

**Chapter 00**  
**Comparing Polarization Measures<sup>1</sup>**

by

**Joan Esteban**

*Instituto de Análisis Económico (CSIC)*

and

**Debraj Ray**

*New York University*

**January 2010**

---

<sup>1</sup>Joan Esteban is a member of the Barcelona GSE Research Network funded by the Government of Catalonia. He gratefully acknowledges financial support from the AXA Research Fund and from the Spanish Government CICYT project n. SEJ2006-00369. Debraj Ray is grateful for support from the National Science Foundation, the Fulbright Foundation, and the Indian Statistical Institute, New Delhi. We thank Xavi Ramos for his comments and to Michelle Garfinkel for the careful editing of this paper. We are specially grateful to Peter Lambert for his deep and detailed comments on a previous version of the Chapter.

## 1. INTRODUCTION

Recent years have seen an increasing interest in the notion and measurement of polarization as an indicator of potential conflict. Early empirical studies on civil conflict focussed on the inequality of the personal distribution of income or of landownership.<sup>2</sup> More recently, with the shift of emphasis of contemporary literature from class to ethnic conflict, the initial presumption has become that ethnic “diversity” is a key factor for conflict.<sup>3</sup> However, both inequality and fractionalization indices fail to account for the fundamental role of deep social cleavages. As Horowitz (1985) noted: “a centrally focused system [with few groupings] possesses fewer cleavages than a dispersed system, but those it possesses run through the whole society and are of greater magnitude. When conflict occurs, the center has little latitude to placate some groups without antagonizing others.”

In line with Horowitz’s emphasis on significant social cleavages, measures of polarization have made their way into empirical studies of conflict. Indeed, Esteban and Ray (1994, 1999) explicitly motivated the measurement of polarization by arguing that such an index could do well as a predictor of social conflict. More recently, Montalvo and Reynal-Querol (2005a) have applied a special case of the Esteban-Ray measure (introduced in Reynal-Querol (2002)) to a cross-section study of ethnic conflict.<sup>4</sup>

But the link between polarization and conflict is not only an empirical regularity, it is also founded on a behavioral model of conflict. Esteban and Ray (1999) discuss the connections between polarization and equilibrium conflict in a model of strategic behavior, explicitly obtaining a measure of polarization from it. In similar fashion, Montalvo and Reynal-Querol (2005b) derive a measure of polarization from a rent-seeking game, while Esteban and Ray (2010) provide a general model that links the level of equilibrium conflict to the fractionalization, Gini and polarization indices, and discuss when each index is likely to be more relevant.

A number of polarization indices have been proposed. The purpose of this paper is to provide a systematic framework against which to locate these alternative polarization measures. The

---

<sup>2</sup>Sen (1972) asserts that “the relation between inequality and rebellion is indeed a close one”.

<sup>3</sup>Alesina et al. (2003), Collier and Hoeffler (2004), Easterly and Levine (1997), Fearon and Laitin (2003), and Miguel, Satyanath and Sergenti (2004) have all used measures of ethnic “fractionalization”.

<sup>4</sup>See also the special issue of the *Journal of Peace Research* edited by Esteban and Schneider (2008), entirely devoted to the links between polarization and conflict.

term itself is of fairly recent use in Economics, but has a long tradition in Political Science. Yet, its precise meaning has remained somewhat ambiguous. In a very broad sense, there is agreement that polarization is designed to capture separation or distance across clustered groups in a distribution. But this broad feature has been specifically implemented in different ways.

One family of measures attempts to describe this separation/clustering over distributions with arbitrary numbers of groupings. The measures of Esteban and Ray (1991, 1994), Duclos, Esteban and Ray (2004) and Reynal-Querol (2002) belong to this family. A second set of measures treats polarization as fundamentally a *two-group* phenomenon, with median income as the divide. This family includes Foster and Wolfson (1992, 2009), Wolfson (1994) and Wang and Tsui (2000).

We begin, then, by drawing a basic distinction between “polarization” and “bipolarization” (Section 2). In Section 3 we introduce a series of properties that capture features that a measure of polarization might be required to possess. These properties are the different axioms that have been put forward by Esteban and Ray (1991, 1994), and Duclos, Esteban and Ray (2004), on one side, and by Foster and Wolfson (1992, 2009) and Wang and Tsui (2000) on another. In both cases the axiomatization of the measure is but partial, as they depend on restricting the measure to belong to a pre-specified class of indices. We do not unpack these black-boxes, but consider three complementary properties or desiderata that the measures might be expected to satisfy. In Section 4 we focus on these polarization measures and check the subset of properties that each measure satisfies or fails to satisfy. The analysis therefore provides a description of what these measures have in common and where they diverge.

It is important to clarify that in this paper, we exclusively focus on uni-dimensional polarization (the leading example being income polarization). The *same* characteristic — say income — is used to define “groups”, within-group dispersion, and across-group distances. It is true that the measures can also be applied to situations in which the groups are given by some other characteristic — such as religion, race or kin — with inter-group “distances” measured by income differences. However, we do not claim that all the axioms listed below apply to this case. Multidimensional measures present new features of interest, which we do not address in the main part of this paper. However, in Section 5, we do discuss multidimensional polarization and attempt to explain why a new approach may be needed for this case.

## 2. POLARIZATION: TWO VIEWS

As mentioned in the Introduction, existing measures of unidimensional polarization can be usefully classified into two broad families. One is designed to capture the formation of any arbitrary number of poles. We shall call them measures of *polarization* proper. The second family sees polarization as the process by which a distribution becomes *bi-polar*. We shall call them measures of *bipolarization*.

Polarization measurement starts with Esteban and Ray (1991), Esteban, Gradín and Ray (1998, 2007) and Duclos, Esteban and Ray (2004) — DER henceforth — for continuous distributions and with Esteban and Ray (1994) — ER — for discrete distributions.

Measures of bipolarization begin with the work of Wolfson (1994) and (1997), based on Foster and Wolfson (1992, 2009), and by Wang and Tsui (2000) — WT thereafter.<sup>5</sup>

Assume (as we do throughout this paper) that we are in the unidimensional case, so that the same characteristic (say, income) does triple-duty in defining a group, measuring inter-group distance *and* measuring within-group dispersion. Then there are three properties<sup>6</sup> that both families seem to consider to be indispensable to a measure of polarization:

- (i) polarization is a matter of groups so that the contribution of isolated individuals should be negligible,
- (ii) with two or more groups, polarization rises when “within-group dispersion” is reduced, and
- (iii) polarization rises when “across-group” distances increase.

Notice that the second claim runs directly against the ordering over distributions generated by second order stochastic dominance. It is plain that the notion of polarization is distinctly different from that of inequality in as much as we require inequality measures to be consistent with Lorenz curve orderings.

---

<sup>5</sup>The measure put forward by Zhang and Kanbur (2001) is a multidimensional index, and we discuss it separately in Section 5.

<sup>6</sup>See Esteban and Ray (1994) for an extensive motivation of these ideas, illustrated with numerous examples.

In order to have a precise idea of the distinguishing properties of each polarization measure, we shall examine whether they satisfy a series of axioms intended to capture essential features of polarization. In the next section we present the collection of axioms.

### 3. AXIOMS FOR POLARIZATION

Except when explicitly stated we shall work with distributions  $F$  in  $\mathfrak{R}_+$  with continuous density functions  $f$ .

We choose to present the different properties of polarization using the approach of DER where the statements of the Axioms are referred to distributions composed of “basic densities”.

**3.1. Basic Densities.** The properties — axioms — we shall deal with will largely be based on domains that are unions of one or more symmetric “basic densities.” The densities will be scaled down and up as they may correspond to varying populations. The building block for these densities we will call *kernels*. These are symmetric, unimodal density functions  $f$  with compact support on the interval  $[0, 2]$ . Their mean is one. By symmetry we mean that  $f(x) = f(2 - x)$  for all  $x \in [0, 1]$ , and by unimodality we mean that  $f$  is nondecreasing on  $[0, 1]$ . We normalize the overall population of a kernel to be unity.

