_f)(\sigma_{kk}/\sigma_{ff})^{1/2} \rho_{fk} \right| \leq \tfrac{1}{2}, \tag{18}$$

where ρ_{fk} is the simple correlation coefficient between two attributes. Condition (18) is more likely to be violated when attribute f is strongly *negatively* correlated with attributes which are relatively more important $(\alpha_k/\alpha_f > 1)$ and/or relatively less equally distributed $(\sigma_{kk}/\sigma_{ff} > 1)$. Hence, negative correlation with other attributes may more than cancel-out the own inequality $(\tfrac{1}{2}\alpha_f^2\sigma_{ff})$ term and reduce overall inequality.

When the distribution of attributes is *not* the log-normal, we may pursue the question at hand by using the following decomposability results established in Maasoumi (1986, proposition 2):

$$T2(S) = \sum_{f=1}^{M} \alpha_f T2(x_f) - \min D_0(S^*, x; \boldsymbol{\alpha}), \tag{19}$$

$$T1(S) = \sum_{f=1}^{M} c_f T1(x_f) - \min D_{-1}(S^*, x; \boldsymbol{c}), \tag{20}$$

[2] We have assumed $\alpha_f \geq 0$ for all f. One may conceive of some attributes with disutility. One example may be the distribution of hours of work. In that case, leisure will be the attribute considered with a nonnegative weight.

where $c_f = \alpha_f T_f / \sum_k \alpha_k T_k$, $T_f = \sum_i X_{if}$, $\sum_f \alpha_f = 1$, $= \sum_f c_f$, and the second terms on the right of (19) and (20) are, respectively, the minimum of $D_0(\ )$ obtained at $S_i = \prod_f x_{if}^{\alpha_f}$ and the minimum of $D_{-1}(\ )$ obtained at $S_i = \sum_f \alpha_f x_{if}$.[3] Since both $D_0(\ )$ and $D_{-1}(\ )$ are nonnegative, it is seen that both $T1$ and $T2$ are bounded above by the weighted *averages* ($\overline{T1}$ and $\overline{T2}$, say) of attribute inequalities. Also, (3a) and (3b) indicate that both $D_0(\ )$ and $D_{-1}(\ )$ are decomposable by attribute contributions. Hence the incremental contribution of an attribute may be identified from (19) and (20) as follows:

$$T2(f) = \alpha_f T2(x_f) - \alpha_f \sum_{i=1}^{N} S_i^* \log(S_i^*/x_{if}), \tag{21}$$

$$T1(f) = c_f T1(x_f) - c_f \sum_{i=1}^{N} x_{if} \log(x_{if}/S_i^*). \tag{22}$$

Expressions (16), (21), and (22) may be employed to measure the incremental contribution to 'inequality' of a variable which is *already included in the aggregation*. The consideration of the effect of extending the set of M variables by an *additional* and possibly questionable variable is somewhat more complicated since we must also allow for a change in attribute weights from $\alpha = (\alpha_1, \ldots, \alpha_M, 0)$ to $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_M, \delta_{M+1})$, say. In general the net effect on inequality of adding this last attribute is indeterminate. Some intuition can be gained, however, by exploring the conditions under which $T2(f) \geq 0$ for $f = M+1$. We note that $\sum_{f=1}^{M} \alpha_f = 1 = \sum_{f=1}^{M+1} \delta_f$ and $\delta_{M+1} = \sum_{f=1}^{M} (\alpha_f - \delta_f) > 0$. Let $W_f = (\alpha_f - \delta_f)/\delta_{M+1}$; it follows that $\sum_{f=1}^{M} W_f = 1$. It may be verified that

$$T2(M+1) = \delta_{M+1}\left[T2(x_{M+1}) - \sum_{f=1}^{M} W_f T2(x_f)\right] + \log(K_{M+1}/K_M), \tag{23}$$

[3]The aggregate attribute used to compute $T1$ and $T2$ (S_i) continues to be defined as a function of the actual quantities, X_{if}'s. This convention was followed since this accords more naturally with the usual interpretation of S_i as a utility function. Consequently the 'ideal' forms for S_i were obtained by minimizing $D(\ )$ defined over X_{if} and S_i rather than the shares x_{if} and S_i^*. The 'ideal' aggregate *shares*, $S_i = \prod_f x_{if}$ and $S_i = \sum_f \alpha_f x_{if}$, can be obtained if we minimize the divergence criteria *as are defined in* (3*a*) *and* (3*b*). If these aggregate *shares* are used to compute $T1$ and $T2$, the decompositions (19) and (20) will be unaffected except that, in $T1(S)$, c_f's should be replaced by α_f's. This is because $T2$ is homogeneous in S_i as well as in every attribute, whereas $T1$ is homogeneous in S_i alone.

