

**ISSN 1352-2469
Programme**

Distributional Analysis Research

**Estimating Welfare Indices: Household Weights
And Sample Design**

by

**Frank Cowell
and
Stephen Jenkins**

**Discussion Paper
No. DARP 48
June 2000**

**Distributional Analysis Research Programme
The Toyota Centre
Suntory and Toyota International
Centres for Economics and
Related Disciplines
London School of Economics
Houghton Street
London WC2A 2AE**

ABSTRACT

The techniques of simple random sampling are seldom appropriate in the empirical analysis of income distributions. Various types of weighting schemes are usually required either from the point of view of welfare-economic considerations (the mapping of household/family distributions into individual distributions) or from the point of view of sample design. The different types of weights have different implications for the sampling distribution of estimators of welfare indices.

Address for correspondence:

F. A. Cowell
STICERD
London School of Economics and Political Science
Houghton Street, London, WC2A 2AE, UK.

Stephen Jenkins
University of Essex
Institute for Social and Economic Research
Wivenhoe Park
Colchester CO4 3SQ
UK

Keywords: Welfare index, inequality, poverty, sample, inference.
JEL Classification: C13, D63.

ACKNOWLEDGEMENT

This research was funded as part of a project on the “Structure of the Income Distribution” financed by the ESRC (grant R000231731). We gratefully acknowledge the intellectual input of Fiona Coulter who discussed in detail with us many of the issues in sections 4 to 6. We wish to thank Joanna Gomulka for comments on an earlier version of this paper and Ceema Namazie for computational assistance. Errors remain our own responsibility.

© Frank A. Cowell and Stephen Jenkins. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

1 Introduction

The analysis of economic welfare sometimes proceeds as though individual income receivers were identical in every respect other than income. There is of course a substantial literature on the problems of comparing income levels between income receivers with different needs' levels, and on the meaning to be given to "income" and "income receiver" in any particular application. However, some of the statistical issues that arise naturally in connection with data on an essentially heterogeneous population have been relatively neglected in the standard approach to the problem of computing welfare indices. The aim of this paper is to show how, for various models of population heterogeneity, these issues may be treated within a unified framework, and why the properties of the estimators will differ between models.

A common approach to welfare measurement in practice is to take a vector of incomes $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and plug it into a welfare index: such an index is essentially an ethical aggregation function such as an explicit social-welfare function (SWF), an inequality measure or a poverty measure.¹ More sophisticated analysis also takes account of the sampling distribution of the welfare statistic expressed as a function of \mathbf{x} , and uses information about the sample to estimate the standard error as well as point estimates of welfare statistics. However, in many economic applications even this more sophisticated approach is inadequate because it is predicated on the assumption that one is sampling from a univariate distribution of incomes. Because of the welfare-theoretic views which one is trying to incorporate, or because of the way in which data on income distribution are typically collected, one very often has to deal with weighted data. In these circumstances the appropriate empirical model is of sampling from a bivariate distribution, where one takes account of the structure of weights in the sample, and of the possible interaction of these weights with income. The purpose of this paper is to examine the problems that arise from treating welfare measurement as a problem of computing estimates based on samples from a bivariate distribution of incomes and associated sampling weights. Although the issues raised are not always explicitly addressed, the problem of weighted estimation is in practice quite common when dealing with data on income distributions, and has to be undertaken for a variety of reasons. The required weights come from a number of sources which should be distinguished from each other when estimating welfare statistics empirically:-

- Welfare-consistent weighting typically arises from the fact that standard welfare economics is founded upon an individualistic approach whereas microdata sources on income are often based on households or families. Given some assumption about the distribution of income within the household

¹For examples of explicit SWFs of the sort discussed below see, for example Jenkins (1997), for inequality indices see Cowell (2000), and for poverty indices see Ravallion (1994).

(typically equal sharing), weighting converts the observed household distribution into the personal distribution required to compute a welfare index.

- Sample-design weights are commonly used in situations where estimation has to be from a non-random sample. A standard example is the procedure of stratified sampling where sub-samples of given size and differing characteristics are merged in order to ensure a sufficiently large subsample of respondents with a given type of attributes.
- In some types of estimation problem the size of the sub-samples may be determined endogenously, on the basis of some observed attribute that characterises members of different groups. An important example of this is “grossing-up” to deal with problems such as differential non-response. This is a technique that is often employed in applied statistics and econometrics, but one where the statistical properties are sometimes inadequately considered. In this case the sample is weighted so that the proportions of observations in the sample with a given attribute correspond to the proportion of the population that is known to possess this attribute: the subsamples are determined by the numbers falling into the specified attribute class.

In practice one may have to deal with all of these features simultaneously within the same model of applied welfare. Previous papers have addressed specific aspects of the weighting problem for particular types of welfare index.² In this paper we adopt a general approach: we examine the rôle of each of these three different types of weights in the estimation problem and investigate the quantitative importance of mis-specification of the type of weights in making empirical welfare comparisons. We shall show that although the formulae for the point estimates of welfare measures are essentially the same across the different types of weighted-data problems, there are important differences in the associated sampling distributions. This can lead to a situation where, for a given welfare index, the differences in the standard errors corresponding to the different types of weights can be quite large.

