J. Econ. (2002) Suppl. 9: 51-89

Journal of Economics
Zeitschrift für Nationalökonomie

Shapley Inequality Decomposition by Factor Components: Some Methodological Issues

Mercedes Sastre and Alain Trannoy

Decomposition analysis provides help in the difficult task of explaining how economics trends and government policies affect the distribution of income. In this context, recent studies have proposed the application of the "Shapley Value" allocation method, a concept from cooperative game theory, to the decomposition of inequality. This paper examines the "Shapley inequality decomposition" by factor components focussing in particular on some methodological issues that cannot be solved from a theoretical point of view. Following an inductive approach, the empirical evidence obtained by applying several variants of the Shapley decomposition to the UK and US income distributions suggests an answer to some of the dilemmas faced by applied economists when implementing the Shapley decomposition technique.

Keywords: Shapley Value, Inequality, Decomposition.

JEL classification: D31.

1 Introduction

Assessing how different income sources affect income inequality certainly appears to be an important issue for interpreting economic trends and designing economic policy. Markets of factors and thus market incomes such as earnings or capital income are differently altered by globalization and the technologies. The impact on economic inequality of the rising share of women in the paid labor force, or of the increasing returns to human capital, are common questions asked of economists. Many economic policies focus on a particular source of income through different tax and benefit instruments, and the relative impact of these targeted measures on economic inequality often remains unknown. Providing a method of estimating how income sources affect income inequality is therefore a central question in the field of inequality measurement.

However, agreement does not exist on the best way to measure how an income source contributes to inequality. The traditional approach to the problem takes two forms. Some researchers have adopted global methods specific to each index. By a global method, we mean that contributions must be computed for all income components defined at the outset. The sum of the components should also add up to the inequality to be "explained". At least in one case, the variance, the global method is based on firm principles. Shorrocks (1982) has proved that the method of decomposition that assigns to each source its covariance with total income[1] is the only one which satisfies six attractive properties. Unfortunately, a by-product of this decomposition is less appealing. The same six axioms imply that the relative contribution of a source—the contribution divided by the inequality of the income distribution—must be independent of the inequality index chosen. In other words the decomposition according to the covariance should apply to all inequality measures. In applied work this solution has been neglected in favor of rather ad-hoc methods. For example, the "natural" Gini decomposition by source is defined by the sum of the product of three terms for each income component: the income share, the component's own Gini, and a ratio of correlations that captures the relationship between the income component and the rank in total income. Although some interesting interpretations have been proposed to justify their use [see for example Lerman and Yitzhaki (1985)], it is difficult to believe that these solutions offer a definite answer to the problem of decomposing inequality by sources.

In contrast, many authors prefer a local approach, aiming only to provide the contribution of a specific income component. A popular method consists in a before-after calculation[2] that assigns to each source the difference between overall income inequality and the inequality obtained if we eliminate the income component whose effect is being assessed (or at least inequality from that source).

In view of its intuitive meaning, we defend the before-after approach not only as a local method but also as a global method. A priori, this goal is relatively challenging since extending this method from a local scale to a global one raises two difficulties: the lack of exactness, i.e. the before-after contributions do not add to the total

[1] The "natural decomposition" of its variance.

[2] See Cancian and Reed (1998) among others for a defense and an application.

amount that needs to be "explained". Another limitation is that—except for cases with only two income sources considered—the contribution of a source depends on the elimination sequence and there is not a unique or natural candidate. A solution to these drawbacks can be provided if the before-after approach is embedded in a more general framework provided by cooperative game theory.

More specifically, by considering the formal similarities between the problem of assigning source contributions to inequality, and the classic question of cost allocation among a set of agents investigated by cooperative game theory [see, for instance, Mirman, Taubman and Zang (1985)], Chantreuil and Trannoy (1999) and Shorrocks (1999) have proposed applications of the Shapley Value allocation method [Shapley (1953)] to the decomposition of inequality by factor components. The Shapley Value is an allocation method that assigns the gains of a coalition of players among its members as a function of what they contribute to the coalition. As the contribution of a player depends on the order in which the player joins the coalition, the Shapley rule weights each possible coalition by its probability and assigns to every player the average of all marginal contributions that this individual can make to all coalitions.

In the inequality decomposition context, this technique implies considering the impact on overall inequality of eliminating each income source. But, since there is no natural order of elimination, we average these impacts over all possible sequences of elimination. Thus, in order to assess the effect of a given income source on overall inequality we apply the before-after concept to the set of all possible combinations of income sources, and take the average of all contributions.

Despite of the internal consistency and attractive interpretation in terms of marginal contribution, empirical applications of the Shapley rule to inequality decomposition raise some dilemmas that cannot be solved solely on theoretical grounds. The purpose of this paper is to focus on these methodological issues. An inductive approach proves to be very useful to reveal the alternatives faced by applied economists. The empirical evidence obtained from several Shapley decomposition variants applied to UK and US income distributions drawn from the LIS database has helped us to propose a way of solving three dilemmas.

1) According to the Shapley decomposition, the contribution of any given source of income to overall inequality can be interpreted as the expected marginal impact of the factor when such an expectation is made over all possible sequences of elimination. This marginalist

interpretation of the Shapley rule can be considered in at least two possible ways: firstly, calculating the difference between overall inequality and the inequality obtained if we eliminate the income component whose effect is being assessed (zero income decomposition). Alternatively inequality from that source of income can be eliminated (equalized income decomposition). Although the second calculation seems to be more in keeping with inequality measurement, the a priori reasons that could favor either one or the other approach are not clear. Empirical analyses can thus help us to clarify the situation and to choose between the two methods.

2) An important shortcoming of the Shapley decomposition rule is that it does not satisfy the principle of independence of the aggregation level. That is to say, the contribution to inequality of a given income source depends on how the rest of the income components are treated. Two different approaches derived from cooperative game theory have been proposed to improve Shapley inequality decompositions in respect of the independence requirement. The first is the Nested Shapley (Chantreuil-Trannoy 1999) and the other is the Owen decomposition (Chantreuil-Trannoy 1997 and Shorrocks 1999). Both methods require a hierarchical structure in which overall income is made up of a set of primary factors. Each of these is then divided into a group of secondary factors. This approach can of course be extended to a sequence of source subgroup partitions. For the exercise to be meaningful it is obviously better if such partitions are given exogenously and if they are relevant from an economic viewpoint. The discussion of the relative merits of the Nested-Shapley versus Owen solution is the second aim of this paper.

3) Whichever of the above solution is chosen, a hierarchy of income sources needs to be defined. The economic relevance of such a tree is, of course, crucial. This paper tries to present the possible options.

This paper is organized as follows. To make the paper self-contained, we start by recalling the definition of the Shapley value in the context of game theory. Section 3 points out the superiority of the equalized income decomposition in respect to the independence requirement. We then discuss the relative merits of the Nested Shapely and Owen decompositions in Section 4. Section 5 analyses the problem of choosing a hierarchical income structure. It also offers a discussion on the merits of several trees that may appear to be obvious candidates. The last section summarizes the results obtained and suggests some rules of thumb. The Appendix contains a detailed description of the data and some additional results for British and American income distributions.

2 The Shapley Value

We start by recalling the meaning of the Shapley value in a cooperative TU-game[3] framework. We are given a set of agents, $N = \{1, , i, , n\}$, and $S \subset N$ denotes a potential subset of agents. The Shapley value attempts to describe a fair way to divide the gains from cooperation, taking for granted the strategic realities. These are captured by a function v called the characteristic function that associates to every coalition S—belonging to the power set 2^S—the payoff $v(S) \in \mathbf{R}$ that the coalition S can achieve by itself. The number $v(S)$ is known as the worth of the coalition S. A game in characteristic form is defined by (N, v).

Several interpretations lead to the Shapley value. We only give the marginalist one which has a direct counterpart in inequality decomposition analyses. For any coalition S and for any agent $i \in S$ the marginal contribution of this agent to the coalition $S \backslash \{i\}$ is defined by $v(S) - v(S \backslash \{i\})$. Coalitions are supposed to occur at random. Suppose now that all agents are arranged in some order, all orders being equally likely. Hence $S \backslash \{i\}$ is the set of agents which precede agent i and $v(S) - v(S \backslash \{i\})$ represents i's marginal contribution to the set of players who precede him. The agent i's payoff according to the Shapley value is simply computed as i's expected marginal contribution to the coalition $S \backslash \{i\}$ when all orderings of agents are held to be equally likely. Defining the Shapley value requires computing the probability that in a random ordering a given coalition $S \subset N, i \in S$, arises as the union of i and its predecessors. This is the product of two probabilities. The first one is the probability that i is in the Sth place which is simply $1/n$. The second one is the probability that $S \backslash \{i\}$ arises when we randomly select[4] $s - 1$ members from the population $N \backslash \{i\}$. This probability is $(n - s)!(s - 1)!/(n - 1)!$.

