Dear Author

Here are the proofs of your article.

- You can submit your corrections online, via e-mail or by fax.
- For online submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and email the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the journal title, article number, and your name when sending your response via e-mail or fax.
- Check the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- Check the questions that may have arisen during copy editing and insert your answers/corrections.
- Check that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the Edited manuscript.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please do not make changes that involve only matters of style. We have generally introduced forms that follow the journal’s style.
- Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections within 48 hours, we will send you a reminder.
- Your article will be published Online First approximately one week after receipt of your corrected proofs. This is the official first publication citable with the DOI. Further changes are, therefore, not possible.
- The printed version will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL:

http://dx.doi.org/10.1007/s10888-012-9238-z

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information, go to:


Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us, if you would like to have these documents returned.
<table>
<thead>
<tr>
<th><strong>Metadata of the article that will be visualized in OnlineFirst</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Please note:</strong> Images will appear in color online but will be printed in black and white.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>Article Title</th>
<th>Reference distributions and inequality measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Article Sub-Title</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Article Copyright - Year</td>
<td>Springer Science+Business Media New York 2013  (This will be the copyright line in the final PDF)</td>
</tr>
<tr>
<td>4</td>
<td>Journal Name</td>
<td>The Journal of Economic Inequality</td>
</tr>
<tr>
<td>5</td>
<td>Corresponding Author</td>
<td>Flachaire</td>
</tr>
<tr>
<td>6</td>
<td>Family Name</td>
<td>Flachaire</td>
</tr>
<tr>
<td>7</td>
<td>Particle</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Given Name</td>
<td>Emmanuel</td>
</tr>
<tr>
<td>9</td>
<td>Suffix</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Organization</td>
<td>Aix-Marseille Université</td>
</tr>
<tr>
<td>11</td>
<td>Division</td>
<td>Greqam</td>
</tr>
<tr>
<td>12</td>
<td>Address</td>
<td>2 rue de la Charité, 13002, Marseille, France</td>
</tr>
<tr>
<td>13</td>
<td>e-mail</td>
<td><a href="mailto:emmanuel.flachaire@univ-cezanne.fr">emmanuel.flachaire@univ-cezanne.fr</a></td>
</tr>
<tr>
<td>14</td>
<td>Author</td>
<td>Cowell</td>
</tr>
<tr>
<td>15</td>
<td>Family Name</td>
<td>Cowell</td>
</tr>
<tr>
<td>16</td>
<td>Particle</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Given Name</td>
<td>Frank A.</td>
</tr>
<tr>
<td>18</td>
<td>Suffix</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Organization</td>
<td>London School of Economics</td>
</tr>
<tr>
<td>20</td>
<td>Division</td>
<td>Sticerd</td>
</tr>
<tr>
<td>21</td>
<td>Address</td>
<td>Houghton Street, London, WC2A 2AE, UK</td>
</tr>
<tr>
<td>22</td>
<td>e-mail</td>
<td>None</td>
</tr>
<tr>
<td>23</td>
<td>Author</td>
<td>Bandyopadhyay</td>
</tr>
<tr>
<td>24</td>
<td>Family Name</td>
<td>Bandyopadhyay</td>
</tr>
<tr>
<td>25</td>
<td>Particle</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Given Name</td>
<td>Sanghamitra</td>
</tr>
<tr>
<td>27</td>
<td>Suffix</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Organization</td>
<td>University of London</td>
</tr>
<tr>
<td>29</td>
<td>Division</td>
<td>Queen Mary</td>
</tr>
<tr>
<td>30</td>
<td>Address</td>
<td>Mile End Road, London, E1 4NS, UK</td>
</tr>
<tr>
<td>31</td>
<td>e-mail</td>
<td>None</td>
</tr>
</tbody>
</table>

| 32 | Schedule | Received 3 May 2011 |
| 33 |            | Revised 7 December 2012 |
Abstract

We investigate a general problem of comparing pairs of distribution which includes approaches to inequality measurement, the evaluation of “unfair” income inequality, evaluation of inequality relative to norm incomes, and goodness of fit. We show how to represent the generic problem simply using (1) a class of divergence measures derived from a parsimonious set of axioms and (2) alternative types of “reference distributions.” The problems of appropriate statistical implementation are discussed and empirical illustrations of the technique are provided using a variety of reference distributions.

Keywords

Divergence measures - Generalised entropy measures - Income distribution - Inequality measurement - D63 - C10

Foot note information
Reference distributions and inequality measurement

Frank A. Cowell · Emmanuel Flachaire · Sanghamitra Bandyopadhyay

Received: 3 May 2011 / Accepted: 7 December 2012 © Springer Science+Business Media New York 2013

Abstract We investigate a general problem of comparing pairs of distribution which includes approaches to inequality measurement, the evaluation of “unfair” income inequality, evaluation of inequality relative to norm incomes, and goodness of fit. We show how to represent the generic problem simply using (1) a class of divergence measures derived from a parsimonious set of axioms and (2) alternative types of “reference distributions.” The problems of appropriate statistical implementation are discussed and empirical illustrations of the technique are provided using a variety of reference distributions.

Keywords Divergence measures · Generalised entropy measures · Income distribution · Inequality measurement

JEL Classification D63 · C10

1 Introduction

There is a broad class of problems in distributional analysis that involve comparing two distributions. This may involve judging whether a functional form is a good fit to an empirical distribution; it may involve computation of the divergence of

F. A. Cowell
Sticerd, London School of Economics, Houghton Street, London WC2A 2AE, UK

E. Flachaire (✉)
GREQAM, Aix-Marseille Université, 2 rue de la Charité 13002 Marseille, France
e-mail: emmanuel.flachaire@univ-cezanne.fr

S. Bandyopadhyay
Queen Mary, University of London, Mile End Road, London E1 4NS, UK
an empirical distribution from a theoretical economic model; it may involve the
ethical evaluation of an empirical distribution with reference to some norm or ideal
distribution. This paper shows how the class of problems can be characterised in a
way that has a natural interpretation in terms of familiar analytical tools.

