On the Measurement of Polarization

Joan-Maria Esteban; Debraj Ray

Econometrica, Vol. 62, No. 4 (Jul., 1994), 819-851.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199407%2962%3A4%3C819%3AOTMOP%3E2.0.CO%3B2-V

Econometrica is currently published by The Econometric Society.

Econometrica, Vol. 62, No. 4 (July, 1994), 819–851

ON THE MEASUREMENT OF POLARIZATION

BY JOAN-MARÍA ESTEBAN AND DEBRAJ RAY[1]

Suppose that a population of individuals may be grouped according to some vector of characteristics into "clusters," such that each cluster is very "similar" in terms of the attributes of its members, but different clusters have members with very "dissimilar" attributes. In that case we say that the society is *polarized*. Our purpose is to study polarization, and to provide a theory of its measurement. Our contention is that polarization, as conceptualized here, is closely related to the generation of social tensions, to the possibilities of revolution and revolt, and to the existence of social unrest in general. We take special care to distinguish our theory from the theory of inequality measurement. We derive measures of polarization that are easily applicable to distributions of characteristics such as income and wealth.

KEYWORDS: Clustering, convergence, inequality, income distribution, polarization, social conflict.

1. INTRODUCTION

THIS PAPER IS CONCERNED WITH THE CONCEPTUALIZATION and measurement of polarization. We shall argue that the notion is fundamentally different from *inequality*, which has received considerable attention in the literature on economic development and elsewhere.[2]

Let Y be a set of *characteristics* or *attributes*. An individual will be described by an element in this set. The prevailing state of affairs in a society can then be captured by a *distribution* over Y, which describes the proportions of the population possessing attributes in any subset of Y.

Consider a particular distribution on Y. Suppose that the population is grouped into significantly-sized "clusters," such that each cluster is very "similar" in terms of the attributes of its members, but different clusters have members with very "dissimilar" attributes. In that case we would say that the society is "polarized." Our purpose is to study these intuitive criteria carefully, and to provide a theory of measurement.

When introducing a new concept, one should start by showing that it is different from other standard summary statistics and by persuading the reader

[1] Research on this project was started when Esteban was visiting the Indian Statistical Institute in 1987. The results reported here were obtained when Ray (then at the I.S.I.) visited the Instituto de Análisis Económico under the Sabbatical Program of the Ministry of Education, Government of Spain. We are grateful for funding from the Ministerio de Educación y Ciencia (Grant No. PB90-0172) and EC FEDER Contract No. 92 11 07 005, and for supportive facilities at the I.S.I. We have benefited from long discussions with David Schmeidler, and have also received useful comments from Anthony Atkinson, Partha Dasgupta, Bhaskar Dutta, Michael Manove, Isaac Meilijson, Nicholas Stern, and participants of numerous seminars and conferences. The current version owes much to the comments of Andreu Mas-Colell and two anonymous referees.

[2] For theoretical studies, see Kolm (1969), Atkinson (1970), Sen (1973), Dasgupta, Sen, and Starrett (1973), Fields and Fei (1978), and the numerous references in the excellent survey by Foster (1985). For empirical studies, the reader may consult Kuznets (1955), Kravis (1960), Adelman and Morris (1973), Weisskoff (1970), Ahluwalia (1976), and Fields (1980), to mention only a few.

that the phenomenon captured by the new concept is indeed relevant. Our task is not easy for at least three reasons. First, one cannot rely on pre-existing intuition or axiomatization. We shall have to introduce the concept and develop its intuition at the same time. Second, the theory of economic inequality—the obvious contender to the theory developed here—is based on a long and venerable tradition in welfare economics. This tradition is by now so deeply rooted in our way of thinking about distributions, that sometimes we take the propositions proven two decades ago as unquestionably intuitive axioms.[3] Finally, while we see polarization as a particularly relevant correlate of potential or open social conflict, "mainstream" economics has thus far paid little attention to this last issue.[4]

We begin with the obvious question: why are we interested in polarization? It is our contention that the phenomenon of polarization is closely linked to the generation of tensions, to the possibilities of articulated rebellion and revolt, and to the existence of social unrest in general. This is especially true if the underlying set of attributes is a variable such as income or wealth. A society that is divided into groups, with substantial intra-group homogeneity and inter-group heterogeneity in, say, incomes, is likely to exhibit the features mentioned above. At the same time, *measured inequality in such a society may be low*. The reader unconvinced of this last point may turn at once to Section 2, where the distinction between polarization and inequality is made via a series of examples.

The idea above is not novel at all. Marx was possibly the first social scientist to give a coherent interpretation of society as characterized by the building up of conflict and its resolution. We are not interested here in his specific theory of the development of the capitalist society, as much as in his view of intergroup conflict dynamics, and the use of a concept such as polarization to describe these dynamics.[5] Deutsch (1971)'s account of the Marxian theory is particularly relevant for our purpose:

> "As the struggle proceeds, 'the whole society breaks up more and more into two hostile camps, two great, directly antagonistic classes: bourgeoisie and proletariat.' The classes *polarize*, so that they become internally more homogeneous and more and more sharply distinguished from one another in wealth and power" (Deutsch (1971, p. 44)).

This way of understanding social conflict has been later taken up by contemporary sociologists (e.g., Simmel (1955) and Coser (1956)) and political scientists

[3] We find particularly revealing the questionnaire results obtained by Amiel and Cowell (1992). The majority of students with an education in Economics ranked distributions in a manner consistent with the Lorenz ordering, while students with other backgrounds produced rankings that violated this ordering.

[4] There is a strand of recent literature that shares our view that social conflict, and in particular distributional conflict, plays a crucial role in the explanation of macroeconomic phenomena. For instance, Alesina et al. (1992), Alesina and Rodrik (1991), Benhabib and Rustichini (1991), and Persson and Tabellini (1991) focus on the relationship between distributional conflict and economic growth.

[5] See also Lenin (1899), Chayanov (1923), and Shanin (1966).

(e.g. Gurr (1970, 1980) and Tilly (1978)).[6] In fact, the broad view that a society split up into two well-defined and separated camps exhibits a high degree of potential (if not open) social conflict is neither new nor specifically Marxist. It has also been shared by such different political analysts as Benjamin Disraeli, the nineteenth century Tory prime minister, or the US Kerner Commission, twenty five years ago.[7]

This literature differs from studies on inequality in one simple yet fundamental way: in the formation of categories or groups, or "subsocieties," *the population frequency in each category also carries weight*.[8]

The following simple example is illustrative. Consider an initially equal society of peasant farmers. Suppose, now, that through a process that redistributes a certain fixed amount of agricultural surplus, a fraction α of these peasants become rich *kulaks* at the expense of the remainder. Given the value of the surplus, for which values of α would we say that a "high" degree of differentiation is taking place? Of course, there is no precise *numerical* answer to this question, but certainly values of α very close to zero or unity would *not* be construed as creating a sharp division among the population. However, note that under *every* known measure of inequality, lower values of α would unambiguously give rise to higher inequality.[9] By effectively neglecting the population frequency in each category, inequality measurement departs from the study of differentiation.

Another way to see the difference is to recognize that the axioms of inequality measurement (or equivalently, second-order stochastic dominance for mean-normalized distributions) fail to adequately distinguish between "convergence" to the global mean and "clustering" around "local means."[10] Consider the following example, related to the question of economic "convergence" in the world economy, a topic that has recently received considerable attention.[11] To be precise, consider the frequency distribution of national growth rates in

[6] Marx's critic R. Dahrendorf (1959) himself does not attack this view of conflict, but rather that Marx overlooked the fact that individual characteristics are multidimensional, so that differences within and similarities between social classes restrained the polarization process (see Deutsch (1971)). Indeed, polarization in some attribute such as income may be counterbalanced by dissimilarities in other attributes, an objection common to the theory of inequality.

[7] Disraeli coined the term "two nations" to describe the English social structure of his time in his novel *Sybil*, published in 1845. The Kerner Commission (1968) expressed concern about the ongoing division in the United States into two societies.

[8] By focussing on this one difference, we do not mean to deny the existence of others. In particular, it is difficult to remain unimpressed by the richness of the political and sociological studies alluded to above. Nevertheless, this very richness opens the door to a possible lack of sharpness at the conceptual level. The theory of inequality can certainly not be criticized on these grounds. Our objective, then, is to retain the sharpness of a formal construction, capturing the new features in the cleanest possible way.

[9] To be precise, this would happen in the case of every Lorenz-consistent inequality measure.

[10] Landsberger and Meilijson (1987) also note (in the context of risk) that second-order stochastic dominance is not necessarily associated with transfers of probability mass from the *center* of the support of a distribution to its tails.

[11] See, for example, Barro (1991), Barro and Sala-i-Martin (1992), and Mankiw, Romer, and Weil (1992).

per capita GDP. Suppose that over time, we were to observe that "low" growth rates were converging to their common mean, while the same is true for "high" growth rates. In terms of growth rates, then, two "growth poles" would be forming—a clear emergence of North and South. Polarization in our sense may well be rising. But *every* inequality measure defined on growth rates, and consistent with the Dalton Transfers Principle would register an unambiguous decline![12]

Indeed, there are a number of social and economic phenomena for which the knowledge of the degree of clustering or polarization can be more telling than a measure of inequality. Quite apart from the distribution of income, wealth, or growth rates, relevant economic examples include labor market segmentation, concepts of the dual economy in developing countries, or distribution of firms by size in a given industry. In the broader arena of the social sciences, there are issues of social class or significant problems concerning racial, religious, tribal, and nationalistic conflicts, which clearly have more to do with the clustering of attributes than with the inequality of their distribution over the population.