A kernel  $f$  can be *population scaled* to any population  $p$  by multiplying  $f$  pointwise by  $p$  to generate a new density  $pf$ . Likewise, any kernel (or density) can undergo a *slide*. A *slide to the right* by  $\delta$  is just a new density  $g$  such that  $g(x) = f(x - \delta)$ . Similarly for a slide to the left. Further, a kernel  $f$  can be *income scaled* to any new mean  $\mu$  as follows:  $g(x) = (1/\mu)f(x/\mu)$  for all  $x$ . This will produce a new density  $g$  with support  $[0, 2\mu]$  and mean of  $\mu$ . Any scaling or slide (or combinations thereof) of a kernel we will call a *basic density*.

We shall also use the notion of a *squeeze*. Let  $f$  be any density with mean  $\mu$  and let  $\lambda$  lie in  $(0, 1]$ . A  $\lambda$ -*squeeze* of the density  $f$  is a transformation of this density as follows:

$$(1) \quad f^\lambda(x) \equiv \frac{1}{\lambda} f\left(\frac{x - [1 - \lambda]\mu}{\lambda}\right).$$

A  $\lambda$ -squeeze is a very special type of *second order stochastic dominance* transformation that “globally contracts” the support of a distribution towards its mean. This has to be contrasted with arbitrary, unrestricted progressive Dalton transfers which can concentrate density around any points in the support of the distribution.

$\lambda$ -squeezes have the following properties:

[P.1] For each  $\lambda \in (0, 1)$ ,  $f^\lambda$  is a density.

[P.2] For each  $\lambda \in (0, 1)$ ,  $f^\lambda$  has the same mean as  $f$ .

[P.3] If  $0 < \lambda < \lambda' < 1$ , then  $f^\lambda$  second-order stochastically dominates  $f^{\lambda'}$ .

[P.4] As  $\lambda \downarrow 0$ ,  $f^\lambda$  converges weakly to the degenerate measure granting all weight to  $\mu$ .

3.2. **Axioms.** Esteban and Ray (1991, 1994), and Duclos, Esteban and Ray (2004), and Foster and Wolfson (1992, 2009), Wolfson (1994) and Wang and Tsui (2000) provide axioms for polarization and bipolarization, respectively.<sup>7</sup> As it turns out, although the spirit of the two sets of axioms is quite similar, they have significantly different implications. In order to stress the similarity in spirit of the different axioms, we shall present the WT counterparts —when they exist— with the same number as in DER, followed by a B for bipolarity, as in 2B, for instance.

It is important to note that the complete axiomatizations by ER and DER, on one side, and by WT, on the other, both make use of an initial restriction that directly begins with a class of measures, different in each case. It is an open problem to disentangle how the initial restriction can in turn be broken down into a collection of simpler statements. We shall instead put forward three desiderata that permit a clear distinction between the different families of polarization measures.

We start by presenting the axioms in DER and in WT. Both papers contain a detailed motivation for these axioms and hence we simply refer the interested reader to the original works for further discussion.

AXIOM 1. If a distribution is composed of a *single* basic density, then a squeeze of that basic density cannot increase polarization.

Axiom 1 is quite uncontroversial. A squeeze, as defined here, corresponds to a compression of the entire basic density towards its mean and we must associate this to no higher polarization.

---

<sup>7</sup>See also Rodriguez and Salas (2003) and Bossert and Schworm (2008).

That said, the axiom does carry bite in narrowing the class of measures. DER motivate their intuition for polarization as the outcome of the interplay between the individual sense of *group identification* combined with the feeling of *inter-personal alienation*. From this point of view, it is clear that Axiom 1 will generate some interesting restrictions on the measurement of polarization. This is so because, on the one hand, a squeeze creates a reduction in inter-individual alienation but, on the other, also serves to raise identification for a positive measure of agents — those located “centrally” in the distribution. The implied restriction is, then, that the positive impact of the latter on polarization must be counterbalanced by the negative impact of the former.

Our next axiom attempts to capture property (ii) above. It considers a situation composed of three disjoint basic densities, derived from identical kernels. The overall distribution is completely symmetric, with densities 1 and 3 having the same total population and with density 2 exactly midway between densities 1 and 3. We also assume that all supports are disjoint. (Note that these additional restrictions make the resulting axiom *weaker*, not stronger.)

AXIOM 2. If a symmetric distribution is composed of three basic densities drawn from the same kernel, with mutually disjoint supports, then a symmetric squeeze of the *side* densities cannot reduce polarization.

This axiom argues that a particular “local” squeeze (as opposed to the “global” squeeze of the entire distribution in Axiom 1) must not bring polarization down. Here we explicitly depart from inequality measurement as it would predict that these local squeezes would reduce inequality, along with the global squeeze.

Property (ii) has been specified by Foster and Wolfson (1992) and by Wang and Tsui (2000) somewhat differently. Bearing in mind that their aim is to measure bipolarization, they take the median income  $m$ ,  $F(m) = \frac{1}{2}$ , as the reference point. Their specification of property (ii) is:

AXIOM 2B. Let two distinct distributions  $F$  and  $G$  have the same mean and the same median and let  $F$  second order stochastically dominate  $G$  separately on  $[0, m]$  and on  $[m, \infty)$ . Then  $F$  should be more bi-polarized than  $G$ .

This is their Axiom 2 (“increased bipolarity”). Notice how Axiom 2B is defined with respect to increased bimodality in the distribution. In contrast, Axiom 2 applies when there is clustering within certain parts of the distribution, but there is no presumption that we are heading “towards” a two-pole distribution.

Our third axiom attempts to capture the idea behind property (iii) concerning an increase in “across-group” inequality. As in Axiom 2 we wish to specify this axiom in the least demanding way. To this effect we restrict to symmetric distributions composed of *four* non-overlapping basic densities, once again all generated by the same kernel.

AXIOM 3. Consider a symmetric distribution composed of four basic densities drawn from the same kernel, with mutually disjoint supports. An equal slide of the two inner densities outwards towards the outer densities makes polarization go up.

Once again Foster and Wolfson (1992, 2009) and Wang and Tsui (2000) present this idea with an additional emphasis on bimodality. Their axiom asserts that if in a distribution everyone’s income is shifted farther away from the *median income*, bipolarization goes up. For continuous distributions their axiom reads as:

AXIOM 3B. Consider two distinct distributions with the same mean and median such that  $|m - F^{-1}(p)| \leq |m - G^{-1}(p)|$ , for every  $p \in [0, 1]$ , then  $G$  has more bipolarization than  $F$ .

Axioms 3 and 3B bear the same relationship to each other as their counterparts Axioms 2 and 2B. Axiom 3 applies to all sorts of changes, while Axiom 3B uses the median income as a reference point. At the same time, Axiom 3 applies to only a specific class of distributions. Like Axioms 2 and 2B, the two axioms are close in spirit, but of course they are not identical.

However, the two sets of axioms can be nested:

OBSERVATION 1. *Axioms 2B and 3B imply Axioms 1–3.*

This perhaps startling and certainly significant result is easily seen. First, consider consider Axiom 1. Note that a global  $\lambda$ -squeeze of a *symmetric* basic density cannot alter its mean nor its median. It is easy to check that such a squeeze satisfies the conditions of Axiom 3B: think of  $G$  as the distribution before the squeeze, and of  $F$  as the distribution after the squeeze. Thus given Axiom 3B, Axiom 1 must be satisfied. Now consider Axiom 2. By the assumed symmetry of



the distribution in that Axiom, the median must lie in the middle density, and in particular the squeeze of the side densities affects neither the mean nor the median of the distribution. Therefore the change satisfies the conditions of Axiom 2B: second-order stochastic dominance occurs on either side of the median. By Axiom 2B and the assumed continuity of the measure, polarization must go up. Finally, consider Axiom 3. Again, use symmetry to argue that the mean and median of the distribution lie between the two inner subdensities, so that the pulling apart of those two subdensities has no effect on the mean and median. It is now easy to see that all the conditions of Axiom 3B are satisfied. Using that Axiom, the change must increase polarization, so Axiom 3 holds. This completes the proof.