This is not some recondite or abstruse topic. Several authors have explicitly
characterised inequality using this two-distribution paradigm: an inequality measure
is defined in terms of the divergence of an empirical income distribution from
an equitable reference distribution [5, 20, 28]. Furthermore several recent papers
have revived interest in the idea of inequality evaluations with reference to “norm
incomes” or a reference distribution; in particular some authors have focused on
replacing a perfectly egalitarian reference distribution with one that takes explicit
account of fairness [3, 19]. In recent contributions Magdalou and Nock [27] have also
examined the concept of divergence between any two income distributions and its
economic interpretation and Cowell et al. [8] show how similar concepts can be used
to formulate an approach to the measurement of goodness of fit.

In this paper we use an a priori approach to the problem that allows one to
construct a distance concept that is appropriate for characterising the divergence
between the empirical distribution function and a proposed reference distribution.
The approach is related to results in information theory and is adaptable to other
fields in economics that make use of models of distributions [10]. The paper is
structured as follows. Section 2 introduces the concept of divergence and a set of
principles for distributional comparisons in terms of divergence; we show how these
principles characterise a class of measures and discuss different concepts of reference
distribution that are relevant for different versions of the generic problem under
consideration. Section 3 discusses issues of implementation and Section 4 performs
a set of experiments and applications using the proposed measures and UK income
data. Section 5 concludes.

2 The approach

The essence of the problem is the characterisation of the divergence of an income
distribution from a reference distribution: this divergence can be seen as an ag-
gregation of discrepancies in different parts of the distribution. To fix ideas we
may interpret this in terms of a classic modelling problem: comparing an Empirical
Distribution Function (EDF) \( \hat{F}(x) = \frac{1}{n+1} \sum_{i=1}^{n} \iota(x_i \leq x) \) and a theoretical reference
distribution \( F_\ast \). For instance, Fig. 1 presents an EDF (dotted line) and a theoretical

\( \iota(S) = 1 \) if statement \( S \) is true and \( \iota(S) = 0 \) otherwise. We use \( \frac{1}{n+1} \)
rather than \( \frac{1}{n} \) to avoid a problem where \( i = n \). Had we used \( \frac{1}{n} \) in Eq. 1 then \( y_n \) would automatically
be set to \( \sup(X) \) where \( X \) is the support of \( F_\ast \).

Almås et al. [4] compare “actual and equalizing earnings”; their work is related to Paglin’s Gini [29]
and Wertz’s Gini [32]. Also see Jenkins and O’Higgins [25] and Garvy [24].

The Cowell et al. [8] approach differs from that developed here in that it deals with the problem of
continuous reference distributions on unbounded support.

\( \iota \) is an indicator function such that \( \iota(S) = 1 \) if statement \( S \) is true and \( \iota(S) = 0 \) otherwise. We use \( \frac{1}{n+1} \)
rather than \( \frac{1}{n} \) to avoid a problem where \( i = n \). Had we used \( \frac{1}{n} \) in Eq. 1 then \( y_n \) would automatically
be set to \( \sup(X) \) where \( X \) is the support of \( F_\ast \).
reference distribution (solid line) on a reverse graph, with \( q \) denoting proportions of the population on the horizontal axis and the income quantiles on the vertical axis. There are two obvious ways to describe discrepancies between the EDF and the reference distribution:

1. The standard approach in the goodness-of-fit literature is based on the “horizontal” differences between \( \hat{F}(x) \) and \( F^*_q(x) \) for any given \( x \).

2. An alternative approach is to compare, for given \( q \) values, the corresponding \( q \)-quantiles given by the EDF and \( F^*_q \). Denote by \( \{x^{(1)}, x^{(2)}, \ldots, x^{(n)}\} \) the members of the sample in increasing order; the corresponding values given by the EDF are the adjusted sample proportion \( q = \{\frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1}\} \) and, for each \( q \), the corresponding value for the reference distribution is equal to

\[
y_i = F^{-1}_*(\frac{i}{n+1}).
\]  

Here we look at the “vertical” differences between the \( x^{(i)} \) and \( y_i \). As we will see below it is this alternative view that is more fruitful for application to the inequality-measurement problem.

2.1 Aggregation of information

A divergence measure for the EDF and the reference distribution aggregates discrepancies between the quantiles of the two distributions, for each value of \( q \). Each discrepancy concerns an income pair \( (x_i, y_i) \)—the \( i \)th quantile in the two distributions—and the profile of these pairs captures all the essential information for the problem to which the divergence measure will be applied. We suggest that an appropriate divergence measure should satisfy the following principles: (1) If two
profiles are equivalent in terms of divergence and the discrepancy for pair \(i\) is the same in each profile, then a local variation at \(i\) simultaneously in the two profiles has no overall effect. (2) If there is zero discrepancy at each of two pairs in the profile then moving \(x\)-income and \(y\)-income simultaneously from one pair to the other has no effect on divergence. (3) Divergence remains unchanged by a uniform scale change to \(x\)-values and \(y\)-values simultaneously. (4) Given two discrepancy profiles with the same aggregate divergence, rescaling all the income discrepancies in each profile by the same factor results in two new profiles that are equivalent in terms of divergence.\(^4\)

Using the result in Theorem 1 (in the Appendix) and normalising on the case where both the observed and the reference distribution exhibit complete equality, these principles gives the following class of measures:

\[
J_\alpha (x, y) := \frac{1}{n\alpha(n - 1)} \sum_{i=1}^{n} \left[ \left( \frac{x_i}{\mu_1} \right)^{\alpha} \left( \frac{y_i}{\mu_2} \right)^{1-\alpha} - 1 \right],
\]

where \(\alpha\) takes any real value\(^5\) and \(J_\alpha (x, y) \geq 0\) for arbitrary \(x\) and \(y\).\(^6\) There is an analogy with the Kullback and Leibler [26] index of relative entropy [10].

So a divergence measure for the EDF and a theoretical reference distribution \(F_\ast\) would be given by replacing \(x_i\) and \(y_i\) in Eq. 2, respectively, by \(EDF^{-1}(\frac{i}{n+1}) = x(i)\) and \(F_\ast^{-1}(\frac{i}{n+1})\):\(^7\)

\[
J_\alpha = \frac{1}{n\alpha(n - 1)} \sum_{i=1}^{n} \left[ \left( \frac{x(i)}{\mu} \right)^{\alpha} \left( \frac{F_\ast^{-1}(\frac{i}{n+1})}{\mu(F_\ast)} \right)^{1-\alpha} - 1 \right], \quad \alpha \neq 0, 1. \quad (3)
\]

The \(J\) index can be used to measure the divergence between an empirical income distribution, given by a sample of individual incomes, and any theoretical reference distribution. It requires the choice of specific values for the parameter \(\alpha\) according to the judgment that one wants to make about the relative importance of different types of discrepancy: choosing a large positive value for \(\alpha\) puts weight on parts of the distribution where the observed incomes \(x_i\) greatly exceed the corresponding values \(y_i\) in the reference distribution; choosing a substantial negative value puts weight on cases where the opposite type of discrepancy arises.