All this suggests a need to develop a theory of the measurement of polarization. Indeed, the theory of inequality measurement is possibly on far weaker ground in its claim to independent existence. Given the underpinnings of inequality theory as simply welfare economics applied to a restricted domain,[13] one might question the value added of this specific attention. A good part of the literature has found its justification in the assertion that one should decompose distributional questions from questions of the aggregate product (or GDP). Even if the social welfare function were to enjoy this curious property,[14] there is still no need to treat the theory of inequality as separate from welfare economics in general, distinguished by properties intrinsic to the theory itself. More closely related to the subject of this paper is the motivation for studying inequality as a phenomenon closely related with social unrest. The opening page of Sen's celebrated book, *On Economic Inequality*, asserts that "the relation between inequality and rebellion is indeed a close one, and it runs both ways." We do not disagree with this position, but it remains to be seen empirically whether this connection is specifically with inequality, or with features that are

[12] See Section 2 for a numerical example. Our preliminary studies using the World Bank data set reveal that the distribution by frequencies over the annual growth indices by countries is multimodal. We compared the behavior of the polarization measures developed in this paper with inequality indices applied to the same data. Specifically, we find that while the Gini coefficient continually diminished over the period 1979–86, polarization falls too in 1981 but increases thereafter. We might add that we plan to apply the ideas developed in this study to the empirical question of "international convergence." But that is the subject of another paper. For interesting recent research that bears on the issue of multimodality in this context, see Quah (1992).

[13] For measures of relative inequality, this is the domain over which the total sum of incomes is constant.

[14] For studies in this area, see Sheshinski (1972), Blackorby and Donaldson (1978), Ebert (1987), and Dutta and Esteban (1992).

common to both polarization and inequality,[15] and perhaps more closely tied to the former in cases of disagreement.[16]

We end our introductory discussion with a remark on the basic set of attributes that we use for the development of our measurement theory. Ideally, this set must include all individual attributes that are relevant (at the social level) for creating differences or similarities between persons, possibly relative to a given announced socioeconomic policy.[17] But we will simplify the analysis enormously by assuming that the attribute can be captured by the values of some *scalar* variable. Indeed, we work with the natural logarithm of income (though this particular interpretation is formally quite unnecessary). The idea is that percentage differences in income are a good proxy for socioeconomic differences or similarities. More generally, even if the attribute space is multidimensional, our analysis is undisturbed if the differences between attributes can be represented by a metric. But such mathematical sleight of hand only sweeps a serious dimensionality issue under the rug. So we do not defend our restriction except on grounds of tractability. We note, moreover, that exactly the same criticisms may be levelled against the theory of inequality measurement.[18]

The rest of the paper is organized as follows. Section 2, which takes the reader through a series of examples, is designed to make two points. First, we argue that there is an analytical distinction between polarization and inequality. Second, we submit that there is no simple (and acceptably rich) partial order (like the Lorenz criterion in inequality theory) that will capture increasing polarization. Specifically, any measure of polarization must be "global" and "nonlinear," in a sense that is made precise in Section 2. In Section 3, we formally develop and study polarization measures. Section 3.1 deduces an allowable class of measures, obtained by combining a behavioral model with certain intuitive assumptions or axioms. Theorem 1 reveals that the allowable class of measures is quite sharp and amenable to empirical use. Section 3.2 (Theorem 2) studies the maximal elements of polarization measures, and reveals

[15] After all, there are a number of situations (see Section 2) where polarization and inequality measures agree in their rankings. In this intersection falls, for instance, the phenomenon of a disappearing middle class, a matter of concern to sociologists and some economists in the context of U.S. society. See, for example, Kosters and Ross (1988), Horrigan and Haugen (1988), and Duncan, Smeeding, and Rodgers (1991).

[16] For instance, the empirical findings of Nagel (1974) show that social conflict is low both under complete equality and under extreme inequality. Indeed, political scientists working on inequality and conflict, such as Midlarski (1988) and Muller, Seligson, and Fu (1989) have already served notice that standard notions of inequality may be inadequate for the study of conflict. Coming across their work after a first draft of this paper was written, we see that their alternative notions of *patterned inequality* and *bifurcated inequality* seem to be motivated by the same arguments we put forward in this paper. A thorough empirical exploration of our contentions is the subject of a forthcoming paper.

[17] Indeed, the idea of using spaces of policies and measuring polarization as a state of affairs that yield high resistance to policies chosen at random forms the basis of an alternative development of the theory. But this is beyond the scope of the present paper.

[18] While multidimensional variants of inequality theory do exist, to go beyond the one-dimensional theory they must impose particular restrictions on how dimensions are to be compared.

them to be degenerate, symmetric bimodal distributions. In Section 3.3 (Theorem 3), we sharpen the class of allowable measures even further by the use of an additional plausible assumption. In Section 3.4, we illustrate how these measures work by applying them to a set of examples, including the motivating examples in Section 2. In Section 4, we extend our class of measures beyond the simple model of Section 3, in a form that might be even more suitable for empirical work. Section 5 describes possible extensions of the basic theory, and concludes the paper. This section also contains remarks on the issue of population and income normalization for the measures we obtain.

2. POLARIZATION: OVERVIEW AND SOME EXAMPLES

The objective of this section is to study a number of examples in an informal way. Some of these examples will be elevated in the sequel to the status of axioms. At present, however, our statements are not meant to be precise. They are meant to align your intuition to ours, to form a preliminary judgement regarding the formal concept we shall later introduce.

Loosely speaking, every society can be thought of as an amalgamation of *groups*, where two individuals drawn from the same group are "similar," and from different groups, are "different" relative to some given set of *attributes* or *characteristics*. The *polarization* of a distribution of individual attributes must exhibit the following basic features.

FEATURE 1: There must be a high degree of homogeneity *within* each group.

FEATURE 2: There must be a high degree of heterogeneity *across* groups.

FEATURE 3: There must be a small number of significantly sized groups. In particular, groups of insignificant size (e.g., isolated individuals) carry little weight.

We start by comparing two alternative distributions of a given attribute (say income) over a population. Example 1 below is designed to illustrate Feature 1.

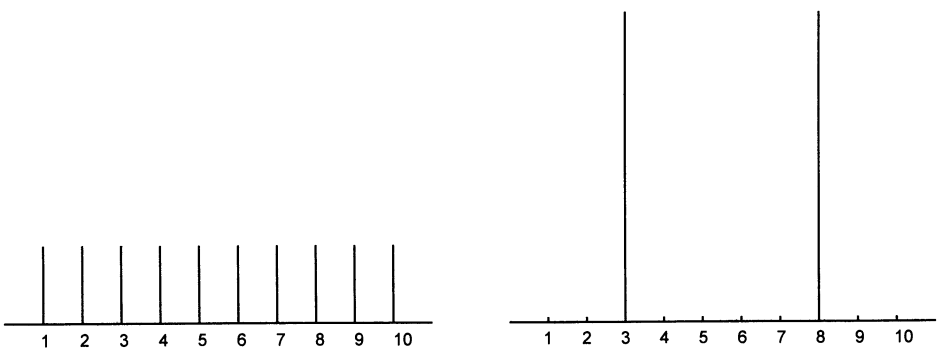


FIGURE 1A FIGURE 1B

FIGURE 1.—Diagrams to illustrate Example 1.

EXAMPLE 1: In Figure 1a, the population is uniformly distributed over ten values of income, spaced apart equally at the points $1, 2, \ldots, 10$. In Figure 1b, we collapsè this distribution into a two-spike configuration concentrated equally on the points 3 and 8. This diagram formalizes the discussion of *per capita* GDP in the Introduction as a specific example, or *per capita* growth rates if the attribute is reinterpreted appropriately.

Which distribution exhibits greater "polarization"? We would argue that it is the latter distribution. Two groups are now perfectly well formed in the second diagram, while in the former diagram the sense of group identity is more fuzzy. Under the second distribution population is either "rich" or "poor," with no "middle class" bridging the gap between the two, and one may be inclined to perceive this situation as more conflictual than the initial one.[19]

If you are still hesitant, think of the attribute as an opinion index over a given political issue, running from right to left. Many would agree that political conflict is more likely under a two-spike distribution—with perhaps not completely extreme political opinions, but sharply defined and involving population groups of significant size—rather than under the uniform distribution.

But the point is this. If you admit the possibility of greater polarization in Figure 1b, you are forced to depart from the domain of inequality measurement. For under *any* inequality measure that is consistent with the widely accepted Lorenz ordering, inequality has come *down* in Figure 1b relative to 1a!

EXAMPLE 2: This example illustrates Feature 2 above. It also shows that there are cases where increasing polarization also conforms to Lorenz-worsening of the underlying distribution. Consider Figures 2a and 2b. Think of the attribute as income. Figure 2a is just the old Figure 1b. In Figure 2b, we also have a two point income distribution, but this time concentrated equally on the incomes 1 and 10. Which has more polarization? Well, as far as *intra*-group homogeneity goes, there is nothing to choose between the two diagrams. But there is greater *inter*-group heterogeneity in Figure 2b. So it should be fairly easy to conclude that there is more polarization in Figure 2b. But note that this time income inequality, too, has worsened in Figure 2b (relative to 2a). In particular, we do not claim that the notion of polarization *always* conflicts with that of inequality.

EXAMPLE 3: With the help of Figure 3, we illustrate Feature 3. In Figure 3a, the population is distributed according to a three-point distribution, with the attributes spaced equally. Think of the central and the right-hand masses as being of approximately the same size. Consider a small shift of population mass from the left extreme to the right extreme. The problem is immediately seen to be more complex than those of the previous examples. While it is true that the shift creates a greater tendency towards forming two sharply defined groups, the

[19] In this context, see Proposition 9 in Coser (1956) (which draws on Simmel (1955)), whereby the internal cohesion of groups is associated with increased conflict between groups.

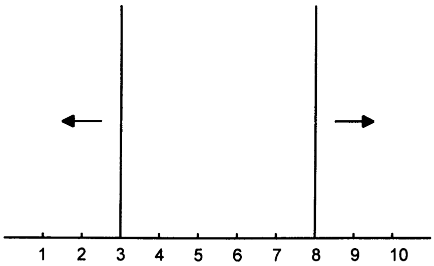


FIGURE 2A

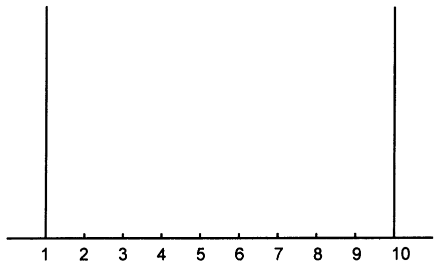


FIGURE 2B

FIGURE 2.—Diagrams to explain Example 2.