Our fourth axiom is a simple population-invariance principle. It states that if one distribution is more polarized than another, it must continue to be so when the populations in both situations are scaled up or down by the same amount, leaving all (relative) distributions unchanged.

AXIOM 4. Let  $F$  and  $G$  be two distributions with possibly different, unnormalized populations such that  $P(F) \geq P(G)$ . Then, for all  $\kappa > 0$ ,  $P(\kappa F) \geq P(\kappa G)$ , where  $\kappa F$  and  $\kappa G$  represent (identical) population scalings of  $F$  and  $G$  respectively.

WT (see also Chakravarty et al. (2007), Chakravarty and D'Ambrosio (2009) and Lasso de la Vega and Urrutia (2006)) propose a scale invariance axiom that serves a similar purpose as Axiom 4. However, while Axiom 4 posits invariance of polarization rankings with respect to population, Wang and Tsui require polarization invariance with respect to a certain type of income scaling.<sup>8</sup> That is:

AXIOM 4B. Let the distributions  $F$  and  $G$  have the same median income  $m$ . Then if  $G$  is more polarized than  $F$ , so is  $G'$  relative to  $F'$ , where  $G'$  and  $F'$  are obtained from  $G$  and  $F$  respectively by dividing all incomes by their median.

---

<sup>8</sup>Their invariance requirement is obviously satisfied by invariance with respect to scale transforms of incomes,  $kx$ . Chakravarty et al. (2007) examine the case of *absolute* bipolarization indices that are invariant to uniform absolute increases of incomes,  $x + k$ , and Chakravarty and D'Ambrosio (2009) and Lasso de la Vega and Urrutia (2006) study the invariance relative to *intermediate* transforms of incomes,  $x + \lambda(\delta x + (1 - \delta)x)$ . These papers propose bipolarization measures that satisfy the invariance requirement.

The last set of restrictions made in this literature concerns the initial restrictions. Think of these as axioms that posit that polarization indices must belong to a particular class. In the ER and DER formulations, we have:

INITIAL RESTRICTION 1. Let  $F$  be a distribution in  $\mathfrak{R}_+$  with a density  $f$ . Then

$$(2) \quad P(F) = \int \int T(f(x), |x - y|) f(x)f(y) dx dy,$$

where  $T$  is some function increasing in its second argument and with  $T(0, a) = T(i, 0) = 0$ .

That is, overall or “aggregate” polarization has to be conceived as the sum of all inter-personal *effective antagonisms*. The interpersonal antagonism  $T$  of a person located at position  $x$  with respect to one located at  $y$  is assumed to result from her own sense of *identity*  $i$  — which in turn depends on the group size  $f(x)$  — and from inter-personal *alienation*  $a$  — which depends on income distance  $|x - y|$ . Hence,

$$(3) \quad T(i, a) = T(f(x), |x - y|).$$

Wang and Tsui (2000) in turn also restrict polarization indices to belong to a specific additive class:

INITIAL RESTRICTION 1B. The polarization measure should be of the form

$$(4) \quad P(F) = \frac{1}{m} \int a(F(x)) x f(x) dx.$$

for some function  $a(\cdot)$  and where  $m$  is the median of  $F$ .

Axioms 1 to 4 — including their B counterparts — consist of simple assertions concerning changes in the distribution. In contrast, Initial Restrictions 1 and 1B are a bit of a black-box. We shall not attempt to decompose them into a collection of simpler statements. We shall instead posit three additional properties that will permit us to identify the differential behavior of the various indices.

**3.3. Additional Properties.** The properties that follow are by no means cast in stone and not to be viewed as axioms. They are possible features that a polarization index might possess, and are simply to be used to obtain a feel for the different measures.

Begin with the leading example that motivates the difference between polarization and inequality. An increase in concentration around the mean of a distribution will reduce both inequality and polarization, while if this concentration takes place around several poles polarization could increase — though inequality still decreases. Our Axiom 1 captures the first type of concentration around the mean income. Axiom 2 captures the second kind of concentration. In what follows, we introduce yet another class of local concentrations to which an analogue of Axiom 2 should reasonably apply.

Suppose that we partition the (bounded) support of a distribution in a number of intervals and that we concentrate the population over incomes within each interval. To avoid ambiguities, it is good to be sure that within each interval population is concentrated around the center (of that interval). This can be achieved by subjecting the “mini-distribution” in each interval to a  $\lambda$ -squeeze. Nevertheless, there is still the possibility that the shape of the mini-distributions may create additional ambiguities, so we will restrict the shape of the distribution as well.<sup>9</sup>

The following case seems to be quite uncontroversial. Take a uniform distribution over some bounded support  $[a, b]$ . Consider now a partition of the support into  $n$  intervals of equal length  $\frac{b-a}{n}$ . Thus, the entire distribution can be seen as  $n$  identical uniform densities with non-overlapping support, each one with a mass  $\frac{1}{n}$ . Suppose that every such uniform density is subject to a  $\lambda$ -squeeze. Intuition clearly says that polarization should increase since we have moved from a uniform distribution — without any pole — to a symmetric, clustered distribution around  $n$  poles or local modes.

The situation described in Axiom 2 is similar to this one. The three basic densities assumed there can be taken to be uniform. However, for the case of three intervals, the assertion in Axiom 2 is somewhat stronger: there only the outer distributions are subject to a squeeze, while here we are squeezing all the three densities. Likewise, in the domain of Axiom 2B, it is only the squeezing of the two outer intervals that is fully compatible with that axiom. The squeezing of all three densities would entail a transfer of income across the median, and the

---

<sup>9</sup>As an example, think of a symmetric distribution with four local modes, and partition it into two intervals at the midpoint of all incomes. A  $\lambda$ -squeeze applied to each interval will move its two modes closer together. The overall effect is genuinely ambiguous as far as overall polarization is concerned: the two central modes move away, while the two extreme modes come closer together. We wish to avoid such cases, as indeed we do in the Axioms when we study unimodal kernels.

conditions of Axiom 2B would not apply. Therefore we have a (presumably) uncontroversial property which could narrow the class of measures somewhat. We state this restriction as

PROPERTY 1. Consider a uniform distribution with support  $[a, b]$ . Partition this support into  $n$  intervals of equal length. Then a  $\lambda$ -squeeze applied to each of the  $n$  uniform sub-densities should not decrease polarization.

Our next property suggests that polarization may be expected to behave in a nonmonotonic way, under a particular class of changing circumstances. Consider a situation similar to the one contemplated in Axiom 3 with a symmetric distribution consisting of four non-overlapping basic densities. Instead of moving the location of the basic densities as in that axiom, symmetrically transfer population mass — a bit at a time — from the inner densities towards the outer densities. We thus start with most of the total population located in the inner densities and end with most of the population located in the outer densities. Clearly, the final distribution is more polarized than the initial one.

However, the interesting question concerns the behavior of polarization through the entire range of transfers. At the very “beginning” of this process, there are just two groups (the inner densities). The start of the process thereby moves us away from a two-group society to a four-group society, as the outer densities begin to grow. To the extent that a smaller number of groups might contribute positively to polarization, we may wish to describe the initial movements of population as *polarization-reducing*. But, of course, once the outer densities have built up “sufficiently”, the increase in alienation must eventually overcome any loss in identification and polarization will “ultimately” be higher than at the start. We formalize these ideas in the following nonmonotonicity property:

PROPERTY 2. Consider a symmetric distribution like the one described in Axiom 3, with four non-overlapping basic densities deriving from the same kernel. For some such configuration, polarization should not vary monotonically with the size of the inner densities.

This is an extremely weak statement which simply asks for there to be *some* four-group symmetric situation under which a polarization measure exhibits nonmonotonicity with respect to the transfer of population from inner to outer groups. But it is of interest to note that such a

property cannot be compatible with Axiom 3B. After all, in the scenario contemplated by Property 2, the sequence of distributions generated by transferring population from the inner to the outer densities satisfies the conditions of Axiom 3B—the distance with respect to the median has been uniformly increased. Thus Axiom 3B requires the sequence of distributions to display monotonically increasing polarization.