\(^4\)See Axioms 4, 5, 6, 7 in the Appendix.

\(^5\)\(J_0 (x, y) = -\frac{1}{n} \sum_{i=1}^{n} \frac{n}{\mu_1} \log \left( \frac{\mu_2}{\mu_1} \right)\), \(J_1 (x, y) = \frac{1}{n} \sum_{i=1}^{n} \frac{\mu_2}{\mu_1} \log \left( \frac{\mu_2}{\mu_1} \right)\).

\(^6\)To see this write \(J_\alpha (x, y)\) as \(\sum_{i=1}^{n} \frac{\mu_2}{\mu_1} \left[ \psi (q_i) - \psi (1) \right]\), where \(q_i := \frac{y_i}{\mu_2} \), \(\psi (q) := \frac{a^q}{q^{\alpha-1}}\). Because \(\psi\) is convex function, for any \((q_1, ..., q_n)\) and any set of non negative weights \((w_1, ..., w_n)\) that sum to 1, \(\sum_{i=1}^{n} w_i \psi (q_i) \geq \psi \left( \sum_{i=1}^{n} w_i q_i \right)\). Letting \(w_i = y_i / [n\mu_2]\) and using the definition of \(q_i\) we can see that \(w_i q_i = x_i / [n\mu_1]\) so we have \(\sum_{i=1}^{n} w_i \psi (q_i) \geq \psi (1)\) and the result follows. See also [6].

\(^7\)The limiting forms are \(J_0 = -\frac{1}{n} \sum_{i=1}^{n} \frac{F_\ast^{-1}(\frac{i}{n+1})}{\mu(F_\ast)} \log \left( \frac{x(i)}{\mu} \right)\), \(J_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{x(i)}{\mu} \log \left( \frac{x(i)}{\mu} \right) \).

\(\odot\) Springer
2.2 Reference distributions

The use of the $J$ index also requires the specification of a reference distribution: there are several possibilities.

**The most equal reference distribution** Let us assume that the most equal income distribution is when the same amount is given to each individual:

$$F_{i}^{-1}\left(\frac{i}{n+1}\right) = \hat{\mu} \quad \text{for } i = 1, \ldots, n$$ (4)

If we use this (egalitarian) distribution as the reference distribution in Eq. 3, then we find the standard Generalised Entropy inequality measure:

$$J_{\alpha} = \frac{1}{n\alpha(\alpha - 1)} \sum_{i=1}^{n} \left[ \left(\frac{x_i}{\hat{\mu}}\right)^{\alpha} - 1 \right], \quad \alpha \neq 0, 1$$ (5)

So the GE inequality measures are divergence measures between the EDF and the most equal distribution, where everybody gets the same income. They tell us how far a distribution is from the most equal distribution. A sample with a smaller index has a more equal distribution.

Figure 2 presents the quantile approach for this case. We can see that the EDF is always above (below) the reference distribution for large (small) values of incomes. It makes clear that large (small) values of $\alpha$ would be more sensitive to changes in high (small) incomes.

---

8The limiting forms are $J_{0} = -\frac{1}{n} \sum_{i=1}^{n} \log \left(\frac{x_i}{\hat{\mu}}\right), J_{1} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{\hat{\mu}} \log \left(\frac{x_i}{\hat{\mu}}\right)$. 

The most unequal reference distribution  Rather than selecting the most equal distribution as a reference distribution, we can reverse the standard approach by using the most unequal distribution as a reference distribution. The most unequal income distribution is when one person gets all the income and the others zero:

\[
F^{-1}_* \left( \frac{i}{n+1} \right) = \begin{cases} 
0 & \text{for } i = 1, \ldots, n-1 \\
\frac{i n \hat{\mu}}{n} & \text{for } i = n 
\end{cases} 
\]  

(6)

If we use this distribution as the reference distribution in Eq. 3, then we have:

\[
J_\alpha = \frac{1}{\alpha(\alpha - 1)} \left[ \left( \frac{x_{(i)}}{n \hat{\mu}} \right)^\alpha - 1 \right], \alpha < 1, \alpha \neq 0 
\]

(7)

This index tells us how far a distribution is from the most unequal distribution. However, two major drawbacks make this index useless in practice:

1. To be comparable for two different samples, the index should use the same reference distribution in both samples. Here \( \max x_i = x_{(n)} \) is an estimate of the \( n/(n+1) \)-quantile in \( F_* \). It follows that, if \( n = 100 \), the reference distribution is when the top 1 % gets all the income, whereas if \( n = 1000 \) it is when the top 0.1 % gets all the income. The reference distribution differs with the sample size.

2. The presence of zero incomes in the reference distribution produce undesirable properties: the index is independent of how the first \( n-1 \) ordered incomes are distributed. The first \( n-1 \) ordered incomes do not appear explicitly in the formula, the index depends on them through the mean only. It follows that, the mean being constant, the distribution of the \( n-1 \) first ordered incomes does not matter. For instance, the two samples \( \{8, 8, 8, 8, 8, 8, 8, 8, 8\} \) and \( \{1, 1, 1, 1, 15, 15, 15, 15, 20\} \) produce the same value of the index.

