MEASURING HEALTH POLARIZATION WITH SELF-ASSESSED HEALTH DATA

BENEDICTE APOUEY*
Paris School of Economics (PSE)

SUMMARY
This paper proposes an axiomatic foundation for new measures of polarization that can be applied to ordinal distributions such as self-assessed health (SAH) data. This is an improvement over the existing measures of polarization that can be used only for cardinal variables. The new measures of polarization avoid one difficulty that the related measures for evaluating health inequalities face. Indeed, inequality measures are mean based, and since only cardinal variables have a mean, SAH has to be cardinalized to compute a mean, which can then be used to calculate an inequality measure. In contrast, the new polarization measures are median based and hence do not require to impose cardinal scaling on the categories. After deriving the properties of these new polarization measures, we provide an empirical illustration using data from the British Household Panel Survey that demonstrates that SAH polarization is also a relevant question on empirical grounds, and that the polarization measures are adequate to evaluate polarization phenomena whereas inequality measures are not adequate in these cases. Copyright © 2007 John Wiley & Sons, Ltd.

INTRODUCTION
Self-assessed health (SAH) data play a prominent role for the analysis of health measures. These data are generated by asking ‘How is your health in general?’, with the response categories ranging from ‘Very poor’ to ‘Excellent’. Self-rating offers several advantages. First, it is one of the most frequently asked questions in epidemiological surveys. Second, it offers a summary of an individual’s general state of health. Third, many longitudinal studies have highlighted that a person’s own appraisal of his or her health is a very good predictor of future mortality and morbidity. The correlation between SAH and mortality remains strong even after controlling for other health variables and for socio-economic variables (Idler and Benyamini, 1997).

The use of self-rating of health has become very common in empirical research. Up to now, when looking at SAH dispersion, people have focused on inequality. By definition, all measures of inequality are mean based. Thus, cardinality must be imposed on ordinal SAH categories in order to calculate a mean, which can then be used to compute an inequality measure such as the Gini coefficient. This approach is interesting (especially when dealing with reporting bias), but it has a shortcoming: imposing cardinality is a supra-ordinal assumption that changes the original properties of the SAH.

In this paper, we provide an axiomatic foundation of dispersion measures that do not require the cardinal scaling of the SAH. We use the median individual as a reference point instead of the mean for studying dispersion of an ordinal variable, because the existence of the median individual and his

*Correspondence to: Paris School of Economics (PSE), 48 Bd Jourdan, 75014 Paris, France. E-mail: apouey@pse.ens.fr
location in the SAH distribution do not depend on a scaling (he is always ‘central’ in the distribution). In the literature investigating income dispersion, an alternative to mean-based inequality measures is median-based polarization measures. Income polarization is the study of the disappearing middle class and the emergence of a divided society. The measures that have been proposed in this literature are only applicable to cardinal data. We take inspiration from these works, and one of the contributions of the current paper lies in applying the concepts of polarization to ordinal variables.

Polarization for ordinal data is not only relevant to theoretical grounds, but can also serve to explain empirical phenomena such as mass relocation from the middle of SAH distribution to the poles. Inequality measures do not necessarily convey an adequate picture of these phenomena. In contrast, they are naturally handled with polarization tools.

More generally, to our knowledge, the notion of polarization has not been widely used so far in health economics. We believe that this concept is also well suited for the analysis of the distribution of cardinal variables such as the body mass index, because the rise of obesity may be seen as a relocation from the middle of the distribution to one tail. The notion could also be useful to know if the income polarization that has been recently observed in a number of countries implies a health polarization.

This paper is organized as follows. The second section reviews the existing methods of measuring inequality of SAH. The third section shows that SAH is an interesting variable for measuring polarization and presents strict partial orderings on polarization. The fourth section contains the axiomatic foundation of the family of polarization indices which do not require supra-ordinal assumption. The fifth section provides an empirical illustration using the British Household Panel Survey (BHPS), and the sixth section concludes.

SAH INEQUALITY

Inequality indices measure the difference between the observed distribution of a variable and the hypothetical distribution where everybody would have an average level of the variable. Thus, inequality measures require the existence of the ‘average level’ of the variable, that is to say the variable should have a mean. The concept of inequality then is well suited to study the dispersion of cardinal variables.

In order to compute a measure of inequality, the ordinal SAH is transformed into a cardinal variable, either discrete or continuous.

When SAH is transformed into a discrete variable, a numerical value is assigned to each category. Thus, two individuals who are in the same category have the same ‘value of health’. On the contrary, when the SAH is transformed into a continuous variable, two individuals who are in the same category do not necessarily have the same ‘value of health’, and there is an individual-level variation in the health. We distinguish three cases:

- There is no continuous latent variable and the SAH is transformed into a discrete variable. For example, one can assign the values \(1, 2, 3, 4, 5\) to the categories.
- There is a continuous latent variable underlying SAH and SAH is transformed into a discrete variable. For example, Wagstaff and Van Doorslaer (1994) assumed that the latent variable is a standard lognormal density function and then scored the SAH using one point of the underlying distribution.
- There is a continuous latent variable and the SAH is transformed into a continuous variable.
  - One can estimate an ordered probit regression in which the SAH is the dependent variable and the latent variable \(y^a\) is a function of health variables (and sometimes of socio-economic variables) \(x\). Then the prediction of the linear index \(x_i\hat{\beta}\) is used as a measure of individual health after rescaling (Groot, 2000).

\[\text{1The works that study this relocation use other tools than the polarization measures for cardinal variables presented below. See Contoyannis and Wildman (2006) and Madden (2006).}\]
In the ECuity project, the latent variable is an empirical index (HUI, SF36). It is assumed that the $q$th quantile of the distribution of the empirical index corresponds to the $q$th quantile of the distribution of the SAH. A technique involving interval-regression enables the authors to attribute to every individual a cardinal measure of health (Van Doorslaer and Jones, 2003).

Some cardinalization methods offer advantages in that they are able to deal with reporting bias in SAH. In fact, individuals with the same true health but different socio-economic characteristics may not answer to the SAH question in the same way. So observations in the same SAH category do not necessarily have the same true level of health. Some methods for cardinalizing SAH allow to take these reporting biases into account, whereas our method does not.

But cardinalization also has shortcomings.

First, if SAH is transformed into a cardinal variable, then the sum of individual levels of health exists. The problem is that we hardly see what it represents.

Second, different cardinalizations are always possible and they induce changes in inequality rankings. If we consider two distributions 1 and 2, we can show that with a certain cardinalization, inequality of 1 is larger than inequality of 2, and with another cardinalization, inequality of 2 is larger than inequality of 1. Allison and Foster (2004) provided such an example using the variance as the inequality measure, but the conclusion remains the same with other inequality measures (the Gini index). Indeed, inequality measures are mean based, and the change of cardinalization leads to a change in the position of the mean in the distribution, which can entail a reversal in the ranking.