We do not point this out to necessarily reject Axiom 3B, but only to show that that axiom is very strongly focused on bipolar movements away from the median. It is unimportant (from the point of view of that axiom) whether there are some processes which are both “bi-polarizing” while at the same time creating new clusters, thereby fragmenting society. Such fragmentation may well reduce tensions, because it lowers deep cleavages. At any rate, this is what Property 2 attempts to capture.

Our third property formalizes the idea that polarization is entirely a feature of groups and inter-group distances, and that the underlying variable has, in a sense, no asymmetry. For instance, we may want to require that the polarization of a distribution with two groups of different sizes should not depend on which of the two groups is rich or poor. Once again, it is not clear that such a feature should be given the status of an axiom,<sup>10</sup> but it is an interesting property to explore because it entirely divorces the measure from notions of poverty or inequality.<sup>11</sup> Moreover, it has particular force in the context of variables such as political positions.

Consider, then, the following property.

**PROPERTY 3.** For any distribution  $F$  with density  $f$  on support  $[a, b]$ , define the distribution  $G$  with density  $g$  such that  $f(x) = g(a + b - x)$ . Then  $F$  and  $G$  have the same polarization.

We use these properties to obtain additional insights into the polarization measures that have been proposed in the literature.

---

<sup>10</sup>Esteban and Ray (1994) do study a one-sided alienation model in which this symmetry property is not satisfied.

<sup>11</sup>In our example, a permutation of incomes across the two groups does affect inequality.

#### 4. COMPARING POLARIZATION MEASURES

**4.1. Measures of Polarization.** Based on the view that polarization is the outcome of interpersonal alienation fueled by the sense of identification, the polarization measure obtained by Esteban and Ray (1991) and Duclos, Esteban and Ray (2004) is

$$(5) \quad P^{DER} = K \int \int f(x)^{1+\alpha} f(y) |y - x| dy dx,$$

where  $\alpha \in [0.25, 1]$  and  $K > 0$  is some arbitrary constant which can be used for normalization.

**PROPOSITION 1.** *A polarization measure satisfies Axioms 1 to 4, and Initial Restriction 1, if and only if it coincides with  $P_{\alpha}^{DER}$  for some  $\alpha \in [0.25, 1]$ .<sup>12</sup> Further,*

(i)  $P_{\alpha}^{DER}$  satisfies Properties 1–3, and Axiom 4B; and

(ii)  $P_{\alpha}^{DER}$  fails to satisfy Axioms 2B, 3B and Initial Restriction 1B.

*Proof.* The first assertion is just the characterization theorem in DER (2004).

Now turn to part (i). In order to prove that  $P^{DER}$  satisfies Property 1, consider a uniform distribution over  $[a, b]$ . Partition this support into  $n$  identical intervals of the form  $[a_i, b_i)$ ,  $i = 1, \dots, n$ , each with length  $\frac{b-a}{n}$  and population of  $\frac{1}{n}$ . Within each interval we have an (unnormalized) uniform distribution with density  $\frac{1}{b-a}$ . Clearly,  $a_1 = a$ ,  $a_{i+1} = a_i + (b-a)/n$ , and  $b_i = a_{i+1}$  for all  $i$ .

Now subject all these  $n$  identical densities to a  $\lambda$ -squeeze. To compute the outcome for each  $\lambda$ , follow the same procedure as in DER. Decompose total polarization into the sum of all the “within-density” polarizations,  $W_i$ , and all the “across-density” effective antagonisms,  $A_{ij}$ . Using Lemma 6 of DER, and setting the normalization constant  $K$  to equal 1, it is possible to

---

<sup>12</sup>Esteban and Ray (2010) prove that  $\alpha = 1$  if the following Axiom is satisfied.

**AXIOM** Suppose that a distribution consists of three equi-spaced uniform basic densities of sizes  $r$ ,  $p$  and  $q$ , each of support  $2\epsilon$ . Assume that  $p = q + r$ . Then there is  $\eta > 0$  such that if  $0 < r < \eta$  and  $0 < \epsilon < \eta$ , any uniform transfer of population mass from  $r$  to  $q$  cannot decrease polarization.

Intuitively, this axiom asserts that if the group of size  $r$  is extremely small, it cannot be contributing much on its own to social tension. If instead the population is transferred from that group to the second largest group which is “equally opposed” to the largest group of size  $p$ , these two groups will become more equally sized and hence polarization cannot come down.

see that

$$\begin{aligned}
\sum_{i=1}^n W_i &= 2^{1+\alpha} [\lambda(b-a)]^{1-\alpha} n^{-2} \int_0^1 f^*(x)^{1+\alpha} \left\{ \int_0^1 f^*(y)(1-y)dy + \int_x^1 f^*(y)(y-x)dy \right\} dx \\
&= 2^{1+\alpha} [\lambda(b-a)]^{1-\alpha} n^{-2} \int_0^1 2^{-1-\alpha} \left\{ \int_0^1 2^{-1}(1-y)dy + \int_x^1 2^{-1}(y-x)dy \right\} dx = \\
&= 2^{-1} (b-a)^{1-\alpha} n^{-2} \lambda^{1-\alpha} \int_0^1 \left\{ \int_0^1 (1-y)dy + \int_x^1 (y-x)dy \right\} dx = \\
&= \frac{1}{3} (b-a)^{1-\alpha} n^{-2} \lambda^{1-\alpha}.
\end{aligned}$$

Let us now compute  $A_{ij}$ . Using Lemma 7 in DER we have

$$\begin{aligned}
A_{ij} &= 2|\mu_i - \mu_j| n^{-2-\alpha} \left( \frac{b-a}{2n} \right)^{-\alpha} \lambda^{-\alpha} \int_0^1 f^*(x)^{1+\alpha} dx = \\
&= |i-j| n^{-3} (b-a)^{1-\alpha} \lambda^{-\alpha}.
\end{aligned}$$

Therefore, adding all the cross antagonisms we obtain

$$A = n^{-3} (b-a)^{1-\alpha} \lambda^{-\alpha} \sum_i \sum_j |i-j|.$$

Thus, for total polarization we finally obtain

$$(6) \quad P^{DER}(\lambda) = W + A = \frac{1}{3} (b-a)^{1-\alpha} n^{-2} \lambda^{1-\alpha} + n^{-3} (b-a)^{1-\alpha} \lambda^{-\alpha} \sum_i \sum_j |i-j|.$$

Differentiating (6) with respect to  $\lambda$  we have

$$\frac{\partial P^{DER}}{\partial \lambda} = (b-a)^{1-\alpha} n^{-2} \lambda^{-1-\alpha} \left( \frac{1-\alpha}{3} \lambda - \alpha \frac{\zeta(n)}{n} \right),$$

where  $\zeta(n) \equiv \sum_i \sum_j |i-j|$ .

The sign of this derivative depends on the term in braces only, which we denote by  $B$ . Clearly

$$B \leq \left\{ \frac{(1-\alpha)}{3} - \alpha \frac{\zeta(n)}{n} \right\}.$$

Observe now that  $\frac{\zeta(n)}{n}$  is increasing in  $n$  and that  $\frac{\zeta(2)}{2} = 1^{13}$ . Hence,  $B \leq (1 - 4\alpha) \leq 0$  for  $\alpha \geq 0.25$ .

It follows that for  $\alpha \geq 0.25$  a  $\lambda$ -squeeze — that is, a decrease in  $\lambda$  — will increase polarization for every partition possible into  $n$  intervals of the support of a uniform distribution. Therefore  $P_\alpha^{DER}$  does satisfy Property 1.

Let us now turn to Property 2. The argument that  $P_\alpha^{DER}$  satisfies this property can be broken into two steps: (i) after having shifted all the population from the inner to the outer densities polarization should be higher than at the beginning of the process, and (ii) when all the population is at the inner densities the beginning of the transfer of population to the outer densities makes polarization come down.

As an overview, we consider a symmetric distribution made of four densities with non-overlapping support. Transfer population from inner to outer densities. We show that  $P_\alpha^{DER}$  is higher at the end of the process, and actually goes up at the very beginning of the process, under some restrictions on the initial distribution.