The Shapley value denoted $Sh(N, v)$ where $Sh(N, v) \in \mathbf{R}^n$ allocates a Shapley value payoff denoted $Sh_i(N, v)$ to the ith player which formula is given by:

[3] Transferable utility (TU) game, there exists a numeraire than can be used to perform transfers of utility across agents.

[4] The number of individuals in a coalition S is denoted s.

$$Sh_i(N,v) = \sum_{S\subset N, i\in S} \frac{(s-1)!(n-1)!}{n!} \left[v(S) - v(S\backslash \{i\})\right] \qquad (2.1)$$

which can also be written:

$$Sh_i(N,v) = \sum_{S\subset N, i\in S} \frac{(s)!(n-s-1)!}{n!} \left[v(S\cup \{i\}) - v(S)\right] \qquad (2.2)$$

The sum of the Shapley value payoffs over N is equal to $v(N)$ which represents the total amount to be distributed among the n agents.

3 Equalized Versus Zero Shapley Decomposition

An income distribution among a set of individuals $N = \{1, ..., i, ..., n\}$, according to a set of income sources $K = \{1, ..., j, ..., k\}$ can be described by a matrix $X = [x_i^j]$. In what follows we will not restrain x_i^j to be positive. Let us assume that a specific inequality index $I : \mathbf{R}^n \to \mathbf{R}$ has been selected.[5] The Shapley decomposition gives the contribution to inequality of any subset of sources. A subset of sources would be denoted S, the set of admissible subsets being the power set 2^k.

Regarding the definition of the Shapley decomposition, at least two calculations seem sensible. They differ in the treatment of components not included in the considered subset. In the first one, defined as the *zero income inequality decomposition*, the components not included in S are removed. In the second calculation, which leads to what is called *equalized income inequality decomposition* by Chantreuil-Trannoy (1999), the inequality from all components not in S is removed.

For the first calculation we can build a distribution of income among subsets of sources $y : 2^K \to \mathbf{R}^n$ such that for all S$\in 2^K, S \neq \emptyset$,

$$y(S) = [\sum_{j\in S}(x_1^j, ..., x_n^j)]_{1XN} \qquad (3.1)$$

and $y(\emptyset) = [0]_{1xN}$ by convention.

[5] We should notice that the Shapley decomposition calculations require agreement not only with the ordinal meaning of an inequality index, but also with its cardinal one.

In the second calculation the distribution of income among subsets of sources is obtained by equalizing complementary sources, i.e. we define: $y^e : 2^K \to \mathbf{R}^n$ such that $y^e(\emptyset) = [0]_{1xN}$, and for all $S \in 2^K, S \neq \emptyset$,

$$y^e(S) = [\sum_{j\in S} x_1^j + \sum_{j\notin S} \mu(x^j), ..., \sum_{j\in S} x_n^j + \sum_{j\notin S} \mu(x^j)]_{1XN}, \tag{3.2}$$

where $\mu(x^j)$ denotes the mean income from source j.

The contribution of source j according to the zero income Shapley decomposition is given by:

$$Sh_j(K, \mathbf{X}, I) = \sum_{\substack{S\subset K\\ j\in S}} \frac{(s-1)!(k-s)!}{k!}[I(y(S)) - I(y(S-\{j\}))] . \tag{3.3}$$

According to this formula, the contribution of any given factor to overall inequality can be interpreted as the expected marginal impact of the factor when the expectation is taken over all the possible elimination sequences. This decomposition rule satisfies some desirable properties. It leads to a perfect[6] symmetric decomposition (in the sense that the contribution assigned to any factor does not depend on the way in which factors are labeled), and it is sensitive to the choice of inequality index. The same comment remains valid for the contribution of source j according to the equalized income Shapley decomposition defined by:

$$Sh_j^e(K, \mathbf{X}, I) = \sum_{\substack{S\subset K\\ j\in S}} \frac{(s-1)!(k-s)!}{k!}[I(y^e(S)) - I(y^e(S-\{j\}))] . \tag{3.4}$$

The distinction between both calculations seems transparent, and maybe the latter one seems more in the spirit of inequality measurement. From a theoretical point of view, the treatment of an equally distributed component of income represents a noticeable difference. Let us consider relative[7] inequality indices; eliminating an equally distributed source from a subset of sources will increase inequality of that subset. Hence the *zero income* Shapley contribution of an equally

[6] The sum of the contributions adds up to the amount of inequality.

[7] An inequality index is relative if a proportional change in all incomes does not alter the level of inequality.

distributed source is negative, suggesting an equalizing effect of such a source. In contrast, the *equalized Shapley* decomposition assigns a zero impact to an equally distributed source; indeed the inequality of any subset of sources including an equally distributed component remains unchanged if we remove this subset of sources in the equalized procedure. The equalized procedure is relevant here since a zero impact of an equally distributed source has been advocated by Shorrocks (1982) as a reasonable property of a decomposition rule. On the other hand, Chantreuil and Trannoy (1999) defend the view that the treatment of an equally distributed source must be different according to the relative or absolute property of the inequality index. Since for absolute inequality indices, the contribution of an equally distributed source is always zero, whatever the approach chosen, the argument here is more in favor of the use of the *zero income* procedure. Thus the theoretical approach is rather inconclusive and we are looking for empirical results allowing us to make a choice on a firmer footing.

By way of example, the relative contributions (the contribution divided by the overall inequality) for the two calculations present marked differences. The empirical exercise has been done using data from the 1995 United Kingdom Family Expenditure Survey and from 1994 United States March Current Population Survey.[8] For the purpose of the presentation and discussion of the results[9], the Gini index has been selected.[10]

For the UK, the data available allow us to disaggregated household income into three main components: (i) Earnings (E): i.e. wages and self-employment income, which represent about 72% of overall household income. (ii) Capital income(C), representing 11.4% of total income. (iii) Transfers (T): including all kind of public and private transfers received by households. The income share of this source is 16.8%. This is only one of all possible partitions, but it is very popular in empirical works. According to equation (3.3), in order to calculate the zero Shapley contribution of the income sources we first

[8] Available through the Luxembourg Income Study. For more information about this database that allows comparisons of income distributions among different countries see http://lissy.ceps.lu.

[9] The unweighted distribution of gross income is non adjusted by family size, as it is usual in the empirical studies that decompose inequality by income sources (see Jenkins 1995 for instance). The unit of analysis is the household. See the appendix for a detailed description of the data. The inequality of a source takes into account the fact that household income is zero for many sources.

[10] Results for other inequality indices (available upon request) exhibit the same pattern. Therefore, the validity of the conclusions does not depend on the choice of the inequality index.

need to compute the inequality for each of the following eight subsets of components S : (E), (C), (T), $(E \cup C)$, $(E \cup T), (C \cup T), (\emptyset)$ and $(E{\cup}C{\cup}T)$.[11] The following relative contributions have been obtained:

$$Sh_E(K,\mathbf{X},I) = 24.9\%$$
$$Sh_C(K,\mathbf{X},I) = 60.9\%$$
$$Sh_T(K,\mathbf{X},I) = 14.2\%$$

to be compared with the following relative contributions in the equalized version:

$$Sh^e_E(K,\mathbf{X},I) = 83.6\%$$
$$Sh^e_C(K,\mathbf{X},I) = 11.8\%$$
$$Sh^e_T(K,\mathbf{X},I) = 4.6\%$$

where the relevant distributions are: $[E + \mu(C \cup T)]$, $[C + \mu(E \cup T)]$, $[T + \mu(E \cup C)]$, $[(E{\cup}T){+}\mu(C)]$, $[(C \cup T) + \mu(E)]$, $[(E \cup C) + \mu(T)]$, and $[(E \cup C \cup T)]$.

These results allow us to establish a relationship between relative contribution of a source and its income share, and to use the ratio of the former to the latter as a more easily interpretable measure of the impact of a source to overall inequality. Thus, in the above example, for the equalized version, the equalized Shapley contribution of earnings is substantially higher than its income share, capital income contributes to inequality in a slightly higher percentage than its weight in total income, and transfers contribute about one forth of its income share. The results are more amazing in the zero approach, since they tell us that earnings contribute slightly more than one third to its income share, the great responsibility of inequality bearing on capital incomes. The story is even more dramatic in the US example since the contribution of earnings to overall inequality is negative! (see second row of Table 2b). Considering the huge difference between both calculations, it seems that in empirical studies we have to select only one of the two.[12]

To decide between the two calculations, the way they perform with respect to the major shortcoming of the Shapley value will be of particular importance. Indeed the Shapley decomposition presents an

---

[11] $([0])_{1xN}) = 0$

[12] The eqialized Shapley contributions are closer to those obtained with standard decomposition methods. Among other studies, see for instance Lerman and Yitzhaki (1985), for a decomposition of the Gini index by sources of income applied to the US and Jenkins (1995) that applies the natural decomposition of the variance to UK income data. Both studies found that except for income transfers, which reduced inequality, the contribution of most sources was similar to their share of total income. See also Tables A.1a and A.1b in the Appendix.