2 The Approach

The essential feature of the problem that we shall discuss is that each person is potentially distinguished not only by his “income” x (used as the indicator of individual economic status in the welfare measure) but also by some other attribute

²See, for example Cowell (1989) on decomposable inequality measures with household (random) weights and Howes and Lanjouw (1998) on poverty indices with weights to allow for stratification and clustering. The poverty case is relatively straightforward because the required index can be expressed simply in terms of means. The inequality case is considerably more complicated because the indices involve (functions of) other moments besides means.

a that may affect the way in which we choose to aggregate the information about the income distribution. To implement this in our model we make the following assumptions:

Axiom 1 *Each member of the population be characterised by a pair (a, x) where $a \in A \subseteq \mathfrak{R}$ and $x \in X \subseteq \mathfrak{R}$.*

Axiom 2 *The joint distribution of attributes and income in the population are given by a function $F : A \times X \mapsto [0, 1]$ belonging to the class \mathfrak{F} of bivariate distribution functions with finite means and variances.*

Axiom 3 *The welfare index is subgroup-decomposable on $\mathfrak{F}(\cdot, \mu)$.*

Note that the specification of the set of attributes A in Axiom 1 encompasses a variety of cases that are relevant to the applied economics of income distribution – for example it could be a doubleton (males and females), a subset of \mathfrak{R}_+ (a monetary measure of needs) – and that this specification can be relaxed without damaging the essential argument which follows. Axiom 2 encompasses both continuous and discrete distributions. Axiom 3 means that, for the subfamily of distributions with a given mean, there is an implied SWF that can be written as an additively separable functional of the income distribution; the welfare index of interest can be written as some function of this SWF that may depend on the mean (Ebert 1987, Blackorby and Donaldson 1978, Blackorby, Donaldson, and Auersperg 1981, Shorrocks 1984). This has a number of practical advantages when dealing with data from a heterogeneous population.³

To implement the decomposable welfare index we introduce two important concepts. The *weight function* $w : A \mapsto \mathfrak{R}_+$ assigns the “importance” of each income-receiving unit in the distribution. Given the class of functions $\wp := \{\phi : X \mapsto \mathfrak{R} \mid \phi \text{ is increasing and concave}\}$ the *evaluation function* $\phi \in \wp$ assigns an importance in terms of welfare to any given income value. The weight function will be determined by the nature of the particular problem of welfare analysis under consideration and the assumptions to be made about the information conveyed by personal attributes. The issue of specifying the set A can be subsumed in the specification of the weight function: in principle, we can think of the heterogeneous population as being distributed either by (attribute, income)-pairs, or by (weight, income)-pairs. Welfare can be represented as a weighted sum of the evaluations of individual incomes.

For any $q, r \in \{0, 1, 2\}$ and $\phi \in \wp$ define the following weighted moments about zero

$$\mu_{q\phi} := \int w(a)^q \phi(x) dF(a, x) \quad (1)$$

$$\mu_{qr} := \int w(a)^q x^r dF(a, x) \quad (2)$$

³See Coulter et al. (1992).

The moments μ_{10}, μ_{11} have a special significance for welfare measurement: the first can be interpreted as the “effective population size”,⁴ and the second as “effective total income”; mean income is given by μ_{11}/μ_{10} . The welfare measure can be taken as a functional defined on the space of distribution functions, $W : \mathfrak{F} \mapsto \mathfrak{R}$. Because we assume that the welfare measure is fully decomposable, and because the desired welfare index may also be a function of mean income, and the implicit population size, a typical welfare index for a distribution F is implicitly defined by a function $W : \mathfrak{R}_+ \times X^2 \mapsto \mathfrak{R}$ thus:

$$W(F) = \psi(\boldsymbol{\mu}) \quad (3)$$

where $\boldsymbol{\mu} := (\mu_{10}, \mu_{11}, \mu_{1\phi})$ is a vector of weighted moments. Indeed the expression (1) can be taken as an elementary welfare index in its own right, given appropriate properties of ϕ : other welfare indices – such as measures of inequality and poverty – can be represented in a form such as (3).

Assume that the attribute a is observable, and that a sample of $n < \infty$ observations $\{(a_i, x_i) : i = 1, 2, \dots, n\}$ has been drawn from F ; denote the sample distribution by $F^{(n)}$.⁵ Our immediate task is to specify estimators of welfare measures, and the sampling variance of the estimates. The estimates of the moments about zero (1), (2), and of welfare (3) from the sample $F^{(n)}$ are given by

$$m_{q\phi} := \sum_{i=1}^n w(a_i)^q \phi(x_i) \quad (4)$$

$$m_{qr} := \sum_{i=1}^n w(a_i)^q x_i^r. \quad (5)$$

$$\hat{W} := W(F^{(n)}) = \psi(\mathbf{m}) \quad (6)$$

where $\mathbf{m} := (m_{10}, m_{11}, m_{1\phi})$. The sample statistic \hat{W} will provide a consistent estimator of $W(F)$ if the sample is randomly drawn.⁶ Using the definitions:

$$\bar{\phi}(x) := x\phi(x) \quad (7)$$

$$\overline{\phi}(x) := \phi(x)^2 \quad (8)$$

we may also state the following asymptotic result on the sampling distribution of the estimate of the welfare index:

⁴If income-receivers are households and if the weight on each observation corresponds to the number of persons in each household, then μ_{10} is exactly the number of persons in the population.