Begin with a symmetric nonoverlapping four-density configuration in which each density is built from the same basic density. Let the mass of each of the inner densities be  $\frac{1}{2} - \delta$ , and that of each of the outer densities be  $\delta$ . Let each density have common width  $2s$ . Let  $D$  denote the distance between the conditional means of the two inner densities, and let  $d$  denote the same distance between an inner and its neighboring outer density.

We wish to examine the behavior of  $P_\alpha^{DER}$  as  $\delta$  increases. In order to compute  $P_\alpha^{DER}$ , we again decompose total polarization into the sum of within-density polarization and across-density antagonisms. Using Lemma 6 in DER as we did before, we have

$$\begin{aligned} P^{DER}(\delta) &= 2W_{\text{inner}}(\delta) + 2W_{\text{outer}}(\delta) + \sum_i \sum_j A_{ij} \\ &= K \left\{ \left[ \left( \frac{1}{2} - \delta \right)^{2+\alpha} + \delta^{2+\alpha} \right] 2s + \left[ \left( \frac{1}{2} - \delta \right)^{1+\alpha} \frac{4\delta d + D}{2} + \delta^{1+\alpha} \frac{2d + D}{2} \right] \psi(f, \alpha) \right\}, \end{aligned}$$

---

<sup>13</sup>Observe first that  $\zeta(n+1) = \zeta(n) + n(n+1)$ . Hence,  $\frac{\zeta(n+1)}{n+1} - \frac{\zeta(n)}{n} = n - \frac{\zeta(n)}{n(n+1)}$ .

We know that  $\zeta(n) \leq n \cdot n^2 = n^3$ . Thus,

$$\frac{\zeta(n+1)}{n+1} - \frac{\zeta(n)}{n} \geq \frac{n}{(n+1)} > 0.$$



where

$$(7) \quad K = 4s^{-\alpha} \int_0^1 f^*(x)^{1+\alpha} \left\{ \int_0^1 f^*(y)(1-y)dy + \int_x^1 f^*(y)(y-x)dy \right\} dx$$

and

$$(8) \quad \psi(f, \alpha) = \frac{\int_0^1 f^*(x)^{1+\alpha} dx}{\int_0^1 f^*(x)^{1+\alpha} \left\{ \int_0^1 f^*(y)(1-y)dy + \int_x^1 f^*(y)(y-x)dy \right\} dx}$$

as in (39) in Lemma 8 of DER.

It is straightforward to verify that  $P^{DER}(\frac{1}{2}) > P^{DER}(0)$ ; this takes care of the first step.

In order to prove the non-monotonicity of  $P^{DER}$  with respect to  $\delta$  it suffices to show that it is strictly decreasing at  $\delta = 0$ , perhaps under some restrictions on  $d$ ,  $D$  and  $s$ . Differentiating  $P_\alpha^{DER}$  with respect to  $\delta$ , we have

$$\begin{aligned} \frac{\partial P_\alpha^{DER}}{\partial \delta} &= K \left[ -(2+\alpha) \left( \frac{1}{2} - \delta \right)^{1+\alpha} + (2+\alpha) \delta^{1+\alpha} \right] 2s \\ &\quad + K \left[ \left( \frac{1}{2} - \delta \right)^\alpha \left( \frac{2d-4\delta d}{2} - (1+\alpha) \frac{4\delta d + D}{2} \right) + (1+\alpha) \delta^\alpha \frac{2d+D}{2} \right] \psi(f, \alpha). \end{aligned}$$

Evaluating this expression at  $\delta = 0$ , we see that

$$\frac{\partial P_\alpha^{DER}}{\partial \delta} \Big|_{\delta=0} = K \left( \frac{1}{2} \right)^\alpha \left\{ -(2+\alpha)2s + \left[ d - (1+\alpha) \frac{D}{2} \right] \psi(f, \alpha) \right\}.$$

Since  $s$  can be arbitrarily small without this imposing any restriction on  $\psi(f, \alpha)$ , we have that  $d < \frac{1+\alpha}{2}D$  (or  $d < \frac{5}{8}D$  for any admissible  $\alpha$ ) is sufficient for polarization to decrease at the early stages of the transferring of population from inner densities towards outer densities.<sup>14</sup> It follows that  $P_\alpha^{DER}$  satisfies Property 2.

It is quite straightforward to show that  $P_\alpha^{DER}$  satisfies Property 3. In this case, we have a new distribution with density  $g(x)$  with the property that  $g(x) = f(2M-x)$ , where  $M = \frac{a+b}{2}$ . Using this equality in (5) and performing the change of variables  $x' = 2M-x$  and  $y' = 2M-y$  it is immediate that  $P_\alpha^{DER}(F) = P_\alpha^{DER}(G)$ .

---

<sup>14</sup>It can be easily computed that  $\frac{\partial P}{\partial \delta} \Big|_{\delta=\frac{1}{4}} < 0$ .

Finally, the rankings generated by  $P_\alpha^{DER}$  are obviously invariant to income-scalings on both sides of the comparison, so Axiom 4B is satisfied. This completes the proof of part (i) of the proposition.

We now address part (ii), which states that  $P_\alpha^{DER}$  fails to satisfy Axioms 2B, 3B and Initial Restriction 1B.

To examine Axiom 2B, consider a symmetric distribution consisting of six nonoverlapping basic densities obtained from the same kernel. From “left to right”, the population masses of each density are  $(\frac{1}{4} - \delta, 2\delta, \frac{1}{4} - \delta, \frac{1}{4} - \delta, 2\delta, \frac{1}{4} - \delta)$ , where  $\delta \in [0, \frac{1}{4}]$ . Take the support of each basic density to be of size  $2s$ , and let  $d$  be the distance between the centers of neighboring basic densities on one side of the mean. Finally, let the distance between the third and fourth densities (on opposite sides of the mean) be given by  $D$ .

When  $\delta = 0$  we have a distribution with the population concentrated on four basic densities (the “starting densities”, which are the first, the third, the fourth and the sixth). When  $\delta = \frac{1}{4}$  we have the population concentrated on two densities (the second and the fifth), each located half way between the two “starting densities” on each side of the median/mean.

Now, as  $\delta$  rises from 0 to  $1/4$ , we create progressive Dalton transfers within the populations below and above the median/mean income. (On each side of the median we decrease the population sizes of the richest and poorest densities, augmenting the group halfway in between.)

As before, compute  $P(\delta)^{DER}$  as the sum of the inner polarization within each density and the sum of the effective antagonisms across densities. We shall continue to use Lemmas 6 and 7 in DER.

We start with the inner polarization  $W$ . There are four identical densities with population mass  $\frac{1}{4} - \delta$  and two with mass  $2\delta$ . Hence, using Lemma 6 we have

$$W = 8s^{1-\alpha} K \left\{ (2\delta)^{2+\alpha} + 2 \left( \frac{1}{4} - \delta \right)^{2+\alpha} \right\}.$$

where  $K$  is defined in (7).

Since the distribution is symmetric, the cross-group antagonism will be twice the sum of antagonisms of the three groups with mean income below the median/mean; call these  $A_1$ ,  $A_2$ , and

$A_3$ . Using Lemma 7 in DER we see that

$$\begin{aligned} A_1 &= 2\left(\frac{1}{4} - \delta\right)^{1+\alpha} s^{-\alpha} K \psi(f, \alpha) \left(2d + \frac{D}{2}\right), \\ A_2 &= 2(2\delta)^{1+\alpha} s^{-\alpha} K \psi(f, \alpha) \left(\frac{3}{2}d - 2d\delta + \frac{D}{2}\right), \\ A_3 &= 2\left(\frac{1}{4} - \delta\right)^{1+\alpha} s^{-\alpha} K \psi(f, \alpha) \left(d + \frac{D}{2}\right). \end{aligned}$$

where  $\psi$  is defined previously in (8).