Gross Income				
Market Income			Transfers	
Sh=94.5			Sh=5.5	
Earnings		Capital	Transfers	
Sh=83.6		Sh=11.8	Sh=4.6	
Earnings		Capital	Replacement Inc.	Non-contributory Benefits
Sh=83.1		Sh=11.4	Sh=3.2	Sh=2.3
Wages	Self-employment	Capital	Transfers	
Sh=65.9	Sh=18.5	Sh=11.5	Sh=4.1	
Wages	Self-employment	Capital	Replacement Inc.	Non-contributory Benefits
Sh=66.5	Sh=17.8	Sh=10.4	Sh=3.2	Sh=2.1

Table 1a: Relative Contributions for the Equalized Shapley Decomposition of the Gini Index. United Kingdom 1995 (Household Non-adjusted Distribution).

important failure: the contribution assigned to any income source is not independent from the level of disaggregation, i.e. it is sensitive to the way in which other sources are clustered.[13] In the following we will explore the intensity of that dependence regarding the alternative selected.

The disaggregation of overall household income presented above is only one of the multiple partitions in which income sources could have been gathered together. Let us go back to that example and focus on the contribution of income from transfers in the UK. Assume that earnings and capital income are combined in a unique source called "market income". In this case, the Shapley contribution of transfers to total inequality jumps from 14.2% to 47.2% in the zero approach while it increases from 4.6% to 5.5% in the equalized version. The following Tables display the kind of problems raised by the lack of independence with a greater detail.

The results exhibited for finer partitions of household income reflect the extremely high volatility of the zero inequality decomposition contributions. An extreme example of this variability is the contribution of transfers in the UK. In the case of splitting income into only two sources, its contribution is about one half of overall inequality. How-

[13]For a detailed discussion of the properties of the Shapley decomposition see Chantreuil and Trannoy (1999) or Shorrocks (1999).

Gross Income				
Market Income			Transfers	
Sh=90.2			Sh=9.8	
Earnings		Capital	Transfers	
Sh=81.2		Sh=10.0	Sh=8.8	
Earnings		Capital	Replacement Inc.	Non-contributory Benefits
Sh=80.8		Sh=9.8	Sh=6.5	Sh=2.9
Wages	Self-employment	Capital	Transfers	
Sh=74.7	Sh=7.1	Sh=9.8	Sh=8.4	
Wages	Self-employment	Capital	Replacement Inc.	Non contributory Benefits
Sh=74.8	Sh=6.7	Sh=9.3	Sh=6.4	Sh=2.8

Table 1b: Relative Contributions for the Equalized Shapley Decomposition of the Gini index. United States 1994 (Household Non-adjusted Distribution).

Gross Income				
Market Income			Transfers	
Sh=52.8			Sh=47.2	
Earnings		Capital	Transfers	
Sh=24.9		Sh=60.9	Sh=14.2	
Earnings		Capital	Replacement Inc.	Non-contributory Benefits
Sh=8.3		Sh=42.3	Sh=21.8	Sh=27.5
Wages	Self-employment	Capital	Transfers	
Sh=4.8	Sh=56.9	Sh=40.8	Sh=2.5	
Wages	Self-employment	Capital	Replacement Inc.	Non contributory Benefits
Sh=5.1	Sh=46.6	Sh=29.4	Sh=11.4	Sh=17.6

Table 2a: Relative Contributions for the Zero Shapley Decomposition of the Gini index. United Kingdom 1995 (Household Non-adjusted Distribution).

<table>
<tr><td colspan="5">Gross Income</td></tr>
<tr><td colspan="3">Market Income</td><td colspan="2">Transfers</td></tr>
<tr><td colspan="3">Sh=30.9</td><td colspan="2">Sh=69.1</td></tr>
<tr><td colspan="2">Earnings</td><td>Capital</td><td colspan="2">Transfers</td></tr>
<tr><td colspan="2">Sh=0.08</td><td>Sh=67.9</td><td colspan="2">Sh=32.2</td></tr>
<tr><td colspan="2">Earnings</td><td>Capital</td><td>Replacement Inc.</td><td>Non-contributory Benefits</td></tr>
<tr><td colspan="2">Sh=15.4</td><td>Sh=49.7</td><td>Sh=34.3</td><td>Sh=31.4</td></tr>
<tr><td>Wages</td><td>Self-employment</td><td>Capital</td><td colspan="2">Transfers</td></tr>
<tr><td>Sh=16.9</td><td>Sh=56.6</td><td>Sh=48.0</td><td colspan="2">Sh=12.3</td></tr>
<tr><td>Wages</td><td>Self-employment</td><td>Capital</td><td>Replacement Inc.</td><td>Non contributory Benefits</td></tr>
<tr><td>Sh=4.8</td><td>Sh=4.8</td><td>Sh=4.8</td><td>Sh=4.8</td><td>Sh=4.8</td></tr>
</table>

Table 2b: Relative Contributions for the Zero Shapley Decomposition of the Gini index. United States 1994 (Household Non-adjusted Distribution).

ever, a further disaggregation of income components makes the contribution of transfers to inequality negative (see forth row of Table 2a). Thus we can conclude that zero contributions are much more dependent on the level of disaggregation than the equalized ones. This empirical fact which is invariant to the inequality index or the country chosen lead us to favor the equalized approach in the following.[14]

Let us comment a little bit more on the results for the equalized case and let us focus on the contribution of capital income for instance. The Shapley contribution of this source is sensitive to the way other factors are handled. Thus, the contribution assigned to capital income is not the same if earnings are considered as a single entity or decomposed in two components: wages and self-employment income. The contribution of capital to overall inequality is also affected by the way transfers are clustered. Indeed, in the UK, the impact of capital income varies from 11.8% when considering three income sources, to 10.4% when disaggregating overall income into five components. Other sources are altered in the same way. For instance the contribution of transfers to overall inequality varies from 4.1% to 5.5% depending on the disaggregation of market income.

[14] Results for other partitions of household income, different inequality indices and data sets reveal a similar pattern of very high dependence of zero contributions on the level of aggregation. These results are available from the authors on request.

As a consequence of its dependence on the level of disaggregation, the Shapley decomposition does not guarantee that the contributions assigned to the components of a given source of income sum up the contribution to inequality of that income source treated as a single unit. For instance it would be interesting to study the impact on inequality of the public system of pensions and unemployment benefits (what we have called "replacement income"[15]), independently from the rest of public transfers programs.[16] By doing so with the Shapley inequality decomposition, two points should be stressed. Firstly, as mentioned previously, the contribution of that source will depend on how market income is treated. Secondly, as can be seen in the last two lines of Tables 2a and 2b, the sum of the contributions of replacement income and the rest of transfers do not equal the contribution of total transfers when it is considered as a single entity in the decomposition that treats market income in the same way.

Although the consequences of the dependence on the aggregation level are relatively moderate—the differences between the *ex ante* contributions and the *ex post* contributions are not higher than 1.4 percentage points—a general trend can be detected regarding the impact of disaggregation on the vector of contributions. On average, the contribution of a given component decreases with the successive disaggregations of complementary sources. However, this cannot be considered as a general property of the Shapley rule: for example the relative contribution of wages to overall inequality is slightly higher when transfers are disaggregated into two components than when they are considered as a single source. Another general empirical statement emerges: the coarser the partition is, the more volatile are the results relative to a finer partition.

An important remark about the interpretation of the Shapley contributions opens the way to a refinement of this decomposition method. As was mentioned before, the contribution of any given factor to overall inequality can be interpreted as the expected marginal impact of the factor when the expectation is taken over all the possible elimination sequences. Thus, it is important that the elimination sequences or, equivalently, the subsets of components considered in the calculation, have an economic appeal. But this is not always the case. Of course, the interpretation of the subset earnings-capital income as

[15]Replacement income is the sum of unemployment, retirement and other public subsidies related to past economic activity.

[16]Including means tested and other public transfers not linked to economic activity, as well as private transfers received from other households or institutions.

market income is straightforward. Nevertheless, the interest in considering the subsets earnings-transfers, and capital-transfers is not that clear, which raises doubts about the interpretation of Shapley contributions. As long as the level of disaggregation increases, the risk of having "unnatural" subsets of sources becomes higher, and interpretation of the results may be puzzling. Introducing an explicit structure on the set of sources can help to mitigate this difficulty as well as to reduce the dependence of source contributions to the way in which other sources are grouped.

