

# Social Polarization: Introducing distances between and within groups.

Iñaki Permanyer\*

Institut d'Anàlisi Econòmica (CSIC), Barcelona.

July 14, 2008

## Abstract

The measurement of social polarization has received little attention from the literature. The only social polarization index that has been used to measure religious or ethnic polarization (the  $RQ$  index) has several shortcomings that are critically discussed in the paper. In particular, that index is not taking into account the existing distance between and within different groups. A couple of axiomatically characterized social polarization indices that overcome these limitations are presented. In the empirical section we show that the rankings of countries according to the levels of polarization change to a great extent when we replace the  $RQ$  index by the indices presented in this paper.

---

\*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 measurement of polarization has witnessed an increasing interest in the past few years both from theorists and practitioners. This is mainly due to the fact that the notion of polarization is closely related to the existence of social tension or conflict phenomena which can not be appropriately predicted by traditional inequality measures. Nowadays, it is widely acknowledged that polarization measures are much more appropriate for the purpose of capturing social tension and conflictivity than inequality measures. The first papers to deal with the measurement of polarization in a rigorous way were those of Esteban and Ray (1994) (from now on ER94) and Wolfson (1994), which laid the foundations of many polarization indices that were to follow with its corresponding empirical applications (see, for example, Duclos, Esteban and Ray (2004) (from now on DER), Wolfson (1997), Alesina and Spalaore (1997), Quah (1997), Wang and Tsui (2000), Esteban, Gradín and Ray (2007), Chakravarty and Majumder (2001), Zhang and Kanbur (2001) and Rodriguez and Salas (2002)). It is important to point out that all these papers focus their attention on the measurement of *income* polarization.

However, the distribution of income or wealth is not always the cause of social tension or conflicts, so it soon became evident that there was a need to extend the notion of polarization to broader contexts. From now on, we will use the term *social polarization* when the factors that determine the tension or conflictivity in a given population are socially driven and do not depend on the distribution of income. Classical examples of these factors are ethnic, religious or nationalistic feelings. In contrast with its importance, the literature has not devoted much attention to the measurement of social polarization. To our knowledge, the only papers that have proposed a “pure” social polarization index are those of Montalvo and Reynal-Querol (2002, 2005a, 2005b) and Duclos, Esteban and Ray (2004) . In the later paper, an attempt is made to propose certain hybrid polarization indices which lie somewhere between the concepts of income and social polarization. These indices are not axiomatically characterized; they are liberal transplants from the pure income polarization index defined in their paper, so, the authors acknowledge, they are somewhat arbitrary and open to legitimate criticism.

Exactly the same can be said about the social polarization index introduced in the papers of Montalvo and Reynal-Querol. The so-called Reynal-

Querol index ( $RQ$ ) used in their papers can be seen as the ER94 income polarization index adapted to the context of social polarization by arbitrarily disregarding the alienation component of the index. The  $RQ$  index is used in Montalvo, Reynal-Querol (2005a) to predict the occurrence of Civil Wars and in Montalvo, Reynal-Querol (2005b) to predict the growth rate of GDP per capita.

The main purpose of this paper is to present axiomatically characterized social polarization indices that overcome the several limitations of the  $RQ$  index, which stands out as the single social polarization index that has been used in practice to explain the occurrence of social conflicts like Civil Wars or the evolution of economic growth. The limitations of that index can be summarized as follows. As mentioned earlier, the index is not axiomatically characterized, so it might be criticized on grounds of arbitrariness<sup>1</sup>. More importantly, the definition of the index assumes that the different population groups in which the society is splitted (say, religious or ethnic groups) are equally alienated *vis-à-vis* each other. In other words, it assumes that the distance between any couple of groups (i.e: the feeling of animosity) is the same no matter which groups we are comparing. In order to justify this decision, the authors contend that “If we want to calculate ethnic (religious) polarization using the index  $P$ , we need to calculate the distance between different ethnic (religious) groups, which is a very difficult task compared to what happens in the case of income or wealth” (Montalvo and Reynal-Querol 2005b, p. 301). However, ruling out the role of alienation in a polarization index misses an *essential* part of its definition. It seems hardly questionable that, fixing the size of the groups, the higher the distance between them, the higher the corresponding level of social tension and polarization. Now, we contend that, regardless of the acknowledged difficulties involved in the calculation of distances between groups, a social polarization index *must* take the alienation factor into account.

In this paper we have proposed a couple of models that try to take into account the notion of alienation/distance between individuals. In the first

---

<sup>1</sup>The importance of providing axiomatic characterizations of social indicators is widely acknowledged and has been discussed elsewhere. The axioms represent the basic properties satisfied by the indices which, joint together, characterize them uniquely. Hence, they are very useful to discriminate between different indices according to the corresponding basic properties they satisfy and their normative implications.

one, it is assumed that individuals belonging to different groups do feel alienated *vis-à-vis* each other according to a distance function that depends on the intensity with which they feel to belong to their particular group. We assume that, other things being equal, the stronger the feeling of individuals' identity with their own group, the higher the corresponding distance/alienation between them. In the second model, the notion of distance between individuals is not only important for the measurement of polarization between groups but also to measure polarization *within* groups. Given the fact that social tension or conflict could arise within the members of a given group, it seems reasonable to make room to capture those phenomena. The choice between these two models for empirical purposes might crucially depend on the accuracy of the underlying hypotheses for the specific context in which they will be used. For both models, an axiomatic characterization for the corresponding social polarization index is provided using different sets of axioms.

As mentioned before, these models can be very useful to measure the levels of religious or ethnic polarization. In the empirical section we compare the values of the  $RQ$  index with the values of the indices introduced in this paper. It turns out that introducing the notion of distance between groups to measure religious polarization is a very relevant issue: the corresponding rankings of countries in terms of religious polarization changes to a great extent. These results suggest that the inclusion of alienation/distance between individuals is a crucial element that can make an important difference when measuring the levels of polarization at the country level.

The paper is structured as follows. In section 2 we examine our first model in which alienation is measured between individuals of different groups only. We present a set of axioms that characterize the corresponding social polarization index. In section 3 we examine our second model in which social tension is allowed to take place both between and within groups. Again, we present a different set of axioms that characterize the corresponding social polarization index. Section 4 contains the empirical results of the paper, in which we compare the values of the different indices. We conclude in section 5. The proofs are relegated to the appendix

## 2 Measuring social polarization between groups

The measurement of social polarization is inspired by the notions introduced in the measurement of income polarization. We will take the ideas and models introduced in ER94 as our starting point. There, it was assumed that polarization could be defined as the sum of all possible effective antagonisms between individuals. Moreover, in the model proposed in that paper, the antagonism felt by individual  $i$  towards  $j$  depends on the so-called identification and alienation components. The alienation component reflects the fact that two individuals can be very different from each other and the greater the difference, the greater the contribution to overall social tension or conflictivity. In ER94, the alienation / distance between individuals is measured as the absolute difference in respective incomes. The identification component is introduced to capture the idea that, the greater the size of the group to which an individual belongs, the greater his/her possibilities of making an effective voicing of his/her alienation. If the population is splitted in  $n$  groups of size  $\pi_i$  ( $i = 1, \dots, n$ ), each of which with an income level of  $y_i$ , Esteban and Ray characterize axiomatically the following income polarization index:

$$P(\boldsymbol{\pi}, \mathbf{y}) = K \sum_{i=1}^n \sum_{j=1}^n \pi_i^{1+\alpha} \pi_j |y_i - y_j| \quad (1)$$

where  $K > 0$  and  $\alpha \in (0, 1.6]$ . Here, the identification component equals  $\pi_i^\alpha$  and the alienation from  $i$  towards  $j$  is  $|y_i - y_j|$ . In a later paper, Duclos, Esteban and Ray provide an analogous result for continuous income distributions  $f(x)$  (see DER). The index is written as

$$P_\alpha(f) \equiv \int \int f(x)^{1+\alpha} f(y) |y - x| dy dx \quad (2)$$

with  $\alpha \in [0.25, 1]$ , which is the continuous analogue of (1).

When these ideas are extrapolated to the context of social polarization one faces an important problem, namely: how to measure the distances between  $N$  different groups (with  $N \geq 2$ ). In Montalvo and Reynal-Querol (2002, 2005a, 2005b) a draconian solution is proposed: to assume that the distances are the same for no matter which group we are comparing. Hence they present what they call a discrete polarization index defined as

$$DP(\alpha, k) = k \sum_{i=1}^N \sum_{j \neq i} \pi_i^{1+\alpha} \pi_j \quad (3)$$

which is a liberal transplant from  $P(\boldsymbol{\pi}, \mathbf{y})$  under the assumption of equal distances between all pairs of groups. Moreover, they show that the only values of  $\alpha, k$  for which  $DP(\alpha, k)$  satisfies certain reasonable properties are 1 and 4 respectively (see Montalvo and Reynal-Querol (2002)). In that case, they obtain the Reynal-Querol index, defined as

$$RQ = 1 - \sum_{i=1}^N \left( \frac{1/2 - \pi_i}{1/2} \right)^2 \pi_i = DP(1, 4) \quad (4)$$

We contend that by excluding the alienation component in the definition of the index, one misses one of the essential aspects of polarization. It seems beyond question that, *ceteris paribus*, the level of social polarization in a given population must be sensitive to the existing feelings of animosity between groups. At this moment, we present a model that takes into account the notion of alienation / distance between individuals.