Adding all the components and rearranging, we have for aggregate polarization

$$P^{DER}(\delta) = 4Ks^{-\alpha} \left\{ \begin{aligned} &2\left(\frac{1}{4} - \delta\right)^{1+\alpha} \left(2\left(\frac{1}{4} - \delta\right)s + \psi(f, \alpha)\frac{1}{2}(3d + D)\right) + \\ &+(2\delta)^{1+\alpha} \left(2\delta s + \psi(f, \alpha)\frac{1}{2}(3d + D - 4d\delta)\right) \end{aligned} \right\}.$$

Differentiating  $P(\delta)^{DER}$  with respect to  $\delta$  and evaluating it at  $\delta = 0$  we finally obtain

$$\frac{\partial P^{DER}(\delta)}{\partial \delta} \Big|_{\delta=0} = 4Km^{-\alpha} \left\{ -\frac{1+\alpha}{2} \left( \frac{s}{2} + \psi(f, \alpha)\frac{1}{2}(3d + D) \right) - s4^{-\alpha} \right\} < 0.$$

Therefore, when  $\delta = 0$ , i.e. at the beginning of the process, the shifting of mass to create an intermediate density does decrease polarization as measured by  $P^{DER}$ , violating the claim in Axiom 2B.

Let us now consider Axiom 3B. The very same argument that shows that Property 2 is satisfied by  $P_\alpha^{DER}$  establishes that Axiom 3B cannot be satisfied. For that argument involves an example in which the conditions of Axiom 3B are satisfied throughout, yet  $P_\alpha^{DER}$  behaves in a nonmonotonic fashion.

Finally, it is obvious from direct inspection that  $P_\alpha^{DER}$  fails Initial Restriction 1B. ■

**4.2. Measures of bipolarization.** We shall examine the bipolarization measures proposed by Wolfson (1994) and (1997) — based on Foster and Wolfson (1992, 2009) — and by Wang and Tsui (2000). These two measures of bipolarization share a common feature: both are different ways of computing the distance from any particular distribution to the degenerate symmetric bimodal distribution located at the extremes of the support.

We shall present the polarization indices for the case of differentiable distributions  $F$ , with  $f$  as the corresponding density. The measure of bipolarization in Wolfson (1994) can be written as

$$(9) \quad P^W(F) = \frac{\mu(F)}{m(F)} \left[ \frac{1}{2} - L(F) - \frac{G(F)}{2} \right],$$

where  $m$  stands for the median income,  $\mu$  for mean income,  $G$  for the Gini index and  $L$  for the value of the ordinate of the Lorenz curve at the median income, all as a function of the distribution function  $F$ .<sup>15</sup>

We now evaluate this measure with respect to the set of axioms and properties introduced in the previous section.

**PROPOSITION 2.** *The polarization measure  $P^W$*

(i) *satisfies Axioms 2B–4B and Axioms 1–4, but*

(ii) *it does not satisfy Properties 1–3.*

*Proof.* Wang and Tsui (2000) have proven that  $P^W$  satisfies Axioms 2B–4B. It follows from Observation 1 that  $P^W$  also satisfies Axioms 1, 2, and 3. Since  $P^W$  is population normalized, it also satisfies Axiom 4.

Let us now turn to part (ii). To examine Property 1, we start with a uniform distribution on  $[a, b]$ . This distribution has a density  $f(x) = \frac{1}{b-a}$  and a mean/median  $\mu = m = \frac{a+b}{2}$ . We partition this uniform distribution into  $n$  equally sized subintervals, and apply a  $\lambda$ -squeeze to each of them. We wish to understand the effect of this change on  $P^W$ .

Let us first suppose that  $n$  is even. In that case, the squeeze would leave the value of  $L$  unchanged, because no income would have been transferred across the mean/median of the distribution, and  $G$  would be reduced. So  $P^W$  must increase, in line with Property 1.

---

<sup>15</sup>An alternative expression for  $P^W$  is

$$P^W = \frac{2\mu}{m} \left[ \frac{\mu_H - \mu_L}{\mu} - G \right],$$

where  $\mu_H$  and  $\mu_L$  are to conditional means of the incomes above and below the median income, respectively.

Now suppose that  $n$  is odd. Now, while  $G$  continues to come down,  $L$  would go up because the squeeze of the interval containing the mean/median income in its interior would transfer income from the higher to the lower incomes in this interval. The two changes would shift  $P^W$  in opposite directions. In order to sign the net effect we need to compute both changes explicitly. To this we turn now.

We shall denote by  $L(n, \lambda)$  the share of total income below the mean/median income when the  $n$  intervals of a uniform distribution are subject to a  $\lambda$ -squeeze. It is immediate that we should focus only on the interval whose conditional mean/median income coincides with the overall mean/median income. Since the support of the overall uniform distribution is  $[a, b]$ , the length of any interval will be  $\frac{b-a}{n}$  and its mass  $\frac{1}{n}$ . We are interested in the share of total income of individuals in the interval  $[\mu - \frac{b-a}{2n}, \mu]$  as the distribution is  $\lambda$ -squeezed.

To this end we first observe that  $L(n, \lambda) = L_- + L_m(\lambda)$ , where the first term stands for the share contributed by the intervals below the median interval and  $L_m$  for the contribution of the median interval up to its median/mean income. When computing  $L_m(\lambda)$  we have to keep in mind that the relevant support is  $[\mu - \frac{b-a}{2n}, \mu]$ , that is  $[\frac{a+b}{2} - \frac{b-a}{2n}, \frac{a+b}{2}]$ . Then we have

$$(10) \quad L = L_- + L_m(\lambda) =$$

$$(11) \quad = \frac{2}{a+b} \sum_{i=1}^{m-1} \frac{\mu_i}{n} + \int_{\frac{a+b}{2} - \frac{b-a}{2n}}^{\frac{a+b}{2}} \frac{\lambda 2x + (1-\lambda)(a+b)}{a+b} \frac{1}{b-a} dx =$$

$$(12) \quad = \frac{2}{a+b} \sum_{i=1}^{m-1} \frac{\mu_i}{n} + \frac{1}{2n} - \frac{\lambda}{4n^2} \frac{b-a}{a+b}.$$

Differentiating with respect to  $\lambda$  we have

$$(13) \quad \frac{\partial L}{\partial \lambda} = -\frac{1}{4n^2} \frac{b-a}{a+b}.$$

In order to compute the effect of the  $\lambda$ -squeeze on  $G$  we can use the results obtained for  $P^{DER}$ . Indeed,  $G$  is equal to  $\frac{1}{2}P^{DER}$  when evaluated at  $\alpha = 0$ . [Notice that for  $P^{DER}$  to satisfy the corresponding axioms it must be that  $\alpha \geq 0.25$ ]. Using (7) evaluated at  $\alpha = 0$ , we have that

$$(14) \quad \frac{\partial G}{\partial \lambda} = \frac{1}{3n^2} \frac{b-a}{a+b}.$$

Differentiating (9) with respect to  $\lambda$  and using (13) and (14) we obtain

$$(15) \quad \frac{\partial P^W}{\partial \lambda} = -\frac{\mu}{m} \left[ \frac{1}{6n^2} \frac{b-a}{a+b} - \frac{1}{n^2} \frac{b-a}{a+b} \right] = \frac{1}{12n^2} \frac{b-a}{a+b} > 0.$$

It follows that, when  $n$  is odd, a  $\lambda$ -squeeze will decrease  $P^W$  violating Property 1.

As we have already pointed out, the fact that  $P^W$  satisfies Axiom 3B implies that it violates Property 2.

We finally show that  $P^W$  does not satisfy Property 3. Consider the following example. We have a degenerate distribution with  $\frac{2}{3}$  of the population with  $y_1 = 0$  and  $\frac{1}{3}$  with  $y_2 = 1$ . It can be readily verified that  $P^W$  is larger than for the distribution resulting from the flip of the initial one. This completes the proof. ■

Thus Wolfson's measure  $P^W$  satisfies the four Axioms which along with the initial Condition 1 characterize  $P^{DER}$ , but does not have any of the Properties 1–3.

It is particularly striking how  $P^W$  reacts to the squeezing of subintervals, as described in Property 1. The effect on the measure appears to hinge crucially on whether the number of subintervals is odd or even. This is hardly surprising: Wolfson's measure can be best understood as a measure of bipolarization. Specifically, it can be seen as a measure of distance with respect to the most bi-polarized distribution: the one with the population concentrated on two spikes located at the extremes of the support of the distribution. The bunching of probability around an odd number of groups takes that distribution “away from” the extreme bimodal distribution.