⁵Formally $F^{(n)} := \frac{1}{n} \sum_{i=1}^n \Delta_{(a_i, x_i)}$ where $\Delta_{\mathbf{y}}$ is a point mass at \mathbf{y} .

⁶This does not require that the sample design be simple random sampling. If the observations are iid according to F then, by the Glivenko-Cantelli theorem, the empirical distribution $F^{(n)}$ tends to F as $n \rightarrow \infty$, and so the index $W(F^{(n)})$ becomes $W(F)$ (Victoria-Feser 1999).

Theorem 1

$$\sqrt{n}[\psi(\mathbf{m}) - \psi(\boldsymbol{\mu})] \sim N(0, nV) \quad (9)$$

where

$$V := \frac{1}{n} \psi'_\mu \boldsymbol{\Sigma} \psi_\mu, \quad (10)$$

$$\boldsymbol{\Sigma} := \begin{bmatrix} \mu_{20} - \mu_{10}^2 & & \\ \mu_{21} - \mu_{10} & \mu_{21} - \mu_{11}^2 & \\ \mu_{2\bar{\phi}} - \mu_{1\bar{\phi}}\mu_{10} & \mu_{2\bar{\phi}} - \mu_{1\bar{\phi}}\mu_{11} & \mu_{2\bar{\phi}} - \mu_{1\bar{\phi}}^2 \end{bmatrix}, \quad (11)$$

$\psi_\mu := \partial\psi/\partial\boldsymbol{\mu}$ and ψ'_μ is its transpose.

Proof. Given the definition of \mathbf{m} and that $F^{(n)}$ is randomly drawn we know that as $n \rightarrow \infty$

$$\sqrt{n}[\mathbf{m} - \boldsymbol{\mu}] \sim N(0, \boldsymbol{\Sigma}) \quad (12)$$

where $\boldsymbol{\Sigma}$ is the covariance matrix with typical element $\Sigma_{ij} = n \text{cov}(m_{1i}, m_{1j})$, $i, j = 0, 1, \bar{\phi}$. Hence, applying the δ -method we have (9), (10) – (Rao 1973, pp 387-388). ■

V is the variance of the welfare statistic \hat{W} under the assumed sampling scheme. Two remarks are appropriate here. First, the result can be extended to cardinalisations of the welfare measure that are transformations of (3) and (6): in effect one modifies the partial derivatives ψ_μ by the slope of the transformation.⁷ Second, the fact that $\boldsymbol{\Sigma}$ is a function of the moments involving $\bar{\phi}$ and $\bar{\phi}$ (derived from 1) means that the value of V in (10) depends on the way in which the individual weights are correlated with income and with the income-evaluation ϕ . A similar point arises if we consider the corresponding sample estimate of V which depends on the modified form of (11) where the μ -terms are replaced by their m counterparts from (4), (5). This correlation of w with x depends upon a number of fundamental issues including sample design and inherent population heterogeneities which one tries to take account of in welfare measurement.

⁷A simple example can be found if one puts $\phi(x) = \log(x)$ in (4) to yield two commonly used and ordinally equivalent inequality measures. The mean logarithmic deviation index of inequality is

$$\log\left(\frac{\mu_{11}}{\mu_{10}}\right) - \frac{\mu_{1\bar{\phi}}}{\mu_{10}}$$

Applying the transformation $T(\cdot) := 1 - 1/\exp(\cdot)$ to this index gives the Atkinson (1970) inequality index with inequality aversion parameter 0:

$$1 - \exp\left(\frac{\mu_{1\bar{\phi}}}{\mu_{10}}\right) \frac{\mu_{10}}{\mu_{11}}$$

Clearly the first derivative of T is $1 - T$. We find that ψ'_μ for the first inequality index is $(-\mu_{10}^{-1}, \mu_{11}^{-1}, \mu_{1\bar{\phi}}\mu_{10}^{-2} - \mu_{10}^{-1})$ and so ψ'_μ for the second inequality index is just $1 - T$ times this, i.e. $\exp(\mu_{1\bar{\phi}}\mu_{10}^{-1})(-\mu_{11}^{-1}, \mu_{10}\mu_{11}^{-2}, \mu_{1\bar{\phi}}\mu_{10}^{-1}\mu_{11}^{-1} - \mu_{11}^{-1})$.

In sections 3 to 6 we consider three important cases that differ in the nature of the weights. Although all three cases yield the same point estimates of the welfare index we will find that they have different sample variances.