## 2.1 The model (1)

Suppose we have  $N$  exogenously given groups, with  $N \geq 2$ . Typical examples might be religious or ethnic groups, but the definition is left as general as possible, so other groupings are also possible. We assume that these groups are cohesive and relevant in defining individuals' sense of identity. Typically, these groups are competing for a share of power or to enforce their own interests. Now, we assume that each individual has a given feeling of identity with the group to which he/she belongs. This feeling of identity is assumed to be closely related to the corresponding degree of involvement when pursuing the interests of the group, which in many cases is a reasonable assumption. We will denote by  $x$  ( $x \in \mathbb{R}_+$ ) the intensity with which individuals feel to belong to their particular group, and call this number *radicalism degree*. 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 as it happens with the  $RQ$  index, since 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  $\mathbb{R}_+$ .

We introduce the following notation. The whole population mass is equal to  $M$  and the population mass of each group equals  $M_i$ . The population share of group  $i$  is  $\pi_i = M_i/M$ . Hence, for each radicalism degree density function one has that  $\int_{\mathbb{R}_+} f_i(x) = M_i$ . From now, the density functions for the whole population will be thought a collection of  $N$  unnormalized density functions (one for each population subgroup), that is  $\mathbf{f}_N : \mathbb{R}_+^N \rightarrow \mathbb{R}_+^N$  where  $\mathbf{f}_N(x_1, \dots, x_N) = (f_1(x_1), \dots, f_N(x_N))$ . 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  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_N)$ .

We can now adapt the identity-alienation framework introduced in ER94 to the present context. For an individual belonging to group  $i \in \{1, \dots, N\}$ , we assume that the identity component depends on the corresponding size  $M_i$ . This is equivalent to say that an individual feels identified with all the members of his/her own group, which is another way of stating that the groups are cohesive and relevant in defining individuals' identity feelings and that two individuals belonging to the same group do not feel alienated *vis-à-vis* each other. Regarding the alienation component for individuals in different groups, let us consider two individuals of radicalism degrees  $x, y$ . Since radicalism degrees are associated with individuals' degree of involvement when pursuing the interests of their group and since the different groups are competing for a share of power, we will simply assume that alienation is monotonically increasing in  $x + y$ . This way, we capture the idea that the higher the involvement of individuals in pursuing the interests of their own groups, the higher the level of social tension. Now, let us define effective antagonism as a nonnegative function  $T(\iota, a)$ , where it is assumed that  $T$  is continuous, increasing in its second argument and  $T(\iota, 0) = T(0, a) = 0$ . Finally, polarization is assumed to be proportional to the sum of all effective antagonisms, so that, if  $\mathbf{f} \equiv (f_1, \dots, f_N)$

$$P_N(\mathbf{f}) = \sum_{i=1}^N \sum_{j=1}^N \int \int T(\iota(x), a(x, y)) f_i(x) f_j(y) dy dx \quad (5)$$

Under the aforementioned identification-alienation assumptions, (5) can be rewritten as

$$P_N(\mathbf{f}) = \sum_{i=1}^N \sum_{j \neq i} \int \int T(M_i, x + y) f_i(x) f_j(y) dy dx \quad (6)$$

We will present some reasonable axioms under which the previous formula can be written down more explicitly in a more operational way.

## 2.2 The axioms (1).

In order to present our axioms, we will use the notions of basic densities and roots presented in DER. 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.

**Axiom 1.** *Consider a distribution in which each population group has the same density function  $f_i(x)$  composed of a single basic density. Suppose, moreover, that one of the population groups has an additional (outer) basic density (with disjoint support sharing the same root as in Figure 1) and that the population mass of the inner one is not smaller than the population mass of the outer one. The population share is the same for all  $N$  groups. Then, if the inner and the outer densities of the two-densities group approach to each other by the same amount slide (while keeping disjoint supports), polarization must not decrease.*

This axiom reflects the following idea. Suppose that within a given group, we have a populous subgroup of low radical individuals and another smaller subgroup of more radicalized individuals. Then, other things being equal, if the first and second subgroups increase and decrease respectively their radicalism by the same amount, polarization should increase because after the transformation, the average level of radicalism of the whole group has increased.

**Axiom 2.** *Consider a distribution in which each population group has a density function  $f_i(x)$  composed of a single basic density. Suppose, moreover, that the population mass of the first group ( $M_1$ ) is not bigger than the*



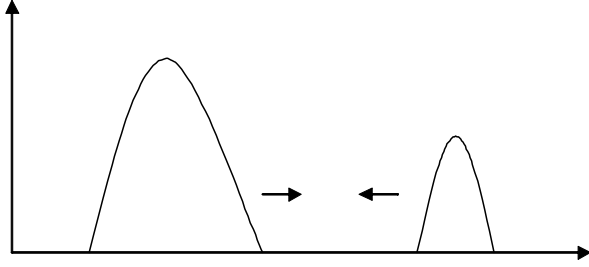


Figure 1: Two basic densities approaching within a group.

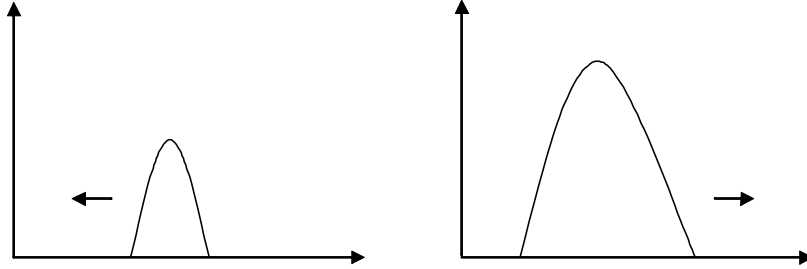


Figure 2: A small group becoming less radical and a radicalizing big group.

*population mass of the second group ( $M_2$ ). Slide the members of the first group inwards and the members of the second group outwards (by the same amount) as in Figure 2. Then polarization must not decrease.*

This axiom captures the intuitive idea that, other things being equal, if there are two groups, the first one being less numerous than the second, with the members of the smaller group becoming less radicalized and the members of the bigger group becoming more radicalized, in overall polarization should increase. In other words, our polarization index must be sensitive to the sizes of the different population groups and give more emphasis to the transformations of the bigger ones. This axiom can be thought as the “between-group version” of axiom 1.

In order to motivate the following axiom we will consider the following

distribution. Suppose that one has three groups, one with a “large” mass and the other two with equally “small” masses. Moreover, it is assumed that the radicalism distributions have the same normalized density functions. In this scenario, polarization is arbitrarily low because population is mostly concentrated in a single group and it is assumed that individuals within a given group do not feel alienated *vis-à-vis* each other. Consider now the process of transferring population mass from the big group to the smaller ones by the same amount. In this context we present an axiom imposing a natural condition on the population transfer process.

**Axiom 3.** *Consider the 3-group distribution where each group has the same normalized density function and with respective population masses  $m, n, n$  ( $m > n > 0$ ). Then, a population mass transfer from the big group to the smaller ones by the same amount without altering the size rank of the groups will not decrease polarization.*

The intuition behind this axiom is straightforward. In the process of transferring population mass from the big group to the smaller ones, the groups become gradually similar, thus equating their relative forces and increasing the tension between them. It seems reasonable to say that, other things being equal, a distribution with three equally populated and equidistant groups is more likely to stir conflict than another one in which one of the groups happens to be more populated than the other two.

**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 well-being 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.

## 2.3 Characterization theorems (1).

**Theorem 1.** *A polarization measure as defined in (6) satisfies axioms 1, 2, 3 and 4 if and only if it is proportional to*

$$P_{N,\alpha}^b(\mathbf{f}) = \sum_{i=1}^N \sum_{j \neq i} M_i^\alpha \int \int f_i(x) f_j(y) (x+y) dy dx \equiv \sum_{i=1}^N \sum_{j \neq i} \pi_i^{1+\alpha} \pi_j (\mu_i + \mu_j) \quad (7)$$

where  $\alpha \in (0, 1]$  and  $\mu_i$  is the mean value of the radicalism distribution  $f_i(x)$ .

This theorem shows that, under the assumptions specified in the model, some mild restrictions completely specify the functional form of  $T(i, a)$ . The polarization index defined in (7) is a generalization of the discrete polarization measure  $DP(\alpha, k)$ , where the absence of an alienation component has now been substituted by a much richer structure that is sensitive to individuals' radicalism distribution. In this sense, the new measure enlarges the previous ones to a richer framework. Clearly, if all  $\mu_i$  happen to be the same,  $P_{N,\alpha}^b(\mathbf{f})$  reduces to the  $DP(\alpha, k)$  index.