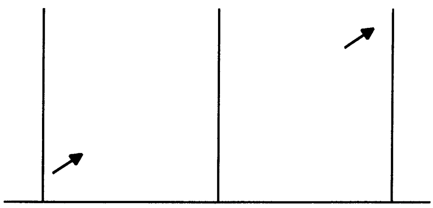


FIGURE 3A

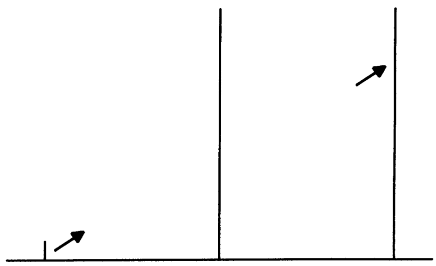


FIGURE 3B

FIGURE 3.—Diagrams to explain Example 3.

left mass may well be instrumental in creating part of the social tensions that do exist, and the net effect is far from clear.

Consider, however, the case illustrated in Figure 3b, where the left mass is very tiny indeed compared to the other two. In this case, the *initial* contribution of the left mass is vanishingly low and it might make sense to argue that the very same move now serves to increase polarization. This is the substance of Feature 3. While we do not adopt the example of Figure 3b as a basic assumption for our main result, we do explore its implications for the choice of measure (see Section 3.3 and Theorem 3). To us, it is reasonable as a basic axiom, though it is of interest to see how far we can go without it.

The examples above are meant to show that the concept of polarization (whatever it is) is logically separate from that of economic inequality, and to bring out certain defining features of the concept. The two remaining examples in this section are meant to illustrate an entirely different set of issues. Specifically, we wish to argue that any reasonable measure of polarization must be *global* in nature, in a way that inequality measures are not. To see this point, recall the Dalton Principle of Transfers that underlies the Lorenz ordering. The principle states that starting from *any* distribution of income, *any* transfer of income from an individual to one richer than him *must* increase inequality. The principle is a *local* one. To apply it, it is unnecessary to take account of the

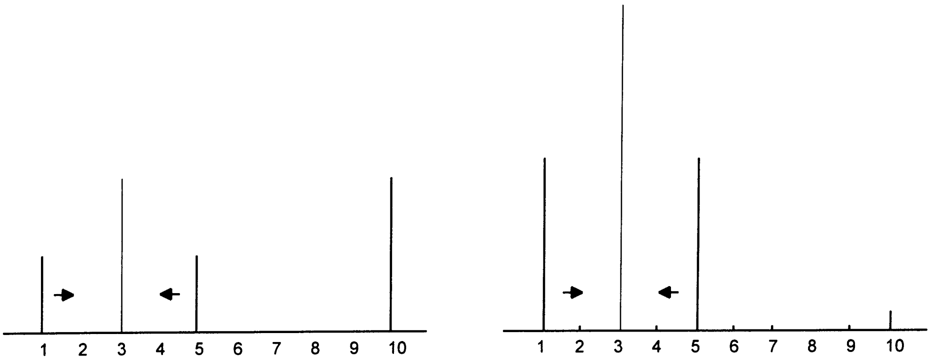


FIGURE 4A FIGURE 4B

FIGURE 4.—Diagrams to illustrate Example 4.

original distribution. In our thinking about the subject, we have found it impossible to come up with a similar local prescription for increasing polarization. The examples that follow try to indicate why this is so.

EXAMPLE 4: Consider the two income distributions in Figures 4a and 4b. Each distribution has point masses at the income levels 1, 5, and 10. In Figure 4a, half the population is equally divided between incomes 1 and 5, and the other half is at income 10. Consider, now, the fusion of the first two groups into a single mass at the income level 3 (see arrows in 4a). Has polarization gone up? Note that the collection of individuals at income levels 1 to 5 no longer have any interpersonal animosity; they are all united at income level 3. They now see as their "common enemy" the sizeable group at income level 10. Moreover, so far as the group at income 10 is concerned, they now see facing them a united group of individuals of similar size. On average, this group is as different from the 10-group as it was before.

It would, indeed, be difficult to defend the view that polarization has decreased in this situation!

But now turn to Figure 4b. Here, the only difference is that almost the entire population is divided between the two income levels of 1 and 5, and there is a tiny fraction of individuals at income level 10. Perform the same experiment as before. You will notice that it is now practically impossible to make the same argument as we made for the preceding case. Because of the small size of the 10-group, "most" of the "polarization" in Figure 4b comes from the fact that there are two groups situated at incomes 1 and 5. Put them together and you have wiped out most of the social tension. It is true that there is a group at 10, but it is a very *small* group, and any notion of continuity[20] would require that the intra-group tension also be small in the new situation, so in Figure 4b the opposite argument—polarization has come *down*—is more convincing.

[20] Simply send ε to 0 in Figure 4b. At $\varepsilon = 0$, polarization *must* come down by the argument of Example 2. So this must also happen for $\varepsilon > 0$ but small!

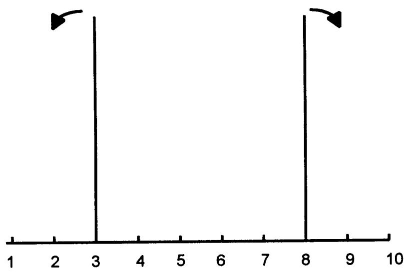


FIGURE 5A

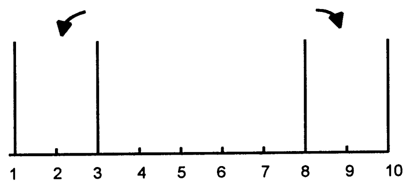


FIGURE 5B

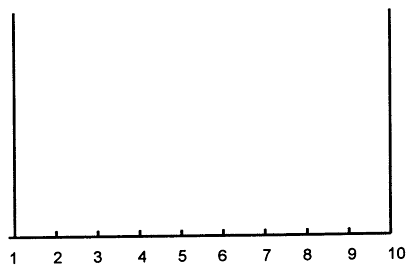


FIGURE 5C

FIGURE 5.—Diagrams to illustrate Example 5.

EXAMPLE 5: Reconsider Example 2. Imagine moving from Figure 2a to Figure 2b by a series of small changes. In each change, equal fractions of people at incomes 3 and 8 are removed and replaced at 1 and 10 respectively. The end result is, of course, Figure 2b. Figures 5a and 5c reproduce Figure 2, with Figure 5b displaying an intermediate scenario where the population is equally divided among the income levels 1, 3, 8, and 10. Note that these changes generate a sequence of regressive Lorenz dominated distributions.

We have already argued that 5c displays more polarization than 5a. But does 5b display more polarization than 5a? We claim that there is no unambiguous answer here. While it is true that there are very different groups in 5b (relative to 5a), there is also little group homogeneity. It all depends on whether one gives more weight to intergroup differences at the expense of within group homogeneity. We are not making any particular claim as to a direction of change, but are only trying to convince you that there is a genuine ambiguity here.

These examples (4 and 5) are meant to illustrate the possibility that the search for a (reasonably rich) partial order for increases in polarization can be a difficult one. For one thing, the effect (on polarization) of a given change may depend on factors that are not directly associated with the change (e.g., the size of the 10-groups in Figures 4a and 4b). This is what we meant above by the *global* nature of the concept. Moreover, the same *directions* of change can be associated with different effects on polarization, depending on the initial configurations. This point is illustrated by both Examples 4 and 5. Both these features are absent in the theory of inequality measurement, embodied in the Lorenz ordering.

In the next section, we present a formal analysis that attempts to capture some of the issues raised here.

3. POLARIZATION MEASURES

3.1. *Axiomatic Derivation of a Class of Measures*

3.1.1. *Distributions and Polarization Measures*

The previous section should have already made clear that the concept of polarization is somewhat complicated, and no *single* simple axiom (such as the Dalton Axiom in the context of inequality measurement) will serve to capture its essence. Accordingly, our aim is limited. We introduce a model of *individual attitudes* in a society that leads naturally to a broad class of polarization measures. On this class, we place certain axioms and narrow down the set of allowable measures considerably. Finally, we show that the measure we obtain performs "well" in the context of the examples earlier described.

Our basic perceptual variable is the natural logarithm of income, denoted by y, and we shall presume that its values lie in $\mathbb{R}$ (zero incomes are not permitted). By an *individual* $\hat{y}$ we shall mean an individual who has income (with log equal to) $\hat{y}$. The choice of this variable is based on the presumption that only percentage differences matter. But any other scalar can be used as the basic perceptual variable with little overall conceptual difference.[21]

In this paper, we consider only those distributions with support on some *finite* set of incomes. This simplification is made for expository convenience. There are nontrivial additional issues involved in dealing with the case of distributions with infinite support. We comment on these briefly below, and refer the reader to Esteban and Ray (1991) for a detailed development of the theory in this case.

We begin with some background definitions and notation. For any positive integer n, $(\pi, \mathbf{y}) \equiv (\pi_1, \ldots, \pi_n; y_1, \ldots, y_n)$ is a *distribution* if $\mathbf{y} \in \mathbb{R}^n$, $y_i \neq y_j$ for all i, j, and $\pi > 0$. The total *population* associated with $(\pi, \mathbf{y})$ is given by $\sum_{i=1}^{n} \pi_i$. Denote by $\mathscr{D}$ the space of all distributions.

A *polarization measure* (PM) is a mapping P: $\mathscr{D} \to \mathbb{R}_+$.

[21] It may be argued, for example, that large, percentage differences in income make little difference to an individual after some threshold. In that case, the basic perceptual variable might be some strictly concave transform of y.

Throughout, we will suppose that the ranking induced by a polarization measure over two distributions is invariant with respect to the size of the population. This homotheticity property is standard in the theory of inequality measurement (see, e.g., Foster (1985)). We refer to such a property as Condition H.