POLARIZATION AND STRICT PARTIAL ORDERINGS

SAH polarization

An ideal measure of dispersion for ordinal data such as SAH may not be mean based. We follow a suggestion of Allison and Foster (2004) to use the median individual as a reference point instead of the mean for ordinal data in order to measure dispersion, because the median individual is central in the distribution (he is such that one half of the population declares lower health than him or equal health to him, and the other half declares equal health to him or better health than him) and his existence does not depend on the scaling (whereas only cardinal variables have a mean). Since the reference point is the median individual, SAH does not have to be transformed and its nature remains unchanged. There already exists polarization measures that are median based, but that are applicable only to cardinal variables. We thus take these measures as a starting point, and we adapt two of their main properties in an ordinal perspective.

Wolfson (1994), Esteban and Ray (1994), and Duclos et al. (2004) have defined polarization concepts for cardinal distributions such as income. In short, the concept of polarization has been used to study the ‘disappearing middle’ of the income distribution, it is adequate to describe a divided society, with increasing shares of rich and poor people. Low polarization corresponds to distributions that are very concentrated around the median. The notion of income polarization is distinct from that of income inequality. We follow Wang and Tsui (2000) and we suppose that the population is divided into two subgroups: those whose income is below the median and those whose income is above the median. The two main properties of polarization measures can then be summarized as follows.

- **Property 1**: Polarization increases for a Pigou–Dalton transfer from someone below the median to someone above the median. There is a spread away from the median, the two subgroups move apart and heterogeneity between them increases. Figure 1 displays a spread away from the median for a hypothetical income distribution in which the population is distributed over 8 and 9 cardinal values of income (10, 20, ..., 90) and the median income equals 50. There is a transfer of amount 10 from someone below the median to someone above the median. Property 1 requires that there is more polarization in Figure 1(b) than in (a).
Property 2: Polarization increases for a Pigou–Dalton transfer between two individuals who are in the same subgroup. Incomes within this subgroup tend to cluster; hence, homogeneity in the subgroup increases which implies that bipolarity increases. Figure 2 contains a Pigou–Dalton transfer below the median, from someone who has 40 to someone who has 10 for a hypothetical income distribution. Property 2 means that polarization measure in (b) is greater than polarization measure in (a).

Thus, Pigou–Dalton transfers from someone above the median to someone below the median decrease polarization (property 1), whereas Pigou–Dalton transfers within a subgroup increase polarization (property 2). In contrast, Pigou–Dalton transfers always decrease inequality for any inequality measure consistent with the Pigou–Dalton principle (e.g. the Gini coefficient); that is the main difference between inequality and polarization.

Strict partial orderings

Like polarization measures, an ideal measure of SAH dispersion could be median based. Indeed, the reference point would preferably not be the mean since the average level of SAH may not exist. We also suggest that our measures satisfy two strict partial orderings that are identical in spirit to the two properties of measures of income polarization. The rest of the section describes these two strict partial orderings that the median-based measures have to satisfy: spread away from the median (S), which is identical in spirit to Property 1, and increased bipolarity (IB), which is identical in spirit to Property 2.

Notations

We use the following notations in the rest of the paper:

- \( C \) is the number of categories. We suppose that \( C \geq 3 \).
- \( n = (n_1, \ldots, n_{C-1}, n_C) \) with \( n_c \) representing the number of people in the \( c \)th category.
- \( N_c = \sum_{i=1}^{c} n_i \) It represents the number of people who are in the first \( c \)th categories.
- \( N = (N_1, \ldots, N_{C-1}, N_C) \).
- \( p = (p_1, \ldots, p_{C-1}, p_C) \) with \( p_c \) representing the proportion for the \( c \)th category.
- \( F_c = \sum_{i=1}^{c} p_i \).
- \( F = (F_1, \ldots, F_{C-1}, 1) \).
- \( m \) is the median category of the distribution. It is such that \( F_{m-1} \leq \frac{1}{2} \) and \( F_m \geq \frac{1}{2} \), or \( N_{m-1} < N_C/2 \) and \( N_{m} \geq N_C/2 \). By definition, the median individual is in the median category.
For an ordinal variable all the information relative to the distribution is given by the distributions of the cumulative distribution functions \((N_1, \ldots, N_C)\) or \((F_1, \ldots, F_{C-1}, 1)\).

\(1\) Spread away from the median, \(S\), Allison and Foster (2004), modified. Allison and Foster (2004) argued that polarization increases when there is a spread away from the median (individual or category).

**Definition 1 (Spread away from the median, \(S\), Allison and Foster (2004), modified)**

Given any two distributions \(F^1\) and \(F^2\), \(F^2\) has greater spread than \(F^1\), \((F^2, F^1) \in S\), if:

- the distributions \(F^1\) and \(F^2\) have the same median state \(m, m(F^1) = m(F^2) = m\), i.e. the median individual remains in the same category \(m\);
- for all \(c < m, F^2_c \geq F^1_c\);
- for all \(c \geq m, F^2_c \leq F^1_c\)

with at least one category \(c < m\) such that \(F^2_c > F^1_c\) or one \(c \geq m\) such that \(F^2_c \leq F^1_c\).

In a very similar way, \((N^2, N^1) \in S\) if:

- \(N^1_C = N^2_C\);
- \(m(N^1) = m(N^2)\);
- for all \(c < m, N^2_c \geq N^1_c\);
- for all \(c \geq m, N^2_c \leq N^1_c\)

with at least one category \(c < m\) such that \(N^2_c > N^1_c\) or one \(c \geq m\) such that \(N^2_c < N^1_c\).

\(S\) are relocations from the middle of the distribution to the tails. Figure 3 displays the \(S\) for an exemplary distribution. The median category remains the ‘Poor’ category for ‘a’ and ‘b’. The \(S\) movement pulls people from ‘Poor’ (P) to ‘Very Poor’ (VP), and from ‘Good’ (G) to ‘Excellent’ (E). The \(S\) partial ordering means that polarization in ‘b’ is larger than polarization in ‘a’.

\(2\) Increased bipolarity, IB, Wang and Tsui (2000), modified. The second strict partial ordering is inspired by Property 2.