What happens with  $P_{N,\alpha}^b(\mathbf{f})$  when  $\alpha = 0$ ? Rewriting expression (7), we would obtain that

$$P_{N,0}^b(\mathbf{f}) \equiv \sum_{i=1}^N \sum_{j \neq i} \pi_i \pi_j (\mu_i + \mu_j).$$

This formula could be interpreted as a weighted fractionalization index<sup>2</sup>, where each weight equals  $\mu_i + \mu_j$ . In particular, when  $\mu_i = \mu_j \forall i \neq j$ ,  $P_{N,0}^b(\mathbf{f})$  reduces to the classical fractionalization index. Since this is *not* what a polarization measure is intended to measure, the value of  $\alpha$  is required to be strictly above 0. As is well known, the value of  $\alpha$  has to be treated as the degree of polarization sensitivity (see ER(1994)) and the larger is its value, the greater is the departure from inequality measurement. However, the fact that  $\alpha$  can approach as much as desired the value of 0 is uncomfortable, since it seems as if the boundaries between polarization and other measures were somewhat fuzzy and not clearly delimited in this context. One possible way of raising this lower bound is to impose the following reasonable axiom.

---

<sup>2</sup>Recall that the fractionalization index is defined as  $FRAC = \sum_i \sum_{j \neq i} \pi_i \pi_j$ . It is interpreted as the probability that two randomly selected individuals from a given population will not belong to the same group, so it has usually been used as a measure of population heterogeneity.

**Axiom 5.** *Consider a symmetric configuration in which all density functions  $f_i(\cdot)$  are the same basic density and in which all population shares are equal. Then,  $P_{N_1}(f) \leq P_{N_2}(f)$  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. 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-Querol (2002, 2005a, 2005b), 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. Imposing this mild restriction, one obtains the following theorem.

**Theorem 2.** *A polarization measure as defined in (6) satisfies axioms 1,2,3,4 and 5 if and only if it is proportional to*

$$P_{N,1}^b(\mathbf{f}) = \sum_{i=1}^N \sum_{j \neq i} M_i \int \int f_i(x) f_j(y) (x+y) dy dx \equiv \sum_{i=1}^N \sum_{j \neq i} \pi_i^2 \pi_j (\mu_i + \mu_j) \quad (8)$$

In this case, the only possible value of  $\alpha$  is reduced to 1. It must be pointed out that fixing the admissible values for  $\alpha$  at a single value simplifies matters greatly for any empirical application. Clearly, this is a generalization of the  $RQ$  index: when  $\mu_i = \mu_j \forall i \neq j$ ,  $P_{N,1}^b(\mathbf{f}) \equiv RQ$ . The social polarization index we have axiomatically characterized is appealing for different reasons. From one side it keeps a simple and intuitive functional form, which makes it easier to understand and to implement empirically. From the other side, it incorporates in a simple way the alienation component which was lacking in  $DP(\alpha, k)$  or  $RQ$ . Hence,  $P_{N,1}^b$  is sensitive not only to the size of the competing groups but also to the existing feeling of animosity between them.

### 3 Measuring social polarization between and within groups

In the previous section, we presented a model in which alienation was assumed to be a between group phenomenon only. However, one might well

argue that alienation can be felt between members within the same group. Take the case, for instance, of a group sharply divided in two subgroups: those who are less radical and those who are more radical. It might be reasonable to say that the social tension existing within this group should affect somehow overall social tension. For this reason, in this section we introduce an alternative model in which room is made for within group alienation.

In this new model, we keep some of the ideas and tools which were introduced in the previous section. In particular, we still assume that the population is partitioned in  $N \geq 2$  groups and that each individual has a radicalism degree of  $x \in \mathbb{R}_+$ . In each group we have a density function  $f_i(x)$  for the corresponding radicalism distribution. When it comes to define alienation between individuals of radicalism degrees  $x, y$  respectively, we distinguish two cases. In the first case, if individuals belong to different groups it is reasonable to assume, as before, that alienation should be monotonically increasing in  $x+y$ . The motivation for this choice is exactly the same, namely: the higher (lower) the involvement of individuals in pursuing the interests of their own groups, the higher (lower) the level of social tension. In the second case, if individuals belong to the same group, it seems reasonable to impose that alienation between them should be monotonically increasing in  $|x - y|$  (this is analogous to the assumption made in the context of income polarization). Under these assumptions, we are implicitly asserting 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 other groups, which in many cases seems to be a plausible hypothesis. Concerning identification, perhaps the most reasonable identification function for an individual belonging to group  $i$  with radicalism degree  $x$  should be  $f_i(x)$ . This choice is plausible because in this model, each individual feels alienated towards all individuals of other groups and all individuals of the same group but with different radicalism degree. Hence, the amount of individuals which are exactly like him/her is simply  $f_i(x)$ . Under these assumptions, total polarization as in equation (5) can be rewritten as

$$P_N(\mathbf{f}) = \sum_{i=1}^N \int \int T(f_i(x), |x-y|) f_i(x) f_i(y) dy dx + \sum_{i=1}^N \sum_{j \neq i} \int \int T(f_i(x), x+y) f_i(x) f_j(y) dy dx \quad (9)$$

Recall that the first component of this equation is the contribution of polarization within groups and that the second component corresponds to

polarization between groups. The index presented in equation (9) is too vague and cannot be used in practice. As before, we present a set of axioms that pins down an explicit and operational formulation of the polarization index.

### 3.1 The axioms (2).

We will use the same concepts and notation as before, namely: basic densities, roots and its transformations. In particular, we will also use the concept of  $\lambda$ -squeeze of a basic density introduced in DER. Given any basic density  $g$  with mean  $\mu$  and  $\lambda \in (0, 1]$ , a  $\lambda$ -squeeze of  $g$  is defined as the mean-preserving transformation

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

Recall that  $g^\lambda$  is nothing but a compression of  $g$  around its mean. Moreover, we will need to introduce the following sets. Define

$$\Delta_N := \{(\pi_1, \dots, \pi_N) \in \mathbb{R}_+^N \mid \sum_i \pi_i = 1\}$$

the standard simplex in  $\mathbb{R}^N$  and

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

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:

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

for some  $\epsilon > 0$ ,  $\|\cdot\|$  being the Euclidean norm.

**Axiom 6.** *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 as in Figure 3. Consider, moreover, that the population shares vector  $\boldsymbol{\pi} \in \mathcal{B}(\epsilon)$  for some arbitrarily small  $\epsilon > 0$ .*

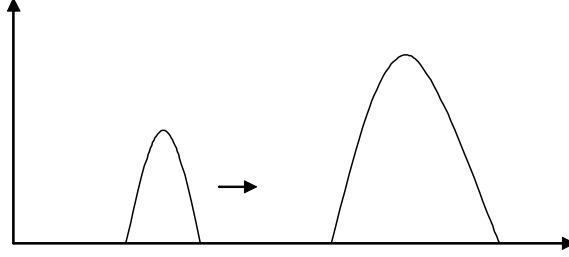


Figure 3: An inner basic density approaching an outer basic density within all groups.

*Slide the inner basic densities outwards, while keeping supports disjoint. Then polarization must increase.*

This axiom 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 1 would decrease polarization rather than increasing it because of the negligible effect of small groups on the final result.

**Axiom 7.** *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 as in Figure 4. Assume that the population shares are exactly the same for all groups (i.e.:  $\pi_i = 1/N$  for all  $i$ ). Then, if all outer distributions are squeezed, polarization must increase.*

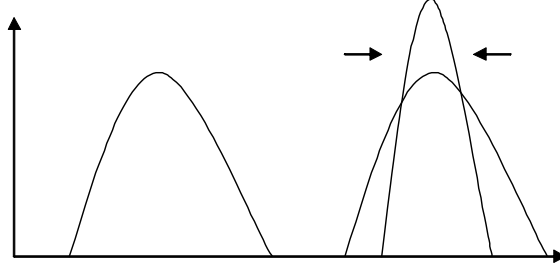


Figure 4: A  $\lambda$ -squeeze of the outer basic densities.

The intuition behind this axiom is the following. Assume that in each group we have two subgroups, one consisting of low radicalized individuals and another with highly radicalized ones. Then, when the most radicalized individuals in each group are made more homogeneous or cohesive, they tend to have a higher voicing power, so polarization is expected to rise.

**Axiom 8.** *Consider a distribution having two equally populated groups with the same basic density  $f(x)$ . Then, if we shift mass from one of the groups to the other, polarization should decrease.*

This axiom captures the intuitive idea that any departure from the equally weighted bipolar case should decrease polarization.

### 3.2 Characterization theorems (2).

**Theorem 3.** *A polarization measure as defined in (9) satisfies axioms 4, 6, 7 and 8 if and only if it is proportional to*

$$P_{N,\alpha}(\mathbf{f}) = \sum_{i=1}^N \int \int f_i^{1+\alpha}(x) f_i(y) |x-y| dy dx + \sum_{i=1}^N \sum_{j \neq i} \int \int f_i^{1+\alpha}(x) f_j(y) (x+y) dy dx \quad (10)$$

where  $\alpha \in [\frac{1}{3N-2}, 1]$ .

