The measurement of social polarization in a multi-group context.

Iñaki Permanyer*
Institut d'Anàlsi Econòmica (CSIC), Barcelona.

April 7, 2008

Abstract

Polarization indices presented up to now have only focused their attention on the distribution of income/wealth. However, in many circumstances income is not the only relevant dimension that might be the cause of social conflict, so it is very important to have a social polarization index able to cope with alternative dimensions. In this paper we present an axiomatic characterization of one of such indices: it has been obtained as an extension of the (income) polarization measure introduced in Duclos, Esteban and Ray (2004) to a wider domain. It turns out that the axiomatic structure introduced in that paper alone is not appropriate to obtain a fully satisfactory characterization of our measure, so additional axioms are proposed. As a byproduct, we present an alternative axiomatization of the aforementioned income polarization measure.

*Address: Institut d'Anàlisi Econòmica (CSIC), Campus de la Universitat Autònoma de Barcelona. 08193 Barcelona (Spain). E-mail: *inaki.permanyer@uab.es*; Phone: +34 935 806 612

1 Introduction

The past few years have witnessed an increasing attention towards the measurement of polarization[1]. The main reason behind this interest is the existing connection between polarization and several social, economic or political phenomena, specially those related to different kinds of social conflict, which can not be properly captured by the classical measures of inequality. Unfortunately, most polarization indices have focused their attention on the measurement of "income polarization" alone, and only a few of them have attempted to focus on what might be broadly referred as "social polarization"[2]. The latter term might be coined when the factors that determine individual's identity are socially driven and do not depend solely on their income. The two classical examples which have been used in related papers (see previous footnote) are ethnic and religious polarization. There is great interest in defining social polarization indices because, in many circumstances, income is not the (only) relevant dimension that might be the cause of social conflict[3]. The limitations of the traditional income polarization measures become also evident when, in particular, one explores the theoretical foundations of the links between polarization and conflict (see, for instance, Esteban and Ray (1999)).

In the paper of Duclos, Esteban and Ray (2004) −DER from now on− there are some attempts to produce such "social polarization" measures which, unfortunately, lack of any axiomatic characterization. As those authors mention, the proposed measures are liberal transplants of their findings in the measurement of income polarization to the different contexts of "Pure Social Polarization" (see equations (14) and (15)) and "Hybrid" polarization measures (see equations (16) and (17)) that take into account the social group

[1]See Esteban and Ray (1994), Duclos, Esteban and Ray (2004), Foster and Wolfson (1992), Wolfson (1997), Alesina and Spolaore (1997), Quah (1997), Wang and Tsui (2000), Esteban, Gradín and Ray (2007), Chakravary and Majumder (2001), Zhang and Kanbur (2001) and Rodriguez and Salas (2002).

[2]See, for example, Duclos, Esteban and Ray (2004) and Garcia-Montalvo and Reynal-Querol (2005).

[3]See Esteban and Ray (1999), Garcia-Montalvo and Reynal-Querol (2005), Collier and Hoeffler (1998) or Easterly and Levine (1997) for some empirical or theoretical works that explore the existing links between polarization and conflict and other related issues.

to which individuals belong *and* their corresponding income levels. However, lacking any kind of characterization result, these indices are open to legitimate criticism on grounds of arbitrariness, so −the authors acknowledge− such characterization result is 'an important subject of future research'. One of the main purposes of this paper is to bridge this gap by extending the ideas presented in DER to a wider domain. We will assume that a population is partitioned into disjoint groups according to exogenously given criteria (*e.g.* religious or ethnic group) and that each individual has a given sense of identification with individuals which are "similar" to him. Our basic assumption is that individuals of a given group tend to feel less alienated with the members of the same group than with the members of other groups. We will show how, using a modified version of the *identification-alienation* framework in our multi-group context, one can obtain an axiomatically characterized natural extension of the income polarization index presented in DER.

The process of characterizing a social polarization index using extensions of the axioms presented in DER to the multi-group case is not straightforward. The axioms 2, 3 and 4 used in DER admit generalizations to a multi-group context, but axiom 1 is intrinsically single dimensional in a deep sense we will later specify, so a completely different axiom must be introduced in its place. Interestingly, it turns out that the single group version of this new axiom can substitute axiom 1 in DER and be used to axiomatically characterize the "pure income" polarization measure in a more plain, intuitive and easily generalizable way.