3 Welfare-consistent Weighting

In this case the weight w is determined by an individual observation in the sample: typically w corresponds to the number of persons in the household who are to be counted as income receivers (Cowell 1989). Although the household or the family may be the formal income-receiving unit, it is the implicit distribution amongst persons which is important in welfare measurement: so, for example, a family of three which receives an income that is equivalent to \$20 000 per person should be counted as three separate incomes of \$20 000 each (Cowell 1984, Danziger and Taussig 1979). We can handle this by allowing the weighting function to be proportional and interpreting the attribute a in a particular fashion. So for all $(a_i, x_i) \in F^{(n)}$ we may write $a_i = w_i$ where a_i is the household size of observation i . Consider the special case of (3) that uses *normalised* weights for population and sample:

$$\hat{w}(a) := \frac{w(a)}{\mu_{10}} \quad (13)$$

$$\hat{w}(a_i) := \frac{w(a_i)}{m_{10}} \quad (14)$$

and take the family of welfare measures W^* given by

$$W^*(F) := \int \hat{w}(a)\phi(x)dF(a, x) . \quad (15)$$

for any $\phi \in \wp$. The corresponding sample estimate will be

$$W^*(F^{(n)})|_{\hat{w}} = \sum_{i=1}^n \hat{w}(a_i)\phi(x_i) . \quad (16)$$

The weight on each observation in equation (16) is determined by the particular value of the a -component of the observation: hence the expression “welfare-consistent weighting” for this type of problem. In some applications the welfare index is defined with reference to incomes that have been standardised relative to the mean, so that the term x in (15) is replaced by $x\mu_{10}/\mu_{11}$. However, this leads to a few extra complications in the variance formulae which follow, without yielding any great additional insight: we give the full formula – for the mean-standardised case – for the specific example reported in section 7 below.

Now consider the sampling distribution of the estimator in (16).

Theorem 2 Under the assumption of welfare-consistent weighting $W^*(F^{(n)})$ is asymptotically distributed $N(W^*(F), V_{\hat{w}})$ where

$$V_{\hat{w}} := \text{var } W^*(F^{(n)})|_{\hat{w}} = \frac{1}{n} \frac{\mu_{1\phi}^2}{\mu_{10}^4} \text{var}(a) - \frac{2}{n} \frac{\mu_{1\phi}}{\mu_{10}^3} [\mu_{2\phi} - \mu_{1\phi}\mu_{10}] + \frac{1}{n} \frac{1}{\mu_{10}^2} [\mu_{2\phi} - \mu_{1\phi}^2] \quad (17)$$

Proof. Given that $F^{(n)}$ is an unbiased sample, it is clear that

$$\mathcal{E} W^*(F^{(n)})|_{\hat{w}} = W^*(F) \quad (18)$$

Moreover, noting that

$$\frac{\partial \psi}{\partial \boldsymbol{\mu}} = \begin{bmatrix} -\frac{\mu_{1\phi}}{\mu_{10}^2} \\ 0 \\ \frac{1}{\mu_{10}} \end{bmatrix} \quad (19)$$

$$\text{var}(a) = \mu_{20} - \mu_{10}^2$$

and applying (9) – (11), Theorem 1 implies (17). ■

Notice the three separate components to (17) including the variance of the attribute (weight), the covariance of the weight with the evaluation of income $\phi(x)$, and the variance of the evaluation of income. Evidently if the data were unweighted the first two of these terms would vanish. It is also evident that the direction of the influence of these two terms on the resulting terms depends on the way in which incomes have been imputed to households with different a -values: if x is taken to be total household income, then presumably there will be a strong positive correlation between \hat{w} and $\phi(x)$ (Cf the second term in equation 17), whereas if x is income per capita this correlation could be negative. Thus the variance of the estimated welfare index in the case of welfare-consistent weighted data may be larger or smaller than the corresponding variance in the unweighted case. For an estimator of (17) from $F^{(n)}$, we use the sample-moment vector \mathbf{m} thus:

$$\hat{V}_{\hat{w}} = \frac{1}{n-1} \left[\frac{m_{1\phi}^2}{m_{10}^4} [m_{20} - m_{10}^2] - 2 \frac{m_{1\phi}}{m_{10}^3} [m_{2\phi} - m_{1\phi}m_{10}] + \frac{1}{m_{10}^2} [m_{2\phi} - m_{1\phi}^2] \right] \quad (20)$$

4 Design Weights: The Problem

For the second and third sub-cases we shall suppose that the population may be treated as though each household consisted of a single income receiver, and

that the observations have been drawn – with or without replacement – from a heterogeneous population in a way that does not correspond to simple random sampling. Instead we assume that the population may be partitioned into internally homogeneous subgroups. The manner in which this partition is done, and the information available about the partition is of central importance.