These two drawbacks lead us to consider the following reference distribution, where the top \( 100k \% \) richest gets \( 100p \% \) of the total income:

\[
F^{-1}_* \left( \frac{i}{n+1} \right) = \begin{cases} 
(1 - p) \hat{\mu}/(1 - k) & \text{for } i = 1, \ldots, [n(1-k)] \\
p \hat{\mu}/k & \text{for } i = [n(1-k)] + 1, \ldots, n 
\end{cases} 
\]

(8)

with \( 0 \leq k \leq 1, 0 \leq p \leq 1 \) and \( \lceil z \rceil \) denotes the smallest integer not less than \( z \). Small values of \( k \) and large values of \( p \) produce very unequal distributions, where a few people get nearly all the income, and the rest get nearly zero. For instance, in setting \( k = 0.01 \) and \( p = 0.99 \) we take the case where the top 1 % richest gets 99 % of the total income. If we use this distribution as the reference distribution in Eq. 3, we obtain:

\[
J_{\alpha,k,p} = \frac{1}{n \alpha(\alpha - 1)} \sum_{i=1}^{n} \left[ \left( \frac{x_{(i)}}{\hat{\mu}} \right)^\alpha c_i^{1-\alpha} - 1 \right], \alpha \neq 0, 1, 
\]

(9)

\(^9\)There is a limiting case \( J_0 = - \log \left( \frac{\max x_i}{n \hat{\mu}} \right) \); but if \( \alpha = 1 \) the index is undefined.
Reference distributions and inequality measurement

Where

\[ c_i = \begin{cases} 
(1 - p)/(1 - k) & \text{if } i \leq \lceil n(1 - k) \rceil \\
p/k & \text{if } i > \lceil n(1 - k) \rceil 
\end{cases} \] (10)

There are two interesting special cases. If \( p = k \), everybody gets the same income value, \( \hat{\mu} \), and the reference distribution is the most equal distribution. If \( k = 1/n \) and \( p = 1 \), only one individual gets all the income, \( n\hat{\mu} \), and the reference distribution is the most unequal distribution.

In practice, \( k \) and \( p \) have to be fixed: (1) to avoid the first drawback, \( k \) and \( p \) should be independent of the sample size, with \( k > 1/n \) and \( p > 1/n \); (2) to avoid the second drawback, zero incomes are not allowed in the reference distribution, that is, if \( p = 1 \) we have \( k = 1 \). Finally, to make our index \( J_{a,k,p} \) useful in practice, we need to use constant values, such that

\[ 1/n < k < 1 \quad \text{and} \quad 1/n < p < 1, \quad \text{or} \quad p = k = 1. \] (11)

In empirical studies, we could use several values of \( k \) and \( p \). For instance, \( k = 1 - p = 0.05, 0.01, 0.005 \), correspond to the reference distributions with the top 5 %, 1 % and 0.5 % getting, respectively, 95 %, 99 % and 99.5 % of the total income.

Other reference distributions Clearly, other reference distributions could be used. For instance, if we assume that productive talents are distributed in the population according to a continuous distribution of talents \( F^* \) and that wages should be related to talent, a situation in which everyone received the same income to everybody might be considered as unfair. In this case one might use \( F^* \) as the reference distribution and make use of the index (Eq. 3): any deviation from \( F^* \) would come from something else than talent. If total income is finite, it makes sense to use a distribution defined on a finite support. For instance, we could use a Uniform distribution or a Beta distribution with two parameters, which can provide a variety of appropriate shapes.

3 Implementation

It is necessary to establish the existence of an asymptotic distribution for \( J_\alpha \) and \( J_{a,k,p} \) in order to justify its use in practice. If the most equal distribution is taken as the reference distribution \( (k = p = 1) \), the index \( J_{a,1,1} \) is nothing but the standard GE inequality measure, which is asymptotically Normal and has well-known statistical properties. If a continuous distribution is taken as the reference distribution, it can be shown that the limiting distribution of \( nJ_\alpha \) is that of

\[ \frac{1}{2\mu_{F^*}} \left[ \int_0^1 \frac{B^2(t)dt}{F^{-1}(t)f^2_x(F^{-1}(t))} - \frac{1}{\mu_{F^*}} \left( \int_0^1 \frac{B(t)dt}{f_x(F^{-1}(t))} \right)^2 \right] \] (12)

The limiting forms are

\[ J_{0,k,p} = -\frac{1}{n} \sum_{i=1}^n c_i \log \left( \frac{\mu_{F^*}}{\hat{\mu}_i} \right), J_{1,k,p} = \frac{1}{n} \sum_{i=1}^n \frac{\mu_{F^*}}{\hat{\mu}_i} \log \left( \frac{\mu_{F^*}}{\hat{\mu}_i} \right). \]

Among others, see Cowell and Flachaire [9], Davidson and Flachaire [12], Schluter and van Garderen [31], Schluter [30], Davidson [11].
where \( f^\ast \) is the density of distribution \( F^\ast \) and \( B(t) \) is a Brownian bridge. This random variable can have an infinite expectation. It is only if \( F^\ast \) has a bounded support that the limiting distribution has reasonable properties—see Cowell et al. [8] and Davidson [11] for more details. If we use a continuous parametric reference distribution, since total income is finite, it makes sense to use a distribution \( F^\ast \) defined on a bounded support only. For instance, one could use a Uniform distribution or a Beta distribution with two parameters, which can provide many different shapes. The same approach can be used for \( nJ_{a,k,p} \), noting that the last statistic is equivalent to the statistic defined in (9) in Cowell et al. [8], where \( 2i/(n+1) \) is replaced by \( c_i \) defined in Eq. 10.

In the two cases, the limiting distribution of \( J_a \) and \( J_{a,k,p} \) exists, but is not tractable. This is enough to justify the use of bootstrap methods for making inference. To compute a bootstrap confidence interval, we generate \( B \) bootstrap samples by resampling from the original data, and then, for each resample, we compute the index \( J \). We obtain \( B \) bootstrap statistics, \( J^b_a, b = 1, \ldots, B \). The percentile bootstrap confidence interval is equal to

\[
CI_{perc} = [c_{0.025}^b; c_{0.975}^b]
\]  

where \( c_{0.025}^b \) and \( c_{0.975}^b \) are the 2.5 and 97.5 percentiles of the EDF of the bootstrap statistics—for a comprehensive discussion on bootstrap methods, see Davison and Hinkley [17], Davidson and MacKinnon [15]. For well-known reasons [14, 17] the number \( B \) should be chosen so that \((B + 1)/100\) is an integer: here we set \( B = 999 \) unless otherwise stated.

To be used in practice, we need to determine the finite sample properties of \( J_a \) and \( J_{a,k,p} \). The coverage error rate of a confidence interval is the probability that the random interval does not cover the true value of the parameter. A method of constructing confidence intervals with good finite sample properties should generate a coverage error rate close to the nominal rate. For a confidence interval at 95 %, the nominal coverage error rate is equal to 5 %. We use Monte-Carlo simulation to approximate the coverage error rate bootstrap confidence intervals in several experimental designs.