Returning to the multi-group case, the characterization theorem arising from the use of the new axioms yields another interesting result: the corresponding lower bound on α (which is the parameter of interest in most polarization measures) depend on the number of groups into which a population is splitted and, moreover, it can approach as much as desired the value of zero as the number of groups increases (see Theorem 1). As we will later see, this is a revealing point which confirms the expected relationship between the concepts of polarization and inequality. On the other hand, the fact of having such varying bounds on α might be uncomfortable because it is not clear at all which is the benchmark lower value of α over which a reasonable social polarization measure can be said to depart from the concept of inequality. In order to overcome these limitations, another axiom that further restricts the class of admissible indices must be imposed (see Theorem 2). These important results emphasize that, when extending the income polarization index presented in DER to a broader multi-group context, the axioms used in that paper are not enough, so additional axioms are required.

In section 2 we will introduce the basic notations used throughout the paper and present the axioms that will be used to characterize our polarization index. In section 3 we show the limitations of using the axioms presented in DER *alone* when characterizing a social polarization index. Then, we introduce an additional axiom that completes our characterization result. The proofs are relegated to the Appendix.

2 Basic notation, assumptions and the axioms.

We will make use of the identification-alienation (IA) framework introduced in Esteban and Ray (1994) and used also in DER but adapted to a multi-group context. From now on, we will consider $N \geq 2$ population groups which are exogenously given, each of which with a population mass $M_i > 0$ and a population share $\pi_i \in (0,1)$. These groups are supposed to be relevant in defining individuals' sense of identity. We will assume that to each individual we can attach a *radicalism degree* $x(\geq 0)$ which can be thought as the intensity with which he/she feels to belong to his/her particular population group. This way, we want our measure to be sensitive to the degree to which individuals feel involved with their own group (and not only to the mere fact of belonging or not belonging to a particular group[4]), as this might greatly influence the polarization levels in a given society. For each population group we will have an unnormalized density function $f_i(x)$ that measures the way in which the radicalism degree is distributed therein. We are assuming that the support of each $f_i(x)$ is $\mathbb{R}_+$. Hence $\int\limits_0^\infty f_i(x) = M_i$ and, if we denote the total population by M, then $M = \sum\limits_{i=1}^N M_i$ and $M_i/M = \pi_i$. From now, the

[4]This is the approach used in the discrete polarization measure presented in DER (equation (14)), or in the Reynal-Querol Index (see Garcia-Montalvo and Reynal-Querol (2005)).

density functions for the whole population will be thought a collection of N unnormalized density functions (one for each population subgroup), that is $\mathbf{f}_N : \mathbb{R}_+^N \to \mathbb{R}_+^N$ where $\mathbf{f}_N(x_1, \ldots, x_N) = (f_1(x_1), \ldots, f_N(x_N))$. When no confusion arises, $\mathbf{f}_N$ will be simply written as $\mathbf{f}$. The population shares vector associated to $\mathbf{f}$ will be denoted by $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$.

When adapting the IA framework to our multi-group context, different more or less reasonable alternatives are feasible: in this paper we will focus on what we think is the most natural of them. Firstly, we assume that an individual belonging to the population group i and with a radicalism degree of x experiences a sense of identification that depends on $f_i(x)$[5]. Concerning alienation, we assume that two individuals belonging to the same population group with radicalism degrees of $x, y \geq 0$ ($x \neq y$) feel alienated, and that this alienation is monotonic in $|x - y|$. From the other side, two individuals belonging to different groups with radicalism degrees of $x, y \geq 0$ are also assumed to feel alienated, the latter being monotonic in $x + y$. Under these assumptions, we are implicitly asserting that individuals belonging to different groups always feel alienated against each other unless $x = y = 0$ (in other words: the groups are truly relevant for individuals' identity feeling) and that, *ceteris paribus*, individuals of a given group tend to feel less alienated with respect to the members of the same group than with the members of the other groups, which in many cases seems to be a reasonable assumption[6]. Using the same notation as in Esteban and Ray (1994) and DER, we measure effective antagonism by means of a nonnegative funtion $T(i, a)$, where it is assumed that T is continuous, increasing in its second argument and $T(i, 0) = T(0, a) = 0$. In those papers, it was assumed that polarization can be defined as the sum of all effective antagonisms. Adapting this

[5]Of course, alternative hypothesis could have been presented. The most simple one would have been to assume that all individuals of a given population group are equally identified between themselves, so that individual's sense of identification would depend only on M_i. It is not difficult to prove that this and other assumptions, together with the suitably chosen axioms, would lead to a "Pure Social Polarization" measure as the one presented in DER, equation (14). However, for the sake of concreteness, this issue will not be pursued here.