4 Nested Shapley Versus Owen

The structure we introduce takes the general form of a nested tree or a hierarchy and the simplest example would be a partition of the set of sources. A partition of the set of income sources K is the set $\mathcal{P}_K = \{S_1, ..., S_l, ..., S_m\}$ such that for all S_h, $S_l \in \mathcal{P}_K$, $S_h \bigcap S_l = \emptyset$ and:

$$\bigcup_{h=1}^{m} S_h = K \tag{4.1}$$

with $1 < m < k$. For the sake of illustration let us consider the following example: suppose that the data available allow us to disaggregate overall household income (X) into four elementary components, and that this partition is given exogenously and is relevant from an economic view:[17] (i) Earnings (E), (ii) Capital income (C), (iii) Replacement income (R) and (iv) Non-contributory benefits (B). These sources may be divided naturally into two main aggregated factors: *Market income* (M), formed by earnings and capital income, M=E∪C, and *Transfers* (T) T= R∪B, which includes all kinds of transfers.

Consider the contribution of an elementary factor, for instance, capital income. With the Shapley decomposition, we have shown in the previous section that its contribution depends on the number of subgroups considered, not only in the disaggregation of market income, but also in the disaggregation of transfers. Therefore, the Shapley value that ignores the additional information about the structure of household income is not an appropriate tool for the decomposition analysis. A desirable requirement would be that the contribu-

[17]We can object that the sources considered are not exogenous from each other (for example transfer income certainly depends on earnings). But any finer partition will face a problem of endogeneity. In quoting Gottshalk and Smeeding (1997, p. 668) "a major drawback of this source decomposition exercice is that they can be easily misinterpreted because they do no make a distinction between endogenous and exogenous factors".

tion of capital income is at least independent from the disaggregation of transfers. In this section, we perform an empirical illustration of two different methods, derived from the Shapley value, satisfying the milder requirement of independence.

4.1 The Nested-Shapley Inequality Decomposition

The components of a partitioned decomposition problem is now the vector $(K, P_K, \mathbf{X}, I)$. The Nested-Shapley procedure [see Chantreuil and Trannoy (1999)] uses the nested structure of a partition and the calculation procedure can be simply described as a two-stage procedure. In a first stage, called the "between stage", the contribution of any subgroup of sources is calculated using the Shapley formula given by (3.4). In other words, each subset of sources $S_l \in P_K$ is considered as an elementary source. If we denote the Nested Shapley contribution of source j in an equalized procedure NSh_j^e, we can write the following identity:

$$NSh_{S_l}^e(K, \mathcal{P}_K, \mathbf{X}, I) \equiv Sh_{S_l}^e(\mathcal{P}_K, \mathbf{X}, \mathbf{I}), \forall \mathbf{S_l} \in \mathcal{P}_\mathbf{K} \qquad (4.2)$$

In a second stage, called the "within stage", the contribution of any elementary source is computed along the Shapley procedure providing that the contributions of the elementary sources belonging to some subset add up to the Nested-Shapley contribution of the subset. Let j be an elementary source belonging to some aggregated factor S_l and let s_l denote the dimensionality of S_l. The general formula for the Nested Shapley contribution of source j is given by:

$$\begin{aligned} &NSh_j^e(K, \mathcal{P}_K, \mathbf{X}, I) = \qquad (4.3)\\ &\sum_{\substack{S \subseteq S_l \\ j \in S}} \frac{(s-1)!(s_l - s)!}{s_l!}[I(y^e(S)) - I(y^e(S - \{j\}))] \\ &+ \frac{1}{s_l}[NSh_{S_l}^e(K, \mathcal{P}_K, \mathbf{X}, I) - I(y^e(S_l))] \end{aligned}$$

The first term on the RHS in expression (4.3) is much the same as the simple equalized Shapley contribution of source j in the sub-decomposition problem of the aggregated factor S_l. The slight difference between both computations comes from the fact that while inequality of aggregated factors different from S_l have been removed in the Nested Shapley approach, aggregated factors are ignored in the simple Shapley contribution of source j in the decomposition of S_l.

Using (4.2) the second term on the RHS in expression (4.3) computed the difference between the Shapley contribution of the aggregated factor S_l and the inequality of this source. An equal split of this difference between all elements of the aggregated factor ensures the exactness of this decomposition method.

The "within stage" only considers subsets of elementary sources belonging to the same aggregated factor. As long as the partition used is relevant from an economic view, the subsets in question will also have an economic interpretation. Here, to keep notation to a minimum, two stage trees have been considered. Obviously this chain calculation can be extented to as many stages as desired.

Under the Nested-Shapley extension the equalized procedure can be viewed in two different ways. The usual one, is the one followed in the above formula, computes the inequality of a subset of sources assuming that inequality is removed in all income components not included in the subset. However, it is conceivable to imagine another calculation that removes inequality only for income components in the same level of the income structure. This second procedure will be termed the *semi-equalized* Nested Shapley method. In the "within stage", the *semi-equalized* distribution of income according to a subset of sources S belonging to the aggregated factor S_l is obtained by equalizing complementary sources belonging to S_l, while income coming from other aggregated sources than S_l is ignored.[18] Let us define y^{se} : $2_l^S \rightarrow \mathbf{R}^n$, such that $y^{se}(\emptyset) = [0]_{1xN}$ and for all $S \in 2_l^S, S \neq \emptyset$,

$$y^{se}(S) = (\sum_{j \in S} x_1^j + \sum_{\substack{j \notin S \\ j \in S_l}} \mu(x^j), ..., \sum_{j \in S} x_n^j + \sum_{\substack{j \notin S \\ j \in S_l}} \mu(x^j)). \qquad (4.4)$$

The contributions of the elementary sources for this semi-equalized version of the Nested Shapley decomposition will be given by:

[18]The example used in Section 4.1 allow us to illustrate the distinction between both approaches. Let S be the subset of income sources that includes earnings (E). The semiequalized distribution $y^{se}(E)$ is given by $y^{se}(E) = (x_1^E + \mu(x^C), ..., x_n^E + \mu(x^C))$, while the equalized distribution is given by: $y^e(E) = (x_1^E + \mu(x^C) + \mu(x^T), ..., x_n^E + \mu(x^C) + \mu(x^T))$.

$$NSh_j^{se}(K,\mathcal{P}_K,\mathbf{X},I) = \sum_{\substack{S\subseteq S_l \\ j\in S}} \frac{(s-1)!(s_l-s)!}{s_l!}[I(y^{se}(S)) - I(y^{se}(S-\{j\}))] + \frac{1}{s_l}[NSh_{S_l}^{e}(K,\mathcal{P}_K,\mathbf{X},I) - I(y^{se}(S_l))] \quad (4.5)$$

Expressions (4.3) and (4.5) are similar except that the various equalized income distributions in equation (4.3) have been substitute for the *semi-equalized* income distributions in expression (4.5). The first term on the RHS in expression (4.5) is now identical to the simple equalized Shapley contribution of source j in the sub-decomposition problem of the aggregated factor S_l. With this procedure, this term does not depend on any data related to sources which do not belong to the same aggregated component. However, as shown by Tables 3a and 3b, the *semi-equalized* procedure produces a negative contribution of capital income for the U.S. and the U.K, which seems an odd result since capital income is one of the more unequally distributed source. For that reason it seems preferable to work with the equalized procedure.

Gross Income			
Market Income		Transfers	
NSh=94.5		NSh=5.5	
Nested Shapley Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=89.0	NSh=5.5	NSh=3.8	NSh=1.7
Nested Shapley Semi-Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=97.4	NSh=2.9	NSh=9.1	NSh=3.6

Table 3a: Relative Contributions to the Equalized and Semi-equalized Nested-Shapley Decompositions of the Gini Index. United Kingdom 1995 (Household Non-adjusted Distribution).

4.2 *The Owen Decomposition*

Chantreuil and Trannoy (1997) and Shorrocks (1999) have considered an application of the Owen Value (1977). A direct application of

Gross Income			
Market Income		Transfers	
NSh=90.2		NSh=9.8	
Nested Shapley Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=89.0	NSh=1.1	NSh=8.2	NSh=1.6
Nested Shapley Semi-Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=97.2	NSh=7.3	NSh=23.4	NSh=13.3

Table 3b: Relative Contributions to the Equalized and Semi-equalized Nested-Shapley Decompositions of the Gini Index. United States 1994 (Household Non-adjusted Distribution).

the Owen[19] formula in an equalized procedure gives the following contribution to inequality of source j:

$$Ow_j^e(K, \mathcal{P}_K, \mathbf{X}, I) = \sum_{\substack{S \subseteq P_k \\ S_l \notin S}} \sum_{\substack{G \in S_l \\ j \notin G}} \frac{g!(s_l - g - 1)!s!(m - s - 1)!}{s_l!m!}$$
$$[I(y^e(S \cup G \cup \{j\})) - I(y^e((S \cup G))] \qquad (4.6)$$

where g denotes the dimensionality of G.

This expression is rather complex at first glance, but it can again be interpreted as the expected marginal contribution of an elementary source j belonging to some aggregated factor S_l. The kind of subset of sources considered here is composed of a union of a subset of sources. The first one, S, is a subset of aggregated factors while the second one, G, is a subset of the elementary sources which belongs to S_l. The coefficient in expression (4.6) can be interpreted as the probability that the subset $S \cup G$ will precede source j in some ordering of the elementary sources compatible with the partition P_K. This probability is the product of two probabilities: $g!(s_l - g - 1)!/s_l!$, which is the probability that G precedes j in the aggregated factor S_l; and $s!(m - g - 1)!/m!$ which is the probability that S precedes S_l in the partition P_K.