CONDITION H: *If* $P(\pi, \mathbf{y}) \geqslant P(\pi', \mathbf{y}')$ *for two distributions* $(\pi, \mathbf{y})$ *and* $(\pi', \mathbf{y}')$, *then for all* $\lambda > 0$, $P(\lambda\pi, \mathbf{y}) \geqslant P(\lambda\pi', \mathbf{y}')$.

We now turn to the study of a behavioral model that will yield a class of polarization measures.

3.1.2. *A Model*

Our first concept deals with the fact that intra-group homogeneity accentuates polarization. We propose that an individual y feels a sense of identification with other individuals who have the same incomes as him. Thus the identification felt by an individual is an increasing function of the number of individuals in the same income class of that individual. Formally, we introduce a continuous *identification function* I: $\mathbb{R}_+ \to \mathbb{R}_+$. Assume $I(p) > 0$ whenever $p > 0$. Formally, no further assumption need be made, though it is no surprise that we will later deduce that $I(p)$ is increasing (and more).

Two remarks are in order. First, we have postulated a particularly sharp form of identification whereby individuals feel identified with people who earn *exactly the same income*, but with no one else. This is defensible only if we think of incomes as point estimates of income *classes* or intervals. It is surely more reasonable to also allow for (a perhaps smaller) identification with neighboring individuals. Indeed, as we point out in detail in Section 4, this postulate leads to a drawback with our derived class of measures. We have nevertheless chosen this route in the interest of a cleaner model.

Second, the sense of identification may depend not only on the *number* of similar individuals but also on the common characteristics (in this case, income) that these individuals possess. We comment on this extension in Section 5.3.

Next, we posit that an individual feels alienated from others that are "far away" from him. This concept will deal with the fact that inter-group heterogeneity accentuates polarization. Let a: $\mathbb{R}_+ \to \mathbb{R}_+$, with $a(0) = 0$, be a nondecreasing continuous function, to be called the *alienation function*. We assume that individual y feels an alienation $a(\delta(y, y'))$ with an individual y', where $\delta(y, y')$ stands simply for absolute distance $|y - y'|$.

We pause here to explicitly take note of an important feature subsumed in the definition. We have assumed that the concept of alienation (as well as identification) is perfectly symmetric. An individual with low income feels the same alienation towards an individual with high income as the latter feels towards the former. This is perhaps inappropriate, as it can be argued that a polarized society arises from the alienation that the poor feel towards the rich, and not the other way around. We discuss an "asymmetric" model that captures these features in Section 5.2.

Now we are in a position to put together these two concepts. What we wish to capture is the *effective* antagonism that y feels towards y'. This is, roughly speaking just the same as y's alienation *vis-a-vis* y', but we wish to allow for the possibility that an individual's feelings of identification may influence the "effective voicing" of his alienation.[22] In the sequel, this influence will turn out to be the critical feature that distinguishes polarization from inequality.

Accordingly, we allow for this influence (though we do not assume it). We postulate that the *effective antagonism felt by y towards y'* is given by a continuous function $T(I, a)$, where $a = a(\delta(y, y'))$ and $I = I(p)$, with p the measure of people in the income class of individual y. This function T is taken to be strictly increasing in a whenever $(I, a) \gg 0$. We assume further that $T(I, 0) = 0$.[23]

Finally, the total polarization in the society is postulated to be the sum of all the effective antagonisms:

$$(1) \qquad P(\pi, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i \pi_j T\big(I(\pi_i), a\big(\delta(y_i, y_j)\big)\big).$$

The postulate (1) embodies two assumptions. First, we record that polarization depends on the "vector" of effective antagonisms in the society and on this *alone*; in this respect it is analogous to a Bergson-Samuelson-type postulate of nonpaternalism. Second it goes further and proposes an analogue of additive utilitarianism. Some justification for the additive postulate may be found along lines suggested by Harsanyi (1953) and others: that an impartial observer who might find himself in any position in the given distribution might use the expected value of his effective antagonism to judge overall polarization.

On the other hand, we have left the structure of the functions $I(\cdot)$, $a(\cdot)$, and $T(\cdot)$ very general indeed. The choice of any one functional form for each will yield a *particular* measure of polarization.

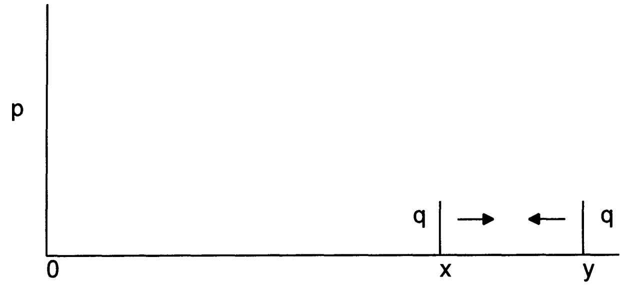
[22] Thus, we are primarily interested in notions of *organized* opinion, protest, etc. The notion of an individual who runs amok, perhaps provoked by his sense of isolation, is an important one that we do not consider here.

[23] Note that these assumptions formally permit identification to *not* have any influence at all ($T(I, a) = a$ is permitted, for instance), or perhaps even have negative influence, though these features will disappear in the sequel. Note, moreover, that it is also possible to weight one's feelings of alienation by invoking the visibility of the group *from* which the individual feels alienated. But there is nothing so far in our discussion and examples which appear to necessitate this. Indeed, whether such a weighting should enter positively or negatively is far from intuitively clear. So we ignore this in the interests of simplicity. See Section 5.3 for more discussion.

3.1.3. *Axiomatic Derivation*

We are interested in narrowing down the class of allowable measures by imposing some "reasonable" axioms.

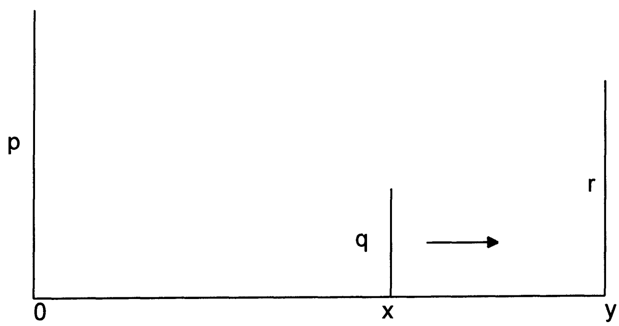
AXIOM 1:



Data: $p, q \gg 0$, $p > q$, $0 < x < y$.
Statement: Fix $p > 0$ *and* $x > 0$. *There exists* $\varepsilon > 0$ *and* $\mu > 0$ *(possibly depending on* p *and* x*) such that if* $\delta(x, y) < \varepsilon$ *and* $q < \mu p$*, then the joining of the two* q *masses at their mid-point,* $(x+y)/2$*, increases polarization.*

The intuition behind this axiom should be very clear. The two right hand masses are individually smaller than p (see Data). Moreover, they are "very close" to each other (see Statement). In such a case, by pooling the two small masses we "identify" them while not changing the average distance from the third mass. This should therefore raise polarization.[24]

AXIOM 2:



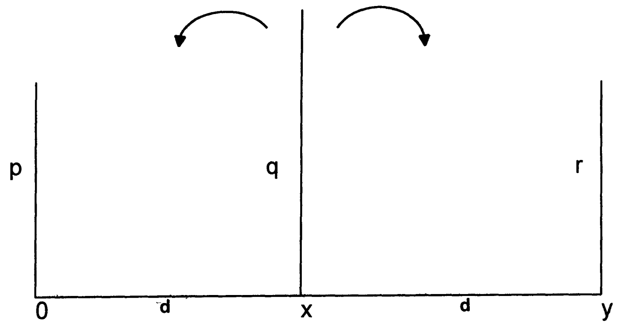
[24] The axiom is actually stated in very weak terms. The axiom is invoked only when the distance between x and y is small, and also only when the q-masses are small relative to p.

Data: $(p, q, r) \gg 0,\ p > r,\ x > |y - x|$.
Statement: There is $\varepsilon > 0$ *such that if the population mass* q *is moved to the right (towards* r*) by an amount not exceeding* ε*, polarization goes up.*

This is also an intuitive axiom. The intermediate point mass q is at least as close to the r-mass as it is to the p-mass. Additionally, the p-mass is larger than the r-mass. So if only small locational changes in the q-mass are permitted, the direction that brings it closer to the nearer *and* smaller mass should raise polarization.

We now turn to our final axiom.

AXIOM 3:



Data: $(p, q) \gg 0,\ x = y - x \equiv d$.
Statement: Any new distribution formed by shifting population mass from the central mass q *equally to the two lateral masses* p*, each* d *units of distance away, must increase polarization.*

This axiom is so intuitive it hardly requires comment. The axiom states that the disappearance of a "middle class" into the "rich" and "poor" categories must increase polarization.

We may now state our main result. To do so, it will be necessary to introduce the function f: $\mathbb{R}^2_+ \to \mathbb{R}$, defined by

$$(2) \qquad f(z, \alpha) \equiv (1 + \alpha)\left[z - \frac{z^\alpha}{2} - z^{1+\alpha}\right] - \frac{1}{2}.$$

It is possible to show that there exists a number $\alpha^* > 0$ such that $\max_{z \geqslant 0} f(z, \alpha) < 0$ (resp. $\leqslant 0$) if and only if $\alpha < \alpha^*$ (resp. $\leqslant \alpha^*$). Moreover, it can be shown that $f(z, \alpha^*) < 0$ for all but one point z^*, where equality holds. It is also possible to show analytically that $\alpha^* \in (1, 2)$. Numerical computation reveals that $\alpha^* \simeq 1.6$. Details of the proofs of these assertions and the numerical computation are available on request.

We now have the following theorem.

THEOREM 1: *A polarization measure P^* of the family defined in* (1) *satisfies Axioms* 1, 2, *and* 3, *and Condition H, if and only if it is of the form*

$$P^*(\pi,\mathbf{y}) = K\sum_{i=1}^{n}\sum_{j=1}^{n}\pi_i^{1+\alpha}\pi_j|y_i-y_j| \tag{3}$$

for some constants $K>0$ and $\alpha\in(0,\alpha^*]$ where $\alpha^*\simeq 1.6$ (see above, equation (2)).