**Definition 2 (Transfer)**

We say that there is a transfer of a number \(\Delta > 0\) or a proportion \(\delta > 0\) from category \(c_1\) to \(c_1 + 1\) and from \(c_2\) to \(c_2 - 1\) if \(n_{c_1} \geq \Delta\) and \(n_{c_2} \geq \Delta\), or \(p_{c_1} \geq \delta\) and \(p_{c_2} \geq \delta\), and \(c_2 - 1 \geq c_1 + 1\).
Given any two distributions \( N_1 \) and \( N_2 \) such that \( N_1 = N_2 = N \) and \( m(N_1) = m(N_2) \), we say that there are Transfers below the median individual, denoted \((N^{2-}, N^{1-}) \in T\), if and only if \((n_1^{2}, \ldots, n_{m-1}^{2}, \frac{1}{2} \sum_{c=1}^{c=n_1^2} n_c^2 - \sum_{c=1}^{c=m-1} n_c^2)\) is derived from \((n_1^{1}, \ldots, n_{m-1}^{1}, \frac{1}{2} \sum_{c=1}^{c=n_1^1} n_c^1 - \sum_{c=1}^{c=m-1} n_c^1)\) via at least one Transfer, and that there are Transfers above the median individual, \((N^{2+}, N^{1+}) \in T\), if and only if \((\sum_{c=1}^{c=n_1^2} n_c^2 - \frac{1}{2} \sum_{c=1}^{c=n_1^1} n_c^1, n_{m+1}^2, \ldots, n_c^2)\) is derived from \((\sum_{c=1}^{c=n_1^1} n_c^1 - \frac{1}{2} \sum_{c=1}^{c=n_1^1} n_c^1, n_{m+1}^1, \ldots, n_c^1)\) via at least one Transfer.

Similarly, given any two distributions \( F_1 \) and \( F_2 \) such that \( m(F_1) = m(F_2) \), we say \((F^{2-}, F^{1-}) \in T\) if and only if \((p_1^{2}, \ldots, p_{m-1}^{2}, \frac{1}{2} - \sum_{c=1}^{c=m-1} p_c)\) is derived from \((p_1^{1}, \ldots, p_{m-1}^{1}, \frac{1}{2} - \sum_{c=1}^{c=m-1} p_c)\) via at least one Transfer, and that \((F^{2+}, F^{1+}) \in T\) if and only if \((\sum_{c=1}^{c=m} p_c^2 - \frac{1}{2}, p_{m+1}^2, \ldots, p_c^2)\) is derived from \((\sum_{c=1}^{c=m} p_c^1 - \frac{1}{2}, p_{m+1}^1, \ldots, p_c^1)\) via at least one Transfer.

**Definition 5 (Increased bipolarity, IB, Wang and Tsui (2000), modified)**

For any \( N_1 \) and \( N_2 \) such that \( N_1 = N_2 \) and \( m(N_1) = m(N_2) \), \((N_2, N_1) \in IB\) if and only if \((N_2, N_1) \in T\) or \((N_2^+, N_1^-) \in T\).

For any \( F_1 \) and \( F_2 \) such that \( m(F_1) = m(F_2) \), \((F_2, F_1) \in IB\) if and only if \((F_2, F_1) \in T\) or \((F_2^+, F_1^-) \in T\).

This ordering means that transfers within the subgroup below the median individual and the subgroup above the median individual increase the polarization. Indeed, transfers imply an increase of homogeneity within the subgroups. Figure 4 displays an IB movement below the median for a hypothetical distribution: the median individual remains in ‘Fair’ category, and there is a transfer from ‘Very Poor’ (VP) to ‘Poor’ (P) and from ‘Fair’ (F) to ‘Poor’ (P). IB strict partial ordering means that this Transfer below the median increases polarization.
POLARIZATION MEASURES

Since $S$ and $IB$ are partial orderings, they do not permit to compare any pair of SAH distributions. To do so, we need polarization measures that give a complete ordering. In this section, we develop measures of polarization which increase when $S$ and IB movements are performed.

Definitions

Since for an ordinal variable, all the information relative to the distribution is given either by the distribution of cumulative numbers $N = (N_1, \ldots, N_{C-1}, N_C)$ or by that of cumulative proportions $F = (F_1, \ldots, F_{C-1}, 1)$, two indices of polarization are conceivable: one is a function of $N$, $P_1(N)$, and the other one is a function of $F$, $P_2(F)$.

Definition 6

Let $D_1$ and $D_2$ denote the sets of cumulative functions $N = (N_1, \ldots, N_C)$ and $F = (F_1, \ldots, F_{C-1}, 1)$. The polarization measures are mappings $P_i : D_i \rightarrow R_+$ ($i = 1, 2$) that satisfy the assumptions and axioms given below.

Definition 7 (Minimum polarization)

There is no polarization when everybody’s health is the same (when $\exists c \in [1, C]/n_c = N_C$ or $p_c = 1$).

For both cases, there are $C$ distributions with no polarization:

$N_{\text{min},1} = (N_C, N_C, \ldots, N_C), \quad N_{\text{min},2} = (0, N_C, N_C, \ldots, N_C), \ldots, N_{\text{min},C} = (0, \ldots, 0, N_C) \quad \text{and} \quad F_{\text{min},1} = (1, \ldots, 1), \ldots, F_{\text{min},C} = (0, \ldots, 0, 1)$.

It does not mean that these $C$ situations correspond to the same social welfare. For instance, the situation in which everybody is in category $C$ is ‘better’ than the situation in which everybody is in category $C - 1$.

Definition 8 (Maximum polarization)

Polarization is maximum when the population is divided into two halves with one half being in the lowest category and the other half being in the top category. This situation corresponds to the distributions $n_{\text{Max}} = (N_C/2, 0, \ldots, 0, N_C/2), \quad N^\text{Max} = (N_C/2, \ldots, N_C/2, N_C), \quad p^\text{Max} = (\frac{1}{2}, 0, \ldots, 0, \frac{1}{2})$ and $F^\text{Max} = (\frac{1}{2}, \ldots, \frac{1}{2}, 1)$.
We illustrate the maximum polarization with the help of Figure 5. There are five categories and one half of the population is in the ‘Very Poor’ (VP) category whereas the other half is in the ‘Excellent’ (E) category. Hence, polarization is maximum.

We can note that the maximum polarization for an ordinal variable differs from maximal dispersion for a nominal variable. Nominal variables are variables with unordered categories. For them, dispersion is maximum when there is the same sample proportion in each category (i.e. when the probability distribution is uniform).

The families of polarization measures

In this subsection, we provide an axiomatic foundation of two families of polarization measures $P_1(N)$ and $P_2(F)$.

We first follow suggestions of Blair and Lacy (2000) to consider the observed distributions $N = (N_1, \ldots, N_C)$ or $F = (F_1, \ldots, F_{C-1}, 1)$ as ordered sets that represent points in $C$-dimensional spaces. We assume like them that polarization measures the distance between the observed distribution $N = (N_1, \ldots, N_C)$, or $F = (F_1, \ldots, F_{C-1}, 1)$, and the distribution of maximum polarization, in which half of the population is in the lowest category and the other half is in the highest category, $N^{\text{Max}} = (N_C/2, \ldots, N_C/2, N_C)$ or $F^{\text{Max}} = (\frac{1}{2}, \ldots, \frac{1}{2}, 1)$.