At the same time, from a multi-group perspective, Property 1 is intuitive. Again, it is intrinsically connected to bipolarization. Indeed, Property 1 does a good job of bringing out this aspect.

The second measure of bipolarization we wish to examine is the one axiomatized by Wang and Tsui (2000). Their measure is

$$(16) \quad P^{WT} = P(F) = k \int \left| \frac{x-m}{m} \right|^r f(x) dx,$$

with  $r \in (0, 1]$ .

Accordingly with this measure, polarization is captured by the average of a concave transformation of the distance with respect to the median income. It is essentially a measure of distance to the polarization minimizing distribution with all the population concentrated at the median income.

**PROPOSITION 3.** *Consider the polarization measure  $P^{WT}$ ,*

*(i) A polarization measure satisfies Axioms 2B to 5B —and hence also Axioms 1, 2 and 3— if and only if it can be written as (16);*

*(ii) The polarization measure  $P^{WT}$  (16) satisfies Axiom 4 but violates Properties 1–3.*

*Proof.* Part (i) of the Proposition has been proven by Wang and Tsui (2000, Proposition 6, p. 360).

Since  $P^{WT}$  is normalized with respect to total population, it does satisfy Axiom 4.

By the same reasons as before,  $P^{WT}$  violates Property 2. It remains to be shown that  $P^{WT}$  violates Properties 1 and 3.

We start with an example of violation of Property 3. Suppose a degenerate distribution with  $\frac{1}{3}$  of the population in each of the three incomes  $(0, 1, 4)$ . It is immediate to compute that  $P^{WT} = \frac{4}{3}$ . Flipping the previous distribution would give incomes  $(0, 3, 4)$  with a polarization of  $P^{WT} = \frac{4}{9} < P^{WT}$ .

We finally turn to the violation of Property 1. Consider the case of the partition of a uniform distribution into three equal intervals. This situation accommodates itself to the scenario described in Axiom 2. The effect of a squeeze on  $P^{WT}$  can be seen as composed by the effect of the squeezing of the two side densities and the effect of squeezing the inner density. By Axiom 2 the former will raise  $P^{WT}$  while the latter —by Axiom 1— will reduce it. We need to show an example where this second effect dominates over the first and makes overall  $P^{WT}$  fall. Consider the following example: a uniform distribution with support  $[0, 12]$  is partitioned into three uniform densities with supports  $[0, 4]$ ,  $[4, 8]$ ,  $[8, 12]$ . We now squeeze them and obtain new uniform densities with supports  $[1, 3]$ ,  $[5, 7]$ ,  $[9, 11]$ . It is a matter of computation to obtain that  $P^{WT}$  will be smaller after the squeeze.<sup>16</sup> This concludes the proof. ■

<sup>16</sup>Note that if the partition of the original distribution had been into an even number of intervals, the squeeze of each uniform density would have produced an increase in  $P^{WT}$ .

Propositions 2 and 3 substantiates the similarity between the measures proposed by Wang and Tsui and by Wolfson in that both satisfy the same set of axioms and violate the same set of properties.

## 5. SOME REMARKS ON MULTIDIMENSIONAL POLARIZATION

A *multidimensional* measure of polarization is any measure in which at least two characteristics enter into the determinants of polarization. One obvious way in which a second characteristic enters is via the determination of alienation and/or identification, the building blocks of Initial Restriction 1 in the ER approach.

A particularly sharp expression of multidimensionality occurs when some characteristic, such as religion, ethnicity or geography is used to *define* the groups, while income clustering (within groups) or income distances (across groups) are used to define identification and alienation, and thereby construct the measure.

An example of such a measure is the one proposed by Zhang and Kanbur (2001).<sup>17</sup> Their approach is based on the family of entropy indices developed by Theil to measure inequality. The Theil inequality index is given by

$$I(f) = \int \log \frac{\mu}{x} f(x) dx.$$

Suppose that population can be partitioned into  $K$  groups on the basis of a non-income characteristic, such as race or geography. Let  $\pi_i$  be the population share of the population in group  $i$ ,  $f_i(x)$  the density of incomes in that group and  $\mu_i$  the mean income within the  $i$ th group. Now, overall inequality *within* the groups is given by

$$I^W \equiv \sum_{i=1}^K \pi_i I(f_i),$$

while inequality *between* groups is given by

$$I^B \equiv \sum_{i=1}^K \log \frac{\mu}{\mu_i} \pi_i.$$

---

<sup>17</sup>See also DER, which extends their axiomatically defined measure to the multidimensional case. The critique in this section will apply to their measure as well.



Zhang and Kanbur view  $I^B$  — the between-group component — as an expression of total alienation, while the *reciprocal* of  $I^W$  — the within-group term — is taken to be a measure of total identification within groups. This generates a polarization multiplicative measure of the form

$$(17) \quad P^{ZK} = \frac{I^B}{I^W} = \frac{\sum_{i=1}^K \log \frac{\mu}{\mu_i} \pi_i}{\sum_{i=1}^K \pi_i I(f_i)}.$$

*Provided that we accept that income can be used to determine identification and alienation in the multidimensional case,  $P^{ZK}$  does well as a measure of polarization. For instance, with squeezes suitably redefined to occur only within the exogenously given groups, one can show that the measure satisfies versions of Axioms 1–4 above.*

However, we feel bound to sound a cautionary note regarding the construction and employment of such polarization measures in the multidimensional case. The warnings apply precisely to the emphasized sentence in the preceding paragraph.

To fix ideas, suppose that there are only two groups in society; let us call them natives and immigrants. Consider two cases: (i) immigrants compete with natives for the same jobs, and (ii) immigrants fill a complementary and different set of jobs. It is easy to argue that the first case will exhibit higher tensions than the second, because the two groups are competing in the same sector. Therefore, given the (noneconomic) definitions of the two groups, it is entirely possible that income *similarities*, rather than differences, will heighten intergroup tensions.

The point here is that inter-group alienation is possibly defined by other features, such as the attitudes towards foreigners or the level of intolerance across religions. We are not entirely rejecting the possibility that income inequalities affect such alienations. But we should be willing to entertain the scenario that income similarities also act as a proxy for competition.<sup>18</sup>

More generally, think of two forms of conflict: one redistributive and the other exclusionary. Redistributive conflict will generally be heightened by increased income differences. But exclusionary conflict may be exacerbated by increased income similarities.

---

<sup>18</sup>Mitra and Ray (2010) empirically study Hindu-Muslim violence in India, and argue that an increase in Muslim per-capita incomes positively affects violence. Because Muslims are poorer than Hindus on average, this boils down to an equalization of inter-group incomes.

Next, fix income differences across the two groups and focus on within-group “identification”. Once again, this is a complex question and it is entirely unclear that within-group *equality* in income unambiguously heightens inter-group tensions. As in the previous example, income does double-duty. Income similarities may be proxies for general cultural and social cohesion between group members, in which case group identity could indeed be enhanced, sometimes at the cost of social conflict across groups. This would be picked up well by a measure such as  $P^{ZK}$ . On the other hand, income differences may make it easier to operationalize conflict. Large within-group inequalities means that some subset of individuals can finance conflict, while another subset can provide the physical labor input.

This latter aspect is emphasized in recent work by Esteban and Ray (2008, 2009). In the first of these two papers, it is shown that conflict might be salient across ethnic lines precisely because there are both rich and poor in each ethnic group, making it easier to engage in “conflict activities” (the poor in each group can be compensated by the rich for providing “conflict labor”). This particular point is made in more detail in Esteban and Ray (2009), and is summarized in their Proposition 4, which argues that a within-group Lorenz-worsening of incomes will fuel inter-group conflict.

The present paper does not study multi-dimensional polarization in any detail. In our opinion it is a rich but complex subject, and represents an open field of inquiry. Zhang and Kanbur (2001) and DER both discuss measures of polarization in which groups are defined exogenously, while income differences (and similarities) are used as proxies for across-group alienation (and within-group identification). For the reasons given in this section, we believe that such measures have to be treated with caution.