This theorem dramatically narrows the class of allowable polarization measures. The only two degrees of freedom are in the constants K and α. The former is simply a multiplicative constant which has no bearing on the order, but which we use for population normalization (see Section 5.1 below). The constant α reflects our deduction that the identification function must be of the form p^α, where $\alpha>0$. Note that no *a priori* assumption of how identification affects effective antagonism had been postulated in the model. The deduction that this effect is indeed positive is not surprising, given the earlier discussion and examples. This is *precisely* what will distinguish polarization measures from inequality measures.

It should be noted, however, that α cannot take on arbitrary positive values, but must be bounded in the way indicated by the theorem. Without this bound, all the axioms *cannot* be satisfied and furthermore, the bound is needed to verify other intuitive properties of the measure (see below).

Indeed, each one of the restrictions implied by the theorem is important in signing the direction of change in intuitive examples. The reader may turn right away to Section 3.2, where a number of examples are worked out and the use of each of these conditions is illustrated.

It should be noted that the measure P^* bears a striking resemblance to the Gini coefficient. Indeed barring the fact that we are using the logarithm of incomes, our measure *would* be the Gini if α were equal to zero. It is precisely the fact that population weights are raised to a power exceeding unity in the formula above that gives rise to distinctly different behavior of a polarization measure.

Thus α may be treated as the degree of "polarization sensitivity" of the derived measure. The larger is its value, the greater is the departure from inequality measurement. For more on this, see the discussion following the statement of Theorem 3, and Section 3.4.2.

With these remarks made, we now turn to a proof.

PROOF: (*Necessity.*) Define θ: $\mathbb{R}_+^2\to\mathbb{R}_+$ by $\theta(\pi,\delta)\equiv T(I(\pi),a(\delta))$. *First, we show that $\theta(\pi,\cdot)$ is linear for all $\pi>0$.* Consider the case depicted in Axiom 1, with $\delta(x,y)<\varepsilon$ and $q<\mu p$. Total polarization is given by

$$\begin{aligned} P^1 \equiv{} & pq[\theta(p,x)+\theta(p,y)] \\ & +pq[\theta(q,x)+\theta(q,y)]+2q^2\theta(q,|y-x|). \end{aligned} \tag{4}$$

When the two q's are put together at $(x+y)/2$, polarization is

$$(5)\qquad P^2 \equiv 2pq\left[\theta\left(p, \frac{x+y}{2}\right)\right] + 2pq\left[\theta\left(2q, \frac{x+y}{2}\right)\right].$$

So by Axiom 1 we have, combining (4) and (5),

$$2p\left[\theta\left(p, \frac{x+y}{2}\right) + \theta\left(2q, \frac{x+y}{2}\right)\right] > p[\theta(p,x) + \theta(p,y)] + p[\theta(q,x) + \theta(q,y)] + 2q\theta(q, |y-x|).$$

Passing to the limit as $q \to 0$, we get for each $p > 0$,

$$(6)\qquad 2\theta\left(p, \frac{x+y}{2}\right) \geqslant \theta(p,x) + \theta(p,y).$$

The above holds, it should be noted, for each $(p,x) \gg 0$ when y is "sufficiently" close to x. Nevertheless, it is possible to show that because $\theta(p,\cdot)$ is continuous for each p, $\theta(p,\cdot)$ *must be concave.* The details of this argument are straightforward but tedious, and are therefore omitted.

Next, pick any $(p,x,x') \gg 0$ with $x > x'$. *We claim that there exists $D > 0$ such that for all $\Delta \in (0, D]$, $\theta(p, x+\Delta) \geqslant \theta(p, x'-\Delta)$.*

To establish this claim, consider Axiom 2. For given $(p,x,x') \gg 0$, choose $(q,r) \gg 0$ such that $p > r$ and y such that $y - x = x'$. Then (p,q,r,x,y) satisfies all the conditions of that axiom. Now, polarization is initially given by

$$(7)\qquad P_\Delta \equiv pr[\theta(p,y) + \theta(r,y)] + pq[\theta(p,x+\Delta) + \theta(q,x+\Delta)] + qr[\theta(q,x'-\Delta) + \theta(r,x'-\Delta)]$$

evaluated at $\Delta = 0$. By the axiom, it must be the case that $P_\Delta > P_0$ for *all* values of $\Delta \in (0,x')$.[25] Using this information in conjunction with (7), we see that $pq[\{\theta(p,x+\Delta) - \theta(p,x)\} + \{\theta(q,x+\Delta) - \theta(q,x)\}] \geqslant qr[\{\theta(q,x') - \theta(q,x'-\Delta)\} + \{\theta(r,x') - \theta(r,x'-\Delta)\}]$ for all $\Delta \in (0, x'/2)$. Passing to the limit as $(q,r) \to (p,p)$, we deduce that

$$\theta(p,x+\Delta) - \theta(p,x) \geqslant \theta(p,x') - \theta(p,x'-\Delta),$$

for all $0 < \Delta \leqslant D \equiv x'/2$. This establishes the claim. Combining this claim with the earlier deduction that $\theta(p,\cdot)$ is concave (and continuous) for all p, we may conclude at once[26] that there exists a continuous function $\phi(\cdot)$ with $\phi(p) > 0$ for all $p > 0$ such that for all $(p,\delta) \geqslant 0$,

$$\theta(p,\delta) = \phi(p)\delta.$$

[25] If this were not true, then keeping (p,q,r,y) fixed, there would exist *some* value $x'' \in (x,y)$ such that Axiom 2 would be violated with x'' in place of x.

[26] We omit the details of this deduction. One easy way is to note that by the concavity of $\theta(p,\cdot)$, θ admits right and left hand derivatives in its second argument. By interpreting this claim as an additional restriction on these derivatives, we may deduce the linearity of θ in its second argument.

We show, now, that $\phi(\cdot)$ must be an increasing function. To do this, first pick $q > 0$ and then recall the scenario of Axiom 1, with (p, x, y) chosen such that $\delta(x, y) < \varepsilon$ and $q < \mu p$. Total polarization is given by

$$P^1 \equiv pq[\phi(p)x + \phi(p)y] + pq[\phi(q)x + \phi(q)y] + 2q^2\phi(q)|y - x|.$$

When the two q's are put together at $(x + y)/2$, polarization is

$$P^2 \equiv 2pq\phi(p)\frac{x+y}{2} + 2pq\phi(2q)\frac{x+y}{2}.$$

By combining these two observations and using Axiom 1, it follows that $\phi(2q) > \phi(q)$. Since q was an arbitrary positive number, the continuous function $\phi(q)$ must be increasing.

We show, next, that $\phi(p)$ must be of the form Kp^α, for constants $(K, \alpha) \gg 0$. To prove this we proceed as follows. Consider any two-point distribution where positive population masses p and q are situated at a distance of one unit from each other. Total polarization is given by $pq[\phi(p) + \phi(q)]$. Now note that for each $(p, q, p') \gg 0$, there exists $q' > 0$ such that

$$(8) \qquad pq[\phi(p) + \phi(q)] = p'q'[\phi(p') + \phi(q')].$$

By Condition H, it follows that for all $\lambda > 0$,

$$(9) \qquad \lambda^2 pq[\phi(\lambda p) + \phi(\lambda q)] = \lambda^2 p'q'[\phi(\lambda p') + \phi(\lambda q')].$$

Combining (8) and (9), we see that

$$(10) \qquad \frac{\phi(p) + \phi(q)}{\phi(\lambda p) + \phi(\lambda q)} = \frac{\phi(p') + \phi(q')}{\phi(\lambda p') + \phi(\lambda q')}.$$

Taking limits in (10) as $q \to 0$ and noting that $q' \to 0$ as a result, we have for all $(p, p', \lambda) \gg 0$,

$$(11) \qquad \frac{\phi(p)}{\phi(\lambda p)} = \frac{\phi(p')}{\phi(\lambda p')}.$$

Put $\lambda = 1/p$ and $r = p'/p$ in (11). Note that p and r can be varied freely over $\mathbb{R}_{++}$. Making the required substitutions in (11), we arrive at the fundamental Cauchy equation

$$(12) \qquad \phi(p)\phi(r) = \phi(pr)\phi(1)$$

for all $(p, r) \gg 0$, where $\phi(\cdot)$ is continuous and strictly increasing (this last observation by our previous argument). The class of solutions to (12) (that satisfy the additional qualifications) is completely described by $\phi(p) = Kp^\alpha$ for constants $(K, \alpha) \gg 0$ (see, e.g., Aczél (1966, p. 41, Theorem 3)).

To summarize, we have proved so far that $\theta(p, \delta) = Kp^\alpha\delta$ for $(p, \delta) \geqslant 0$. It remains to deduce the bounds on α. To do so, consider the case of Axiom 3. Polarization is given by

$$(13) \qquad P(\Delta) \equiv 2d\left[(p - \Delta)^{1+\alpha}(q + 2\Delta) + (1 + 2\Delta)^{1+\alpha}(p - \Delta)\right] + 4d(p - \Delta)^{2+\alpha},$$

evaluated at $\Delta = 0$. The axiom implies that the derivative of P with respect to Δ must be nonpositive for *all* values of $(p,q) \gg 0$. Performing the necessary calculation and defining z as the ratio p/q, we see that this requirement is equivalent to the condition that $f(z,\alpha) \leqslant 0$ for all z, where $f(z,\alpha)$ is defined in (2). But this implies that $\alpha \leqslant \alpha^*$.

(*Sufficiency*.) Clearly P^* as described in the statement of the theorem satisfies Condition H. It remains only to verify Axioms 1, 2, and 3.

(*Verifying Axiom* 1): Consider p, q, x, y as in the Axiom. Initially, polarization is

$$P = (p^{1+\alpha}q + q^{1+\alpha}p)(x+y) + q^{2+\alpha}|y-x|.$$

Putting the two q's together at $(x+y)/2$,

$$\begin{aligned} P' &= (2p^{1+\alpha}q + 2^{1+\alpha}q^{1+\alpha}p)(x+y)/2 \\ &= (p^{1+\alpha}q + q^{1+\alpha}p)(x+y) + (2^{\alpha}-1)(x+y)q^{1+\alpha}p, \end{aligned}$$

so that $P' > P$ whenever $(2^{\alpha}-1)(x+y)p > 2q|y-x|$. It is easy, therefore, to find $(\varepsilon, \mu) \gg 0$ so that Axiom 1 is satisfied.