This implies that $P_1$ is a function of $(|N_1 - N_C/2|, \ldots, |N_{C-1} - N_C/2|, |N_C - N_C|) = (|N_1 - N_C/2|, \ldots, |N_{C-1} - N_C/2|, 0)$; since the last difference is a constant, we can omit it, so that $P_1$ is a function of $(|N_1 - N_C/2|, \ldots, |N_{C-1} - N_C/2|)$. It also implies that $P_2$ is a function of $(|F_1 - \frac{1}{2}|, \ldots, |F_{C-1} - \frac{1}{2}|)$ for the same reasons.

We assume that total polarization is equal to the sum of the contributions of each category, and that the contribution of one category $c$ is a function of $|N_c - N_C/2|$ or $|F_c - \frac{1}{2}|$: $g_1(|N_c - N_C/2|)$ or $g_2(|F_c - \frac{1}{2}|)$.

Thus, we assume that there exist two continuous functions $g_1$ and $g_2$ such that

$$P_1(N) = \sum_{c=1}^{C-1} g_1 \left( \left| N_c - \frac{N_C}{2} \right| \right)$$

$$P_2(F) = \sum_{c=1}^{C-1} g_2 \left( \left| F_c - \frac{1}{2} \right| \right)$$

---

2But this is not equivalent. 
To highlight the meaning of the assumption, we develop an example. Suppose that the distribution is 
p_1 = (0.2, 0.2, 0.2, 0.1, 0.3). If 10% of individuals move from category 3 to category 4, the distribution becomes 
p_2 = (0.2, 0.2, 0.1, 0.2, 0.3) and the contribution of category 3 to polarization changes, but the contributions of the other categories are not affected by this movement. Thus, the assumption implies that the contributions of the categories are independent.

This assumption also implies that \( P_1(N) \) and \( P_2(F) \) have two interesting properties (slide invariance and symmetry), which are presented with other properties below, and which are important for polarization measures for the ordinal data.

It is clear that these functional forms are equivalent to the existence of two continuous functions \( \phi \) and \( \psi \) such that

\[
P_1(N) = M_1 - \frac{1}{C-1} \sum_{c=1}^{C-1} \phi \left( \left| N_c - \frac{N_C}{2} \right| \right)
\]

with

\[
M_1 = \begin{cases} 
\max_N \frac{1}{C-1} \sum_{c=1}^{C-1} \phi \left( \left| N_c - \frac{N_C}{2} \right| \right) \\
0 \leq N_1 \leq \cdots \leq N_C
\end{cases}
\]

\[
P_2(F) = M_2 - \frac{1}{C-1} \sum_{c=1}^{C-1} \psi \left( \left| F_c - \frac{1}{2} \right| \right)
\]

with

\[
M_2 = \begin{cases} 
\max_F \frac{1}{C-1} \sum_{c=1}^{C-1} \psi \left( \left| F_c - \frac{1}{2} \right| \right) \\
0 \leq F_1 \leq \cdots \leq 1
\end{cases}
\]

The following axiom indicates that the two polarization measures satisfy the first strict partial ordering (S).

\textit{Axiom (Spread away from the median, S)}

\( P_1 \) and \( P_2 \) are increasing in spreads away from the median.

\textit{Proposition 1}

\( P_1(N) \) and \( P_2(F) \) satisfy S if and only if \( \phi(u) \) and \( \psi(u) \) are strictly increasing in \( R_+ \).

Since \( \phi \) and \( \psi \) are increasing,

\[
M_1 = \frac{1}{C-1} \sum_{c=1}^{C-1} \phi \left( \frac{N_C}{2} \right) = \phi \left( \frac{N_C}{2} \right) \quad \text{and} \quad M_2 = \psi \left( \frac{1}{2} \right)
\]

The next axiom is a technical axiom which is needed in what follows. It states that when SAH distribution exhibits maximum polarization, \( N^{\text{Max}} \) and \( F^{\text{Max}} \), then \( P_1 \) and \( P_2 \) equal \( M_1 \) and \( M_2 \).

\textit{Axiom (Normalisation, NM)}

\( P_1(N^{\text{Max}}) = P_1(N_{C/2}, \ldots, N_{C/2}, N_C) = M_1 = \phi(N_C/2) \) and \( P_2(F^{\text{Max}}) = P_2(\frac{1}{2}, \ldots, \frac{1}{2}, 1) = M_2 = \psi(\frac{1}{2}) \).

\textit{Proposition 2}

\( P_1(N) \) and \( P_2(F) \) satisfy NM if and only if \( \phi(0) = 0 \) and \( \psi(0) = 0 \).
The next axiom indicates that the two polarization measures induce the same ranking of SAH distributions.

Axiom (Compatibility, COMP (Foster and Shorrocks, 1991; Wang and Tsui, 2000))

The polarization indices \( P_2 \) and \( P_1 \) are compatible if for all \( NA = (N_1, \ldots, N_C) \) and \( NB = (N_1', \ldots, N_C') \) such that \( N_{1C} = N_{1C}' = NC \), \( P_1(NA) \geq P_1(NB) \) if and only if \( P_2(N_1/N_C, \ldots, N_{C-1}/N_C, 1) \geq P_2(N_1'/N_C, \ldots, N_{C-1}'/N_C, 1) \).

Proposition 3

The indices satisfy SYM, \( S \), NM and COMP if and only if

\[
P_1(N) = K_1 \left[ \frac{(NC)}{2} - \frac{1}{C - 1} \sum_{c=1}^{C-1} \left| NC \left( \frac{N_c}{NC} - \frac{1}{2} \right) \right|^2 \right] = K_1 N_C^2 \left[ \frac{(1/2)}{2} - \frac{1}{C - 1} \sum_{c=1}^{C-1} \left| N_c - \frac{1}{2} \right|^2 \right]
\]

\[
P_2(F) = K_2 \left[ \frac{(1/2)}{2} - \frac{1}{C - 1} \sum_{c=1}^{C-1} \left| F_c - \frac{1}{2} \right|^2 \right]
\]

where \( K_1 \) and \( K_2 \) are strictly positive constants and \( \alpha > 0 \).

The following axiom states that the two polarization measures satisfy the second strict partial ordering (IB).

Axiom (IB)

\( P_1 \) and \( P_2 \) are strictly increasing in IB.

This axiom enables us to find the bounds of the parameter \( \alpha \).

Proposition 4

IB implies that \( \alpha \in [0, 1[ \).

Propositions 3 and 4 contain the main result of the paper, that is to say the two polarization measures \( P_1 \) and \( P_2 \) and the bounds of \( \alpha \).

The rest of the section develops the interpretation of \( \alpha \) and the properties of the measures, which are identical for \( P_1 \) and \( P_2 \). We thus choose to focus on \( P_2 \).