[6]Recall that these assumptions would not make sense in a (multi-group) context of *income* polarization (see, for instance, equation (17) in DER). In that case, one would have that two individuals with respective incomes x, y belonging to different population groups but with $x = y$ would not feel alienated *vis-à-vis* each other, whereas in our social polarization context, this would not be the case.

reasonable assumption to our multi-group context, social polarization can be defined as

$$P_N(\mathbf{f}) := \sum_{k=1}^{N}\sum_{l=1}^{N}\int\int T(i(x), a(x,y)) f_k(x) f_l(y) dy dx. \tag{1}$$

Under the aforementioned identification-alienation assumptions, equation (1) can be rewritten as

$$\sum_{k=1}^{N}\int\int T(f_k(x), |x-y|) f_k(x) f_k(y) dy dx + \sum_{k=1}^{N}\sum_{l\neq k}\int\int T(f_k(x), x+y) f_k(x) f_l(y) dy dx, \tag{2}$$

Clearly, when $N = 1$, one has that $\pi_1 = 1$ and $P_N(\mathbf{f})$ reduces to the classical polarization index defined for a single population in DER. One of the main purposes of this paper is to present some reasonable axioms under which $P_N(\mathbf{f})$ can be written down more explicitly, so that equation (2) becomes more operational.

2.1 The axioms

In order to present our axioms, we will use the notions of basic densities, roots and λ-squeezes presented in DER. Here, we will only present the definitions; for more details and discussion the interested reader is referred to that paper. A basic density is a density function which is unnormalized by population size, symmetric, unimodal, with compact and connected support. A root is a basic density with mean 1 and support $[0, 2]$ with population size set to unity. Given any basic density g with mean μ and $\lambda \in (0, 1]$, a λ-squeeze of g is defined as the mean-preserving transformation

$$g^{\lambda}(x) := \tfrac{1}{\lambda} g\left(\tfrac{x-(1-\lambda)\mu}{\lambda}\right).$$

Moreover, we will need to introduce the following sets. Define $\Delta_N := \{(\pi_1, \ldots, \pi_N) \in \mathbb{R}_+^N | \sum_i \pi_i = 1\}$ the standard simplex in $\mathbb{R}^N$ and

$$\mathcal{B} := \{(\pi_1, \ldots, \pi_N) \in \Delta_N | \pi_i = \pi_j = \tfrac{1}{2} \text{ for some } i \neq j \in \{1, \ldots, N\}\}.$$

The set $\mathcal{B}$ contains the population shares in which only two population groups have (the same) positive mass, that is: it contains the equally weighted bipolar distributions. Now, we will define the set of population shares which are arbitrarily close to any of the two equal sized group share distributions:

$$\mathcal{B}(\epsilon) := \{\boldsymbol{\pi} \in \Delta_N | \|\boldsymbol{\pi} - \widetilde{\boldsymbol{\pi}}\| < \epsilon \text{ for some } \widetilde{\boldsymbol{\pi}} \in \mathcal{B}\}$$

for some $\epsilon > 0$, $\|.\|$ being the Euclidean norm. Finally, we will define

$$\mathcal{S} := \{(\pi_1, \ldots, \pi_N) \in \Delta_N | \pi_i = 1 \text{ for some } i \in \{1, \ldots, N\}\}.$$

The set $\mathcal{S}$ contains the population shares in which there is a single population group. Analogously, we can define the set

$$\mathcal{S}(\epsilon) := \{\boldsymbol{\pi} \in \Delta_N | \|\boldsymbol{\pi} - \widetilde{\boldsymbol{\pi}}\| < \epsilon \text{ for some } \widetilde{\boldsymbol{\pi}} \in \mathcal{S}\},$$

which contains the population shares which are arbitrarily close to any single population distribution. Recall that in this paper we are *not* interested in considering those cases in which *only* a single group has a positive mass, since then we would be exactly in the same context as in DER: the cases at stake here are the ones in which at least two population groups have positive mass. Hence, even if we are not interested in the configurations for which the corresponding population shares belong to $\mathcal{S}$, those belonging to $\mathcal{S}(\epsilon)\backslash\mathcal{S}$ will be included in our domain.

In order to axiomatically characterize our polarization measure, whenever possible we propose natural extensions to our multi-group context of the axioms presented in DER (compare axioms 2, 3 and 4 below with their analogues in DER). However, when it comes to generalize axiom 1 in DER to the multi-group context, an important limitation arises. It turns out that, if one introduces more population groups, the corresponding axiom is not straightforwardly satisfied, so it must be formulated in a careful way that makes it loose its intuitive force[7]. For this reason, our first axiom will be completely different, having no analogue in the DER axiomatization.