## 6. CONCLUSIONS

We have reviewed some alternative measures of unidimensional polarization, grouped into two families: polarization and bipolarization measures. Taking as a base the set of axioms that characterize the  $P^{DER}$  measure of polarization, we have demonstrated that the four basic axioms in DER are satisfied by all the measures of either family. It follows that these axioms capture what

appears to be common to all the polarization measures.<sup>19</sup> But, using three further Properties introduced in this paper, we also show that the two families behave quite differently.

Where is the source of the departure of the two families of polarization measures? We cannot provide a complete answer to this question. However, this paper goes one step towards a better understanding of the differences between the two families. In view of our results, there are two lines of departure. One line is that the bipolarization measures satisfy axioms that are stronger than the ones proposed in DER, albeit responding to the same intuition. The second line of inquiry is that there are basic differences in the assumed “classes of measures” as described by Initial Restriction 1 and Initial Restriction 1B. These are essentially black boxes. ER (and DER) provide a “behavioral” motivation for Initial Restriction 1, based on the interaction of identification and alienation in the building of interpersonal antagonism. But this is no substitute for a precise axiomatization. Thus Initial Restrictions 1 and 1B need to be decomposed into a collection of simpler axioms.<sup>20</sup> Only then we will be able to clearly see the specific departure points.

In a way, this is partially the role played by Properties 1–3, which takes us back to the first line of argument above. We do not call these “axioms” because we do not see them as self-evident truths. They are secondary properties which are attractive and worth exploring. Properties 1 and 2 point in the same direction: the emergence of polarization need not occur around two poles only, it can be around any number of poles. Of course the smaller the number of poles — but no less than two — the higher the polarization.

Property 1 captures exactly this as it considers the formation of an arbitrary number of poles. It is compatible with the axioms of DER but not with the stronger axioms that describe bipolarization. Accordingly, Property 1 is not satisfied by existing bipolarization measures. This is not surprising given the multiple-poles intuition upon which the property rests. The same is true of Property 2, in which the intuition of many poles, not just two, plays an important role.

Finally, Property 3 creates a symmetry between high and low values of the characteristic, arguing that “flipping” the characteristics distribution should have no impact on polarization.

---

<sup>19</sup>However, the same cannot be said of the two essential axioms for bipolarization —Axioms 2B and 3B — because  $P^{DER}$  fails to satisfy them.

<sup>20</sup>Bossert and Schworm (2008) make a valuable contribution along these lines for bipolarization measures.

Once again, we put it forward as a possible interesting distinction between inequality and polarization, but not as an unquestionable desideratum. All measures of inequality will concur in stating that a distribution with 9 people with an income of 0 and 1 with an income of 10 is more unequal than one with 1 with income 0 and 9 with 10. However, if we are interested on how “polarized” the two distributions are, it may be a plausible position to claim that the level of polarization is the same.

In the last section of the paper, we make some brief observations on multi-dimensional polarization, something that we do not consider in the main analysis. Our main reservations have to do with the use of income as proxies for identification and alienation within and across groups, when the groups themselves are defined by characteristics other than income. We argue that such proxies may lead to misleading results. The measurement of multidimensional polarization is an open and important area of research.

## 7. REFERENCES

- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT AND R. WACZIARG (2003) “Fractionalization”, *Journal of Economic Growth* **8** 155-194.
- BOSSERT, W AND W. SCHWORM(2008) “A class of two-group polarization measures” *Journal of Public Economic Theory* **10**, 1169-1187.
- CHAKRAVARTY S.R. AND C. D’AMBROSIO(2009) “Polarization ordering of income distributions” *Review of Income and Wealth*, Forthcoming.
- CHAKRAVARTY, S.R. AND A. MAJUMDER(2001), “Inequality, Polarization and Welfare: Theory and Applications,” *Australian Economic Papers* **40**, 1–13.
- CHAKRAVARTY, S.R., A. MAJUMDER AND S. ROY(2007), “A Treatment of Absolute Indices of Polarization,” *Japanese Economic Review* **58**, 273-293.
- COLLIER, P. AND A. HOEFFLER (2004), “Greed and Grievance in Civil War” *Oxford Economics Papers* **56**, 563-595.
- DUCLOS, J-Y., J. ESTEBAN AND D. RAY(2004), “Polarization: Concepts, Measurement, Estimation” *Econometrica* **72** 1737-1772.

EASTERLY W. AND R. LEVINE (1997), *Africa's Growth Tragedy: Policies and Ethnic Divisions*, *Quarterly Journal of Economics* **111**, 1203-1250.

ESTEBAN, J., C. GRADÍN AND D. RAY (1998) An Extension of a Measure of Polarization, with an application to the income distribution of five OECD countries, LIS working paper.

ESTEBAN, J., C. GRADÍN 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**, 1-19.

ESTEBAN, J. AND D. RAY (1991) "On the Measurement of Polarization," Boston University, Institute for Economic Development, Working Paper 18.

Esteban, J. and D. Ray (1994) "On the Measurement of Polarization," *Econometrica* **62**, 819–852.

ESTEBAN, J. AND D. RAY (1999) "Conflict and Distribution," *Journal of Economic Theory* **87**, 379–415.

ESTEBAN, J. AND D. RAY (2008a), "Polarization, Fractionalization and Conflict," *Journal of Peace Research* **45**, 163-182.

ESTEBAN, J. AND D. RAY (2008b), "On the Saliency of Ethnic Conflict," *American Economic Review* **98**, 2185–2202.

ESTEBAN, J. AND D. RAY (2009) "A Model of Ethnic Conflict", forthcoming, *Journal of the European Economic Association*.

ESTEBAN, J. AND D. RAY (2010) "Linking Conflict to Inequality and Polarization", forthcoming, *American Economic Review*.

FEARON, J. AND D. LAITIN (2003), "Ethnicity, Insurgency, and Civil War," *American Political Science Review* **97**, 75–90.

FOSTER, J. AND M.C. WOLFSON (1992), "Polarization and the Decline of the Middle Class: Canada and the US", .

FOSTER, J. AND M.C. WOLFSON (2009), "Polarization and the Decline of the Middle Class: Canada and the US", *Journal of Economic Inequality* forthcoming.

HOROWITZ, D.L. (1985), *Ethnic Groups in Conflict*, Berkeley, CA: University of California Press.

LASSO DE LA VEGA, M.C. AND A.M. URRUTIA (2006) "An alternative formulation of the Esteban-Gradín-Ray extended measure of bipolarization" *Journal of Income Distribution* **15**, 42-54.

- MIGUEL, E., SATYANATH, S. AND E. SERGENTI (2004), "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy* **112**, 725–753.
- MITRA, A. AND D. RAY (2010), "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India," mimeo., Department of Economics, New York University.
- MONTALVO, J. G. AND M. REYNAL-QUEROL (2005a), "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review* **95**, 796–816.
- MONTALVO, J. G. AND M. REYNAL-QUEROL (2005b), "Ethnic Diversity and economic development" *Journal of Development Economics* **76**, 293-323.
- REYNAL-QUEROL, M. (2002), "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution* **46**, 29–54.
- RODRÍGUEZ, J.G. AND R. SALAS (2003), "Extended bipolarization and Inequality Measures", *Research on Economic Inequality* **9**, 69-83.
- SEN, A. K. (1973), *On Economic Inequality*, Clarendon Press, Oxford.
- WANG, Y.Q. AND K.Y. TSUI (2000), "Polarization Orderings and New Classes of Polarization Indices," *Journal of Public Economic Theory* **2**, 349–363.
- WOLFSON, M.C. (1994), "When Inequalities Diverge", *American Economic Review P&P* **94** 353-358.
- WOLFSON, M.C. (1997), "Divergent Inequalities: theory and empirical results", *Review of Income and Wealth* **43** 401-421.
- ZHANG, X. AND R. KANBUR (2001), "What Difference Do Polarisation Measures Make? An Application to China," *Journal of Development Studies* **37**, 85–98.