Let K be a non-empty index set, and suppose that the space of attributes A is partitioned into a finite collection of subsets $\{A_k : k \in K\}$, where $\bigcup_{k \in K} A_k = A$ and, for any distinct $k, k' \in K$, $A_k \cap A_{k'} = \emptyset$. Again assume that a person's attribute a is observable, and write the probability of any person's having an attribute $a \in A_k$ thus:

$$\pi_k := \Pr \{a \in A_k\} = \int_{a \in A_k} dF(a, x) , \quad (21)$$

$$\mu_{\phi}^k := \int_{a \in A_k} \phi(x) dF(a, x) , \quad (22)$$

$$\mu_{\phi}^k := \int_{a \in A_k} \phi(x)^2 dF(a, x) . \quad (23)$$

Because the attributes are observable it is possible to design a sampling frame conditioned on the attribute values in the population. The number of observations with attributes that fall in class k is n_k ; assume that sampling is random within each class k so that for any population member j with income-attribute pair (a_i, x_i) :

$$\Pr \{(a_i, x_i) \in F^{(n)} \mid a_i \in A_k\} = \frac{1}{n_k} \quad (24)$$

Furthermore we shall write the probability that a sample observation i has an attribute that falls in any one of the K attribute classes A_k thus:

$$p_k := \Pr \{a_i \in A_k\} \quad (25)$$

– the same for all i . By definition we have

$$\pi_k \geq 0, p_k \geq 0, \sum_{k \in K} \pi_k = 1, \sum_{k \in K} p_k = 1 \quad (26)$$

Note that each observation belongs to exactly one attribute class, and that π_k refers to population probabilities while p_k refers to sample probabilities. Since sampling is random within each class the sample moments

$$m_{\phi}^k := \sum_{a_i \in A_k} \phi(x_i) \quad (27)$$

$$m_{\phi}^k := \sum_{a_i \in A_k} \phi(x_i)^2 \quad (28)$$

will be unbiased estimates of their population counterparts (22), (23). Writing the distribution of x within attribute class k as F_k it is evident that we may express the sample distribution of x (for any observation in the sample) as

$$\bar{F}(x) := \sum_{k \in K} p_k F_k(x) \quad (29)$$

which will not be the same as the distribution of x in the population, since in general the probabilities π_k and p_k are different. What of the sample probabilities themselves? The specification of these, and the specification of appropriate estimators for them, will depend upon the method by which the membership – and hence the size – of the attribute classes is to be determined.

The main variants of this – cases two and three mentioned above – are discussed separately in sections 5 and 6.

5 Design Weights: Exogenous Attribute Classes

In this case the size of each attribute class within the sample (in other words the set of sample observations that lies in the subset A_k) has been exogenously fixed *a priori*, so that the numbers n_k are non-stochastic.⁸ This is akin to simple stratified sampling, and should be distinguished from the case of grossing-up weights discussed below. Using (21) and (25) we may define the stratification weight s_i for each observation i with reference to the attribute class within which i falls:

$$s_i := \frac{\pi_k}{p_k} \text{ iff } a_i \in A_k \quad (30)$$

for all i, k . Information about the π_k may be available independently or these probabilities may themselves have to be estimated, a point to which we shall return; for the moment we shall assume that they are known. Because the sub-sample sizes are independently fixed:

$$p_k = \frac{n_k}{n} . \quad (31)$$

These weights are non-stochastic and so, given that the sample has been randomly drawn within each of the A_k , the weights s_i will automatically sum to n . Contrast this with the case of welfare-consistent weighting where the sum-of-weights is stochastic.

Now consider the problem of estimating the welfare measure $W^*(F)$ in this case. We require the statistic

$$W^* (F^{(n)}) \Big|_s = \sum_{i=1}^n s_i \phi(x_i) . \quad (32)$$

⁸See for example Beach and Kaliski (1986).

In view of (30), (32) implies

$$W^* (F^{(n)})|_s = \frac{1}{n} \sum_{k \in K} \frac{\pi_k}{p_k} \sum_{a_i \in A_k} \phi(x_i) . \quad (33)$$

Formula (32) for the point estimate of the welfare measure in this case is similar to that for the welfare-consistent weighted case (16): one replaces the $\hat{w}(a_i)$ terms by s_i s. However the sampling distribution of this statistic is different in an important respect. Theorem 1 is not applicable, as it stands, because sampling is no longer random here. Instead we have:

Theorem 3 *Under exogenous-attribute class weighting $W^* (F^{(n)})$ is asymptotically distributed $N(W^*(F), V_s)$ where*

$$V_s = \frac{1}{n} \left[\sum_{k \in K} \frac{\pi_k^2}{p_k} \mu_{\phi}^k - [\mu_{1\phi}]^2 \right] - \frac{n-1}{n} \sum_{k \in K} \frac{1}{n_k} [\pi_k \mu_{\phi}^k]^2 . \quad (34)$$

Proof. See Appendix.

Notice that the expected value of the estimate may be written in terms of the conditional moments (22) as

$$\mathcal{E} W^* (F^{(n)})|_s = \mu_{1\phi} = \sum_{k \in K} [\pi_k \mu_{\phi}^k] , \quad (35)$$

and that an estimator for V_s can be found by replacing n by $n-1$, and the population conditional moments by the sample counterparts, given in (27), (28), thus:

$$\hat{V}_s = \frac{1}{n-1} \left[\sum_{k \in K} \frac{\pi_k^2}{p_k} m_{\phi}^k - [m_{1\phi}]^2 \right] - \sum_{k \in K} \frac{1}{n_k} [\pi_k m_{\phi}^k]^2 . \quad (36)$$

We shall consider further the components of this expression for the variance in more detail after we have introduced the case in section 6.