We now turn to the interpretation of \( \alpha \). The contribution of category \( c \in [1, C - 1] \) to polarization is given by

\[
g_2 \left( \left| F_c - \frac{1}{2} \right| \right) = K_2 \left[ \frac{1}{2^{\alpha(C - 1)}} - \frac{1}{C - 1} \left| F_c - \frac{1}{2} \right|^2 \right]
\]

Hence, the relative contribution of category \( c \) to polarization is

\[
\frac{g_2(|F_c - \frac{1}{2}|)}{P_2(F)} = \frac{1 - |2F_c - 1|^{2\alpha}}{C - 1 - \sum_{c=1}^{C-1} |2F_c - 1|^2}
\]
Thus,

- If \( F = F_{\text{Max}} \) or \( F = F_{\text{min,c}} \) \((c = 1, \ldots, C)\), the choice of \( \alpha \) does not have any effect on the contributions of the categories.
- In all other cases, when \( \alpha \to 0 \), the relative contribution of the median category increases whereas the relative contributions of the other categories decrease. When \( \alpha \to 1 \), the relative contribution of the median category decreases whereas the relative contributions of the other categories increase.

Thus, \( \alpha \) reflects the importance that is given to the median category.

Since \( K_2 \) has no bearing on the ranking, we impose \( K_2 = 2^\alpha \) in the rest of the paper, so that \( P_2(F) \in [0, 1] \) with \( P_2(F_{\text{min,c}}) = 0 \) and \( P_2(F_{\text{Max}}) = 1 \):

\[
P_2(F) = 1 - \frac{2^\alpha}{C-1} \sum_{c=1}^{C-1} \left| F_c - \frac{1}{2} \right|^\alpha
\]

We propose a calibration of \( \alpha \). We believe that when there is the same number of individuals in the \( C \) categories, then polarization is medium. Thus, the uniform distribution \( p = p_{\text{U}} = (1/C, \ldots, 1/C) \) could represent an intermediate polarization between minimum polarization \( p_{\text{min,c}} \) and maximum polarization \( p_{\text{Max}} \). We thus propose to set \( \alpha \) such that \( P_2(F_{\text{U}}) = \frac{1}{2} \). Table I contains the calibrated \( \alpha^* \) (the mathematical expression when possible, or an approximate value) for the different values of \( C \).

**Some properties of \( P_2 \)**

The indices have important features. Since the properties of \( P_1 \) are the same as those of \( P_2 \), we only present them for \( P_2 \).

**Property 1 (Continuity)**

The index is continuous: small changes in the distribution \( F \) do not induce large changes in \( P_2(F) \).

**Property 2 (Slide)**

**Definition 9 (Slide)**

We say that \( F^2 \) is obtained from \( F^1 \) through a slide to the right if

\[
\begin{align*}
F^1_{C-1} &= 1 \\
F^1_1 &= 0 \\
F^2_c &= F^1_{c-1} \quad \text{if } c \in [2, C]
\end{align*}
\]

<table>
<thead>
<tr>
<th>( C )</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha^* )</td>
<td>( \frac{\ln 2}{\ln 3} )</td>
<td>( \frac{\ln 4 - \ln 3}{\ln 2} )</td>
<td>( \approx 0.73 )</td>
<td>( \approx 0.66 )</td>
<td>( \approx 0.78 )</td>
<td>( \approx 0.75 )</td>
<td>( \approx 0.82 )</td>
<td>( \approx 0.81 )</td>
</tr>
</tbody>
</table>

Copyright © 2007 John Wiley & Sons, Ltd.  
DOI: 10.1002/hec
Property 2 is slide invariant, which means that if a distribution $F^1$ is obtained from a distribution $F^2$ through a slide, then $P_2(F^1) = P_2(F^2)$.

Table II presents a slide to the right for a hypothetical distribution. Since $P_2$ is slide invariant, polarization remains the same: $P_2(F^1) = P_2(F^2)$. This property underlines that polarization only takes into account the order of the proportions in the categories, that is to say the differences between individuals, and not welfare levels.

Property 3 (Symmetry)

Definition 10

A polarization measure $P(F)$ is symmetric when for all $F$, $P(F_1, F_2, \ldots, F_{C-1}, 1) = P(1 - F_{C-1}, 1 - F_{C-2}, \ldots, 1 - F_1, 1)$.

In other words, a polarization measure is symmetric if it does not change when the categories are in the reverse order, for instance, if instead of ‘Very Poor’ in category 1, ‘Excellent’ in category 5, we set ‘Excellent’ in category 1, ‘Very poor’ in category 5.

Property 4 (Distance)

In properties 4 and 5, we denote $\forall q \in [0, 1], \forall (c_1, c_2) \in [1, C]^2$ with $c_1 < c_2$,

$$F(c_1, c_2, q) = (F_1, \ldots, F_C)$$

with

- $\forall c \in [1, c_1[, \quad F_c = 0$
- $\forall c \in [c_1, c_2[, \quad F_c = q$
- $\forall c \in [c_2, C[, \quad F_c = 1$

Table II. Slide

<table>
<thead>
<tr>
<th></th>
<th>Very poor</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^1$</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>$F^1$</td>
<td>0.1</td>
<td>0.4</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p^2$</td>
<td>0.0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>$F^2$</td>
<td>0.0</td>
<td>0.1</td>
<td>0.4</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table III. Symmetry

<table>
<thead>
<tr>
<th></th>
<th>Very poor</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^1$</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>$F^1$</td>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>$p^2$</td>
<td>0.4</td>
<td>0.0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>$F^2$</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Let $F(c_1, c_2, q)$ and $F(c'_1, c'_2, q)$ be two distributions

$$
\begin{align*}
\begin{cases}
  c'_1 \geq c_1 & \Rightarrow P_2(F(c'_1, c'_2, q)) \leq P_2(F(c_1, c_2, q)) \\
  c'_2 \leq c_2
\end{cases}
\end{align*}
$$

This property is inspired by an axiom for income polarization (Esteban and Ray, 1994; Milanovic, 2000). When the population is divided into two clusters, if we shorten the distance between the clusters then polarization is reduced. Table IV illustrates this case for a hypothetical distribution.

**Property 5 (Frequency)**

$P_2$ also satisfies a frequency property which is inspired by an axiom for income polarization (Milanovic, 2000):

$$
\forall q \in [0, 1], P_2(F(c_1, c_2, q)) \leq P_2(F(c_1, c_2, \frac{1}{2}))
$$

When the population is divided into two groups, then $P_2$ decreases as the distribution of individuals between the two peaks diverges from half and half. Table V is designed to illustrate this idea with a hypothetical distribution: $P_2(F^2) \leq P_2(F^1)$.