[7] The most natural candidate to extend axiom 1 in DER to the multi-group context would read as follows: "**Axiom 1'**: *Consider a distribution in which each population group has the same normalized density function $f_i(x)$ composed of a single basic density. Suppose that some population group concentrate an arbitrarily large proportion of the whole*

Axiom 1. *Consider a distribution in which each population group has the same normalized density function* $f_i(x)$ *composed of a single basic density. Suppose that some couple of population groups have the same population share and concentrate an arbitrarily large proportion of the whole population mass (that is:* $\boldsymbol{\pi} \in \mathcal{B}(\epsilon)$ *for any arbitrarily small* $\epsilon > 0$*). Then, if we shift mass from one of the big groups to the other, polarization should decrease.*

This axiom captures the intuitive idea that the equally weighted bipolar distributions maximize polarization. It is important to point out that the single group version of this axiom[8] would serve the same purpose as axiom 1 in DER, that is: it would (upper) bound the values of α to 1. In particular, this implies that the polarization measure presented in DER could also be axiomatically characterized substituting their "single-squeeze axiom" by the single group version of our new axiom (see proof of Lemma 2 in the appendix). We contend that this alternative way of characterizing the income polarization measure is preferable on two grounds: first, it is even more plain and intuitive[9] and second, it allows for a straightforward generalization to the multi-group case. Clearly, it is preferable to characterize a measure with axioms that can be easily extended to broader contexts than with dimension-specific ones. From the other side, it is reassuring that both axioms yield almost the same restriction on the bounds of α (see later).

Axiom 2. *Consider a symmetric distribution in which each population group has the same density function* $f_i(x)$ *composed of two basic densities with disjoint support sharing the same root. Then, if all outer distributions*

population mass (that is: $\boldsymbol{\pi} \in \mathcal{S}(\epsilon)\backslash\mathcal{S}$ *for some arbitrarily small* $\epsilon > 0$*). Then, if we squeeze the density of the large group, polarization should decrease."* However, it can be verified that axiom 1' holds only when the size of the small groups and the extent to which the bigger group density is squeezed (measured by $\lambda \in (0, 1]$) are related in a certain way. When a certain value of λ is fixed, one can find sufficiently small (but depending on λ) population masses for the other groups making axiom 1' hold true. Since this might not be intuitively compelling, it has not been included as a characterizing axiom.

[8] Clearly, the single group version would be: "**Axiom 1"**: *Consider a distribution composed of the same two basic densities, with the same population mass and disjoint supports. Then, if we shift mass from one group to the other, polarization must decrease*".

[9] Recall that the "single-squeeze" axiom introduced in DER seems quite intuitive but is introducing some non-trivial restrictions. In particular, it states that when a basic density is globally compressed, the effect of reduced inter-individual alienation on polarization counterbalances the effect of increasing identification for the individuals located at the center of the distribution. Stated in this way, this restriction is not particularly compelling.

are squeezed, polarization must increase.

This axiom is a straightforward extension of axiom 2 used in DER to the multi-group case. Its intuition is completely analogous: when the most radicalized individuals in each group are made more homogeneous or "compact", then polarization is expected to rise. This is the axiom that distinguishes polarization from inequality: recall that an inequality measure would decrease under the transformation presented in axiom 2.

Axiom 3. *Consider a distribution in which each population group has the same normalized density function $f_i(x)$ composed of two basic densities with disjoint support sharing the same root. Consider, moreover, that the population shares vector $\boldsymbol{\pi} \in \mathcal{B}(\epsilon)$ for some arbitrarily small $\epsilon > 0$. Slide the inner basic densities outwards, while keeping supports disjoint. Then polarization must increase.*

This axiom is a plain and intuitive generalization of axiom 3 used in DER. It tries to capture the idea that if the different groups are made more homogeneous, then polarization should increase. After this transformation, individuals feel less alienated with respect to the members of the same group but more alienated with respect to the others. A couple of remarks are in order at this point. First, recall that this axiom captures the idea that alienation between individuals of the same group is less important than alienation between individuals of different groups, which we consider a reasonable assumption. Second: in order to ensure that the axiom makes sense, the population shares are imposed to be arbitrarily close to any of the equal-sized two groups shares. If no restriction were imposed on the population shares distribution, the axiom could make no sense at all: imagine, for example, a distribution in which a single population concentrates the most part of the mass and that the other $(N-1)$ groups had a negligible mass. In that case, an outward slide transformation as proposed in axiom 3 would decrease polarization rather than increasing it because of the negligible effect of very small groups on the final result.

Axiom 4. *If $P_N(\mathbf{f}) \geq P_N(\mathbf{g})$ and $p > 0$ then $P_N(p\mathbf{f}) \geq P_N(p\mathbf{g})$, where $p\mathbf{f}$ and $p\mathbf{g}$ represent population scalings of $\mathbf{f}$ and $\mathbf{g}$ respectively.*

This population invariance axiom is very common in the literature of polarization or inequality measurement. It states that if polarization is higher in one situation than in another, it must continue to be so when populations in both situations are scaled up or down by the same amount.

3 Statement of the main results and some remarks.

We will start with the following characterization result.