In our experiments, samples are drawn from a lognormal distribution. For fixed values of \( \alpha, k, p \) and \( n \), we draw 10,000 samples. For each sample we compute \( J_a \) or \( J_{a,k,p} \) and its confidence interval at 95 %. The coverage error rate is computed as the proportion of times the true value of the inequality measure is not included in the confidence intervals. Confidence intervals perform well in finite samples if the coverage error rate is close to the nominal value, 0.05. Table 1 presents the coverage error rate of bootstrap confidence intervals at 95 % of \( J_a \) and \( J_{a,k,p} \) for several reference distributions. The standard GE measures use the most equal reference distribution, it corresponds to \( J_{a,1,1} \). When “the top 1 % richest gets 99 % of the income” is the reference distribution we use the index \( J_{a,0.01,0.99} \); when “the top 5 % richest gets 99 % of the income” is the reference distribution we use \( J_{a,0.05,0.99} \).

In addition, we examine \( J_a \) with two continuous (bounded) parametric reference distributions, the Beta(1,1) distribution which is equal to the Uniform(0,1), and the lognormal distribution in our experiments.

\[ x(i) \text{ in Eq. 3 by } F^{-1}(i/(n+1)) \], where \( F \) is the distribution of \( x \), that is, the lognormal distribution in our experiments.
Reference distributions and inequality measurement

Table 1 Coverage error rate of bootstrap confidence intervals at 95% of $J_\alpha$ and $J_{\alpha,k,p}$, 10,000 replications, 499 bootstraps, and $x \sim \text{Lognormal}(0, 1)$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
<th>$n = 500$</th>
<th>$n = 1000$</th>
<th>$n = 2000$</th>
<th>$n = 5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.0753</td>
<td>0.0747</td>
<td>0.0669</td>
<td>0.0658</td>
<td>0.0567</td>
<td>0.0557</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0734</td>
<td>0.0667</td>
<td>0.0673</td>
<td>0.0606</td>
<td>0.0565</td>
<td>0.0562</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0832</td>
<td>0.0713</td>
<td>0.0716</td>
<td>0.0642</td>
<td>0.0620</td>
<td>0.0606</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0912</td>
<td>0.0785</td>
<td>0.0781</td>
<td>0.0709</td>
<td>0.0658</td>
<td>0.0672</td>
</tr>
<tr>
<td>0.20</td>
<td>0.1166</td>
<td>0.1024</td>
<td>0.0983</td>
<td>0.0878</td>
<td>0.0831</td>
<td>0.0809</td>
</tr>
</tbody>
</table>

Beta(1,1) which is a symmetric inverted-U-shape distribution. Table 1 shows that the finite sample properties of the indices with alternative reference distributions are not very different from those of the standard GE measures, except for $J_{0.01,0.99}$ when $n \leq 500$. The coverage error rate is close to 0.05 for very large samples. For small and moderate samples, further investigation is required to improve the finite sample properties, with, for instance, a fast double or triple bootstrap [13, 16].
4 Application

Let us compare the performance of $J_{\alpha,k,p}$ with that of conventional GE inequality measures using UK income data as a case study. Table 2 presents the results of indices $J_{\alpha}$ and $J_{\alpha,k,p}$ estimated with three different types of reference distribution, along with bootstrap confidence intervals at 95%.

Equality The top panel of Table 2 presents estimates of $J_{\alpha,k,p}$ using an “equality” reference distribution. Clearly, when we select the most equal distribution as the reference distribution, i.e. $k = p = 1$, the index $J_{\alpha,k,p}$ is reduced to the standard GE inequality measure. Estimates for standard GE measures, $J_{\alpha,1,1}$ are tabulated in the first row, for values of $\alpha$ ranging from $-1$ to 2. When $\alpha = 1$, $J_{\alpha,1,1}$ is the Theil index. For values of $\alpha = 0.5, 1, 1.5, 2$, $J_{\alpha,1,1}$ represents transformed Atkinson indices [7]. All estimates of standard GE measures increase considerably between 1979 and 1988, suggesting a significant rise in inequality in the 80s.

Extreme inequality The key point highlighted earlier was that changing the reference distribution from which we measure the distance of the empirical distribution opens up the possibility for researchers to choose the exact distribution from which they wish to measure distance of the empirical distribution. While standard GE indices tell us about the distance of the empirical distribution from an equal reference distribution, one can change the focus to that of its distance from an unequal reference distribution. The second panel of Table 2 presents estimates of $J_{\alpha,k,p}$ using several “extreme inequality” reference distributions. The interpretation of the size of the $J_{\alpha,k,p}$ index now is the reverse of the interpretation of standard GE measures. For a standard GE inequality measure, a small value of $J_{\alpha,1,1}$ corresponds to the empirical distribution being close to the equal reference distribution compared to that of a large value of $J_{\alpha,1,1}$. However, for an unequal reference distribution a small value of $J_{\alpha,k,p}$ corresponds to the empirical distribution being close to the particular “extreme inequality” reference distribution that has been specified.

To illustrate we focus on two different unequal reference distributions: one, where the top 1% of the income distribution receive 99% of the income, and second, where the top 5% of the income distribution receive 99% of the income. From Table 2 we can see that, with one exception, the values of $J_{\alpha,k,p}$ have dropped between years 1979 and 1988: in other words, it is almost always true that the distance from the “extreme inequality” reference distribution has decreased. The exception is the case $(k = 0.05, p = 0.99, \alpha = 2)$ where the movement relative to the reference distribution is not significant. The implication is that UK inequality grew during the 1980s whether one interprets this in terms of distance from equality, or as distance from a reference unequal distribution, except for one case. This case concerns top-

---

13 The application uses the “before housing costs” income variable of the Family Expenditure Survey for years 1979 and 1988 [18], deflated and equivalised using the McClements adult-equivalence scale. We exclude households with self-employed individuals as reported incomes are known to be misrepresented. The years 1979 and 1988 have been chosen to represent the maximum recorded difference in inequality across the available years, post-1975.