**Statistical inference**

We use a formula given by Agresti (1990) which is presented in Blair and Lacy (2000) for statistical inference. Using the multivariate delta method, we find that the sample estimator for the variance of $P_2$, which will be used to compute confidence intervals for $P_2$, is:

$$
\hat{s}^2[P_2] = \frac{x^2 a^2 [\sum_{c=1}^{C-1} p_c a^2_c - (\sum_{c=1}^{C-1} p_c a_c)^2]}{N(C-1)^2}
$$

---

3Details are available on request. The formula cannot be used if in the sample distribution the cumulative function for the median category is exactly equal to one half ($F_m = \frac{1}{2}$). Inference in this particular case has not been developed yet.
with

\[ a_c = - \mathbf{1}_{(c \leq m-1)} \sum_{i=c}^{M} \left( \frac{1}{2} - F_i \right)^{z-1} + \mathbf{1}_{(c \leq m-1)} \mathbf{1}_{(m \leq C-1)} \sum_{i=m}^{C-1} \left( F_i - \frac{1}{2} \right)^{z-1} \]

\[ + \mathbf{1}_{(c > m-1)} \sum_{i=m}^{C-1} \left( F_i - \frac{1}{2} \right)^{z-1} \]

where

\[ \mathbf{1}_{(x \leq y)} = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{if } x > y \end{cases} \quad \text{and} \quad \mathbf{1}_{(x > y)} = \begin{cases} 1 & \text{if } x > y \\ 0 & \text{if } x \leq y \end{cases} \]

with \( N \) the sample size, \( C \) the number of categories, \( m \) the median category, \( p = (p_1, \ldots, p_C) \) and \( F = (F_1, \ldots, F_C) \) the proportions and cumulative proportions in the sample.

**EMPIRICAL ILLUSTRATION**

In order to illustrate the usefulness of the measure of polarization, we apply it to the BHPS. By doing so, we highlight the empirical relevance of the concept of SAH polarization. This application also demonstrates that inequality and polarization are empirically different.

The data are taken from the BHPS which is an annual and longitudinal survey of private households in Great Britain. It contains SAH and other self-reported measures of health. The wave 1 data were collected in 1990–1991, and the last wave data (wave 14) were collected in 2003–2004. We do not use wave 1 and wave 9 because they do not have data on specific health problems, which are important for the cardinalization of SAH, or on the SAH we use.⁴ In a first approach, we have decided to study an unbalanced sample of men who were in the original sample in 1991 and that we follow over time, and to ignore the problem of attrition. This has two advantages: first, since we follow individuals, the results can be compared for different years (which is not the case in a cross-sectional approach where new individuals are added every year); and second, the sample is larger than a balanced panel of individuals who are observed in all the waves (longitudinal approach).⁵ The sample size decreases over time because of attrition (3836 individuals in 1992 and 2341 in 2004), and the working sample contains 36,095 observations.

The key variable, SAH, represents health status over the last 12 months, and it is recorded in five categories: ‘Very poor’ (VP), ‘Poor’ (P), ‘Fair’ (F), ‘Good’ (G) and ‘Excellent’ (E).

In this section, we contrast the Gini coefficient as an example of an inequality measure with the polarization measure \( P_2 \).

To measure polarization, we simply use the given SAH distributions and compute \( P_2 \) for different values of \( z \) (\( z = 0.1, 0.5, 0.9 \), and the calibrated value \( z^* = 0.73 \)).

In contrast, to measure inequality, we also use SAH but we transform it into a cardinal variable, and then calculate the Gini index. Since the BHPS does not contain an empirical distribution function as HUI, we construct a cardinal health variable using an ordered probit approach. We estimate SAH as a function of specific health problems.⁶ These objective variables are binary variables, each indicating the presence of one of the following health problems: arms, legs or hands, sight, hearing, skin conditions or allergies, chest/breathing, heart/blood pressure, stomach or digestion, diabetes, anxiety or depression,

---

⁴The wording of the SAH question is different in wave 9 from that of the other waves.

⁵The results for the cross-sectional and longitudinal approaches, and for women, are available on request.

⁶We could have also estimated SAH as a function of health problems and socio-economic characteristics. As this is a first approach, we only use health problems.
alcohol or drugs, epilepsy or migraine, or ‘other’. Denoting SAH the ordinal variable, $y^*$ the latent health variable underlying SAH, and $c$ a SAH category, we estimate the following model for each year:

$$\text{SAH}_t = c \text{ if } \mu_{c-1} < y^*_t \leq \mu_c, \quad c = 1, 2, \ldots, 5$$

with

$$y^*_t = x_t \alpha + \epsilon_t$$

with $\epsilon_t \sim N(0, 1), \mu_0 = -\infty, \mu_c \leq \mu_{c+1}$ and $\mu_5 = \infty$. $x_t$ includes objective health problem variables.\(^7\)

We then compute the linear predicted values of health. In order to avoid negative values of health, the predictions are rescaled to the [0,1] interval using the same method as Van Doorslaer and Jones (2003). Denoting $y^1_t$ the predicted linear value, and $y^{\text{max}}_t$ and $y^{\text{min}}_t$ the largest and the smallest individual predictions, the rescaled variable is $y^2_t = (y^1_t - y^{\text{min}}_t)/(y^{\text{max}}_t - y^{\text{min}}_t)$. This cardinalization and rescaling enable us to compute the Gini index.

Table VI presents SAH distributions, that is to say the probability distribution function $p^{\text{year}}$ and the cumulative distribution function $F^{\text{year}}$, the four polarization indices, the inequality measure described above, the rankings of the years for the indices (the larger the rank, the higher polarization or inequality measures are) and the 95% confidence interval for the indices.\(^8\)

(1) In a first step, we highlight that data show polarization phenomena. Since the two strict partial orderings (spreads away from the median and IB) correspond to the intuitive idea of polarization, we focus on SAH distributions whose level of polarization can be compared using strict partial orderings only. Polarization measures satisfy these strict partial orderings for all $x$, it is thus logical that for all $x$ polarization measures increase when the strict partial orderings are observed (but the increase is not necessarily significant). We also show that the strict partial orderings do not always increase the inequality measure, and thus inequality measures are not adequate to deal with polarization phenomena.

One aspect of polarization is mass relocation from the middle of the distribution to the poles, i.e. spreads away from the median ($S$). Table VI contains several $S$ movements: the distribution of 1997 is a $S$ of those of 1994, 1995 and 1996; the distributions of 2000 and 2002 are $S$ of that of 2003; the distribution of 2001 is a $S$ of that of 1995, and the distribution of 2002 is a $S$ of that of 2004. Thus, polarization phenomena seem to be frequent in data.

More precisely, the distribution of 2001 is a $S$ of that of 1995 because the median category is category 4 (‘Good’) in 1995 and 2001, and $F^c_{\text{new}} < F^c_\text{old}$ for $c < 4$, and $F^c_{\text{new}} > F^c_\text{old}$ for $c \geq 4$. The axiomatic foundation of our polarization measures says that spreads away from the median increase polarization measures, and here, the four polarization indices are significantly larger in 2001 than in 1995. In contrast, the Gini coefficient is smaller in 2001 than in 1995 (the decline is not significant), which underlines that the relocation from the middle to the pole is not handled by the Gini coefficient.