The discussion of the differences between the Owen contribution and the Nested-Shapley one will be pursued through the example

[19]For a generalization of the Owen decomposition that considers situations in which the set of sources is decomposed into a level structure see Chantreuil (1998).

given in Section 4.1. The Owen contribution of earnings (E) according to this decomposition rule is given by:

$$\begin{aligned} Ow^e_E(K,\mathcal{P}_K,\mathbf{X},I) &= \frac{1}{4}\left[I(y^e(E)) + I(y^e(M)) - I(y^e(C))\right] \\ &+ \frac{1}{4}\Big[I(y^e(E\cup T)) - I(y^e(T)) + I(y^e(X)) \\ &\quad -I(y^e(C\cup T))\Big] \end{aligned} \tag{4.7}$$

There are four admissible subsets of sources relative to which we have to compute the marginal contribution of earnings: the empty set, capital income (C), transfers (T), and transfers augmented by capital income ($T \cup C$).

A direct application of formula (9) gives for the Nested Shapley contribution of earnings:

$$\begin{aligned} NSh^e_E(K,\mathcal{P}_K,\mathbf{X},I) &= \frac{1}{2}\left[I(y^e(M)) + I(y^e(E)) - I(y^e(C))\right] \\ &+ \frac{1}{2}\Big\{\frac{1}{2}[I(y^e(X)) + I(y^e(M)) \\ &\quad -I(y^e(T))] - I(y^e(M))\Big\}. \end{aligned} \tag{4.8}$$

The first term is simply the equalized Shapley contribution of earnings in the sub-decomposition of market income while the second term stands for half the difference between the equalized Shapley contribution of market income and the market income inequality.

To understand the general difference between the Owen and the Nested-Shapley methods, we need to focus on the subgroups of components considered in the calculation. The Nested-Shapley rule only takes account of subgroups of sources at the same level, *i.e.* either subgroups of elementary income sources or collections of aggregated factors. The interpretation of those subgroups is straightforward if the partition P_K is relevant from an economic point of view. The Owen rule, however, consideres subgroups that mix elementary as well as aggregated components of income. For example, to obtain the contribution of earnings to inequality according to the Owen procedure, we have to compute inequality of the subset formed by earnings and all kinds of transfers ($E \cup T$), as well as inequality of the subset formed by capital income and transfers ($C \cup T$). These two sets of sources have a doubtful economic meaning. This feature reduces the interest of the latter decomposition rule, and favors the Nested-Shapley approach, that will be used in the following section.

Tables 4a and 4b present the differences between the values of the Owen and Nested-Shapley contributions for the British and American income distributions. These differences are significant (particularly important for the contribution of capital) and show the importance of the selection of the decomposition method.

Gross Income			
Market Income		Transfers	
NSh=Ow=94.5		NSh=Ow=5.5	
Nested Shapley Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=89.0	NSh=5.5	NSh=3.8	NSh=1.7
Nested Shapley Semi-Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
Ow=83.1	Ow=11.4	Ow=3.2	Ow=2.3

Table 4a: Relative Contributions to the Equalized Nested-Shapley and Owen Decompositions of the Gini Index. United Kingdom 1995 (Household Non-adjusted distribution).

Gross Income			
Market Income		Transfers	
NSh=Ow=90.2		NSh=Ow=9.8	
Nested Shapley Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
NSh=89.0	NSh=1.1	NSh=8.2	NSh=1.6
Nested Shapley Semi-Equalized			
Earnings	Capital	Replacement Inc.	Non-contributory Benefits
Ow=80.7	Ow=9.5	Ow=6.9	Ow=3.0

Table 4b: Relative Contributions to the Equalized Nested-Shapley and Owen Decompositions of the Gini Index. United States 1994 (Household Non-adjusted distribution).

5 Alternative Grouping Solutions

In order to apply the Nested-Shapley rule to real data, it is necessary to define an adequate income structure which leads to a meaningful decomposition of inequality by income components. We first focus on the disaggregation of gross income before tackling the question raised by the introduction of taxes. In this section, definite solutions are not offered. We aim only to expose some additional problems arising in the implementation of the Shapley procedure.

5.1 Gross Income Decomposition

In most countries, microeconomic datasets allow us to disaggregate household income into at least eleven elementary sources whose impact on inequality are interesting in some respect: wages and salaries of household heads, wages of other household members, self-employment income, private pensions, cash property income, social retirement or public pensions, other replacement income, means-tested benefits, other benefits, private transfers and direct taxes.[20] However, among this set of sources, the hierarchy of factors seems ambiguous and somewhat different hierarchical or tree structures can be defended depending on the criteria considered to split income sources. Furthermore the nature of income (gross versus net) influences the structure we have in mind. This paves the way to differences in results concerning the contribution of income sources to overall inequality. The ideal structure would be one that takes into account all relevant economic links but does not include any non pertinent subset of income sources. The following graph indicates which links between income sources seem relevant. The impossibility of reducing the structure of economic links to a hierarchy emerges quite naturally from the picture.

Consider for instance the case of replacement income. As it forms part of public sector transfer programs, and, at least in some countries, also includes a redistributive component, a first solution would be to place it alongside purely redistributive transfers (means-tested and other benefits). Although the redistributive feature of replacement income cannot be denied, some authors like Bourguignon (1999) prefer to focus on their social insurance roots. Since in most countries social retirement and unemployment benefits are linked to the earnings obtained by workers when active through social security contributions, they can be viewed as a form of delayed salaries. Moreover,

[20] A detailed description of each income source is given in the Appendix.

differences in the coverage of social security programs among developed countries, along with the substitution between public and private insurance, have driven researchers to limit redistributive analysis to non-contributory social benefits and taxes.[21]. This line of reasoning leads to a second solution where it is convenient to include replacement income within the "linked to labor" income subset. But this option has some limitations: it would be impossible to assess the joint impact of a subset traditionally analyzed in this kind of study, namely "market income". The same kind of difficulty arises when analyzing the effect of income from private pensions. On the one hand, it can be seen as replacement or insurance income; on the other hand, since it comes from savings, it seems quite natural to include it with the rest of property income. It is difficult to decide between these two alternatives.

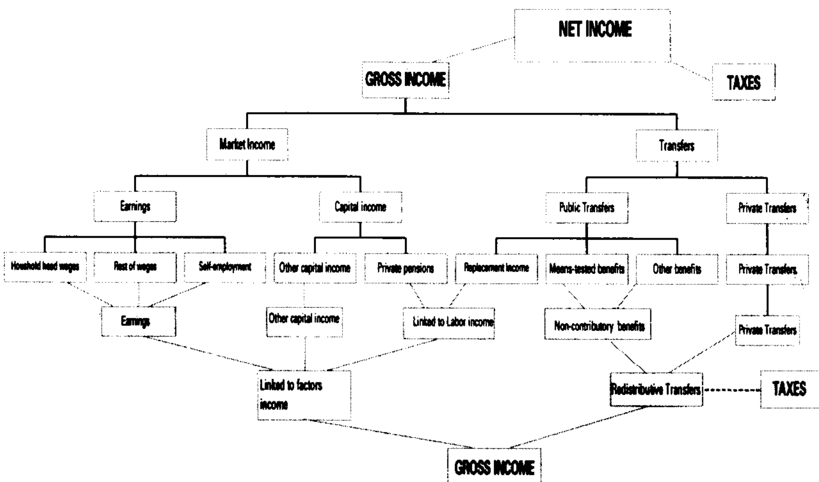

Fig. 1: Relevant links between elementary sources

[21]See Bourguignon (1999) for further details.

These are the kind of decisions we have to face when trying to define an income tree for an empirical decomposition of inequality. The applied analysis requires the careful definition of the subsets considered more relevant to our objectives and the selection of the tree structure consistent with them. As there is no ideal income structure, it is important to assess the extent to which the selection of the tree affects the results.

Nested-Shapley inequality contributions to households gross income inequality have been computed for ten elementary sources according to three appealing income structures. The first tree follows the traditional approach that splits household income in two main sources: market income and transfers. Then, market income divides into earnings and property income, and we have distinguished between replacement income (made of social retirement benefits and unemployment and other replacement income) and redistributive transfers (subdivided into public—or non contributory benefits—and private transfers). See Figures A.1a and A.1b in the Appendix. The other proposed trees adopt a restrictive definition of redistribution limited to non-contributory benefits, and consider replacement income as a "linked to labor" income. They differ regarding the treatment given to private pensions. In the second tree, income from pensions funds is combined with replacement income (see Figures A.2a and A.2b in the Appendix), while the third tree emphasizes the link between that source and the rest of property income (see Figures A.3a and A.3b in the appendix).