where

$$-\log K_M = -\log \sum_{i=1}^{N} \left(\prod_{f=1}^{M} x_{if}^{\alpha_f} \right) = \min D_0(S^*, x; \alpha), \tag{24}$$

$$-\log K_{M+1} = -\log \sum_{i=1}^{N} \left(\prod_{f=1}^{M+1} x_{if}^{\delta_f} \right) = \min D_0(S^*, x_a; \delta), \tag{25}$$

where x_a is matrix x augmented by x_{M+1} and $K_M, K_{M+1} \leq 1$. From (23) it is seen that, if $K_{M+1} \geq K_M$, we have

$$T2(M+1) \geq 0 \quad \text{if} \quad T2(x_{M+1}) \geq \sum_{f=1}^{M} W_f T2(x_f). \tag{26}$$

That is, inequality will be no less if the inequality in the $(M+1)$th attribute is no less than a weighted *average* of the inequalities in the first M attributes. The size of this increment depends on the weight given to the $(M+1)$th attribute (δ_{M+1}). For a 'normal' additional attribute it is plausible to assume $W_f \geq 0$ since $\alpha_f \geq \delta_f$, $f = 1, \ldots, M$, is likely to characterize the new weight scheme.

The condition $K_{M+1} \geq K_M$ is not of course a necessary condition for $T2(M+1) \geq 0$. According to the information criterion $D_0(\ )$, $K_{M+1} \geq K_M$ holds when the fit of S^* to the available data *improves* with the addition of the $(M+1)$th attribute [see (24) and (25)]. It may be verified that a *sufficient* condition for $K_{M+1} \geq K_M$ is that

$$x_{i,M+1} \geq \prod_{f=1}^{M} x_{if}^{w_f} \quad \text{for all} \quad i \in [1, N]. \tag{27}$$

When $\alpha_f \geq \delta_f$, $\forall f \in [1, M]$, it is easily verified that the reverse of (27) cannot hold for *all* $i \in [1, N]$. An example, using the Michigan panel data on components of income, schooling, and a proxy for 'wealth', serves to clarify these conditions. Employing these data for the years 1968, 1975, and 1979, Maasoumi and Nickelsburg (1988) find that in all cases the quantities $\log K_2$ and $\log K_3$ are very small indeed, and condition (26) above provides a reliable and simple guide. In particular, adding schooling to income and wealth reduces the *joint* inequality in the latter two variables. Schooling inequality is far less than any weighted average of income and 'wealth' inequalities (the actual weights were $\frac{1}{2}$ and $\frac{1}{2}$, and $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ when schooling was added). 'Wealth', on the other hand, is found to be generally the least equally distributed attribute in these data. It's addition is found to increase multivariate inequality in income and schooling as could have been clearly predicted by condition

(26). These observations are confirmed in an extensive study of international data on per capita GNP, Physical Quality of Life Indices (PQLI) and Basic Needs (BN) Fulfillment indices reported in Maasoumi and Jeong (1985).[4]

5. Concluding remarks

Under the assumption of jointly log-normally distributed attributes and Cobb–Douglas individual utility functions, the multivariate versions of Theil's inequality measures were derived and analyzed. The formulae demonstrate the important part played by attribute correlations in the multivariate context. Relaxing the log-normality assumption, approximate formulae for both indices are given which are applicable in general for any aggregate function of the attributes.

The incremental contribution of attributes to inequality were derived with certain sufficient conditions which determine their sign.

Other distributions than the multivariate log-normal may be considered. Examples are the multivariate Beta and certain multivariate Dirichlet distributions. For certain utility functions, some inequality indices may be derived analytically since only the moments of such distributions are generally required [e.g., see Basmann et al. (1982) and Slottje (1987) for the Gini coefficients of a Beta distribution]. The complexity in the attendant estimation requirements increases rather rapidly, however.

In addition to the choice of a particular inequality measure and the implicit value judgements thereof, several new decision problems are posed by the multivariate approach. Among these the choice of an aggregate or utility function is the most immediate. When measurement of inequality is the primary objective, we think the information theory approach outlined in section 2 provides a sensible solution to the *functional form* problem. These functional forms also satisfy the more traditional properties (e.g., concavity, quasi-concavity, etc.) required of utility functions. The unknown parameters in these aggregate functions have to be determined in practice. When market prices exist for *all* the attributes, data-based 'solutions' are available through the estimation of systems of demand equations. Some well-known identification problems not withstanding, this situation is as ideal, however, as it is rare since market prices do not generally exist for many qualitative and/or nontraded (e.g., public) goods. Faced with this problem, another data-based 'solution' exists only for the values of the α_f's which, however, does not appear to have any welfare-theoretic justification. This 'solution' is a variant of the method of Principal Components (PC) proposed by Ram (1982). Using