Theorem 1 *A polarization measure as defined in (2) satisfies axioms 1,2,3,4 if and only if it is proportional to*

$$P_{N,\alpha}(\mathbf{f}) = \sum_{i=1}^{N} \int\int f_i^{1+\alpha}(x) f_i(y)|x-y| dydx + \sum_{i=1}^{N}\sum_{j\neq i} \int\int f_i^{1+\alpha}(x) f_j(y)(x+y) dydx, \tag{3}$$

where $\alpha \in [\frac{1}{3N-2}, 1)$. *In particular, when* $N = 2$, α *can reach the admissible value of* 1.

According to Theorem 1, the lower bound of α is strictly positive and depends on the number of groups we are taking into account[10] in such a way that, as those increase, the lower bound increasingly approaches the value of 0. As is well known, the value of α has to be treated as the degree of polarization sensitivity (see Esteban and Ray (1994)) and the larger is its value, the greater is the departure from inequality measurement. The fact that the lower bound of α can approach 0 as much as desired by considering a large number of groups is somewhat uncomfortable, but, as we will now see, this is a natural result that confirms the expected relationship between the concepts of polarization and inequality. Keeping other things equal, an increase of the number of groups leads to a decrease in individuals' sense of identification, as there are fewer individuals in each group. Pushing this argument to its limits, one would end up with a population in which the identification component would be the same for all individuals (that is: it

[10] Recall that the bounds for α found in Theorem 1 coincide with those proposed in DER only when the number of groups is two.

would be irrelevant) and the only relevant contribution to polarization would come from the alienation component. Recall that when the identification component plays no role in the IA framework, we are back to the conceptual foundations of inequality. Hence, it should not be surprising that, in those extreme cases, α is allowed to approach the value of 0.

Be that as it may, it must be acknowledged that the fact of having varying admissible lower bounds for α is uncomfortable to a considerable extent, because it is not clear at all which is the benchmark lower value of α over which a reasonable social polarization measure can be said to depart from the concept of inequality. Moreover, the analysis of sensitivity of social polarization to different values of α (equivalent to the empirical analysis presented in DER, section 3.2) is conceptually flawed if we pretend to compare populations which are splitted into different number of groups, as the corresponding ranges of admissible values of α will not coincide. One possible way of avoiding these problems is to impose the following reasonable axiom.

Axiom 5. *For any population of fixed mass M consider symmetric configurations in which all density functions $f_i(.)$ are the same basic density. Then, $P_{N_1}(\mathbf{f}_{N_1}) \leq P_{N_2}(\mathbf{f}_{N_2})$ for any $N_1 \geq N_2 \geq 2$.*

This axiom captures the widespread idea that, other things being equal, the larger the number of groups, the lower the corresponding polarization[11]. Some authors have used this idea or very similar ones in the study of conflict and polarization (see, for example, Esteban and Ray (1994,1999) or Montalvo and Reynol-Queral (2005), who trace this idea from the seminal works of Horowitz (1985)). It is important to recall that this axiom would not make sense if our purpose were to measure bipolarization, as is the case, for example, of the Wolfson Index (see Wolfson (1994)). Imposing this mild restriction, one obtains the following theorem.

Theorem 2 *A polarization measure as defined in equation (2) satisfies axioms 1-5 if and only if it is of the form presented in equation (3) with the additional restriction that $\alpha \geq 1/2$.*

[11] Recall that axiom 5 cannot be faithfully translated to the single group context used in DER because of the way in which the corresponding domains have been defined. In the multi-group case, all group densities are at the same distance between them whereas in the single group case this cannot happen by construction.

This way, we are ruling out the uncomfortable dependecy of the range of admissible values of α on the values of N. With the inclusion of axiom 5, the new bounds of α make clear which is the range of values for which a reasonable polarization measure is obtained. Theorems 1 and 2 make clear that when passing from the single group to the multi-group context, the axioms used to characterize the polarization index presented in DER are not enough and that further restrictions must be imposed. We contend that, in some instances, the social polarization index presented in this paper can be better suited than its pure income counterpart to explore, either from a theoretical or an empirical point of view, the intertwined relationships between polarization and social conflict.

4 Appendix: Proof of the theorems.

4.1 Proof of theorem 1.

The proof of theorem 1 is lengthy and technically involved. However, its structure is analogous to the proof of the main characterization theorem presented in DER (theorem 1). Firstly one must establish that axioms 1-4 imply (3). Then, in proving the sufficiency part, one can establish the bounds for α. Since the underlying ideas in each of the different steps are similar and there are only some technical differences, the proof will not be shown here, but is available upon request. The only part of the proof we will reproduce here is the one concerning the use of axiom 1(which is the only one which is conceptually different from its single group counterpart shown in DER). In Lemma 1 we establish the upper bound of α. Moreover, in the proof of Lemma 1 we show that the characterization of the income polarization measure presented in DER can be improved by using the single group version of axiom 1.