Nested-Shapley contributions according to these three trees exhibit moderate volatility. As expected, the sources treated differently in the alternative structures, and the sources directly related to them are most affected. Thus, the Nested-Shapley contributions of the different types of earnings are very little affected by the disaggregation of the rest of factors. However, the impact is higher for replacement income, redistributive transfers and property income. For instance, for the United Kingdom, the contribution of replacement income varies from -1.8% when this source is regarded as income-linked to production factors, to 3.8% when they are merged with the rest of transfers. The contribution of "cash property income" to inequality oscillates between 2.3 and 0.6%, while that of private pension contributions varies from 4.9% to 7.6%. Even if the magnitude of the changes is not huge in absolute terms, the picture is less favorable in relative terms since the components in question have a rather small income shares. Therefore, it seem impossible to escape the conclusion that the choice of a tree has a critical impact on the decomposition results.

5.2 *Introducing Taxes: A Difficulty*

In all the previous examples, income has been defined in gross terms, i.e., before deducting direct taxes. We now investigate the consequences of including direct taxes[22] in the analysis. Are the contributions of income sources to net income inequality different from the contributions obtained when decomposing gross income inequality?. Results of the decomposition of net income are displayed in Tables 5a and 5b.

Net Income				
Gross Income				Taxes
NSh=98.8				NSh=1.2
Market Income		Transfers		
NSh=109.1		NSh=10.3		
Earnings	Capital	Replacement Inc.	Non-contributory Benefits	
NSh=111.8	NSh=2.7	NSh=3.7	NSh=6.6	

Table 5a: Relative Contributions to the Equalized Nested-Shapley Decompositions of the Gini Index. United Kingdom 1995 (Household non-adjusted distribution).

Net Income				
Gross Income				Taxes
NSh=100.8				NSh=0.8
Market Income		Transfers		
NSh=106.0		NSh=5.1		
Earnings	Capital	Replacement Inc.	Non-contributory Benefits	
NSh=115.0	NSh=9.0	NSh=2.1	NSh=7.2	

Table 5b: Relative Contributions to the Equalized Nested-Shapley Decompositions of the Gini Index. United States 1994 (Household non-adjusted distribution).

Including taxes influences the sign as well as the magnitude of the relative contributions. Comparison of the Table 5a and 5b results with

[22]According to LIS, taxes include: income taxes (personal income tax liabilities), mandatory contributions for self-employed and mandatory employee contributions.

those in Tables 4a and 4b shows us that the relative contribution of transfers becomes negative, while the relative contribution of market income exceeds 100. Moreover the results concerning the contribution of capital income are perturbed in an unpleasant way. The contribution of capital income becomes negative, a rather bizarre result since capital is one of the more unequally distributed source of income (see Table 0 in the Appendix). These changes are a by-product of the lack of independence of the Shapley approach from the degree of disaggregation. More specifically Nested Shapley contributions are not invariant to adding a starting stage in the tree of components. Inspection of formula (9) helps us to understand why this is so. Suppose that source j is capital income. The second term on the RHS in expression (9) depends on the Nested Shapley contribution of the aggregated factor to which capital income belongs, i.e., market income. But adding an initial stage changes the Nested Shapley contribution of market income and consequently the contribution of capital income. In view of this flaw of the Nested Shapley method, it might be sensible to require of a decomposition method to respect the following property: the ratio of relative contributions at some stage must be independent from what happens on previous stages of the tree.

6 Concluding Comments

As the results have shown, mechanically applying the Shapley decomposition rule may give odd results [for a rough application to French data see Auvray and Trannoy (1992)]. The major shortcoming comes from the lack of independence from the level of disaggregation. Some empirical solutions can circumvent the problem but not cure the illness. As far as the decomposition of gross income is concerned the proposed solutions are the following:

1) To avoid the use of the *zero income decomposition*. Indeed our results obtained by the zero approach show that the contributions of income sources to overall inequality are highly volatile. These contributions are much more dependent on the level of aggregation than those obtained through the *equalized* method.

2) To apply the Nested Shapley rule. In both, Nested Shapley and Owen decompositions, the contribution to overall inequality of a secondary factor is dependent on the treatment given to the primary factor it belongs to. Nevertheless, it is independent of the remaining factors' disaggregation. Although the Nested Shapely and Owen rules satisfy this milder independence requirement, the Owen contribution

needs to consider the coalitions of elementary sources to aggregated sources to which they do not belong. Hence, the Owen contribution intrinsically considers some subsets whose meaning is doubtful. The Nested-Shapley decomposition therefore seems to be more adequate.

3) Careful design of the income tree. Once the Nested-Shapley procedure is chosen, it is necessary to define a hierarchical income structure that allows a meaningful marginalist interpretation of the decomposition rule. We show that the structure of economic links between income sources cannot be easily reduced to a tree. An ideal tree is consequently difficult to imagine. Nevertheless, the differences between the alternatives are not that great. Introducing a more structural model generating the hierarchy of sources before applying the Shapley procedures could be useful.

Once these principles have been respected, one may hope to have sensible results that can be compared to those obtained with "natural" decompositions. In a companion paper (see Sastre and Trannoy 2000) we apply the lessons gathered here and compare the Shapley decomposition results to those obtained with more standard methods. For the decomposition of net income, there is still a difficulty which can be solved if we are looking for methods requiring an additional property: the ratio of relative contributions at some stage must be independent from what happens on previous stages of the tree (see the above quoted paper for more details).

Acknowledgements

This document was completed when Mercedes Sastre was visiting THEMA. The financial support of the European Commission through the Contract No. ERBFMRXCT980248 is grateful acknowledged. We would like to thank F. Chantreuil, S. P. Jenkins, A. F. Shorrocks and two anonymous referees for their helpful comments.

References

Auvray, C., and Trannoy, A. (1992): "Décomposition par source de l'inégalité des revenus à l'aide de la valeur de Shapley." Journées de Microéconomies Appliquées. Sfax.

Bourguignon, F. (1999): "Fiscalité et Redistribution." *Rapport au Conseil d'Analyse Economique* 11: 11–50.

Cancian, M., and Reed, D. (1998): "Assessing the Effects of Wives Earnings on Family Income Inequality."*Review of Economics and Statistics* pages 73–79.

Chantreuil, F. (1998): "Axiomatics of Level Structure Values." In *Power Indices and Coalition Formation*, edited by M. J. J. Holler, and G. Owen. : .

Chantreuil, F., and Trannoy, A. (1997): "Inequality Decomposition Values." Université de Cergy-Pontoise. presented in Bordeaux TMR Network *Distribution and Redistribution of Income.* Conference.

Chantreuil, F., and Trannoy, A. (1999): "Inequality Decomposition Values: The Trade-off Between Marginality and Consistency." - THEMA. DP 99-24.

Gottsshalk, P., and Smeeding, T. (1997): "Cross-National Comparisons of Earnings and Income Inequality."*Journal of Economic Literature* 35: 633–687.

Jenkins, S. P. (1995): "Accounting for Inequality Trends: Decomposition Analyzes for the UK 1971-86."*Economica* 62: 29–64.

Lerman, R. (1999): "How Income Sources Affect Income Inequality?" In *Handbook on Income Inequality Measurement.*, edited by J. Silber. Boston, Dordrecht, London: Kluwer Academic Publishers.

Lerman, R., and Yitzhaki, S. (1985): "Income Inequality by Income Source: a New Approach and Applications to the United States."*Review of Economics and Statistics* 57,1.

Mirman, L., Taubman, Y., and Zang, I. (1985): "On the Use of Game-theoretic Concepts in Cost Accounting." In *Cost Allocation: Methods, Principles, Applications*, edited by H. P. Young. North Holland: Elsevier Science.

Owen, G. (1977): "Values of Games With Priori Unions." In *Essays in Mathematical Economics and Game Theory*, edited by R. Heim, and O. Moeschlin. New York: Springer Verlag.

Sastre, M., and Trannoy, A. (2000): "A Marginalist Approach to Inequality Decomposition By Factor Components: An Application to OECD Countries using the LIS database." THEMA.

Shapley, L. S. (1953): *A Value for N-person Games.* In Kuhn, H. W., and Tucker, A. W. (eds.), *Annals of Mathematics Studies.* volume 28. pages 307–317. Princeton University Press. Contributions to the Theory of Games, Vol. 2.

Shorrocks, A. F. (1982): "Inequality Decomposition by Factor Components."*Econometrica* 50: 193–211.

Shorrocks, A. F. (1999): *Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value.* University of Essex: mimeo.