[4]Our weights were generally based on the principal components of the variables. As is amply clear from the definition of $T2(M+1)$, one can conceive of weighting schemes which change its direction and magnitude.

this method, the vector α is the normalized form of the characteristic vector of the attributes' covariance matrix corresponding to its largest characteristic root. These are the (normalized) weights given to the attributes in their first PC which accounts for the largest fraction of *variability* in the data. We propose these weights be computed to provide a *benchmark* in a subjective determination of the unknown parameters. One advantage of the decompositions given in (19) and (20) and in Proposition 2 of Maasoumi (1986) is that they make explicit any subjective choice of these parameters. Hence, observer's valuations of the attributes have easily traceable consequences for the assessment of multivariate inequality which are distinct from other subjective differences arising from the choice of inequality measures (or the SWF's) and/or the aggregate functions.

The condition which produces the convenient decomposability formulae (19) and (20) is quite restrictive and deserves some examination. Once a member of the Generalized Entropy family of inequality measures is selected (e.g., *T2*), the imposition of the 'attribute decomposability' requirement [e.g., (19)] determines the constant elasticity of substitution in the individual utility (aggregate) functions (e.g., $\sigma = 1$ for the Cobb–Douglas). This is because the decomposability condition forces the 'individualistic' SWF to be 'specific' as well [see Kolm (1977)]. Equivalently, a knowledge of the observer's constant elasticity of substitution in his valuation of individual utility (S_i) *and* the requirement of the attribute decomposability property together determine a unique member of the GE family of inequality measures. In general, when the decomposability condition is not imposed, a value for the elasticity of substitution parameter must be estimated (when possible) or assigned. Clearly, it is worthwhile to investigate and to report the robustness of inequality measurements with respect to reasonable changes in the unknown parameter values.

The ranking of certain distributions, however, can be shown to be less dependent on specific utility functions and inequality measures. For instance, if $\bar{X}$ and X are two attribute matrices such that $\bar{X} = BX$, with B a bistochastic matrix which, as is well known, leaves the attribute totals (T_f's) unchanged and implements certain 'averaging' redistributions, we may assert that

$$I(\bar{S} = h(\bar{X})) \leq I(S = h(X)),$$

for *all* inequality measures, $I(\ )$, such that $-I(\ )$ is Schur-concave, and *any* positive, real-valued, and *concave* function $S_i = h(X_i)$ [see Maasoumi (1986, proposition 3)].

Appendix

In order to provide some numerical evidence in general, and on the performance of the approximation formulae (13) and (14), table 1 below

Table 1

1979 random panel – PC weights.

	1	2	3	4	5	6	7	8	9	10
	$G(S)$	$T2(S)$	$T2(In)$[a]	$T2(H)$[b]	$T2(Ed)$[c]	Sample size	Sample share	Income share	Housing share	Education share
All	0.610	0.181	0.335	0.422	0.093	3555	1.00	1.00	1.00	1.00
1st G	0.640	0.087	0.234	0.323	0.011	1005	0.28	0.258	0.137	0.300
2nd G	0.629	0.106	0.238	0.358	0.017	959	0.27	0.249	0.279	0.284
3rd G	0.621	0.117	0.263	0.258	0.032	1081	0.30	0.372	0.426	0.296
4th G	0.563	0.147	0.331	0.346	0.073	510	0.14	0.119	0.159	0.118
Between-term		0.071	0.077	0.105	0.065					
Within-term		0.110	0.258	0.317	0.028					
PC weights			0.382	0.339	0.278		$\rho_{I,H}$	$\rho_{I,Ed}$	$\rho_{H,Ed}$	
$\bar{T}2 = 0.296$				$\bar{T}2 - T2(S) = 0.115$			0.229	0.303	0.157	
γ_1		−0.202	0.235		−2.497					
γ_2		−0.296	0.939		13.442					
$\frac{1}{2}\sigma^2$		0.137	0.284	0.348	0.039					
Approx. $T1(\ )$		0.124	0.288		0.031					
Approx. $T2(\ )$		0.131	0.279		0.033					

[a] In (income) is adjusted for family size.
[b] H is categorized net equity in housing, with 1 ≡ no equity and 10 ≡ greater than 45,000 dollars in equity.
[c] Ed is average years of schooling for household adults, categorized from $\frac{1}{2}$ to 17 in increments of $\frac{1}{2}$ year.