Lemma 1. *Given that $P_{N,\alpha}(\mathbf{f})$ is of the form (3), axiom 1 is satisfied if and only if* $\alpha < 1$ for $N > 2$ and $\alpha \leq 1$ for $N = 2$.

Proof: Consider a configuration as given in axiom 1. Without loss of generality, we will assume that the whole population mass (M) is normalized at 1. Moreover, we assume that the first two population groups are the ones concentrating an arbitrarily large proportion of the whole population mass, so $\pi_1 = \pi_2 = \pi$ and π can be arbitrarily close to (and smaller than) 1/2. Each (normalized) $f_i(x)$ is the same basic density with mean μ on support $[a, b]$ and root f^*. Let $m = \mu - a$ and let d be the distance between the means of two basic densities. By symmetry, one has that $d = 2(m + a)$. According to equation (3), and using lemmas 6 and 7 in DER, total polarization in this context is equal to

$$P_{N,\alpha}(\mathbf{f}) = (4km^{1-\alpha}\psi_1(f^*, \alpha)) \sum_{i=1}^{N} \pi_i^{2+\alpha} + (2kdm^{-\alpha}\psi_2(f^*, \alpha)) \sum_{i=1}^{N} \sum_{j \neq i} \pi_i^{1+\alpha}\pi_j.$$

Since we want to control for the effects of a transfer of population mass (Δ) from one of the big groups to the other, we will rewrite the last expression as

$$P_{N,\alpha}(\mathbf{f}, \Delta) = (4km^{1-\alpha}\psi_1(f^*, \alpha)) \left((\pi + \Delta)^{2+\alpha} + (\pi - \Delta)^{2+\alpha} + \sum_{i>2}^{N} \pi_i^{2+\alpha} \right) +$$
$$2kdm^{-\alpha}\psi_2(f^*, \alpha)[(\pi + \Delta)^{1+\alpha}(\pi - \Delta) + (\pi - \Delta)^{1+\alpha}(\pi + \Delta) +$$
$$\left((\pi + \Delta)^{1+\alpha} + (\pi - \Delta)^{1+\alpha}\right) \sum_{i>2}^{N} \pi_i + 2\pi \sum_{i>2}^{N} \pi_i^{1+\alpha} + C], \tag{4}$$

where C includes terms not depending on Δ. Now, according to axiom 1, $P_{N,\alpha}(\mathbf{f}, \Delta)$ should have a maximum at $\Delta = 0$. Hence, we need to compute the first and second derivatives of $P_{N,\alpha}(\mathbf{f}, \Delta)$ with respect to Δ. Computing $\frac{\partial P_{N,\alpha}(\mathbf{f},\Delta)}{\partial \Delta}$ we obtain

$$4km^{1-\alpha}\psi_1(f^*, \alpha)(2 + \alpha)\left((\pi + \Delta)^{1+\alpha} - (\pi - \Delta)^{1+\alpha}\right) +$$
$$2kdm^{-\alpha}\psi_2(f^*, \alpha)[(\pi - \Delta)^{1+\alpha} - (\pi + \Delta)^{1+\alpha} + (1 + \alpha)$$
$$\left((\pi^2 - \Delta^2)^{\alpha}\left((\pi - \Delta)^{1-\alpha} - (\pi + \Delta)^{1-\alpha}\right) + ((\pi + \Delta)^{\alpha} - (\pi - \Delta)^{\alpha}) \sum_{i>2}^{N} \pi_i \right)].$$

Clearly, $\frac{\partial P_{N,\alpha}(\mathbf{f},\Delta=0)}{\partial \Delta} = 0$, so $\Delta = 0$ is a critical point of $P_{N,\alpha}(\mathbf{f}, \Delta)$. Now, $\frac{\partial^2 P_{N,\alpha}(\mathbf{f},\Delta)}{\partial \Delta^2}$ is equal to

$$4km^{1-\alpha}\psi_1(f^*, \alpha)(2 + \alpha)(1 + \alpha)\left((\pi + \Delta)^{\alpha} + (\pi - \Delta)^{\alpha}\right) +$$
$$2kdm^{-\alpha}\psi_2(f^*, \alpha)(1 + \alpha)$$
$$[\alpha((\pi + \Delta)^{\alpha-1}(\pi - \Delta) + (\pi - \Delta)^{\alpha-1}(\pi + \Delta)) - 2\left((\pi + \Delta)^{\alpha} + (\pi - \Delta)^{\alpha}\right) +$$
$$\alpha((\pi + \Delta)^{\alpha-1} + (\pi - \Delta)^{\alpha-1}) \sum_{i>2}^{N} \pi_i].$$