(*Verifying Axiom* 2): Take p, q, r, x, y satisfying the conditions of Axiom 2. Polarization is given by

$$(14) \qquad P = x(p^{1+\alpha}q + q^{1+\alpha}p) + (y-x)(q^{1+\alpha}r + r^{1+\alpha}q) + y(p^{1+\alpha}r + r^{1+\alpha}p).$$

Consider a small increase in x by Δ. Then, using (14), the change in polarization is $\Delta\{(p^{1+\alpha}q + q^{1+\alpha}p) - (q^{1+\alpha}r + r^{1+\alpha}q)\}$.

It is easy to verify that this expression is positive if $p > r$.

(*Verifying Axiom* 3): Consider the case depicted in Axiom 3. Recall (14). One can verify that it will be sufficient to prove that for *every* $(p,q) \gg 0$, the derivative of $P(\Delta)$ evaluated at $\Delta = 0$ is nonpositive, and that it is *strictly* negative for all but at most one ratio of p to q. For each ratio $z \equiv p/q$, this derivative is given by $f(z,\alpha)$. Now refer to the assertions following (2) to complete the verification.

This completes the proof of the theorem. *Q.E.D.*

3.2. *Maximal Elements of the Polarization Measure*

This subsection is devoted to examining the partial orderings generated by the class of measures given by Theorem 1. Moreover, we show that for given income classes $y_1, y_2, \ldots, y_n$ and total population, the distribution that assigns half the population to the lowest income class and the remainder to the highest income class is more polarized than any other distribution, under any of the measures given by Theorem 1. In what follows, we shall take as fixed the income classes and identify distributions by the vector of population masses.

THEOREM 2: *The bimodal distribution* $(\Pi/2, 0, \ldots, 0, \Pi/2)$ *is more polarized than any other distribution* $(\pi_1, \ldots, \pi_n)$ *with total population* Π, *under any measure* P^* *in the class given by* (3).

PROOF: Because of the continuity of any given measure P^* in the class given by Theorem 1, and given the compactness of the space of distributions, it suffices to show that any distribution *not* equal to $(\Pi/2, 0, \ldots, 0, \Pi/2)$ can be dominated by another in terms of polarization as measured by P^*. In the arguments to follow, we will consider measures with K normalized to unity: no loss of generality is involved. The following lemmas will be useful.

LEMMA 1: *Suppose that there exists k, $1 < k < n$, with the property that $\pi_k > 0$ and $\{\sum_{j=1}^{k-1}\pi_j - \sum_{j=k+1}^{n}\pi_j\}$ and $\{\sum_{j=1}^{k-1}\pi_j^{1+\alpha} - \sum_{j=k+1}^{n}\pi_j^{1+\alpha}\}$ are nonzero and of the same sign. Then, if the sign is positive (resp. negative), the new distribution created by moving all mass from k to $k+1$ (resp. to $k-1$) raises polarization under P^*.*

PROOF OF LEMMA 1: In what follows, we will denote $|y_i - y_j|$ by $a(ij)$.

Suppose that the condition of the lemma is met with sign positive (the negative case uses a symmetric argument). Denote by P and P' the values of polarization before and after the change. By singling out the indices k and $k+1$, we can write P as

$$(15) \qquad P \equiv \sum_{\substack{i=1 \\ i \neq k, k+1}}^{n} \sum_{\substack{j=1 \\ j \neq k, k+1}}^{n} \pi_i^{1+\alpha}\pi_j a(ij) + \pi_k^{1+\alpha}\sum_{j=1}^{n}\pi_j a(kj)$$
$$+ \pi_k \sum_{j=1}^{n} \pi_j^{1+\alpha} a(jk)$$
$$+ \pi_{k+1}^{1+\alpha} \sum_{\substack{j=1 \\ j\neq k}}^{n} \pi_j a(k+1, j)$$
$$+ \pi_{k+1} \sum_{\substack{j=1 \\ j\neq k}}^{n} \pi_j^{1+\alpha} a(k+1, j).$$

Now observe that for $j > k$, $a(kj) = a(k+1, j) + a(k+1, k)$, while for $j < k$, $a(kj) = a(k+1, j) - a(k+1, k)$. Consequently, we may rewrite (15) as

$$(16) \qquad P \equiv \sum_{\substack{i=1 \\ i \neq k, k+1}}^{n} \sum_{\substack{j=1 \\ j \neq k, k+1}}^{n} \pi_i^{1+\alpha}\pi_j a(ij) + \left(\pi_k^{1+\alpha} + \pi_{k+1}^{1+\alpha}\right)\sum_{\substack{j=1 \\ j\neq k}}^{n}\pi_j a(k+1, j)$$
$$+ (\pi_k + \pi_{k+1}) \sum_{\substack{j=1 \\ j\neq k}}^{n} \pi_j^{1+\alpha} a(k+1, j)$$
$$+ \pi_k^{1+\alpha} a(k+1, k)\left(\sum_{j=k+1}^{n} \pi_j - \sum_{j=1}^{k-1}\pi_j\right)$$
$$+ \pi_k a(k+1, k)\left(\sum_{j=k+1}^{n} \pi_j^{1+\alpha} - \sum_{j=1}^{k-1}\pi_j^{1+\alpha}\right).$$

Recall now that P' is the measure achieved by shifting all mass from k to $k+1$.

Therefore

$$P' \equiv \sum_{\substack{i=1\\ i\neq k,k+1}}^{n} \sum_{\substack{j=1\\ j\neq k,k+1}}^{n} \pi_i^{1+\alpha}\pi_j a(ij) + (\pi_k+\pi_{k+1})^{1+\alpha} \sum_{\substack{j=1\\ j\neq k}}^{n} \pi_j a(k+1,j) + (\pi_k+\pi_{k+1}) \sum_{\substack{j=1\\ j\neq k}}^{n} \pi_j^{1+\alpha} a(k+1,j). \tag{17}$$

Subtracting (16) from (17), we have

$$\begin{aligned} P'-P = &\left[(\pi_k+\pi_{k+1})^{1+\alpha} - (\pi_k^{1+\alpha}+\pi_{k+1}^{1+\alpha})\right] \sum_{j=1}^{n} \pi_j a(k+1,j) \\ &+ \pi_k^{1+\alpha} a(k+1,k)\left(\sum_{j=1}^{k-1}\pi_j - \sum_{j=k+1}^{n}\pi_j\right) \\ &+ \pi_k a(k+1,k)\left(\sum_{j=1}^{k-1}\pi_j^{1+\alpha} - \sum_{j=k+1}^{n}\pi_j^{1+\alpha}\right). \end{aligned} \tag{18}$$

Note that because $\alpha>0$, $(\pi_k+\pi_{k+1})^{1+\alpha} > (\pi_k^{1+\alpha}+\pi_{k+1}^{1+\alpha})$. Combining this observation with (18) and the conditions of the lemma, we see that $P'>P$. *Q.E.D.*

LEMMA 2: *Suppose that an income distribution has at least four nonzero mass points. Then there exists k, $1<k<n$, such that the conditions of Lemma 1 are met.*

PROOF OF LEMMA 2: Let k_1 be the second nontrivial mass point counting from the left, and k_2 be the second nontrivial mass point counting from the right. Then, because there are at least four nonzero mass points, $k_1<k_2$. We claim that the conditions of Lemma 1 hold for at least one of these two indices.

Clearly, for either $k=k_1$ or $k=k_2$, $\sum_{i=1}^{k-1}\pi_i - \sum_{i=k+1}^{n}\pi_i \neq 0$.

Case 1: Suppose that $k=k_1$. Consider the subcase where

$$\sum_{i=1}^{k_1-1}\pi_i - \sum_{i=k_1+1}^{n}\pi_i > 0. \tag{19}$$

Because $\pi_i>0$ for at most one index i for $1\leqslant i<k_1$, we have $(\sum_{i=1}^{k_1-1}\pi_i)^{1+\alpha} = \sum_{i=1}^{k_1-1}\pi_i^{1+\alpha}$. Consequently, by (19),

$$\sum_{i=1}^{k_1-1}\pi_i^{1+\alpha} = \left(\sum_{i=1}^{k_1-1}\pi_i\right)^{1+\alpha} > \left(\sum_{i=k_1+1}^{n}\pi_i\right)^{1+\alpha} \geqslant \left(\sum_{i=k_1+1}^{n}\pi_i^{1+\alpha}\right).$$

This means that the conditions of Lemma 1 are met for $k=k_1$.

Next, consider the subcase where (19) fails. Then it certainly must be the case that (19) holds when k_1 is replaced by k_2. Now carry out exactly the same argument for k_2 as we did for k_1.

Case 2: Suppose that $k = k_2$. Then use exactly the same argument as in Case 1, with the roles of k_1 and k_2 reversed. *Q.E.D.*

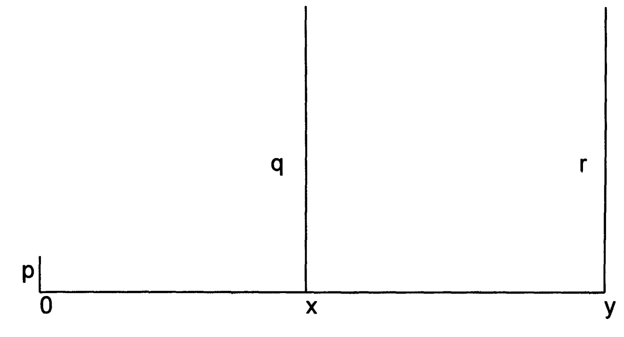
We may now complete the proof of the theorem. If the number of mass points is at least 4, use Lemmas 1 and 2. If the number of mass points is 3, use Axiom 3. Finally, if the number of mass points is 2, then use the computation in Section 3.4.1 below. *Q.E.D.*

Lemmas 1 and 2 are of independent interest for polarization analysis. Lemma 1 establishes a criterion of polarization dominance which generates a partial ordering. Roughly speaking, this criterion says that polarization is raised if we shift a population mass to a neighboring class whenever the distribution of the population on the side to which this neighbor belongs satisfies the following two conditions: total population is less, *and* more "dispersed" than on the other side. Lemma 2 demonstrates that for every distribution there always is an income class such that the previous conditions are satisfied. For additional discussion of this point, see Section 3.4.3, item (3).