This theorem characterizes axiomatically a social polarization index that takes into account the existing social tensions between and within groups.



As is well known, the value of  $\alpha$  has to be treated as the degree of polarization sensitivity (see ER(1994)) and the larger is its value, the greater is the departure from inequality measurement. Looking at the statement of Theorem 3, we can check that the lower bound of  $\alpha$  is strictly positive but that it depends on the number of population groups ( $1/(3N - 2)$ ). This is an “uncomfortable” result, as it states that the set of permissible values for  $\alpha$  widens as the number of groups gets larger and that, by taking an arbitrarily large number of groups, the admissible values of  $\alpha$  could approach the non-desirable value of 0 as much as desired. One possible way of raising this lower bound is to impose axiom 5. In that case we would obtain the following characterization result.

**Theorem 4.** *A polarization measure as defined in equation (9) satisfies axioms 4, 5, 6, 7 and 8 if and only if it is of the form presented in equation (10) with the additional restriction that  $\alpha \geq 1/2$ .*

This way, by adding a reasonable assumption, we are ruling out two uncomfortable facts: 1. The dependency of the lower bound of  $\alpha$  to the values of  $N$  and 2. The fact that, for an arbitrary large number of groups, the admissible values for  $\alpha$  could approach too much the non-desirable value of 0.

At this moment, it will be interesting to compare the social polarization indices that have arisen from the two models presented in this paper and which are axiomatically characterized in Theorems 1 to 4. It can be noted that  $P_{N,\alpha}(\mathbf{f})$  and  $P_{N,\alpha}^b(\mathbf{f})$  provide different generalizations of the discrete polarization measure  $DP(\alpha, k)$ , and that each of which has its own advantages. From one side  $P_{N,\alpha}(\mathbf{f})$  is a very detailed measure that takes into account the degree to which each individual feels alienated *vis-à-vis* the other individuals of the population regardless of the whether they belong to the same population subgroup or not. However, this might be very demanding from the information availability side, because in many cases it might be very difficult to measure the exact degree of radicalism of each individual. From the other side, this problem is overcome to a certain extent by using  $P_{N,\alpha}^b(\mathbf{f})$ , where the only information that is needed is the *mean* radicalism degree of each population group. Even if this measure is less sensitive than  $P_{N,\alpha}(\mathbf{f})$  with respect to certain transformations of the radicalism degrees dis-

tributions which might be of interest<sup>3</sup>, it has the great advantage of making use of aggregated data, which, in general, will be more readily available. The formulation of  $P_{N,\alpha}^b(\mathbf{f})$  is not so cumbersome as that of  $P_{N,\alpha}(\mathbf{f})$ , so it might be intuitively easier to understand. This might prove to be very useful for any empirical application of the index.

## 4 An empirical illustration.

In this section, we will use empirical data to compare the values of the  $RQ$  index with those of  $P_{N,\alpha}^b$  and  $P_{N,\alpha}$ . For that purpose we will use data from the World Value Surveys (WVS). These surveys have been conducted in many countries all over the world since 1981, but for this empirical exercise we will only use the surveys included in the fourth wave (WVS2000) to ensure comparability. In overall, data is available for 79 countries. These surveys contain a detailed questionnaire on religious issues, so it is easy to compute an individual-level index of religious radicalism that can be used for our polarization indices (a higher level of the index corresponds to a higher involvement in religious-related issues and *vice versa*). Unfortunately, these surveys do not contain enough information to construct an ethnic polarization index<sup>4</sup>, so this interesting empirical exercise can not be carried out by the moment.

In order to measure the level of intensity of individuals' religiosity feelings we have used their reported answers to the question: "How important is God in your life?". That question has been asked in the 79 countries included in our dataset. It could be answered in a  $[0, 10]$ -scale, which we have used as a proxy of individuals' religiosity feelings.

---

<sup>3</sup>For example, the  $\lambda$ -squeeze transformation introduced in axiom 7 is a mean-preserving transformation, so it is not detected by  $P_{N,\alpha}^b(\mathbf{f})$ .

<sup>4</sup>Using the core questionnaire it is possible to find out the ethnic distribution within a given country. However, there is not enough information to derive the corresponding radicalism degree distribution, which would tell us the extent to which individuals feel identified with their ethnic group.

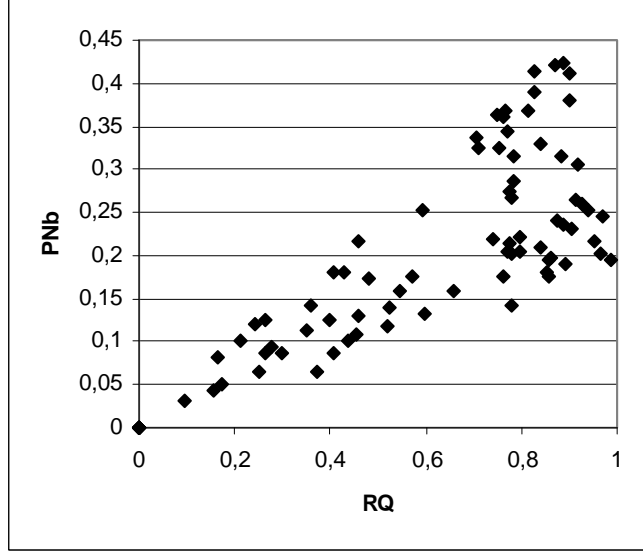


Figure 5: Scatterplot of  $RQ$  versus  $P_{N,\alpha}^b$ .

#### 4.1 Comparing religious polarization indices.

In this section we want to compare the values of  $RQ$  with the values of  $P_{N,\alpha}^b$  and  $P_{N,\alpha}$  when using WVS data at the country level. In particular, we are interested in the similarity/dissimilarity between the country rankings corresponding to the use of one index or another. Let us start by comparing  $RQ$  with  $P_{N,\alpha}^b$  : in Figure 5 we present a scatterplot of its values.

As expected, there is a positive association between the values of  $RQ$  and  $P_{N,\alpha}^b$ . However, it must be pointed out that the consistency between both rankings tends to fade away as the values of  $RQ$  increase. That is: for the set of countries which are highly polarized according to  $RQ$  (say,  $RQ$  above 0,7), the corresponding ranking according to  $P_{N,\alpha}^b$  can be very different. Clearly, this must be attributed to the role of religious intensity in  $P_{N,\alpha}^b$ . The Spearman rank correlation coefficient for the whole set of countries is 0,742. Interestingly, if we restrict our attention to the countries for which  $RQ$  is above 0,7, the Spearman rank correlation drops dramatically to -0,131. This illustrates the lack of consistency between the corresponding rankings. Let

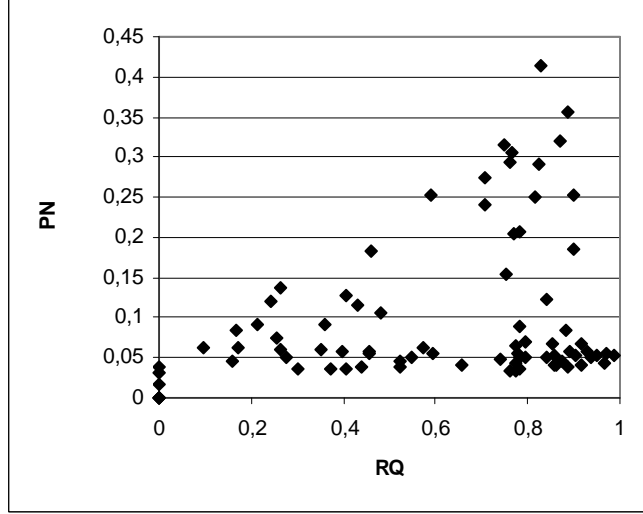


Figure 6: Scatterplot of  $RQ$  versus  $P_{N,\alpha}$ .

us now compare the values of  $RQ$  and  $P_{N,\alpha}$ : a scatterplot is presented in Figure 6.

Here too we find the expected positive relation between both indices. However, the lack of consistency between the corresponding country rankings is even more evident than before. As the values of  $RQ$  increase, the variance of  $P_{N,\alpha}$  increases too. Hence, for the set of countries which are highly polarized according to  $RQ$  (say,  $RQ$  above 0,7), the corresponding values of  $P_{N,\alpha}$  have wide variations. The Spearman rank correlation coefficient for the whole set of countries is only 0,12. Restricting our attention to the set of countries for which  $RQ$  is above 0,7, the Spearman rank correlation drops to -0,18, illustrating again the lack of consistency between both rankings.

## 5 Conclusions.

The measurement of social polarization has not received much attention from the literature. To the present date, most efforts have concentrated on the

concept of *income* polarization, but nothing is said when the causes of conflict are not related to the distribution of income or wealth. Up to now, the *RQ* index is the only one which, to our knowledge, has been introduced to measure religious or ethnic polarization (see Montalvo and Reynal-Querol 2002, 2005a, 2005b). However, this index does only take into account the sizes of the competing groups, but nothing is said about the feeling of alienation between them. Confronted with the difficult problem of defining and operationalizing a distance function of alienation between groups, the authors have opted for a simplifying compromise solution in which all groups are assumed to feel equally alienated *vis-à-vis* each other. However, it is intuitively clear that, other things being equal, the higher the feeling of animosity between groups, the higher the corresponding level of social tension and the probability of conflict.