Hence, one has that $\frac{\partial^2 P_{N,\alpha}(\mathbf{f},\Delta=0)}{\partial\Delta^2}$ is equal to

$$4km^{1-\alpha}\psi_1(f^*,\alpha)(2+\alpha)(1+\alpha)2\pi^{\alpha}+$$
$$2kdm^{-\alpha}\psi_2(f^*,\alpha)(1+\alpha)[2\alpha\pi^{\alpha}-4\pi^{\alpha}+(1-2\pi)2\alpha\pi^{\alpha-1}]$$

since $\sum_{i>2}^{N}\pi_i = 1-2\pi$. Inspecting the last expression, we see that the first term is always positive and that the second one can be negative. Hence, and given the fact that $d = 2(m+a)$ (so one has that $2kdm^{-\alpha}\psi_2(f^*,\alpha) = 4km^{1-\alpha}\psi_2(f^*,\alpha)+4kam^{-\alpha}\psi_2(f^*,\alpha)$), a necessary and sufficient test case to test whether $\frac{\partial^2 P_{N,\alpha}(\mathbf{f},\Delta=0)}{\partial\Delta^2} \leq 0$ is to impose that $a = 0$. Moreover, by lemma 10 in DER, one has that, for any $\alpha > 0$, $\psi_2(f^*,\alpha) = \Gamma\psi_1(f^*,\alpha)$ for some $\Gamma \geq 3$. In that case, after some computations we can rewrite $\frac{\partial^2 P_{N,\alpha}(\mathbf{f},\Delta=0)}{\partial\Delta^2}$ as

$$4km^{1-\alpha}\psi_1(f^*,\alpha)(1+\alpha)2\pi^{\alpha-1}\left[\pi(2+\alpha)(1-\Gamma)+\alpha\Gamma\right].$$

Finally, we have to check which are the values of α for which some $\pi \leq 1/2$ can be found such that $[\pi(2+\alpha)(1-\Gamma)+\alpha\Gamma] \leq 0$. If the last restriction must hold true, one must have that

$$\alpha \leq \tfrac{2\pi(\Gamma-1)}{\Gamma+\pi-\Gamma\pi},$$

Now, since $2\pi(\Gamma-1)/(\Gamma+\pi-\Gamma\pi)$ is an increasing function in Γ when $\pi > 0$ and $\Gamma \geq 3$, from the last expression we deduce that

$$\alpha \leq \tfrac{4\pi}{3-2\pi}.$$

According to the configuration stated in axiom 1, $\boldsymbol{\pi} \in \mathcal{B}(\epsilon)$, so one must have that $1/2 - \pi < \epsilon$ for some arbitrarily small ϵ. Using this fact, if one inspects the last bound on α one deduces that $\alpha \leq 1$ when $\pi = 1/2$ (in which case there are only two groups with positive mass ($N = 2$)) and that $\alpha < 1$ when $1/2 - \epsilon < \pi < 1/2$. This proves the lemma.

Q.E.D.

Recall that the income polarization measure presented in DER can be axiomatically characterized substituting the "single-squeeze axiom" by the single group version of axiom 1 (see footnote 10). In order to do so, it suffices to follow an analogous argument as in Lemma 1 but working only with two population groups ($N = 2$) and setting $\pi = 1/2$.

4.2 Proof of theorem 2.

According to theorem 1, a polarization index as defined in (2) is proportional to $P_{N,\alpha}(\mathbf{f})$ if and only if axioms 1 to 4 are satisfied. Let us now check what happens when axiom 5 is imposed. Without loss of generality, we will assume that the whole population mass (M) is normalized at 1 so, by symmetry, the population shares π_i will be equal to $1/N$. Total polarization can be decomposed as

$$P_{N,\alpha}(\mathbf{f}) = NP^w + N(N-1)P^b,$$

where P^w is the internal polarization within groups and P^b is the polarization between groups. Now, if we use lemma 6 (DER) with $\lambda = 1$ we have that

$$P^w = 4k\left(\tfrac{1}{N}\right)^{2+\alpha} m^{1-\alpha}\psi_1(f^*, \alpha)$$

where k is a positive constant, m is the distance between the mean and the lower tail of the basic density. From the other side, using lemma 7 (DER) with $\lambda = 1$ we obtain

$$P^b = 2k(2\mu)\left(\tfrac{1}{N}\right)^{2+\alpha} m^{-\alpha}\psi_2(f^*, \alpha),$$

where μ is the mean of the basic density. Recall that, by definition, $m \leq \mu$. Substituting the last two expressions into $P_{N,\alpha}(\mathbf{f})$ we obtain

$$P_{N,\alpha}(\mathbf{f}) = 4km^{-\alpha}\left(\tfrac{1}{N}\right)^{1+\alpha}(m\psi_1(f^*, \alpha) + (N-1)\mu\psi_2(f^*, \alpha)),$$