Another aspect of polarization is the emergence of poles through transfers within at least one half of the distribution, i.e. IB movements. IB and $S$ are sometimes combined: for instance, the distribution of 1998 can be obtained from the distribution of 1995 through $S$ and IB movements. Consider the distribution $p^{\text{new}} = (0.0148, 0.0530, 0.1954, 0.4954, 0.2414)$; an IB below the median (with $\epsilon_1 = 1, \ \epsilon_2 = 4$ and $\delta = 0.0008$) yields $p^{\text{new}} = (0.0140, 0.0538, 0.1962, 0.4946, 0.2414)$ and then the distribution of 1998 is a $S$ of the distribution ‘new’. Thus, polarization measures increase, they are significantly larger in 1998 than in 1995, whereas inequality decreases (but the decrease is not significant).

---

\(^7\)The results of the 12 regressions are available on request and omitted for space reasons.

\(^8\)The confidence intervals for the polarization measures are obtained using statistical inference developed in the fourth section. The Gini coefficient and its confidence intervals are obtained using the INEQERR routine in STATA, this procedure requires 1000 bootstrap repetitions.
Table VI. Polarization and inequality measures, BHPS 1992–2004

<table>
<thead>
<tr>
<th>Year</th>
<th>VP</th>
<th>P</th>
<th>F</th>
<th>G (m)</th>
<th>E</th>
<th>0.1 Index (Rkg)</th>
<th>0.5 Index (Rkg)</th>
<th>(\alpha = 0.73) Index (Rkg)</th>
<th>0.9 Index (Rkg)</th>
<th>Gini Index (Rkg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>0.0148</td>
<td>0.0513</td>
<td>0.1779</td>
<td>0.4703</td>
<td>0.2856</td>
<td>0.0408 (10)</td>
<td>0.1783 (9)</td>
<td>0.2420 (9)</td>
<td>0.2831 (9)</td>
<td>0.0667 (1)</td>
</tr>
<tr>
<td>1993</td>
<td>0.0115</td>
<td>0.0598</td>
<td>0.2520</td>
<td>0.7260</td>
<td>1</td>
<td>0.0398 (7)</td>
<td>0.1743 (7)</td>
<td>0.2367 (6)</td>
<td>0.2770 (6)</td>
<td>0.0662 (2)</td>
</tr>
<tr>
<td>1994</td>
<td>0.0101</td>
<td>0.0612</td>
<td>0.1825</td>
<td>0.4833</td>
<td>0.2629</td>
<td>0.0394 (5)</td>
<td>0.1735 (5)</td>
<td>0.2362 (5)</td>
<td>0.2770 (5)</td>
<td>0.0774 (5)</td>
</tr>
<tr>
<td>1995</td>
<td>0.0148</td>
<td>0.0530</td>
<td>0.1954</td>
<td>0.4954</td>
<td>0.2414</td>
<td>0.0383 (1)</td>
<td>0.1694 (1)</td>
<td>0.2312 (1)</td>
<td>0.2716 (2)</td>
<td>0.0792 (8)</td>
</tr>
<tr>
<td>1996</td>
<td>0.0117</td>
<td>0.0559</td>
<td>0.1888</td>
<td>0.4851</td>
<td>0.2585</td>
<td>0.0391 (4)</td>
<td>0.1722 (4)</td>
<td>0.2546 (4)</td>
<td>0.2752 (4)</td>
<td>0.0786 (6)</td>
</tr>
<tr>
<td>1997</td>
<td>0.0194</td>
<td>0.0523</td>
<td>0.1987</td>
<td>0.4638</td>
<td>0.2659</td>
<td>0.0418 (12)</td>
<td>0.1830 (12)</td>
<td>0.2474 (12)</td>
<td>0.2908 (12)</td>
<td>0.0785 (4)</td>
</tr>
<tr>
<td>1998</td>
<td>0.0140</td>
<td>0.0616</td>
<td>0.1961</td>
<td>0.4744</td>
<td>0.2538</td>
<td>0.0407 (9)</td>
<td>0.1789 (10)</td>
<td>0.2433 (10)</td>
<td>0.2851 (10)</td>
<td>0.0784 (7)</td>
</tr>
<tr>
<td>2000</td>
<td>0.0161</td>
<td>0.0556</td>
<td>0.2067</td>
<td>0.4880</td>
<td>0.2335</td>
<td>0.0395 (6)</td>
<td>0.1738 (6)</td>
<td>0.2367 (7)</td>
<td>0.2777 (7)</td>
<td>0.0891 (10)</td>
</tr>
<tr>
<td>2001</td>
<td>0.0155</td>
<td>0.0587</td>
<td>0.1983</td>
<td>0.4809</td>
<td>0.2472</td>
<td>0.0401 (8)</td>
<td>0.1769 (8)</td>
<td>0.2400 (8)</td>
<td>0.2815 (8)</td>
<td>0.0756 (3)</td>
</tr>
<tr>
<td>2002</td>
<td>0.0138</td>
<td>0.0615</td>
<td>0.2056</td>
<td>0.4715</td>
<td>0.2475</td>
<td>0.0411 (11)</td>
<td>0.1799 (11)</td>
<td>0.2445 (11)</td>
<td>0.2862 (11)</td>
<td>0.0946 (12)</td>
</tr>
<tr>
<td>2003</td>
<td>0.0138</td>
<td>0.0653</td>
<td>0.2098</td>
<td>0.4936</td>
<td>0.2280</td>
<td>0.0387 (3)</td>
<td>0.1706 (3)</td>
<td>0.2325 (3)</td>
<td>0.2728 (3)</td>
<td>0.0918 (11)</td>
</tr>
<tr>
<td>2004</td>
<td>0.0098</td>
<td>0.0620</td>
<td>0.2022</td>
<td>0.4951</td>
<td>0.2308</td>
<td>0.0384 (2)</td>
<td>0.1696 (2)</td>
<td>0.2313 (2)</td>
<td>0.2715 (1)</td>
<td>0.0886 (9)</td>
</tr>
</tbody>
</table>

Note: 95% confidence intervals for the indices are displayed in parentheses below the indices. Rankings for the indices are displayed in parentheses next to the indices.
Second, we look at cases where polarization measures are necessary to compare SAH distributions because partial orderings are not sufficient. Then the ranking of the distributions may depend on the value of the parameter \( \alpha \). So we underline that the four polarization measures induce rankings of the years (according to the level of polarization) that are not very different, thus the choice of \( \alpha \) does not seem to be very important, at least for these data.

Again we observe that polarization and inequality measures do not behave similarly: for example, compared to the distribution of 1992, the distributions of 2003 and 2004 have significantly smaller polarization and significantly larger inequality.

There are also cases when polarization and inequality move in the same direction, but the examples have highlighted that it is not always true. Since polarization measures and the Gini coefficient do not behave similarly in data, the difference between inequality and polarization is not only a theoretical question.

CONCLUSION

SAH is one of the most frequently used health measures. Thus, it is important to be able to measure dispersion in SAH. By definition, inequality measures are mean based, and since only cardinal variables have a mean, SAH has to be cardinalized to compute an inequality index. But cardinalization is a supra-ordinal assumption.