Addresses of authors:
Mercedes Sastre, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain. email: msastre@ccee.ucm.es.
Alain Trannoy, THEMA, Université de Cergy-Pontoise, 33 Bvd. du Port, 95011 Cergy-Cedex, France. email: trannoy@u-cergy.fr

Appendix

INCOME SOURCE*	DEFINITION
	TOTAL EARNINGS (1)
Wages	Includes all forms of cash wage and salary income, including employer bonuses, 13th month bonus...gross of employee social insurance contributions/taxes but net of employer social insurance contribution/taxes
Self-employment Income	Gross of social insurance contributions
	REPLACEMENT INCOME (2)
Social Retirement	Social retirement benefits (Cash social security benefits for old age or survivors (widows/widowers) and public sector pensions (these include pensions for public employees and do not include amounts coming from social security benefits for the aged or survivors).
Other replacement income	Cash sickness insurance benefits (sick pay). Accident pay (cash accidents or injury payments, only short term public stipends for injured workers are included). Disability pay (long-term cash benefits for partial or total permanent disability of permanent injury). Unemployment compensation (cash payments insurance benefits in case of unemployment: LIS excludes means-tested unemployment benefits). Maternity allowances (LIS excludes means-tested or non mandatory employer provided benefits). Military/vet/ war benefits (Cash veteran's or military benefits for old age, military disability, etc. LIS also includes cash benefits provided to dependents of the military as long as they are non means-tested).

	Other social insurance (other cash or near cash benefits that are not included in the more specific cash benefit variables, such as other social insurance, e.g. education training of retraining allowances as long as they are non means-tested, also scholarships).
PROPERTY INCOME (3)	
Cash property income	Includes cash interest, rent, dividends, annuities, royalties, etc, but excludes capital gains, lottery winnings, inheritances, insurance settlements and all other forms of lump sum payments.
Private pensions	Employer payments for retirement that may (or may not) supplement social security transfers. Self-employment pension plans are included if they are designed to supplement social security, e.g. individual retirement accounts.
LINKED TO FACTORS INCOME (1+2+3)	
NON-CONTRIBUTORY BENEFITS (4)	
Means-tested benefits	Means-tested cash benefits (means-tested or so called "emergency" benefits and benefits for long term unemployed if means-tested). LIS includes also mandatory cash transfers not tied to some form of in-kind benefits, e.g. not tied to food or education). All near cash benefits (includes all forms of transfers that are, in a strict sense, in kind payments) but have a cash equivalent value equal or nearly equal to the market value, including near-cash housing benefits.
Other benefits	Child or family allowances (cash payments for child of family allowances. This may include refundable tax credits as long as they are not means-tested. Value of food, housing, medical, heating and educations benefits.

PRIVATE TRANSFERS (5)	
Private transfers	Alimony or child support received (these are counted apart from child or family allowances even if government mandated). Other regular private income (regular, cash private inter-household transfers, including from friends and relatives, but not including one time cash gifts). Other cash income (rest of variables in which LIS puts all cash income that can not be classified in one of the previous variables).
TAXES (6)	
Taxes	Mandatory contributions for self-employed (all forms of social insurance contributions paid by the self-employed: social security, medical insurance, unemployement, etc. This information is rarely available). Income taxes (personal income tax liabilities). Mandatory employee contributions (payroll taxes from wage and salary workers for all forms of social insurance: social security, health plans, unemployement insurance, etc.)
GROSS INCOME* (1+2+3+4+5)	
NET INCOME* (1+2+3+4+5-6)	

*All income variables are recorded as yearly amounts. LIS imputes yearly amounts if not provided in original survey.

	Income Share(%)	Gini index	Coeff. Variation	N.Decom. Variance	N.Decom. Gini
Household head wages	40.0	0.70	274.4	40.0	50.1
Rest of wages	19.4	0.80	334.3	21.8	28.2
Self-employment	12.3	0.94	808.4	32.8	20.5
Cash property income	4.7	0.90	731.8	7.4	5.3
Private Pensions	6.7	0.90	528.2	2.9	3.6
Social retirement	6.2	0.74	304.9	-2.7	-4.2
Other replacement income	3.2	0.90	483.8	-0.5	-0.1
Means-tested benefits	4.7	0.86	394.7	-3.0	-4.6
Other benefits	1.5	0.75	296.0	0.3	0.5
Private transfers	1.0	0.97	1435.3	0.8	0.7
Gross Income	100	0.42	164.5	100	100

Table A.1a: Basic Statistics: Income Share, Inequality and Sources Relative Contributions (%) to Inequality According to Natural Decompositions of Variance and Gini. United Kingdom 1995 (Household Nonadjusted Distribution)

	Income Share (%)	Gini index	Coeff. Variation	N.Decom. Variance	N.Decom. Gini
Household head wages	48.8	0.61	5450.3	54.5	55.4
Rest of wages	22.3	0.74	7452.1	25.9	29.6
Self-employment	5.3	0.96	22104.2	9.2	7.2
Cash property income	5.3	0.89	15970.3	10.1	7.5
Private Pensions	2.1	0.95	20415.8	1.4	1.6
Social retirement	8.2	0.81	8476.6	0.3	0.7
Other replacement income	1.3	0.92	19663.2	0.2	0.3
Means-tested benefits	1.3	0.93	15555.5	-0.8	-1.5
Other benefits	4.7	0.75	6924.0	-1.0	-1.4
Private transfers	0.7	0.97	33877.6	0.2	0.3
Gross Income	100	0.41	3574.9	100	100

Table A.1b: Basic Statistics: Income Share, Inequality and Sources Relative Contributions (%) to Inequality According to Natural Decompositions of Variance and Gini. United States 1994 (Household Nonadjusted Distribution)

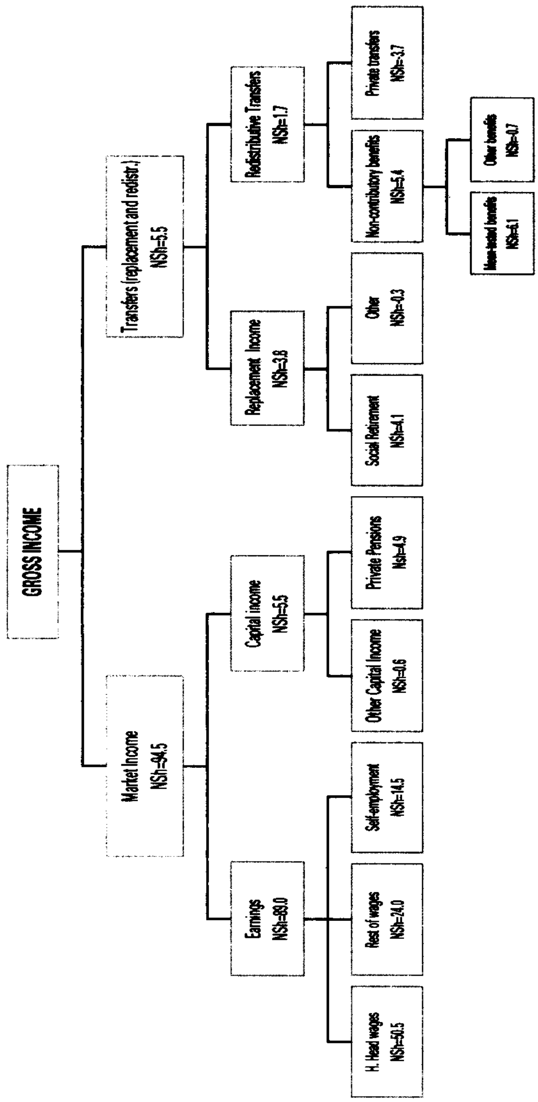

Fig. A.1a: Nested Shapley Equalized Decomposition. Gini Index. United Kingdom 1995 (Household Nonadjusted Distribution)

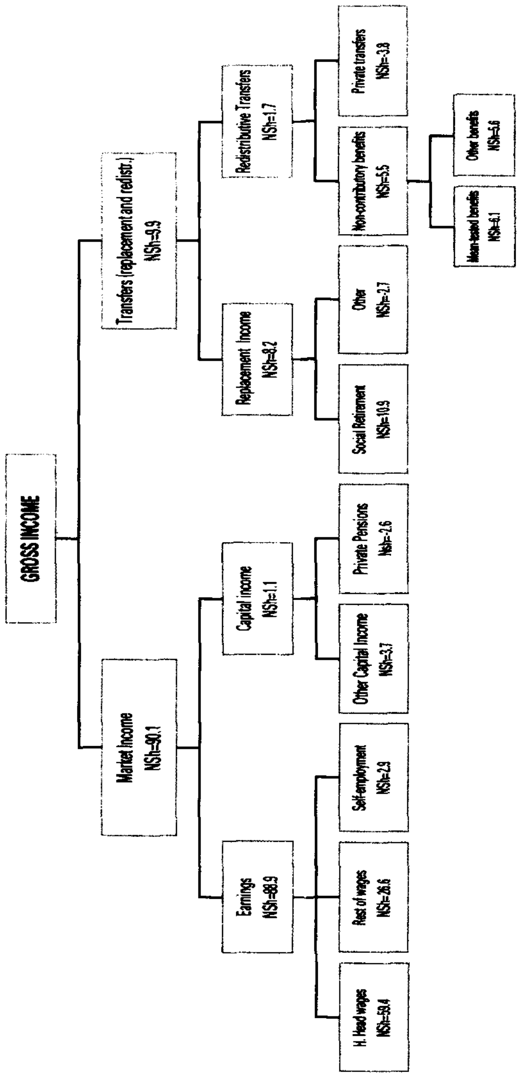

Fig. A.1b: Nested Shapley Equalized Decomposition. Gini Index. United States 1994 (Household Nonadjusted Distribution)