If axiom 5 has to be satisfied, one must have that $\frac{\partial P_{N,\alpha}(f)}{\partial N} \leq 0$. Differentiating the last expression with respect to N, we obtain

$$-(1+\alpha)\left(\tfrac{1}{N}\right)^{2+\alpha}(m\psi_1(f^*, \alpha) + (N-1)\mu\psi_2(f^*, \alpha)) + \left(\tfrac{1}{N}\right)^{1+\alpha}\mu\psi_2(f^*, \alpha),$$

where we have dropped the constant term $4km^{-\alpha}$. Rearranging the last expression, one obtains

$$\frac{\partial P_{N,\alpha}(\mathbf{f})}{\partial N} \equiv \left(\tfrac{1}{N}\right)^{2+\alpha}\left[(1+\alpha)(-m\psi_1(f^*, \alpha) + \mu\psi_2(f^*, \alpha)) - \alpha N\mu\psi_2(f^*, \alpha)\right].$$

Observe that the term $(-m\psi_1(f^*,\alpha)+\mu\psi_2(f^*,\alpha))$ must be positive, because $m \leq \mu$ and, (by lemma 8 (DER)) $\psi_2(f^*,\alpha) = \Gamma\psi_1(f^*,\alpha)$ for some $\Gamma \geq 3$. Hence, a necessary and sufficient test case to prove the theorem is to consider the lowest possible value of N (which is 2) and the highest possible value of m (which is μ). In that case, one should impose that

$$(1+\alpha)(\psi_2(f^*,\alpha)-\psi_1(f^*,\alpha)) - 2\alpha\psi_2(f^*,\alpha) \leq 0.$$

Manipulating a little bit, we see that this is satisfied when

$$\alpha \geq \frac{\psi_2(f^*,\alpha)-\psi_1(f^*,\alpha)}{\psi_2(f^*,\alpha)+\psi_1(f^*,\alpha)} = \frac{\Gamma-1}{\Gamma+1}.$$

Clearly, the lowest possible value for this function when $\Gamma \geq 3$ is 1/2, so the theorem is proven.

Q.E.D.

References

[1] ALESINA, A., AND E. SPALAORE (1997):"On the number and size of nations", *Quarterly Journal of Economics*, 113, 1027-1056.

[2] CHAKRAVARTY, S.R., AND A. MAJUMDER (2001):"Inequality, polarization and welfare: theory and applications", *Australian Economic Papers*, 40,1-13.

[3] COLLIER, P., AND A. HOEFFLER (2001):"Greed and grievance in civil war", Mimeo, World Bank.

[4] DUCLOS, J-Y, ESTEBAN J. AND D. RAY (2004):"Polarization: Concepts, Measurement, Estimation", *Econometrica*, 72, 1737-72.

[5] EASTERLY, W., AND R. LEVINE (1997):"Africa's growth tragedy: policies and ethnic divisions", *Quarterly Journal of Economics*, 112,1203-50.

[6] ESTEBAN J., GRADIN, C., AND D. RAY (2007): "An estension of a measure of polarization, with an application to the income distribution of five OECD countries", *Journal of Economic Inequality*, 5 (1), 1-19.

[7] ESTEBAN J., AND D. RAY (1994): "On the Measurement of Polarization", *Econometrica*, 62, 819-852.

[8] ESTEBAN J., AND D. RAY (1999): "Conflict and distribution", *Journal of Economic Theory*, 87, 379-415.

[9] FOSTER, J.E. AND M.C. WOLFSON (1992):"Polarization and the decline of the middle class: Canada and the US", Mimeo, Vanderbilt University.

[10] HOROWITZ, D.L. (1985): "Ethnic groups in conflict", Berkeley: University of California Press.

[11] MONTALVO, J.G., AND M. REYNAL-QUEROL (2005):"Ethnic polarization, potential conflict and civil wars", *American Economic Review*, 95, 796-815.

[12] QUAH, D. (1997):"Empirics for growth and distribution: stratification, polarization and convergence clubs", *Journal of Economic Growth*,2,27-59.

[13] RODRIGUEZ, J.G., AND R. SALAS (2002):"Extended bi-polarization and inequality measures", Mimeo, Universidad Complutense de Madrid.

[14] WANG, Y.Q., AND K.Y.TSUI (2000):"Polarization orderings and new classes of polarization indices", *Journal of Public Economic Theory*,2,349-363.

[15] WOLFSON, M.C. (1994): "When inequalities diverge", *American Economic Review*, 84, Papers and Proceedings, 353-358.

[16] WOLFSON, M.C. (1997):"Divergent inequalities: Theory and Empirical Results", *Review of Income and Wealth*, 43, 401-421.

[17] ZHANG, X., AND R. KANBUR (2001):"What difference do polarization measures make? An application to China", *Journal of Development Studies*,37,85-98.