14 A large value of $\alpha$ implies greater weight on parts of the distribution where the observed incomes are vastly different from the corresponding values in the reference distribution.
### Table 2  Inequality indices $J_{\alpha,k,p}$ and $J_{\alpha}$ computed with different reference distributions

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$J_{\alpha,k,p}$</th>
<th>$J_{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1979</td>
<td>1988</td>
</tr>
<tr>
<td>Equal reference distribution</td>
<td>[0.1119;0.1355]</td>
<td>[0.1685;0.2018]</td>
</tr>
<tr>
<td>Standard GE measures ($k = p = 1$)</td>
<td>[0.1016;0.1097]</td>
<td>[0.1460;0.1634]</td>
</tr>
<tr>
<td>$t_2.3$</td>
<td>0.1218</td>
<td>0.1836</td>
</tr>
<tr>
<td>1979</td>
<td>0.1056</td>
<td>0.1544</td>
</tr>
<tr>
<td>$t_2.4$</td>
<td>[0.1005;0.1086]</td>
<td>[0.1460;0.1634]</td>
</tr>
<tr>
<td>1988</td>
<td>0.1046</td>
<td>0.1543</td>
</tr>
<tr>
<td>$t_2.5$</td>
<td>[0.1023;0.1111]</td>
<td>[0.1508;0.1728]</td>
</tr>
<tr>
<td>$t_2.6$</td>
<td>[0.1132;0.1271]</td>
<td>[0.1843;0.2381]</td>
</tr>
<tr>
<td>Unequal reference distributions</td>
<td>[0.1005;0.1086]</td>
<td>[0.1460;0.1634]</td>
</tr>
<tr>
<td>Top 1 % gets 99 % of the income ($k = 0.01, p = 0.99$)</td>
<td>[0.1023;0.1111]</td>
<td>[0.1508;0.1728]</td>
</tr>
<tr>
<td>$t_2.7$</td>
<td>15.29</td>
<td>3.370</td>
</tr>
<tr>
<td>1979</td>
<td>3.370</td>
<td>2.795</td>
</tr>
<tr>
<td>$t_2.8$</td>
<td>[14.46;16.21]</td>
<td>[3.315;3.427]</td>
</tr>
<tr>
<td>1988</td>
<td>2.887;2.926</td>
<td>4.390;4.419</td>
</tr>
<tr>
<td>$t_2.9$</td>
<td>11.70</td>
<td>3.086</td>
</tr>
<tr>
<td>1979</td>
<td>2.795</td>
<td>4.341</td>
</tr>
<tr>
<td>$t_2.10$</td>
<td>[10.59;12.792]</td>
<td>[2.749;2.836]</td>
</tr>
<tr>
<td>1988</td>
<td>4.300;4.378</td>
<td>57.32;58.66</td>
</tr>
<tr>
<td>$t_2.11$</td>
<td>22.71</td>
<td>3.768</td>
</tr>
<tr>
<td>1979</td>
<td>3.768</td>
<td>44.34</td>
</tr>
<tr>
<td>$t_2.12$</td>
<td>[2.254;2.288]</td>
<td>[3.747;3.789]</td>
</tr>
<tr>
<td>1988</td>
<td>44.34</td>
<td>44.48</td>
</tr>
<tr>
<td>$t_2.13$</td>
<td>1.915</td>
<td>2.151</td>
</tr>
<tr>
<td>1979</td>
<td>3.631</td>
<td>44.14</td>
</tr>
<tr>
<td>$t_2.14$</td>
<td>[2.123;2.175]</td>
<td>[3.591;3.665]</td>
</tr>
<tr>
<td>1988</td>
<td>43.47;44.73</td>
<td></td>
</tr>
<tr>
<td>Continuous reference distributions</td>
<td>[0.0391;0.0421]</td>
<td>[0.0444;0.0536]</td>
</tr>
<tr>
<td>Beta(1,1) or Uniform(0,1)</td>
<td>[0.0465;0.0501]</td>
<td>[0.0538;0.0655]</td>
</tr>
<tr>
<td>$t_2.15$</td>
<td>0.0320</td>
<td>0.0339</td>
</tr>
<tr>
<td>1979</td>
<td>0.0406</td>
<td>0.0418</td>
</tr>
<tr>
<td>$t_2.16$</td>
<td>[0.0465;0.0501]</td>
<td>[0.0538;0.0655]</td>
</tr>
<tr>
<td>1988</td>
<td>0.0483</td>
<td>0.0541</td>
</tr>
<tr>
<td>$t_2.17$</td>
<td>[0.0589;0.0638]</td>
<td>[0.1014;0.1276]</td>
</tr>
<tr>
<td>$t_2.18$</td>
<td>0.0613</td>
<td>0.1457</td>
</tr>
<tr>
<td>1979</td>
<td>0.1014</td>
<td>0.1276</td>
</tr>
<tr>
<td>$t_2.19$</td>
<td>[0.0589;0.0638]</td>
<td>[0.1014;0.1276]</td>
</tr>
<tr>
<td>1988</td>
<td>0.1014</td>
<td>0.1276</td>
</tr>
</tbody>
</table>
Q5  

**Table 2** (continued)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta(2,2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t2.3 1979</td>
<td>0.0115</td>
<td>0.0132</td>
<td>0.0143</td>
<td>0.0158</td>
</tr>
<tr>
<td>t2.4 [0.0105;0.0127]</td>
<td>[0.0121;0.0146]</td>
<td>[0.0131;0.0158]</td>
<td>[0.0144;0.0175]</td>
<td>[0.0181;0.0222]</td>
</tr>
<tr>
<td>t2.5 1988</td>
<td>0.0210</td>
<td>0.0243</td>
<td>0.0267</td>
<td>0.0299</td>
</tr>
<tr>
<td>t2.6 [0.0180;0.0242]</td>
<td>[0.0204;0.0283]</td>
<td>[0.0221;0.0316]</td>
<td>[0.0242;0.0362]</td>
<td>[0.0300;0.0524]</td>
</tr>
<tr>
<td>Beta(2,5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t2.7 1979</td>
<td>0.0116</td>
<td>0.0138</td>
<td>0.0153</td>
<td>0.0173</td>
</tr>
<tr>
<td>t2.8 [0.0109;0.0124]</td>
<td>[0.0129;0.0147]</td>
<td>[0.0143;0.0163]</td>
<td>[0.0162;0.0185]</td>
<td>[0.0219;0.0256]</td>
</tr>
<tr>
<td>t2.9 1988</td>
<td>0.0121</td>
<td>0.0142</td>
<td>0.0157</td>
<td>0.0175</td>
</tr>
<tr>
<td>t2.10 [0.0100;0.0145]</td>
<td>[0.0116;0.0172]</td>
<td>[0.0127;0.0192]</td>
<td>[0.0141;0.0217]</td>
<td>[0.0177;0.0294]</td>
</tr>
</tbody>
</table>

t2.11 Data are from the Family Expenditures Surveys in UK. Bootstrap confidence intervals at 95 % are given in brackets.
Reference distributions and inequality measurement

sensitive inequality where, in terms of “distance from maximum inequality,” the change in the distribution is inconclusive.