3.3. *An Additional Restriction*

The conclusion of Theorem 1 may be sharpened further by the inclusion of an additional assumption, which we find plausible.[27] This assumption concerns the case studied in Example 3 of Section 2, which, it will be recalled, was designed to illustrate the insignificance of small groups in the concept of polarization.

AXIOM 4:



Data: $(p, q, r, \Delta) \gg 0$ *and* $q > r$.
Statement: There is $\mu > 0$ *such that if* $p \leqslant \mu r$ *and* $q - r < \mu$, *then a transfer of population mass from* p *to* r *will not decrease polarization.*

[27] We decided against using this assumption directly in Theorem 1. Theorem 1 does achieve a dramatic simplification, and it may be useful to explicitly see which assumptions are driving the result obtained there.

Note that Axiom 4 only applies if p is *small* relative to the masses q and r, and if $q \simeq r$. We note, too, that mass is transferred from p to only the *smaller* of the two remaining masses, so the Axiom is quite weak. We refer the reader to Section 2, Example 3, for a discussion of this point.

THEOREM 3: *Under Condition* (H) *and Axioms* 1–4, *every polarization measure must be of the form described in Theorem* 1, *with the additional restriction that* $\alpha \geqslant 1$.

We note that Theorem 1 is deficient in one respect. Because of the weak nature of the axioms employed, Theorem 1 fails to rule out the possibility that α may be arbitrarily close to zero. As noted above, smaller values of α indicate an increasing lack of "polarization sensitivity," and a greater degree of concordance with inequality measurement. The limiting case of $\alpha = 0$, as we have already observed, *is* one of inequality measurement, with our measure converging to the Gini coefficient (defined on log incomes). *En route*, to the limit, Axioms 1 to 3 are satisfied, but in progressively weaker terms (with failure occurring, of course, only at the limit). The additional restriction achieved in Theorem 3 rules out these cases altogether, bounding the permissible values of α below by unity, and imparting a minimal degree of "polarization sensitivity" to the entire class of measures. See Section 3.4.2 for additional illustration.

PROOF OF THEOREM 3: For given (p, q, r, Δ) as in Axiom 4, polarization is given by

$$P \equiv \Delta\{[p^{1+\alpha}q + q^{1+\alpha}p] + [q^{1+\alpha}r + r^{1+\alpha}q] + 2[p^{1+\alpha}r + r^{1+\alpha}p]\}. \tag{20}$$

Differentiating (20) with respect to p under the understanding that $dq = 0$ and $dr = -dp$, we see that

$$\frac{1}{\Delta}\frac{dP}{dp} = (1+\alpha)p^{\alpha}q + q^{1+\alpha} - q^{1+\alpha} - (1+\alpha)r^{\alpha}q + 2(1+\alpha)p^{\alpha}r - 2p^{1+\alpha} - 2(1+\alpha)r^{\alpha}p + 2r^{1+\alpha}. \tag{21}$$

Using Axiom 4, it follows that evaluated at $p = 0$ and $q = r$, $(1/\Delta)(dP/dp) \leq 0$. Passing to the limit as $p \to 0$ and $r \to q$ in (21), we see that

$$\frac{1}{\Delta}\frac{dP}{dp}(p = 0, r = q) = -(1+\alpha)q^{1+\alpha} + 2q^{1+\alpha}.$$

So the required inequality holds if and only if $\alpha \geqslant 1$. This completes the proof of the proposition. *Q.E.D.*

3.4. *Properties of the Polarization Measure in Particular Cases*

In this subsection, we apply the polarization measure that we have obtained to some simple and (intuitively) not-so-simple cases, to examine how the measure behaves. These cases include an analysis of the various "desirable" features of a polarization measure, discussed in Section 2. In all the examples that follow, we normalize population to a unit mass and set $K=1$ in (3).

3.4.1. *Two-point Distributions*

Our first example studies aspects of Example 2 (Section 2). Consider the two-point income distribution with weights π and $1-\pi$ concentrated at incomes x and y. Polarization is given by

$$P \equiv \left[\pi^{1+\alpha}(1-\pi)+(1-\pi)^{1+\alpha}\pi\right](y-x).$$

It is immediate that pulling apart the two groups will increase polarization. Furthermore, by using the expression above, it can be verified that for given (x, y), polarization is maximized at $\pi = 0.5$. This last feature uses the condition (proved in Theorem 1) that $\alpha \leqslant \alpha^*$.[28]

3.4.2. *Three-point Distributions*

We now take a look at different three-point distributions. Without loss of generality, denote by $(0, x, y)$ the incomes $(0<x<y)$ and by (p,q,r) the corresponding population weights. The general form of the measure now reads

$$P \equiv \left[p^{1+\alpha}q+q^{1+\alpha}p\right]x+\left[p^{1+\alpha}r+r^{1+\alpha}p\right]y + \left[q^{1+\alpha}r+r^{1+\alpha}q\right](y-x).$$

One special case is where $p=r$. This yields a symmetric distribution. As population weight on this distribution is shifted from the center to the sides, we expect polarization to rise. This is exactly the sentiment of Axiom 3, which our polarization measure is known to satisfy.

Another special case is that illustrated by Example 4 in Section 2. Here $q=r$. The idea is to examine what happens when the two q-masses are fused at the midpoint of their incomes. We argued in Section 2 that polarization should rise if p is "large" relative to q, but fall if the relative sizes were reversed. Denoting the polarization values before and after the change by P and P' respectively, we can reproduce exactly the same argument in the proof of Theorem 1 to see that $P>P'$ if and only if

$$\frac{p}{q} > \frac{2(y-x)}{(2^\alpha-1)(x+y)}.$$

[28] Indeed, this feature holds as long as $\alpha < 2$.

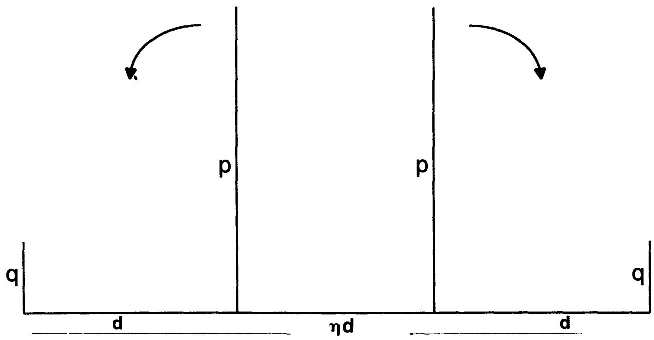


FIGURE 6.—Weight-shifting from center to sides.

The behavior of the measure therefore reinforces our intuition. In particular, if p is large relative to q, polarization rises, while exactly the opposite is true if p is small.[29] This is precisely the kind of global sensitivity that we referred to in Section 2.

Apropos the discussion following Theorems 1 and 3, note how clearly α acts as a measure of "polarization sensitivity." In particular, the smaller the value of α the higher must be the ratio of p to q before polarization can be said to have increased under the measure.

3.4.3. *Distributions with Four or More Mass Points*

(1) We start by reproducing Example 5 of Section 2. This example, it will be recalled, was designed (along with Example 4), to show the global nature of the measure. Consider Figure 6.

Imagine that weight is being transferred equally from the two mass points nearer the center towards the sides: thus, we have initially the configuration $(0, \frac{1}{2}, \frac{1}{2}, 0)$ and finally, the more polarized distribution $(\frac{1}{2}, 0, 0, \frac{1}{2})$. Our purpose is to investigate intermediate behavior.

Use Figure 6 to examine the formula to be given below. Initially, polarization is given by

$$\begin{aligned} P &\equiv 2\left[p^{1+\alpha}q + q^{1+\alpha}p\right]d + 2p^{2+\alpha}\eta d \\ &\quad + 2\left(p^{1+\alpha}q + q^{1+\alpha}p\right)(1+\eta)d + 2q^{2+\alpha}(2+\eta)d \\ &= 2p^{2+\alpha}\eta d + 2\left[p^{1+\alpha}q + q^{1+\alpha}p + q^{2+\alpha}\right](2+\eta)d \\ &= 2p^{2+\alpha}\eta d + 2\left[p^{1+\alpha}q\tfrac{1}{2}q^{1+\alpha}\right](2+\eta)d, \end{aligned}$$

using the fact that $p+q=1/2$. With $dp/dq = -1$, the derivative of this

[29] Of course, the precise threshold values depend on the values of x and y. But that is exactly as it should be.

expression with respect to q is

$$-2(2+\alpha)p^{1+\alpha}\eta d + 2\left[p^{1+\alpha} - (1+\alpha)p^{\alpha}q + \tfrac{1}{2}(1+\alpha)q^{\alpha}\right](2+\eta)d.$$

We wish to examine the sign of this as q ranges over $[0, \frac{1}{2}]$. First note that *at* $q = \frac{1}{2}$ (and $p = 0$), the derivative is unambiguously positive. This reflects the intuitively appealing feature that when the population is already largely bunched at the two extreme points, further bunching will serve to accentuate polarization.

At $q = 0$ (and $p = \frac{1}{2}$), the derivative has the value $d[2-(1+\alpha)\eta]2^{-\alpha}$. This could be of either sign, depending on the values of η and α. For instance, if η is "small," then the initial polarization is relatively small and all moves to the extremes raise polarization. However, if $\eta > 2/(1+\alpha)$, initial polarization is relatively large. In this case, the above analysis shows that as population is moved away from the central masses to the extremes, polarization first goes down and then goes up,[30] ultimately attaining a higher level than the initial one. Under the conditions yielding Theorem 3, this requirement is very weak, saying only that $\eta > 1$. The reader may wish to compare this with our intuitive discussion of Example 5.