presents an analysis based on the 1979 data taken from the Michigan Panel Study of Income Dynamics. There are 3555 randomly selected households in this sample which we divide into four distinct age groups. These are, respectively, age ≤ 30, $30 < \text{age} \leq 45$, $45 < \text{age} \leq 65$, and age > 65. Three attributes are considered as follows: *In* denotes income adjusted for family size and includes asset income as well as such transfer payments as unemployment and food stamps. *H* denotes bracketed net equity in housing as a factor component of wealth. *Ed* denotes averaged years of schooling of household adults (see the explanations below the table). Column 1 reports the Gini coefficient of inequality in the Cobb–Douglas function S_i for all the sample (row 1) and each subgroup. The next four columns report Theil's second measure for S_i and each of the three attributes. Between-group inequality is distinguished from the (population shares) weighted average of the within-group inequalities. The Gini coefficient does not permit this type of useful decomposition. A more extensive study in Maasoumi and Nickelsburg (1988) demonstrates that between-group inequalities have declined substantially compared to 1968 and 1975, whereas within-group inequalities have either increased slightly or remained stable. The overall measure $[T2(S)]$ is more representative of the general picture as well as of each attribute, whereas each attribute (e.g., income) can misrepresent both the overall picture and the pattern of inequality in the other attributes. The weights given the three attributes were based on the PC method as described in the conclusion to the paper. The pattern of inequality is quite robust with respect to changes in this valuation scheme $((\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ and equal weights were also studied).

To shed some light on the performance of the approximation formulae (13) and (14), the last two rows of table 1 report the approximate values of Theil's two measures for S, income and education. Housing distribution provides very similar results to that of income and is not reported in order to facilitate the comparison between the two extremes observed for income and S on the one hand and schooling distribution on the other. γ_1 and γ_2 suggest that, while the log-normal distribution fits the S and income distributions reasonably well, the distribution of schooling is very flat. The latter is very equally distributed with a very small variance. A comparison on the entries in the last three rows suggest that, in all cases, the estimates based upon the formulae (13) and (14) are practically the same as the estimated inequality assuming the log-normal distribution $(\frac{1}{2}\sigma^2)$. Approximate *T1* and *T2* values are practically identical. All three estimates are less than the corresponding entries in the first row of the table. However, the entries in the first row of the table are themselves discrete approximations (estimates) of 'true' inequalities. The most plausible observation seems to be that, for highly equally distributed attributes (near-zero inequalities) such as schooling, the approximations do not perform as well as might be hoped. In other cases the performance appears quite good relative to the available data.

References

Aitchison, J. and J.A.C. Brown, 1957, The log-normal distribution, Monograph 5 (Cambridge University Press, Cambridge).

Atkinson, A.B., 1970, On the measurement of inequality, Journal of Economic Theory 2, 244–263.

Atkinson, A.B. and F. Bourguignon, 1982, The comparison of multi-dimensioned distributions of economic status, Review of Economic Studies 12, 183–201.

Basmann, R.L., D.J. Molina, M. Rodarte, and D.J. Slottje, 1982, Some new methods of predicting changes in economic inequality associated with trends in growth and development, in: Issues in 3rd world development (Westview Publ. Company, Denver, CO).

Blackorby, C. and D. Donaldson, 1978, Measurement of relative inequality and their meaning in terms of social welfare, Journal of Economic Theory 18, 59–80.

Bourguignon, F., 1979, Decomposable income inequality measures, Econometrica 47, 901–920.

Cowell, F.A. and K. Kuga, 1981, Inequality measurement: An axiomatic approach, European Economic Review 15, 287–305.

Foster, J.E., 1983, An axiomatic characterization of the Theil measure of income inequality, Journal of Economic Theory 31, 105–121.

Kendall, M. and A. Stuart, 1969, The advanced theory of statistics, Vol. 1, 3rd ed. (C. Griffin and Company, London).

Kolm, S.-ch., 1977, Multi-dimensional Egalitarianism, Quarterly Journal of Economics 91, 1–13.

Maasoumi, E., 1986, Measurement and decomposition of multi-dimensional inequality, Econometrica 54, 991–997.

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

Maasoumi, E. and J.-H. Jeong, 1985, International inequality in the distribution of several welfare attributes: A multidimensional approach (Department of Economics, Indiana University, Bloomington, IN).

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, 327–334.

Ram, R., 1982, Composite indices of physical quality of life, basic needs fulfillment, and income: A 'principal component' representation, Journal of Development Economics 11, 227–247.

Shorrocks, A.F., 1980, The class of additively decomposable inequality measures, Econometrica 48, 613–625.

Shorrocks, A.F., 1984, Inequality decomposition by population subgroups, Econometrica 52, 1369–1385.

Slottje, D.J., 1987, Relative price changes and inequality in the size distribution of various components of income: A multidimensional approach, Journal of Business and Economic Statistics 5, 19–26.

Theil, H., 1967, Economics and information theory (North-Holland, Amsterdam).