It is possible that this supra-ordinal assumption is not necessary in the case of measures of polarization. We presented two alternative (polarization) indices \( P_1 \) and \( P_2 \), which do not rely on this assumption. They are median based and satisfy important axioms which are identical in spirit to properties of polarization measures for cardinal variables. Using BHPS data, we then highlighted that the Gini coefficient and our polarization measures are substantially different.

The inequality and polarization approaches are complementary in analysing dispersion in SAH for several reasons. First, the polarization measures we propose have drawbacks and advantages different from those of inequality measures. In fact the polarization indices for ordinal data avoid the difficulties of inequality measures (related to cardinalization), while they do not allow to take reporting bias into account, as certain cardinalization methods and inequality measures do. Above all, polarization and inequality are conceptually different, which is reflected in data. Consequently, a complete analysis of the empirical phenomena could comprehend both approaches.

While the polarization measures derived in this paper represent a first step to deal with the problem of cardinalization in measuring health dispersion, the strong assumption concerning the functional form of our measures could be an avenue for future research.

Finally, the family of polarization measures can be applied to characterize many self-assessed measures, for example, happiness, well-being and political opinions.

ACKNOWLEDGEMENTS

Data from the British Household Panel Survey (BHPS) were supplied by the ESRC Data Archive. Neither the original collectors of the data nor the Archive bear any responsibility for the analysis or interpretations presented here. I am grateful to two anonymous referees, Ramses H. Abul Naga, Lionel Da Cruz, my supervisor Pierre-Yves Geoffard, Mark Hanly, Andrew M. Jones, Sebastian Linnemayr, Owen O’Donnell, Pedro Rosa Dias, Amedeo Spadaro, Eddy Van Doorslaer, and participants to the Fifteenth European Workshop on Econometrics and Health Economics for helpful comments.
APPENDIX

The proofs for $P_1(N)$ and $P_2(F)$ are the same and we only present them for $P_1(N)$.

Proof (Proposition 1)

Step 1: ⇒ For $m \in [1, C - 1]$ and for any $a \geq 0, b > 0$ such that $0 \leq a + b \leq N_C/2$, there exist $N^1$ and $N^2$ such that

$$m(N^1) = m(N^2) = m$$

$$N_i^1 = N_i^2 \text{ for any } i = 1, \ldots, C - 2, C$$

$$N_{C-1}^1 = N_{C-1}^2 = \frac{N_C}{2} + a$$

and

$$N_{C-1}^2 = \frac{N_C}{2} + a + b$$

Then polarization for $N^1$ is greater than polarization for $N^2$:

$$P_1(N^1) > P_1(N^2) \iff \phi(a + b) > \phi(a)$$

Then $\phi$ is strictly increasing on $R_+$. The case where $m = C$ is left to the reader. Thus, $P_1$ satisfies the $S$ axiom only if $\phi$ is a strictly increasing function on $R_+$.

Step 2: ⇐ Obvious.

Proof (Proposition 2)

$$P_1(N_C/2, \ldots, N_C/2, N_C) = \phi(N_C/2) - \phi(0) = \phi(N_C/2) \iff \phi(0) = 0.$$
In the case where \( C = 3 \) we get (*):
\[
\frac{1}{2} \sum_{c=1}^{2} \phi \left( N_3 \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right) = H(3) \left[ \frac{1}{2} \sum_{c=1}^{2} \psi \left( \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right), N_3 \right]
\]
for all \( N_1, N_2, N_3 \) such that \( 0 \leq N_1 \leq N_2 \leq N_3 \) for a function \( H(3)[u, N_3] \) continuous and increasing in \( u \).

If \( N_1 = N_2 = N_x \) it follows that
\[
\phi \left( N_3 \left| \frac{N_x}{N_3} - \frac{1}{2} \right| \right) = H(3) \left[ \psi \left( \left| \frac{N_x}{N_3} - \frac{1}{2} \right| \right), N_3 \right] \quad \text{for all } N_x, N_3 \text{ such that } 0 \leq N_x \leq N_3
\]

Thus, substitution into (*) gives:
\[
\frac{1}{2} \sum_{c=1}^{2} H(3) \left[ \psi \left( \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right), N_3 \right] = H(3) \left[ \frac{1}{2} \sum_{c=1}^{2} \psi \left( \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right), N_3 \right]
\]
The solution of this Jensen equation is \( H(3)(u, N_3) = g(N_3)u + h(N_3) \). Thanks to NM and S, \( H(3)[0, N_3] = 0, \) thus \( h(N_3) = 0 \).

\( H(3)(u, N_3) = g(N_3)u \) is equivalent to
\[
\phi \left( N_3 \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right) = g(N_3)\psi \left( \left| \frac{N_c}{N_3} - \frac{1}{2} \right| \right) \quad \text{for all } N_c, N_3 \text{ such that } 0 \leq N_c \leq N_3
\]
and consequently
\[
\phi(u) = K_1u^\alpha \quad \text{and} \quad \psi(u) = K_2u^\beta
\]
with \( K_1 \) and \( K_2 \) are constants such that \( K_1 > 0, K_2 > 0. \) \( \phi \) strictly increasing implies that \( \alpha > 0. \)

**Step 2:** Obvious.

**Proof (Proposition 4)**

We show that IB implies that \( \phi \) is strictly concave on \( R_+ \). Suppose that \( P_2(N) \) satisfies IB. For any \((a, b, e, \Delta)\) such that \( a, b, e \geq 0, a \geq \Delta \geq 0, a + b + e + \Delta \leq N_C/2 \) and \( b + e > 0 \), there exist \( N^1 \) and \( N^2 \) such that
\[
m(N^1) = m(N^2) \in [1, C - 2]
\]
\[
N^1_c = N^2_c \quad \text{for } c \in [1, C - 3] \quad \text{if } C \geq 4
\]
\[
N^1_C = N^2_C = N_C \quad \text{(with } C \geq 3)\]
\[
N^1_{C-2} = \frac{N_C}{2} + a
\]
\[
N^1_{C-1} = \frac{N_C}{2} + a + b
\]
We perform an IB movement above the median and we set \( c_1 = C - 2, c_2 = C: \)
\[
N^2_{C-2} = \frac{N_C}{2} + a - \Delta
\]
\[
N^2_{C-1} = \frac{N_C}{2} + a + b + e + \Delta
\]
Thus $P(N^2) > P(N^1)$ that is to say that
\[
\phi(a - \Delta) + \phi(a + b + e + \Delta) < \phi(a) + \phi(a + b + e)
\]
\[
\iff \phi(a + b + e + \Delta) - \phi(a + b + e) < \phi(a) - \phi(a - \Delta)
\]
Then \( \phi \) is strictly concave on \( R_+ \), thus \( \alpha \in ]0, 1[ \).

REFERENCES