In this paper we have presented a couple of social polarization indices that take into account the feeling of alienation between individuals. We assume that alienation depends on the degree to which individuals feel identified with their own social group. Our basic hypothesis is that the level of social tension or the probability of conflict are closely and monotonically related with individuals' feeling of involvement with their own group. In our first model, it is assumed that individuals within the same group do not feel alienated between themselves whereas in the second model, room is made to consider the existing tension within a given group. For both models we present the respective axiomatically characterized social polarization indices,  $P_{N,\alpha}^b$  and  $P_{N,\alpha}$ , which can be seen as a generalization of the *RQ* index.

From one side,  $P_{N,\alpha}$  is an interesting measure that takes into account the existing alienation within *and* between groups. However, it must be acknowledged that it is a computationally expensive measure, requiring a great amount of data at the individual level, which sometimes is difficult to obtain in certain empirical applications. From the other side,  $P_{N,\alpha}^b$  is a simpler measure which is somehow halfway between *RQ* and  $P_{N,\alpha}$ . Its greatest advantages are that it is a conceptually simple measure (it can be seen as the *RQ* index with an attached between-group distance function) which is not computationally expensive, as it only uses the *average* feeling of alienation of one group towards another. Hence, it might be a more attractive measure for empirical purposes when detailed micro data is not available.

In the empirical section, we show that the religious polarization indices

presented in this paper give very different country rankings when compared with the rankings associated with the values of  $RQ$  using data from the World Value Surveys. This means that introducing the notion of distance between groups does make an important difference when measuring the level of social polarization. Again, this points out the shortcomings of the  $RQ$  index as a social polarization index, which must be attributed to its lack of sensitivity to the existing distances between competing groups. Finally, it would be interesting to test the performance of  $P_{N,\alpha}^b$  and  $P_{N,\alpha}$  as predictors of the occurrence of Civil Wars (as is done in the paper of Montalvo and Reynal-Querol (2005a)) and compare it with the corresponding performance of the  $RQ$  index. However, this interesting research must await a future paper.

## 6 Appendix.

### 6.1 Proof of Theorem 1.

The structure of the proof is as follows: Lemmas 1 and 2 show that  $T(\iota, a)$  must be linear in  $a$  for every  $\iota > 0$ . Lemma 3 shows that  $T(\iota, a)$  must be of the form  $k\iota^{1+\alpha}a$ , for some positive constants  $\alpha, k$ . Finally, by lemma 4 we can see that  $\alpha$  must be lower than 1.

**Lemma 1.** *The function  $T(i, a)$  must be concave in  $a$  for every  $i > 0$ .*

**Proof:** We will show that axiom 1 implies that  $T(i, a)$  must be concave in  $a$  for every  $i > 0$ . For a single population group, consider the distribution shown in Figure 7 corresponding to the basic densities as in axiom 1. There are two basic densities, each of which being a transformation of a uniform basic density. The first one, containing the less radicalized individuals, is centered at  $x$ , its width is  $2\epsilon$  and its height  $h(\epsilon)(> 0)$ . Since we want the population mass of this subgroup to remain constant in a process in which  $\epsilon$  will be arbitrarily small, we must have that  $2\epsilon h(\epsilon) = h_0 \in \mathbb{R}_+$ . Moreover, one must have that  $x > 0$  and  $x > \epsilon > 0$ . The outer density, containing the more radicalized individuals, is centered at  $y$ , its width is  $2\delta$  and its height  $i(\delta)(> 0)$ . As before, since the population mass of this subgroup must remain

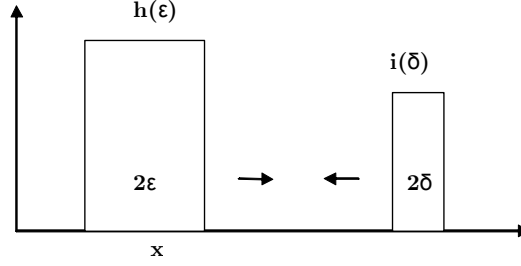


Figure 7: Two uniform densities approaching to each other within a group.

constant in a process in which  $\delta$  will be arbitrarily small, we must have that  $2\delta i(\delta) = i_0 \in \mathbb{R}_+$ . Moreover, one must have that  $y > x$  and  $y - \delta > x + \epsilon$  to ensure that the supports are disjoint. Moreover, one must have that  $h_0 \geq i_0$  to ensure that the population mass of the inner density is not smaller than the population mass of the outer one. An outer slide of the inner density corresponds to an increase of the value of  $x$ . In order to ensure that the supports remain disjoint throughout the slide (of amount  $\Delta$ ), one must have that  $y - \delta - \overline{\Delta} > x + \epsilon + \overline{\Delta}$  for some  $\overline{\Delta} > 0$ . The other population groups have a single basic density centered at  $x$ , with width  $2\gamma$  and height  $k(\gamma)(> 0)$ , with  $2\gamma k(\gamma) = k_0 \in \mathbb{R}_+$ . Since the population shares of the different groups is the same, one must have that  $k_0 = h_0 + i_0$ .

The total polarization formula as in (6) can be decomposed in terms of  $\Delta$  as follows:

$$\begin{aligned}
& (N-1) \left[ \int_{x+\Delta-\epsilon}^{x+\Delta+\epsilon} \int_{x-\gamma}^{x+\gamma} T(h_0 + i_0, b + b') h(\epsilon) k(\gamma) db' db + \right. \\
& \int_{y-\Delta-\delta}^{y-\Delta+\delta} \int_{x-\gamma}^{x+\gamma} T(h_0 + i_0, b + b') i(\delta) k(\gamma) db' db + \int_{x-\gamma}^{x+\gamma} \int_{x+\Delta-\epsilon}^{x+\Delta+\epsilon} T(k_0, b + b') h(\epsilon) k(\gamma) db' db + \\
& \left. \int_{x-\gamma}^{x+\gamma} \int_{y-\Delta-\delta}^{y-\Delta+\delta} T(k_0, b + b') i(\delta) k(\gamma) db' db \right] + C,
\end{aligned}$$

where  $C$  includes all those terms not depending on  $\Delta$ . Axiom 1 requires that  $P(\Delta) \geq P(0)$  for all  $\Delta \in (0, \overline{\Delta})$ . This is equivalent to require that  $(P(\Delta) - P(0))/(k(\gamma)(N-1)) \geq 0$ . In particular, one must have that

$$\begin{aligned}
& h(\epsilon) \left[ \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x + \Delta + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + \Delta + b + b') db' db \right] + \\
& i(\delta) \left[ \int_{\delta}^{\delta} \int_{-\gamma}^{\gamma} T(k_0, x + y - \Delta + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\delta}^{\delta} T(k_0, x + y - \Delta + b + b') db' db \right] \geq \\
& h(\epsilon) \left[ \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + b + b') db' db \right] + \\
& i(\delta) \left[ \int_{\delta}^{\delta} \int_{-\gamma}^{\gamma} T(k_0, x + y + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\delta}^{\delta} T(k_0, x + y + b + b') db' db \right].
\end{aligned}$$

In the last expression, we can consider the specific case in which  $\epsilon = \delta$ . Moreover, by continuity this expression must also hold true as  $h(\epsilon) \rightarrow i(\epsilon)$  so one must have that

$$\begin{aligned}
& \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x + \Delta + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + \Delta + b + b') db' db + \\
& \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, x + y - \Delta + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, x + y - \Delta + b + b') db' db \geq \\
& \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + b + b') db' db + \\
& \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, x + y + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, x + y + b + b') db' db.
\end{aligned}$$

Dividing both sides by  $\epsilon$  and taking  $\epsilon$  to zero and then doing the same with  $\gamma$ , the last expression is equivalent to

$$T(k_0, x' + \Delta) + T(k_0, y' - \Delta) \geq T(k_0, x') + T(k_0, y'),$$

where  $x' = 2x, y' = x + y$ . Hence, one has that

$$T(k_0, x' + \Delta) - T(k_0, x') \geq T(k_0, y') - T(k_0, y' - \Delta),$$

for any  $x' < y'$  such that  $x' + \Delta < y' - \Delta$  for some  $\Delta > 0$ . The last expression implies that  $T(k_0, a)$  must be concave in  $a$  for any  $k_0 > 0$ .

Q.E.D.