6 Endogenous Attribute Classes

In this case membership of the attribute classes A_k is not determined by the sample design, but rather by inspection of observations in the sample. So, although the sample size n is exogenously fixed, the numbers n_k (which sum to n) are themselves stochastic, and the probability of inclusion in the sample p_k is no longer automatically equal to the sample proportions n_k/n , as in the case of exogenous attribute classes, above.

Using (21) and (25) we may then define the grossing-up weight g_i for each observation i with reference to the population probabilities (π_k) and the sample

probabilities (p_k) for the attribute class within which i happens to fall, for all i, k :

$$g_i := \frac{\pi_k}{p_k} \text{ iff } a_i \in A_k . \quad (37)$$

However, the sample probabilities p_k will be unknown in this case, because we do not have an exogenously determined sampling frame, as we did in the case considered in section 5. However the numbers n_k – which are random variables – may be used to estimate them. We shall assume that the design of the sample is such that:

$$\forall k \in K : \text{plim} \left(\frac{n_k}{n} \right) = p_k . \quad (38)$$

Thus, when we aggregate over the grossing-up weights:

$$\sum_{i=1}^n g_i := \sum_{k \in K} n_k \frac{\pi_k}{p_k} \quad (39)$$

$$\simeq n , \quad (40)$$

with equality for large samples.

Now let us consider estimation of the welfare measure $W^*(F)$ in this case. We require the statistic

$$W^* (F^{(n)}) \Big|_g = \sum_{i=1}^n g_i \phi(x_i) \quad (41)$$

which is again of the same form as (16) and (32) – the corresponding equation in the case of individual weights. In view of (37), (41) also implies (33).

Theorem 4 *Under endogenous-attribute class weighting $W^* (F^{(n)})$ is asymptotically distributed $N(W^*(F), V_g)$ where*

$$V_g := \frac{1}{n} \left[\sum_{k \in K} \frac{\pi_k^2}{p_k} \mu_{\phi}^k - [\mu_{1\phi}]^2 \right] . \quad (42)$$

Proof. See Appendix.

Using definitions (4), (27), (28) and also the estimator

$$\hat{g}_i := \frac{n\pi_k}{n_k} \text{ iff } a_i \in A_k \quad (43)$$

an appropriate estimator of (42) is

$$\hat{V}_g = \frac{1}{n-1} \left[\sum_{i=1}^n \hat{g}_i^2 \phi(x_i)^2 - [m_{1\phi}]^2 \right] \quad (44)$$

$$= \frac{1}{n-1} \left[m_{2\phi} - [m_{1\phi}]^2 \right] \quad (45)$$

$$= \frac{1}{n-1} \left[\sum_{k \in K} \frac{n}{n_k} \pi_k^2 m_{\phi}^k - [m_{1\phi}]^2 \right] . \quad (46)$$

7 Comparison of Variance Formulae

We now have three different formulae for the variance of the estimate of the welfare index in the three cases of the weights. Let us see how they compare.

First, notice that the variance expression for the grossing-up case (42) contains fewer terms than the corresponding formula for the welfare-consistent weighted case – equation (17) – since the variance-of-attributes and the attributes-income-evaluation covariance components are not relevant to the “grossing-up” case. We have:

$$V_{\hat{w}} - V_s = \frac{1}{n} \frac{\mu_{1\phi}^2}{\mu_{10}^4} [\mu_{20} - \mu_{10}^2] - \frac{2}{n} \frac{\mu_{1\phi}}{\mu_{10}^3} [\mu_{2\phi} - \mu_{1\phi}\mu_{10}] . \quad (47)$$

Now consider the relationship between the sampling variance of the welfare statistic in the case of endogenous attribute classes with the exogenous case discussed in section 6 above. Comparing expression (42) with (34) we find, in general:

$$w_k := \frac{\pi_k}{p_k} \quad (48)$$

$$V_g - V_s = \frac{n-1}{n^2} \sum_{k \in K} p_k [w_k \mu_{\phi}^k]^2 . \quad (49)$$

In the case of both stratification and grossing-up weights one needs information about the population proportions π_k . It may be that the π_k can be obtained independently: for example, if the partition $\{A_k : k \in K\}$ corresponds to household size then we may have to hand census data on the relative proportions of different-sized households in the population (which would yield the π_k) and which can be compared with the proportions of different-sized households that appear in the sample.⁹ However, it is sometimes the case that the π_k have to be estimated, along with the p_k . In such circumstances then clearly there will be an additional component to the terms in (34) and (42).

7.1 Example

The implications of this can be seen in an example using the simple welfare index $\mu_{1\phi}$. Suppose the data are in the form of Table 1, where the first ν observations have an attribute value 1, and the remaining $n - \nu$ have attribute value 2, and consider two alternative stories underlying this situation.

- A: the two groups of observations refer respectively to 1-person and to 2-person families, and the x s refer to income per person; we are interested in income distribution amongst persons, and for this reason we want to take into account this information about family size.