Theoretical distribution  Finally let us consider how inequality changed using a continuous reference distribution $F_s$. The last panel of Table 2 tabulates the results for three $F_s$ from the Beta distribution family. Did UK income inequality, interpreted as a distance from a Beta-family reference distribution increase? We can see that the values of $J_\alpha$ are not statistically different between 1979 and 1988 when the Beta(1,1) (uniform) or Beta(2,5) (unimodal, right skewed) is used as the reference distribution distribution, while they are statistically different when the Beta(2, 2) (unimodal symmetric) is used as the reference distribution.

The estimates of the standard GE inequality measures $J_{\alpha,1,1}$ and of those of $J_{\alpha,k,p}$ and $J_\alpha$ in Table 2 provide us with different information about divergence of the empirical distribution from the chosen reference distribution. By varying the values of $k$ and $p$, one can specify the exact skewness of the reference distribution one would like to measure distance of the empirical distribution from. Likewise, by varying the values of $\alpha$ one can focus on different parts of the income distribution. A large value of $\alpha$ implies a greater weight on parts of the distribution where the observed incomes are vastly different from the corresponding values in the reference distribution. Finally, one can choose specific parametric distributions which correspond to the relevant reference distribution that the researcher is interested in.

5 Conclusion

The problem of comparing pairs of distributions is a widespread one in distributional analysis. It is often treated on an ad-hoc basis by invoking the concept of norm incomes and an arbitrary inequality index.

Our approach to the issue is a natural generalisation of the concept of inequality indices where the implicit reference distribution is the trivial perfect-equality distribution. Its intuitive appeal is supported by the type of axiomatisation that is common in modern approaches to inequality measurement and other welfare criteria. The axiomatisation yields indices that can be interpreted as measures of divergence and are related to the concept of divergence entropy in information theory [10]. Furthermore, they offer a degree of control to the researcher in that the $J_\alpha$ indices form a class of measures that can be calibrated to suit the nature of the economic problem under consideration. Members of the class have a distributional interpretation that is close to members of the well-known generalised-entropy class of inequality indices.

In effect the user of the $J_\alpha$-index is presented with two key questions:

1. the income discrepancies underlying inequality are with reference to what?
2. to what kind of discrepancies do you want the measure to be particularly sensitive?

As our empirical illustration has shown, different responses to these two key questions provide different interpretations from the same set of facts.
294 Appendix: Axiomatic foundation

295 For convenience we work with \( z := (z_1, z_2, \ldots, z_n) \), where each \( z_i \) is the ordered pair \((x_i, y_i)\), \( i = 1, \ldots, n \) and belongs to a set \( Z \), which we will take to be a connected subset of \( \mathbb{R}_+ \times \mathbb{R}_+ \). For any \( z \in Z^n \) denote by \( z(\zeta, i) \) the member of \( Z^n \) formed by replacing the \( i \)-th component of \( z \) by \( \zeta \in Z \). The divergence issue focuses on the discrepancies between the \( x \)-values and the \( y \)-values. To capture this introduce a discrepancy function \( d : Z \to \mathbb{R} \) such that \( d(z_i) \) is strictly increasing in \( |x_i - y_i| \). The solution is in two steps: (1) characterise a weak ordering \( \succeq \) on \( Z^n \) (\( z \succeq z' \) means “the income pairs in \( z \) are at least as close according to \( \succeq \) as the income pairs in \( z' \)”) (2) use the function representing \( \succeq \) to generate the index \( J \).

304 Axiom 1 (Continuity) \( \succeq \) is continuous on \( Z^n \).

305 Axiom 2 (Monotonicity) If \( z, z' \in Z^n \) differ only in their \( i \)-th component then \( d(x_i, y_i) < d(x'_i, y'_i) \iff z > z' \).

307 Axiom 3 (Symmetry) For any \( z, z' \in Z^n \) such that \( z' \) is obtained by permuting the components of \( z: z \sim z' \).

309 Axiom 3 implies that we may order the components of \( z \) such that \( x_1 \leq x_2 \leq \ldots \leq x_n \) and \( y_1 \leq y_2 \leq \ldots \leq y_n \).

311 Axiom 4 (Independence) For \( z, z' \in Z^n \) such that: \( z \sim z' \) and \( z_i = z'_i \) for some \( i \) then \( z(\zeta, i) \sim z'(\zeta, i) \) for all \( \zeta \in [z_{i-1}, z_{i+1}] \cap [z'_{i-1}, z'_{i+1}] \).

313 Axiom 5 (Zero local discrepancy) Let \( z, z' \in Z^n \) be such that, for some \( i \) and \( j, x_i = y_i, x_j = y_j, x'_i = x_i + \delta, y'_i = y_i + \delta, x'_j = x_j - \delta, y'_j = y_j - \delta \) and, for all \( k \neq i, j, x'_k = y'_k \). Then \( z \sim z' \).

316 Axiom 6 (Income scale irrelevance) For any \( z, z' \in Z^n \) such that \( z \sim z' \), \( tz \sim tz' \) for all \( t > 0 \).

318 Axiom 7 (Discrepancy scale irrelevance) Suppose there are \( z_0, z'_0 \in Z^n \) such that \( z_0 \sim z'_0 \). Then for all \( t > 0 \) and \( z, z' \) such that \( d(z) = td(z_0) \) and \( d(z') = td(z'_0) \): \( z \sim z' \).

320 Lemma 1 Given Axioms 1 to 5 \((a) \succeq \) is representable by \( \sum_{i=1}^n \phi_i(z_i) \), \( \forall z \in Z^n \) where, for each \( i, \phi_i : Z \to \mathbb{R} \) is a continuous function that is strictly decreasing in \( |x_i - y_i| \) and \((b) \phi_i(x, x) = a_i + b_i x \).