**Lemma 2.** *The function  $T(i, a)$  must be convex in  $a$  for every  $i > 0$ .*



**Proof:** We will show that axiom 2 implies that  $T(i, a)$  must be convex in  $a$  for every  $i > 0$ . Let us start by considering the following basic densities configuration. For the first group we have a uniform density centered at  $x > 0$ , with width  $2\epsilon$  and height  $h(\epsilon)$ . Since we want the population mass of this group to remain constant in a process in which  $\epsilon$  will be arbitrarily small, we must have that  $2\epsilon h(\epsilon) = h_0 \in \mathbb{R}_+$ . For the second group we have a uniform density centered at  $y > 0$ , with width  $2\delta$  and height  $i(\delta)$ . Since we want the population mass of this group to remain constant in a process in which  $\delta$  will be arbitrarily small, we must have that  $2\delta i(\delta) = i_0 \in \mathbb{R}_+$ . Moreover, one must have that  $h_0 \leq i_0$  to ensure that the population mass of the first group is not bigger than the population mass of the second one. Finally, the other  $N - 2$  population groups are assumed to have a uniform density centered at  $\min\{x, y\} > 0$ , with width  $2\gamma$ , height  $k(\gamma)$  and  $2\gamma k(\gamma) = k_0 \in \mathbb{R}_+$ . From now on, we will assume that  $x \leq y$ , so  $\min\{x, y\} = x$ . The other part of the proof, when  $x \geq y$ , is completely analogous and will not be presented here. If the amount of the slides described in axiom 2 is equal to  $\Delta$ , the total polarization formula as in equation (6) can be decomposed as

$$\begin{aligned}
(N - 2) & \left[ \int_{x-\Delta-\epsilon}^{x-\Delta+\epsilon} \int_{x-\gamma}^{x+\gamma} T(h_0, b + b') h(\epsilon) k(\gamma) db' db + \int_{y+\Delta-\delta}^{y+\Delta+\delta} \int_{x-\gamma}^{x+\gamma} T(i_0, b + \right. \\
& \quad \left. b') i(\delta) k(\gamma) db' db + \right. \\
& \quad \left. \int_{x-\gamma}^{x+\gamma} \int_{x-\Delta-\epsilon}^{x-\Delta+\epsilon} T(k_0, b + b') h(\epsilon) k(\gamma) db' db + \int_{x-\gamma}^{x+\gamma} \int_{y+\Delta-\delta}^{y+\Delta+\delta} T(k_0, b + b') i(\delta) k(\gamma) db' db \right] + \\
& \quad \int_{x-\Delta-\epsilon}^{x-\Delta+\epsilon} \int_{y+\Delta-\delta}^{y+\Delta+\delta} T(h_0, b + b') h(\epsilon) i(\delta) db' db + \int_{y+\Delta-\delta}^{y+\Delta+\delta} \int_{x-\Delta-\epsilon}^{x-\Delta+\epsilon} T(i_0, b + b') h(\epsilon) i(\delta) db' db + C
\end{aligned}$$

where  $C$  includes all those terms not depending on  $\Delta$ . Axiom 2 requires that  $P(\Delta) \geq P(0)$  for all  $\Delta \in (0, x - \epsilon)$ . This is equivalent to require that  $(P(\Delta) - P(0))/k(\gamma) \geq 0$ . In particular, one must have that

$$\begin{aligned}
& h(\epsilon) \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(h_0, 2x - \Delta + b + b') db' db + i(\delta) \int_{\delta}^{\delta} \int_{-\gamma}^{\gamma} T(i_0, x + y + \Delta + b + b') db' db + \\
& h(\epsilon) \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x - \Delta + b + b') db' db + i(\delta) \int_{-\gamma}^{\gamma} \int_{\delta}^{\delta} T(k_0, x + y + \Delta + b + b') db' db \geq \\
& h(\epsilon) \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(h_0, 2x + b + b') db' db + i(\delta) \int_{\delta}^{\delta} \int_{-\gamma}^{\gamma} T(i_0, x + y + b + b') db' db +
\end{aligned}$$

$$h(\epsilon) \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + b + b') db' db + i(\delta) \int_{-\gamma}^{\gamma} \int_{\delta}^{\delta} T(k_0, x + y + b + b') db' db$$

In the last expression, we can consider the specific case  $\epsilon = \delta$  and choose  $k_0 = i_0$ . Moreover, by continuity this expression must also hold true for any  $\epsilon > 0$  as  $h(\epsilon) \rightarrow i(\epsilon)$  (in which case  $h_0 \rightarrow i_0$ ) so one must have that

$$\begin{aligned} & \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x - \Delta + b + b') db' db + \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, x + y + \Delta + b + b') db' db + \\ & \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x - \Delta + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, x + y + \Delta + b + b') db' db \geq \\ & \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, 2x + b + b') db' db + \int_{\epsilon}^{\epsilon} \int_{-\gamma}^{\gamma} T(k_0, x + y + b + b') db' db + \\ & \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, 2x + b + b') db' db + \int_{-\gamma}^{\gamma} \int_{\epsilon}^{\epsilon} T(k_0, x + y + b + b') db' db \end{aligned}$$

Dividing both sides by  $\epsilon$  and taking  $\epsilon$  to zero and then doing the same with  $\gamma$ , the last expression is equivalent to

$$T(k_0, y' + \Delta) - T(k_0, y') \geq T(k_0, x') - T(k_0, x' - \Delta),$$

where  $x' = 2x$ ,  $y' = x + y$  ( $0 \leq x \leq y$ ), which implies that  $T(k_0, a)$  must be convex in  $a$  for any  $k_0 > 0$ .

Q.E.D.

**Lemma 3.**  $\phi(i)$  must be of the form  $ki^\alpha$ , for some positive constants  $\alpha, k$ .

**Proof:** We want to prove that  $\phi$  satisfies the fundamental Cauchy equation  $\phi(p)\phi(p') = \phi(pp')\phi(1)$  for every  $p, p' > 0$ . Let us fix  $p$  and  $p'$  and define  $r = pp'$ . From now on we will assume that  $p \geq r$ . Assume that we have the following configuration: for the first two population groups we have a uniform density centered at 1 of width  $2\epsilon$ . The heights are  $p$  and  $h$  respectively. The other  $N - 2$  population groups have negligible mass, so by continuity they can be dispensed with. In this distribution, total polarization is equal to

$$P = P_{12}^b + P_{21}^b,$$

where  $P_{ij}^b$  is the total effective antagonism felt by the individuals of population subgroup  $i$  towards the individuals of population subgroup  $j$ . Hence, total polarization is proportional to

$$\int_{1-\epsilon}^{1+\epsilon} \int_{1-\epsilon}^{1+\epsilon} \phi(p) (b' + b) ph db' db + \int_{1-\epsilon}^{1+\epsilon} \int_{1-\epsilon}^{1+\epsilon} \phi(h) (b' + b) ph db' db$$

After some computations, total polarization can be rewritten as

$$P = 4\epsilon^2 ph(\phi(p) + \phi(h)).$$

We can now change the height of the densities centered at 1 to  $r$ . One can check that for any  $\epsilon$ , there exists a height  $h(\epsilon)$  for the densities centered at 2 such that the polarizations of the two configurations can be equated. Hence, using the last expression of  $P$  we can say that

$$ph(\phi(p) + \phi(h)) = rh(\epsilon)(\phi(r) + \phi(h(\epsilon)))$$

By axiom 4, for all  $\lambda > 0$ ,

$$\lambda^2 ph(\phi(\lambda p) + \phi(\lambda h)) = \lambda^2 rh(\epsilon)(\phi(\lambda r) + \phi(\lambda h(\epsilon))).$$

Notice that, as  $\epsilon \rightarrow 0$ ,  $h(\epsilon) \rightarrow h'$  for some  $h' > 0$ . By continuity of  $\phi$ , when we pass to the limit, the last two equations can be rewritten as

$$ph(\phi(p) + \phi(h)) = rh'(\phi(r) + \phi(h'))$$

and

$$\lambda^2 ph(\phi(\lambda p) + \phi(\lambda h)) = \lambda^2 rh'(\phi(\lambda r) + \phi(\lambda h')).$$

Combining the last two expressions one has that

$$\frac{\phi(p) + \phi(h)}{\phi(\lambda p) + \phi(\lambda h)} = \frac{\phi(r) + \phi(h')}{\phi(\lambda r) + \phi(\lambda h')}.$$

Taking limits as  $h \rightarrow 0$  (therefore  $h' \rightarrow 0$ ), we have that for any  $\lambda > 0$ , the last expression reduces to

$$\frac{\phi(p)}{\phi(\lambda p)} = \frac{\phi(r)}{\phi(\lambda r)}.$$

Putting  $\lambda = 1/p$  and recalling that  $r = pp'$ , the last expression yields the required Cauchy equation. The class of solutions to this kind of equation is given by  $\phi(p) = kp^\alpha$  for some positive constants  $\alpha, k$  (see Aczél (1966)).

Q.E.D.