⁹See, for example Atkinson, Gomulka, and Sutherland (1988).

weight	income	
1	x_1	
1	x_2	
1	x_3	
...	...	
1	x_ν	
2	$x_{\nu+1}$	
2	$x_{\nu+2}$	
2	$x_{\nu+3}$	
...	...	
2	x_n	

Table 1: Hypothetical distribution of (a, x)

- B: the structure of the sample is such that those with $a = 1$ are more likely to be included in the sample than those with $a = 2$; type-1 observations are likely to be over-represented in the sample by a factor of two when we check the relative proportions of type-1 and type-2 cases in the population.

In each story the weight is equal to the numerical value of the attribute a . The weights and the within-group moments are identical in each of the two cases; and so the point estimate of the welfare measure is identical in the two cases – see equations (16) and (41). However, when we estimate the variance of the welfare statistic we will find that cases A and B differ by an amount given by (47), where $\mu_{10} = 2 - \lambda$, $\mu_{20} - \mu_{10}^2 = \lambda - \lambda^2$ and $\lambda := \nu/n$. If we further assume that the distribution of income is the same in each of the two groups we find

$$V_{\hat{w}} - V_s = -\frac{1}{n} \mu_{1\phi}^2 \frac{\mu_{20} - \mu_{10}^2}{\mu_{10}^4} \quad (50)$$

$$= -\frac{1}{n} \mu_{1\phi}^2 \frac{\lambda - \lambda^2}{[2 - \lambda]^4}. \quad (51)$$

The result generalises if instead of the population consisting only of families of size 1 and size 2 we consider a population consisting only of families of size 1 and size H . Then (51) becomes

$$V_{\hat{w}} - V_s = -\frac{1}{n} \mu_{1\phi}^2 [H - 1]^2 \frac{\lambda - \lambda^2}{[H - [H - 1]\lambda]^4}. \quad (52)$$

Figure 1 shows, for three values of H , the bias in (52) expressed as a proportion of $n^{-1} \mu_{1\phi}^2$, where $\mu_{1\phi}$ is the true value of the welfare index. To interpret Figure 1 note the following:

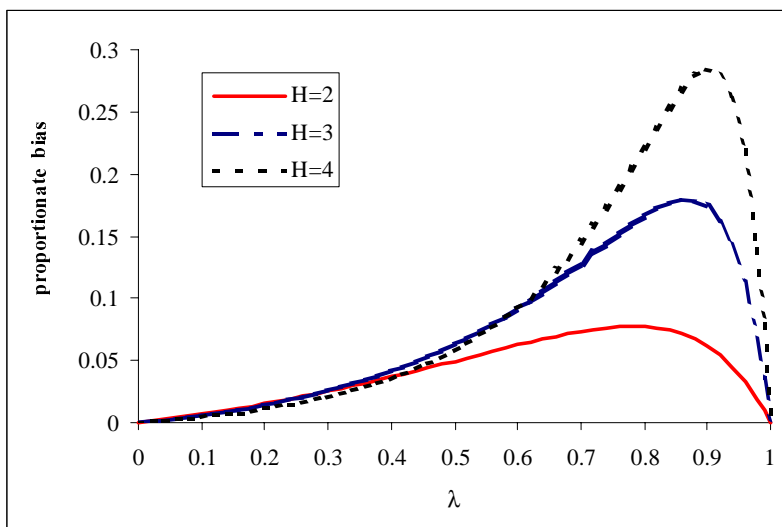


Figure 1: Composition of sample (λ) and the bias on estimated s.e. as a proportion of $n\mu_{1\phi}^2$

- The absolute implied bias depends on the heterogeneity of the population by size. In our example the maximum bias clearly increases with H .
- The bias is asymmetric in λ : it is particularly large if the proportion of single-person households is in the range (0.7-0.9)
- If the sample variance of $m_{1\phi}$ is about two to three times $\frac{1}{n}\mu_{1\phi}^2$ (a reasonable assumption for many practical cases) then a bias of 0.15 in Figure 1 would imply an error of about 5-7½ percent on the sample variance.

8 Conclusions

The types of weights that we have to use can significantly affect the distribution of an estimated welfare index. We have to distinguish between the situation of random sampling with welfare-consistent weighted data, and that of simple data in a non-random sample. The latter case requires explicit evaluation of the numbers p_k , the probability of inclusion in different subsamples, which will depend on the sampling procedure adopted: it matters considerably whether the sample proportions are fixed, or are probabilistic.

Our paper is in the nature of cautionary tale, drawing attention to potential problems. There are of course further issues that we have not addressed within this framework. For example, in practice, one would also need to allow for the

interaction of clustering and weighting and the use of weights to handle simultaneously the issues of welfare consistency, sample design, and grossing-up. We defer these topics to future research.