323 Proof Axioms 1 to 5 imply that \( \succeq \) can be represented by a continuous function \( \Phi : Z^n \to \mathbb{R} \) that is increasing in \( |x_i - y_i|, i = 1, \ldots, n \). Part \((a)\) of the result follows from Axiom 4 and Theorem 5.3 of [23]. Now take \( z' \) and \( z \) as specified in Axiom 5: \( z \sim z' \) if and only if \( \phi_i(x_i + \delta, x_i + \delta) - \phi_i(x_i, x_i) - \phi_j(x_j + \delta, x_j + \delta) + \phi_j(x_j, x_j) = 0 \) which implies

\[
\phi_i(x_i + \delta, x_i + \delta) - \phi_i(x_i, x_i) = f(\delta)
\]

for any \( x_i \) and \( \delta \). The solution to this Pexider equation implies \((b)\).
Lemma 2 Given Axioms 1 to 6 $\succeq$ is representable by $\phi\left(\sum_{i=1}^{n} x_i h_i\left(\frac{x_i}{y_i}\right)\right)$ where $h_i$ is a real-valued function.

Proof Using the function $\Phi$ introduced in the proof of Lemma 1 Axiom 6 implies $\Phi(z) = \Phi(z')$ and $\Phi(tz) = \Phi(tz')$; since this has to be true for arbitrary $z, z'$ we have

$$\Phi(z) = \Phi(z') = \psi(t)$$

where $\psi$ is a continuous function $\mathbb{R} \to \mathbb{R}$. Hence, using the $\phi_i$ given in Lemma 1 we have for all $i : \phi_i(tz_i) = \psi(t)\phi_i(z_i)$ or, equivalently $\phi_i(tx_i, ty_i) = \psi(t)\phi_i(x_i, y_i)$. In view of Aczél and Dhombres [2], page 346 there must exist $c \in \mathbb{R}$ and a function $h_i : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\phi_i(x_i, y_i) = x_i^c h_i\left(\frac{x_i}{y_i}\right). \quad (14)$$

From Lemma 1 and Eq. 14: $\phi_i(x_i, x_i) = x_i^c h_i(1) = a_i + b_i x_i$, which, if $\phi_i(x, x)$ is non-constant in $x$, implies $c = 1$. Noting that Lemma 1 implies that $\succeq$ is also representable by $\phi\left(\sum_{i=1}^{n} \phi_i(z_i)\right)$ (where $\phi : \mathbb{R} \to \mathbb{R}$ is continuous and monotonic increasing, and taking Eq. 14 with $c = 1$ gives the result.

Theorem 1 Given Axioms 1 to 7 $\succeq$ is representable by $\phi\left(\sum_{i=1}^{n} x_i^\alpha y_i^{1-\alpha}\right)$ where $\alpha \neq 1$ is a constant.

Proof (Cf [21]) Take the special case where, in distribution $z'_0$ the income discrepancy takes the same value $r$ for all $n$ income pairs. If $(x_i, y_i)$ represents a typical component in $z_0$ then $z_0 \sim z'_0$ implies

$$r = \psi\left(\sum_{i=1}^{n} x_i h_i\left(\frac{x_i}{y_i}\right)\right) \quad (15)$$

where $\psi$ is the solution in $r$ to

$$\sum_{i=1}^{n} x_i h_i\left(\frac{x_i}{y_i}\right) = \sum_{i=1}^{n} x_i h_i(r) \quad (16)$$

In Eq. 16 we take the $x_i$ as fixed weights. Using Axiom 7 in Eq. 15 requires

$$tr = \psi\left(\sum_{i=1}^{n} x_i h_i\left(t\frac{x_i}{y_i}\right)\right), \text{ for all } t > 0. \quad (17)$$

Using Eq. 16 we have

$$\sum_{i=1}^{n} x_i h_i\left(\frac{t}{\psi\left(\sum_{i=1}^{n} x_i h_i\left(\frac{x_i}{y_i}\right)\right)}\right) = \sum_{i=1}^{n} x_i h_i\left(t\frac{x_i}{y_i}\right) \quad (18)$$

Introduce the following change of variables

$$u_i := x_i h_i\left(\frac{x_i}{y_i}\right), i = 1, ..., n \quad (19)$$
and write the inverse of this relationship as
\[ \frac{x_i}{y_i} = \psi_i(u_i), \quad i = 1, \ldots, n \] (20)

Substituting Eqs. 19 and 20 into Eq. 18 we get
\[ \sum_{i=1}^{n} x_i h_i \left( t \psi \left( \sum_{i=1}^{n} u_i \right) \right) = \sum_{i=1}^{n} x_i h_i \left( t \psi_i(u_i) \right). \] (21)

Also define the following functions
\[ \theta_0(u, t) := \sum_{i=1}^{n} x_i h_i (t \psi(u)), \quad \theta_i(u, t) := x_i h_i (t \psi_i(u)), \quad i = 1, \ldots, n. \]

Substituting these into Eq. 21:
\[ \theta_0 \left( \sum_{i=1}^{n} u_i, t \right) = \sum_{i=1}^{n} \theta_i(u_i, t) \]

which has as a solution
\[ \theta_i(u, t) = b_i(t) + B(t) u_i, i = 0, 1, \ldots, n \]

where \( b_0(t) = \sum_{i=1}^{n} b_i(t) \) ([1], p. 142). Therefore we have
\[ h_i \left( t \frac{x_i}{y_i} \right) = \frac{b_i(t)}{x_i} + B(t) \frac{x_i}{y_i}, \quad i = 1, \ldots, n \] (22)

From Eichhorn [22], Theorem 2.7.3 the solution to Eq. 22 is of the form
\[ h_i(v) = \begin{cases} \beta_i v^{\alpha - 1} + \gamma, & \alpha \neq 1 \\ \beta_i \log v + \gamma, & \alpha = 1 \end{cases} \] (23)

where \( \beta_i > 0 \) is arbitrary. Lemma 2 and Eq. 23 give the result. \( \Box \)

References

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q1. Please check author’s affiliation if captured correctly.
Q2. Please check provided running title if appropriate.
Q3. For figures 1 and 2 we proceed using the separate pdf. Please check if appropriate.
Q4. Please check capturing of paragraphs if correct.
Q5. Please check presentation of Tables 1 and 2 if captured correctly.
Q6. Please check reference items [5], [8], [10] if captured correctly.