**Lemma 4.** *Given that  $P_{N,\alpha}^b(f)$  is of the form shown in equation (7), axiom 3 is satisfied if and only if  $\alpha \leq 1$ .*

**Proof:** Without loss of generality, we will assume that the whole population mass ( $M$ ) is normalized at 1. Hence, the population mass distribution for the three groups will be simply  $1 - 2x, x, x$ , where  $x$  denotes the amount of population mass that is transferred from the big group to the smaller ones (by the same amount). When  $x = 0$  there is only a single group, when  $x \in (0, 1/3)$  there is a big group and two smaller groups and when  $x = 1/3$  we have the uniform distribution. Recall that, since the size rank of the groups is not allowed to be reversed in the transfer process,  $x$  can *not* be greater than  $1/3$ . Polarization can be written in terms of  $x$  as:

$$P(x) = 2\mu [(1 - 2x)^{1+\alpha}x + x^{1+\alpha}(1 - 2x) + x^{2+\alpha}] \quad (11)$$

where  $\mu > 0$  is the mean radicalism degree of each group. According to axiom 6,  $P(x)$  must be a non-decreasing function in  $x$  for all  $x \in [0, 1/3]$ . Let us compute the first derivative:

$$\frac{\partial P}{\partial x} \equiv (1 - 2x)^{1+\alpha} + \alpha x^{1+\alpha} + (1 + \alpha)[x^\alpha(1 - 2x) - 2x(1 - 2x)^\alpha].$$

One must find which are the values of  $\alpha$  for which  $(\partial P / \partial x)(x)$  is non-negative for all  $x \in [0, 1/3]$ . At this moment, we define the following function:

$$f(x) := \frac{\frac{\partial P}{\partial x}(x)}{(1 - 2x)^{1+\alpha}} = 1 + \alpha \left( \frac{x}{1 - 2x} \right)^{1+\alpha} + (1 + \alpha) \left[ \left( \frac{x}{1 - 2x} \right)^\alpha - 2 \left( \frac{x}{1 - 2x} \right) \right].$$

Recall that if  $x \in [0, 1/3]$ ,  $(1 - 2x)^{1+\alpha} > 0$ , so the sign of  $(\partial P / \partial x)(x)$  is the same as the sign of  $f(x)$ . Renamig variables, we define  $z = x/(1 - 2x)$ , so  $f(x)$  can be rewritten in terms of  $z$  as

$$g(z) = 1 + \alpha z^{1+\alpha} + (1 + \alpha)[z^\alpha - 2z].$$

It is straightforward to check that for  $x \in [0, 1/3]$ , the function  $x/(1-2x)$  increases monotonically and is bounded between 0 and 1. Hence, we have to find the values of  $\alpha$  for which  $g(z)$  is non-negative for all  $z \in [0, 1]$ .

Observe that  $g(z)$  is a continuous function with  $g(0) = 1$  and  $g(1) = 0$ . Moreover, if we compute the derivative of  $g$  with respect to  $z$  we obtain

$$\frac{\partial g}{\partial z} = \alpha(\alpha + 1)z^\alpha + (\alpha + 1)[\alpha z^{\alpha-1} - 2].$$

In particular,  $(\partial g/\partial z)(1) = 2(\alpha^2 - 1)$ . This means that for  $\alpha > 1$ ,  $(\partial g/\partial z)(1) > 0$ . By continuity of  $g$ , for some arbitrarily small  $\epsilon > 0$ , one must have that  $g(1 - \epsilon) < 0$ . Hence, axiom 3 does not hold whenever  $\alpha$  happens to be strictly bigger than 1.

We will now prove that the sign of  $g(z)$  for  $z \in [0, 1]$  when  $\alpha \in (0, 1]$  can *not* be negative. Assume the contrary: that is, assume that  $g(z)$  can be negative for a certain range of values in  $[0, 1]$ . Since  $g(0) = 1$ ,  $g(1) = 0$  and  $(\partial g/\partial z)(1) \leq 0$ , by continuity of  $g$  it must follow that  $g(z)$  has at least 3 roots for  $z \in [0, 1]$ . Now, this is impossible because  $g$  is defined as a linear combination of three potential functions whose powers are between 0 and 2. Hence, we have proven that when  $\alpha \in (0, 1]$ ,  $g(z) \geq 0$ , so  $(\partial P/\partial x)(x) \geq 0$  for all  $x \in [0, 1/3]$ . This is the range of values of  $\alpha$  for which axiom 3 holds. This proves the lemma.

This completes the necessity part of the theorem. The sufficiency part is completely straightforward, so it will not be shown here (it is available upon request).

Q.E.D.

## 6.2 Proof of Theorem 2.

In theorem 1 we saw that  $P_N(\mathbf{f})$  is proportional to  $\sum_i \sum_{j \neq i} \pi_i^{1+\alpha} \pi_j (\mu_i + \mu_j)$  if and only if axioms 1, 2, 3 and 4 are satisfied. We only have to check the effect of imposing axiom 5 to this measure. According to axiom 5, the total population mass ( $M$ ) is divided into  $N$  equally populated groups. Moreover,

for any  $i \in \{1, \dots, N\}$ , the mean radicalism degree of group  $i$  equals  $\mu$ . Total polarization can be written in terms of  $N$  as

$$P(N) = \sum_i \sum_{j \neq i} \left(\frac{M}{N}\right)^{2+\alpha} 2\mu = 2\mu \left(\frac{M}{N}\right)^{2+\alpha} N(N-1)$$

What axiom 5 says is that  $P(N)$  should be a non-increasing function of  $N$ . Let us simply compute the first derivative:

$$\begin{aligned} \frac{\partial P}{\partial N} &= (2\mu M^{2+\alpha}) \frac{(2N-1)N^{2+\alpha} - (2+\alpha)(N^2-N)N^{1+\alpha}}{N^{4+2\alpha}} = \\ &= \frac{2\mu M^{2+\alpha}}{N^{3+\alpha}} ((2N-1)N - (2+\alpha)(N^2-N)). \end{aligned} \quad (12)$$

Manipulating the last expression a little bit and imposing that  $(\partial P / \partial N)(N) \leq 0$  one obtains that  $N + \alpha N(1-N) \leq 0$ . This is equivalent to impose

$$\alpha \geq \frac{1}{N-1}.$$

Since the number of groups ( $N$ ) is greater or equal than 2 and  $(N-1)^{-1}$  is a decreasing function in  $N$ , a sufficient test case is to take  $N = 2$ . In that case, one obtains that  $\alpha \geq 1$ .

The sufficiency part is straightforward, so the theorem is proved.

Q.E.D.

### 6.3 Proof of Theorem 3.

The proof of theorem 3 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 4, 6, 7 and 8 imply the functional form shown in equation (10). Then, in proving the sufficiency part, one can establish the bounds for  $\alpha$ . 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 bounds of  $\alpha$ . In Lemmas 5 and 6 we establish the upper and lower bounds of  $\alpha$  respectively.

**Lemma 5.** *Given that  $P_{N,\alpha}(f)$  is of the form shown in equation (10), axiom 8 is satisfied if and only if  $\alpha \leq 1$ .*

**Proof.** Without loss of generality, we will assume that the whole population mass ( $M$ ) is normalized at 1. The first basic density, which is denoted by  $f(x)$ , has a support  $[a, b]$  with mean  $\mu$  and root  $f^*$ . Let  $m = \mu - a$ . The second basic density is the same as the first one but translated  $d$  units away. In order to have disjoint supports, one must have that  $d > 2m$ . The amount of population mass that is transferred from one group to the other is denoted by  $\Delta$ . Using lemmas 6 and 7 in DER we can write polarization in terms of  $\Delta$  as follows:

$$P_\alpha(f, \Delta) = (4km^{1-\alpha}\psi_1(f^*, \alpha)) \left( \left( \frac{1}{2} + \Delta \right)^{2+\alpha} + \left( \frac{1}{2} - \Delta \right)^{2+\alpha} \right) + \\ 2kdm^{-\alpha}\psi_2(f^*, \alpha) \left[ \left( \frac{1}{2} + \Delta \right)^{1+\alpha} \left( \frac{1}{2} - \Delta \right) + \left( \frac{1}{2} - \Delta \right)^{1+\alpha} \left( \frac{1}{2} + \Delta \right) \right].$$

Now, according to axiom 8,  $P_\alpha(f, \Delta)$  should have a maximum at  $\Delta = 0$ . Hence, we need to compute the first and second derivatives of  $P_\alpha(f, \Delta)$  with respect to  $\Delta$ . Computing  $\frac{\partial P_\alpha(f, \Delta)}{\partial \Delta}$  we obtain

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

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

$$4km^{1-\alpha}\psi_1(f^*, \alpha)(2 + \alpha)(1 + \alpha) \left( \left( \frac{1}{2} + \Delta \right)^\alpha + \left( \frac{1}{2} - \Delta \right)^\alpha \right) + \\ 2kdm^{-\alpha}\psi_2(f^*, \alpha)(1 + \alpha)$$

$$\left[ \alpha \left( \left( \frac{1}{2} + \Delta \right)^{\alpha-1} \left( \frac{1}{2} - \Delta \right) + \left( \frac{1}{2} - \Delta \right)^{\alpha-1} \left( \frac{1}{2} + \Delta \right) \right) - 2 \left( \left( \frac{1}{2} + \Delta \right)^{\alpha} + \left( \frac{1}{2} - \Delta \right)^{\alpha} \right) \right].$$

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

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

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  $d > 2m$ , a necessary and sufficient test case to test whether  $\frac{\partial^2 P_{\alpha}(f, \Delta=0)}{\partial \Delta^2} \leq 0$  is to impose  $d = 2m$ . Moreover, by lemma 8 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  $\frac{\partial^2 P_{\alpha}(f, \Delta=0)}{\partial \Delta^2}$  as

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

Finally, we have to check which are the values of  $\alpha$  for which  $[2 + \alpha + \Gamma(\alpha - 2)] \leq 0$ . If the last restriction must hold true, one must have that

$$\alpha \leq 2 \left( \frac{\Gamma - 1}{\Gamma + 1} \right),$$

Now, since  $(\Gamma - 1)/(\Gamma + 1)$  is an increasing function in  $\Gamma$  and  $\Gamma \geq 3$ , from the last expression we deduce that  $\alpha \leq 1$ . This proves the lemma.