References

- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory* 2, 244–263.
- Atkinson, A. B., J. Gomulka, and H. Sutherland (1988). Grossing-up FES data for tax-benefit models. In A. B. Atkinson and H. Sutherland (Eds.), *Tax-Benefit Models*, pp. 223–253. London School of Economics, London: STICERD.
- Beach, C. M. and S. F. Kaliski (1986). Lorenz Curve inference with sample weights: an application to the distribution of unemployment experience. *Applied Statistics* 35(1), 38–45.
- 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.
- Blackorby, C., D. Donaldson, and M. Auersperg (1981). A new procedure for the measurement of inequality within and among population subgroup. *Canadian Journal of Economics* 14, 665–685.
- Coulter, F. A. E., F. A. Cowell, and S. P. Jenkins (1992). Equivalence scale relativities and the extent of inequality and poverty. *Economic Journal* 102, 1067–1082.
- Cowell, F. A. (1984). The structure of American income inequality. *Review of Income and Wealth* 30, 351–375.
- Cowell, F. A. (1989). Sampling variance and decomposable inequality measures. *Journal of Econometrics* 42, 27–41.
- Cowell, F. A. (2000). Measurement of inequality. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Chapter 2. Amsterdam: North Holland.
- Danziger, S. and M. K. Taussig (1979). The income unit and the anatomy of income distribution. *Review of Income and Wealth* 25, 365–375.
- Ebert, U. (1987). Size and distribution of incomes as determinants of social welfare. *Journal of Economic Theory* 41, 25–33.
- Howes, S. R. and J. O. Lanjouw (1998). Poverty comparisons and household survey design. *Review of Income and Wealth* 44, 99–108.
- Jenkins, S. P. (1997). Trends in real income in Britain: a microeconomic analysis. *Empirical Economics* 22(4), 483–500.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. New York: John Wiley.
- Ravallion, M. (1994). *Poverty Comparisons: a guide to concepts and methods*. Chur, Switzerland: Harwood Academic Publishers.

- Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. *Econometrica* 52, 1369–1385.
- Victoria-Feser, M.-P. (1999). The sampling properties of inequality indices: Comment. In J. Silber (Ed.), *Income Inequality Measurement: From Theory to Practice*, pp. 260–267. Dordrecht: Kluwer.

A Proofs

A.1 Theorem 3

If the sample has been randomly drawn within each of the A_k then again we find

$$\mathcal{E} W^* (F^{(n)})|_s = W^*(F). \quad (53)$$

Cf. equation (18). Evaluating the required variance we have

$$\begin{aligned} \text{var } W^* (F^{(n)})|_s &= \frac{1}{n^2} \sum_{i=1}^n \sum_{k \in K} \Pr \{a_i \in A_k\} \left[\mathcal{E} (s_i^2 \phi(x_i)^2 | a_i \in A_k) \right. \\ &\quad + \sum_{\substack{j=1 \\ j \neq i}}^n [\Pr \{a_j \in A_k\} \mathcal{E} (s_i s_j \phi(x_i) \phi(x_j) | a_i, a_j \in A_k) \\ &\quad + \sum_{\ell \in K \setminus \{k\}} \Pr \{a_j \in A_\ell\} \mathcal{E} (s_i s_j \phi(x_i) \phi(x_j) | a_i \in A_k, a_j \in A_\ell)] \left. \right] \\ &\quad - [\mathcal{E} W^* (F^{(n)})]^2 \end{aligned} \quad (54)$$

which, taking into account the definitions of the conditional moments (4), (5) and of the stratification weights (30), reduces to:

$$\begin{aligned} &\frac{1}{n} \sum_{k \in K} \frac{\pi_k^2}{p_k} \mu_{\phi}^k + \frac{n-1}{n} \sum_{k \in K} \left[1 - \frac{1}{n_k} \right] [\pi_k \mu_{\phi}^k]^2 \\ &+ \frac{n-1}{n} \sum_{k \in K} \sum_{\ell \in K \setminus \{k\}} \pi_k \pi_\ell \mu_{\phi}^k \mu_{\phi}^\ell - [\mathcal{E} W^* (F^{(n)})]^2 \end{aligned} \quad (55)$$

from which (34) follows immediately. ■

A.2 Theorem 4

If the sample has been randomly drawn within each of the A_k then once again

$$\mathcal{E} W^* (F^{(n)})|_g = W^*(F). \quad (56)$$

Cf. equation (18). Turning to the sampling variance of the estimate, we have by definition:

$$\text{var } W^* (F^{(n)})|_g = \mathcal{E} \left(\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n g_i g_j \phi(x_i) \phi(x_j) \right) - [\mathcal{E} W^* (F^{(n)})]^2 \quad (57)$$

the right-hand side of which, in view of the structure of the sample, may be written

$$\begin{aligned} & \frac{1}{n} \sum_{k \in K} \Pr \{a_i \in A_k\} \mathcal{E} (g_i^2 \phi(x_i)^2 \mid a_i \in A_k) \\ & + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \mathcal{E} (s_i \phi(x_i)) \mathcal{E} (s_j \phi(x_j)) - [\mathcal{E} W^* (F^{(n)})]^2 . \end{aligned} \quad (58)$$

Using definitions (22), (23) and (37), we find that (58) yields

$$\text{var } W^* (F^{(n)}) \Big|_g = \frac{1}{n} \sum_{k \in K} p_k \left[\frac{\pi_k}{p_k} \right]^2 \mu_{\phi}^k - \frac{1}{n} [\mathcal{E} W^* (F^{(n)})]^2 \quad (59)$$

which immediately gives (42) ■