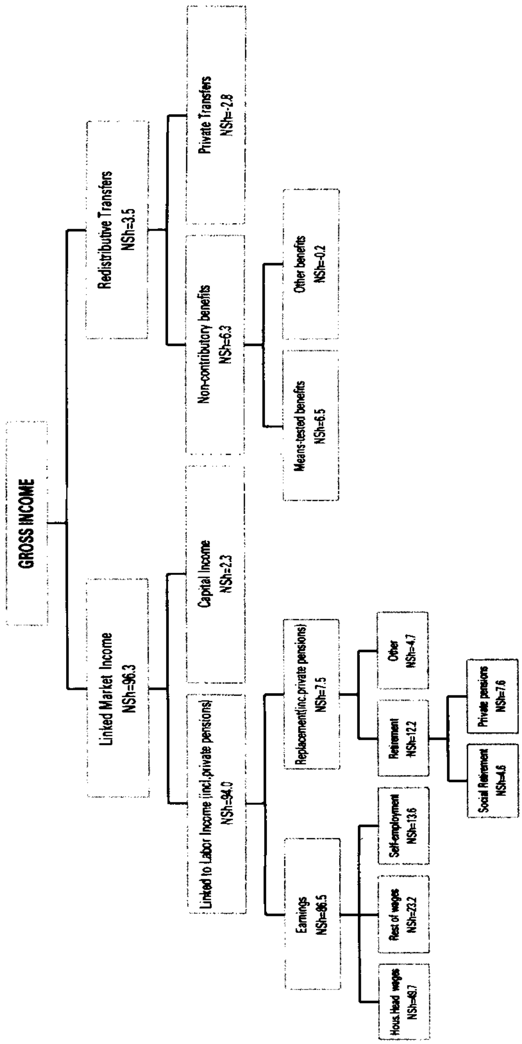

Fig. A.2a: Nested Shapley Equalized Decomposition. Gini Index. United Kingdom 1995 (Household Nonadjusted Distribution)

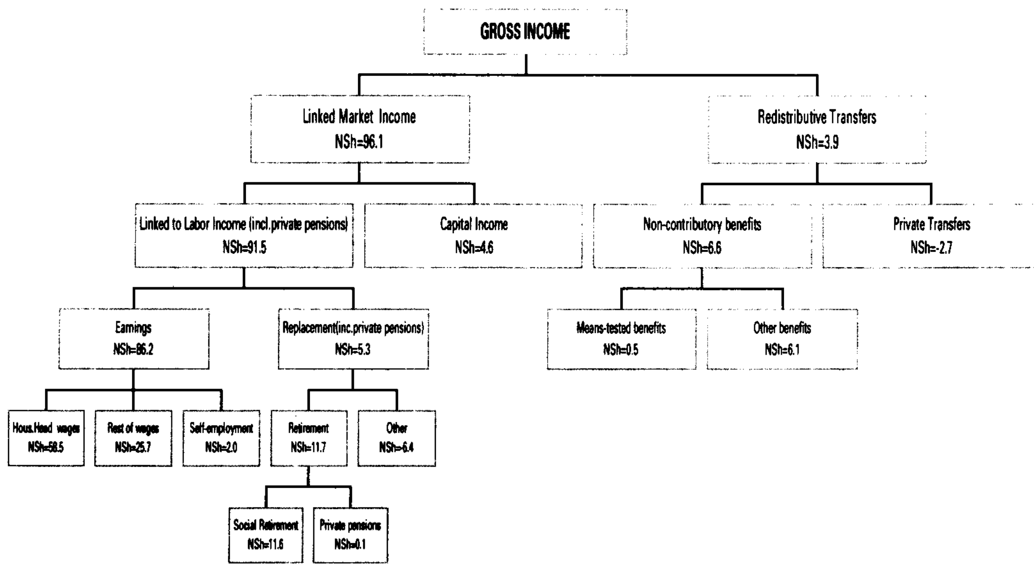

Fig. A.2b: Nested Shapley Equalized Decomposition. Gini Index. United Kingdom 1995 (Household Nonadjusted Distribution)

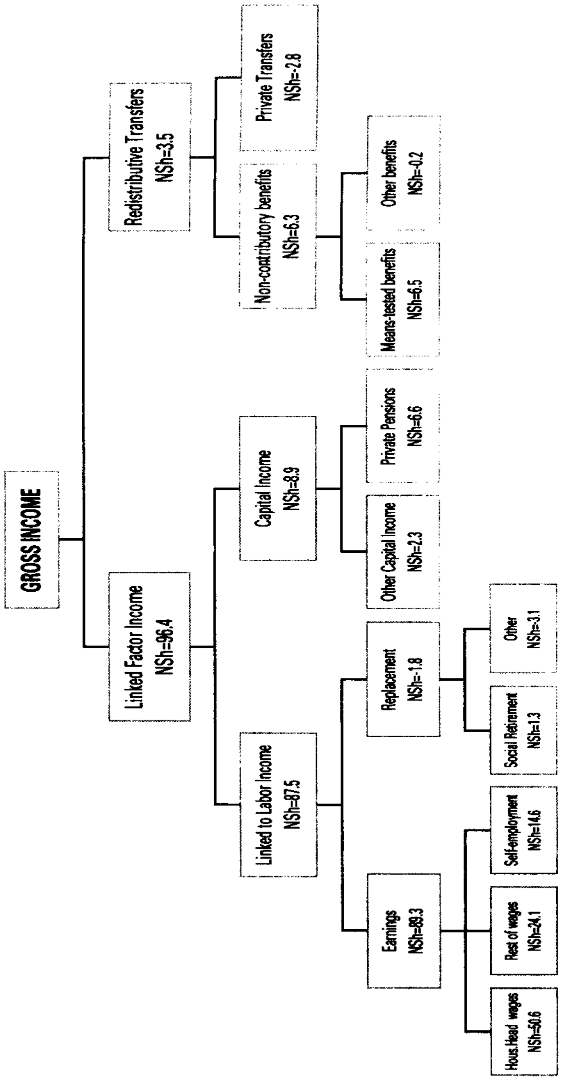

Fig. A.3a: Nested Shapley Equalized Decomposition. Gini Index. United Kingdom 1995 (Household Nonadjusted Distribution)

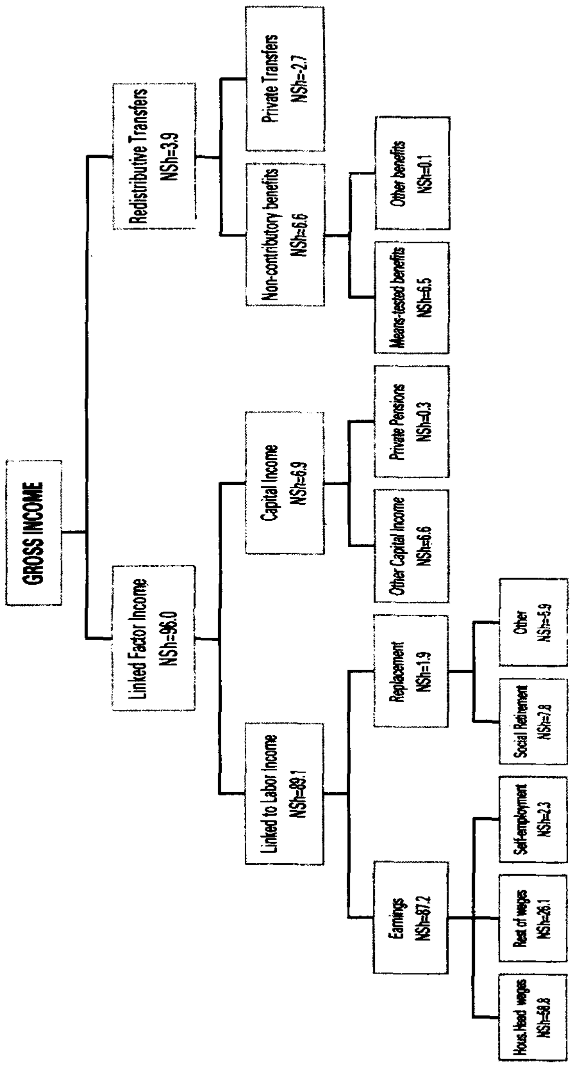

Fig. A.3b: Nested Shapley Equalized Decomposition. *Gini Index. United States 1994* (Household Nonadjusted Distribution)