Q.E.D.

**Lemma 6.** *Given that  $P_{N,\alpha}(f)$  is of the form shown in equation (10), axiom 7 is satisfied if and only if  $\alpha \geq \frac{1}{3N-2}$ .*

**Proof:** Consider a configuration as given in axiom 7. Each  $f_i(x)$  is composed of the same two basic densities with disjoint support. The inner density will be denoted by  $g$ . Let  $n$  be the difference from its mean to its lower support and  $q$  its population mass. Its mean will be denoted by  $\mu$ . The outer density will be denoted by  $h$ . Let  $m$  be the difference from its mean to its



lower support and  $p$  its population mass. We will denote the distance between the means of the two basic densities (within each group) by  $d$ . Moreover, we are assuming that all groups are equally numerous, i.e:  $\pi_i = 1/N$  for all  $i \in \{1, \dots, N\}$ . Hence, one should have that  $p + q = 1/N$ . Total polarization in this configuration is simply

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

where  $P^w$  is the internal polarization within each group and  $P^b$  is the total antagonism felt between the members of any couple of groups. We will now decompose  $P^w$  and  $P^b$  in terms of  $\lambda$ , where the terms insensitive to  $\lambda$  will not be explicitly described. Clearly, one has the usual decomposition

$$P^w(\lambda) = P_g + P_h(\lambda) + P_{gh}^w + P_{hg}^w(\lambda)$$

which, by Lemmas 6 and 7 in DER can be rewritten as

$$P^w(\lambda) = P_g + 4 \left(\frac{p}{N}\right)^{2+\alpha} (m\lambda)^{1-\alpha} \psi_1(h^*, \alpha) + P_{gh}^w + 2d \left(\frac{p}{N}\right)^{1+\alpha} \left(\frac{q}{N}\right) (m\lambda)^{-\alpha} \psi_2(h^*, \alpha).$$

Recall that the definitions of  $\psi_1(\cdot, \alpha), \psi_2(\cdot, \alpha)$  were presented in lemma 10 (DER). Analogously

$$P^b(\lambda) = P_{gg}^b + P_{hh}^b(\lambda) + P_{gh}^b + P_{hg}^b(\lambda)$$

which, by lemma 7 in DER can be rewritten as

$$P^b(\lambda) = P_{gg}^b + 2 \left(\frac{p}{N}\right)^{2+\alpha} (2d + 2\mu)(m\lambda)^{-\alpha} \psi_2(h^*, \alpha) + P_{gh}^b + 2 \left(\frac{p}{N}\right)^{1+\alpha} \left(\frac{q}{N}\right) (d + 2\mu)(m\lambda)^{-\alpha} \psi_2(h^*, \alpha).$$

This way, we can write total polarization in terms of  $\lambda$  as:

$$P_{N,\alpha}(\mathbf{f}, \lambda) = C + D \left[ 2\lambda^{1-\alpha} + \psi(h^*, \alpha)\lambda^{-\alpha} \left[ \frac{qd}{pm} + (N-1) \left[ \frac{q(d+2\mu)}{pm} + \frac{2d+2\mu}{m} \right] \right] \right]$$

where  $C$  includes all those terms not depending on  $\lambda$ ,  $D$  is a positive constant equal to  $\left(\frac{1}{N}\right)^{1+\alpha} 2m^{1-\alpha} \psi_1(h^*, \alpha)$  and

$$\psi(h^*, \alpha) = \frac{\psi_2(h^*, \alpha)}{\psi_1(h^*, \alpha)}$$

Hence, in order to prove the lemma we have to show that

$$2\lambda^{1-\alpha} + \psi(h^*, \alpha)\lambda^{-\alpha} \left[ \frac{qd}{pm} + (N-1) \left[ \frac{q(d+2\mu)}{pm} + \frac{2d+2\mu}{m} \right] \right]$$

is nonincreasing in  $\lambda$  over  $(0, 1]$ . Observing the last expression, we can see that, since  $\lambda^{-\alpha}$  and  $\lambda^{1-\alpha}$  are decreasing and increasing functions in  $\lambda$  respectively (recall that  $\alpha$  is known to be smaller or equal to 1), a necessary and sufficient test case to prove the lemma is to consider values of  $q$  and  $\mu$  arbitrarily close to 0. Using the same argument, one should make  $d/m$  as small as possible: its smallest possible value is 1 (otherwise, the disjoint support hypothesis would not be satisfied). Hence, we need to show that, for every root  $h^*$ ,

$$\lambda^{1-\alpha} + (N-1)\psi(h^*, \alpha)\lambda^{-\alpha}$$

is nonincreasing in  $\lambda$  over  $(0, 1]$ . Differentiating the last expression we have that

$$(1-\alpha)\lambda^{-\alpha} - \alpha(N-1)\psi(h^*, \alpha)\lambda^{-\alpha-1}$$

must be nonpositive for  $\lambda \in (0, 1]$ . This is equivalent to impose that

$$\alpha \geq \frac{1}{1+(N-1)\psi(h^*, \alpha)}.$$

By lemma 8 in DER, we know that  $\psi(h^*, \alpha)$  attains its minimum value when  $h^*$  corresponds to a uniform distribution, and that this minimum value is 3. Hence,  $\alpha$  must be greater or equal than  $(3N-2)^{-1}$ .

Q.E.D.

## 6.4 Proof of theorem 4.

According to theorem 3, a polarization index as defined in (7) is proportional to  $P_{N,\alpha}(\mathbf{f})$  if and only if axioms 4, 6, 7, 8 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  $\pi_i$  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

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

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

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

where  $\mu$  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

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

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

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

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

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

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  $\mu$ ). 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. SPALAOORE (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] DUCLOS, J-Y, ESTEBAN J. AND D. RAY (2004): “Polarization: Concepts, Measurement, Estimation”, *Econometrica*, 72, 1737-72.
- [4] EASTERLY, W., AND R. LEVINE (1997): “Africa’s growth tragedy: policies and ethnic divisions”, *Quarterly Journal of Economics*, 112, 1203-50.
- [5] ESTEBAN J., GRADIN, C., AND D. RAY (2007): “An extension of a measure of polarization, with an application to the income distribution of five OECD countries”, *Journal of Economic Inequality*, 5 (1), 1-19.
- [6] ESTEBAN J., AND D. RAY (1994): “On the Measurement of Polarization”, *Econometrica*, 62, 819-852.
- [7] ESTEBAN J., AND D. RAY (1999): “Conflict and distribution”, *Journal of Economic Theory*, 87, 379-415.
- [8] HOROWITZ, D.L. (1985): “Ethnic groups in conflict”, Berkeley: University of California Press.
- [9] MONTALVO, J.G., AND M. REYNAL-QUEROL (2002): “Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth.” Universitat Pompeu Fabra, Economics Working Paper: No. 660, 2002.
- [10] MONTALVO, J.G., AND M. REYNAL-QUEROL (2005a): “Ethnic polarization, potential conflict and civil wars”, *American Economic Review*, 95, 796-815.
- [11] MONTALVO, J.G., AND M. REYNAL-QUEROL (2005b): “Ethnic diversity and economic development”, *Journal of Development Economics*, 76, 293-323.

- [12] PERMANYER, I. (2008): “The measurement of social polarization in a multi-group context”, Barcelona Economics Working Paper Series, Working Paper n<sup>o</sup> 341 (<http://www.barcelonaeeconomics.org/onepaper.php?id=341>).
- [13] QUAH, D. (1997): “Empirics for growth and distribution: stratification, polarization and convergence clubs”, *Journal of Economic Growth*, 2, 27-59.
- [14] RODRIGUEZ, J.G., AND R. SALAS (2002): “Extended bi-polarization and inequality measures”, Mimeo, Universidad Complutense de Madrid.
- [15] WANG, Y.Q., AND K.Y.TSUI (2000): “Polarization orderings and new classes of polarization indices”, *Journal of Public Economic Theory*, 2, 349-363.
- [16] WOLFSON, M.C. (1994): “When inequalities diverge”, *American Economic Review*, 84, Papers and Proceedings, 353-358.
- [17] WOLFSON, M.C. (1997): “Divergent inequalities: Theory and Empirical Results”, *Review of Income and Wealth*, 43, 401-421.
- [18] ZHANG, X., AND R. KANBUR (2001): “What difference do polarization measures make? An application to China”, *Journal of Development Studies*, 37, 85-98.