(2) We turn next to the question of Dalton transfers that are equalizing, to check whether polarization can rise in this case. Numerical examples are easy to find. Consider, for instance, the case of Example 1 in Section 2. To be concrete, think of incomes in this example as absolute values (and not logarithms), to retain an explicit notion of Lorenz domination. It is possible to verify by numerical methods that polarization will rise as long as α exceeds approximately 0.3. This condition is more than amply satisfied under Theorem 3.

Indeed, it is easy to provide a general formula that checks whether the "bunching" of incomes initially uniformly distributed over a finite number of income classes, into a smaller number of income points, increases polarization. Apart from being of possible interest for its own sake, this exercise shows that our polarization measure is readily amenable to analytical calculation.

Consider n (logarithmic) income points spaced uniformly with pairwise distance d, and suppose that the population is uniformly distributed on these points. Defining $P(n,d)$ to be the value of polarization thus obtained, we see that

$$P(n,d) = dn^{-(2+\alpha)}\sum_{k=1}^{n}\left\{\frac{k(k-1)}{2} + \frac{(n-k)(n-k+1)}{2}\right\},$$

where in deriving the above expression, we recall that the sum $1+2+\cdots+m = m(m+1)/2$ for any positive integer m. Using, in addition, the formula $1+2^2+\cdots+m^2 = m(m+1)(2m+1)/6$ for any positive integer m, the above simplifies to yield

$$P(n,d) = dn^{-(2+\alpha)}n(n^2-1)/3. \tag{22}$$

[30] It can be checked that a higher degree of nonlinearity is not possible in this case.

Consider positive integers M and K. Suppose, now, that we start with MK income points spaced at uniform distance, normalized to unity. We "collapse" this distribution by taking adjacent sets of K points each and concentrating these at the average income levels of these sets. We now have a *new* uniform distribution over M points spaced K units apart. Using (22), we may conclude that polarization has increased if

$$(KM)^{-(2+\alpha)}(K^2m^2-1)KM/3 > KM^{-(2+\alpha)}(M^2-1)M/3,$$

which on simplification yields the condition

$$K^{2+\alpha}(M^2-1) > K^2M^2-1. \tag{23}$$

It is immediate that for (23) to be satisfied for any given K, $\alpha > 0$. This is intuitive. Recall that for $\alpha = 0$, the polarization measure essentially reduces to a Gini measure of inequality (in the perceptual variable log-income), so that all bunching of incomes will bring the measure down by the Dalton principle of transfers. Note, too, that if bunching results in a *single* group ($M = 1$), polarization *cannot* go up, and this too is reflected by the condition (23).

Next, observe that for any K and $M \neq 1$, condition (23) is most stringent when $M = 2$. It follows, therefore, that for *given* K, if polarization is to rise under all such circumstances which involve the bunching of K groups, a necessary and sufficient condition for this is

$$3K^{2+\alpha} > 4K^2-1. \tag{24}$$

Note that if the polarization measure satisfies Axiom 4, then $\alpha \geqslant 1$ and (24) is automatically satisfied for all values of $K \geqslant 2$.

(3) We end this section of the paper by observing that our polarization measure combines a preference for numerical equality at different income points with a preference for creating bunched populations. Thus, provided that groups can somehow be defined, "intra-group" *equality* enhances polarization. One way of examining this is to generalize the conditions of Axiom 2. Consider a particular income class k and examine the distributions on either side of this class. Suppose that the following two conditions hold: (i) total population to the left of this class exceeds total population to the right of this class, and (ii) the income distribution on the space of incomes to the left of this class is Lorenz-dominated by the income distribution to the right of this class. Informally, higher incomes are more scarcely populated *and* more bunched together. Then a shift of population weight from income class k to class $k+1$ must raise polarization as defined by (3). The proof of this observation is implied by Lemma 1 and the fact that $\alpha > 0$, and is omitted.

4. CROSS-IDENTIFICATION: AN EXTENDED CLASS OF MEASURES

It is appropriate at this stage to reconsider a serious assumption that drives the derivation of the polarization measure (3) in the previous section. It will be recalled that our model posited a particularly sharp form of the identification

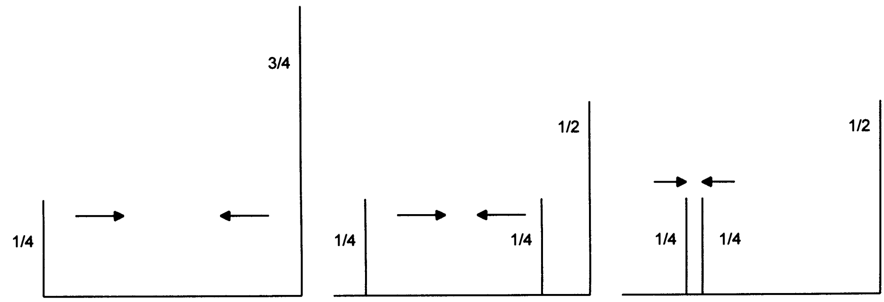


FIGURE 7A FIGURE 7B FIGURE 7C

FIGURE 7.—Diagrams to illustrate the discontinuity problem.

function; namely, that an individual identifies with another in exactly the same income class *and with no one else*. This extreme postulate is echoed in the measure in the form of an unsatisfactory discontinuity.[31]

Figure 7 reconsiders Example 4 in more detail. Begin with two population masses, say $1/4$ and $3/4$ at two income points, as shown in Figure 7a. Now suppose that the first mass and a subset of the second mass (of measure $1/4$) begin to move towards each other, leaving the remaining half of the population where it is. Figures 7b and 7c illustrate this process. Initially, we would expect polarization to fall and finally, when the two population masses are fused (just as in Example 4) we would have polarization rise. All this is satisfied by our derived measure. The point is that when the masses are already close (say, as in Figure 7c) but not exactly superimposed, we would *also* expect polarization to be increasing in the described movement (supporting this by an argument that the identification between the two $1/4$-masses is growing). But our measure does not incorporate cross-identification between groups, and so *continues* to fall, only to jump up "in the limit" as the two masses exactly fuse.

The problem is easily resolved by permitting identification across income groups that are "sufficiently" close.[32] But this resolution is not costless. Admitting cross-identification opens up a set of new issues. In particular, it is impossible to specify (or deduce) a priori the domain over which a sense of identification prevails.

Nevertheless, such an extended class of measures (indexed by the choice of the identification zone) can be defined for the purpose of empirical work, and can be shown to satisfy the axioms in the previous section, as well as the properties discussed in Section 3.4. *The choice of this zone must be left to the empirical researcher, for it depends on the degree of grouping or averaging in each of the "income classes" of the available data.*

[31] We are grateful to a referee for making this point, and for suggesting a discussion of it in the paper.

[32] Indeed, an earlier version of our paper (Esteban and Ray (1991)) does just this.

To illustrate an extended class, let $D > 0$ be such that if an income y' is within D of an income y, then there is some identification between the two incomes. Formally, let $w(d)$ be a positive weighting scheme on $[0, D]$ such that $w(\cdot)$ is a decreasing function with $w(D) = 0$. For a given distribution $(\pi, \mathbf{y})$, define the identification felt by any member of income class i as

$$I_i \equiv \sum_{j:|y_i - y_j| \leqslant D} \pi_j w(y_j).$$

The new measure induced by the weighting function w is then some scalar multiple of

$$P_w(\pi, \mathbf{y}) \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i \pi_j I_i^{\alpha} \max\{|y_i - y_j| - D, 0\}. \tag{25}$$

A complete characterization of such measures is beyond the scope of this paper. But note that these measures do resolve the discontinuity described above.

5. OTHER EXTENSIONS AND CONCLUDING REMARKS

5.1. *Income and Population Normalization*

As stated earlier, we use the logarithm of incomes as the perceptual variable in our model. While this may appear nonstandard in the theory of inequality measurement, where absolute differences in income are noted (and subsequently normalized by mean income), we find our approach suitable for two reasons. The first is conceptual. We find it natural that individuals react to percentage differences in income rather than absolute differences, and we wish to keep these bilateral differences symmetric.[33] The second reason is one of convenience. A number of our axioms and examples involve the movement of *population* from one income class to another, which then alter aggregate income. With absolute income, the measure would have to be renormalized at each step before the axioms dictate how we want the measure to behave. This is unnecessary in the present formulation. A corollary is that the measure comes prenormalized in incomes.

The same is not true of population. The natural normalization to choose, of course, is to replace the population weights π_i, $i = 1, \ldots, n$, by the population *frequencies*. This is tantamount to choosing $K = [\sum_{i=1}^{n} \pi_i]^{-(2+\alpha)}$ in (3). But it is important to be clear as to what such a normalization implies in the context of our model.

In our model, polarization among small populations is smaller than among larger groups, ceteris paribus. But a careful reading of the model will make it

[33] Thus, for example, the percentage difference in incomes, normalized by the income of the individual concerned, is not symmetric. While more complicated formulations are symmetric (such as the absolute difference of income divided by the sum of the two incomes), we felt that the logarithm would do just as well.

clear that this refers to the relative importance of the small population in a context where it is embedded within a larger population. Thus polarization in a small region of a large country may mean little relative to polarization occurring *in that same country* at a national level. But if this small region *is* an entirely separate country, then the comparison of the *two* countries requires a scaling of the population. Population homogeneity of degree zero may then be legitimately invoked without violating the basic tenets of the model.

5.2. *Asymmetric Polarization Measures*

Earlier, we noted that a perfectly symmetric model of polarization may not be entirely appropriate, especially in a context where the variable of interest is income or wealth. While the poor feel alienation from the rich, it may be argued that the corresponding sense of distance felt by the rich from the poor should not enter in a symmetric way into the polarization measure. We briefly discuss an extreme version of this observation: the alienation function for individual y registers positive values only for income values y' that are *greater* than that of individual y.

Fortunately, we can exploit the formal structure above with no change of notation if we don't mind changing the concept a bit. Interpret $\delta(y, y')$ now as $\max\{y'-y, 0\}$ (instead of $|y'-y|$). The general class of measures studied in Section 3 may now be also put forward for this case. As before, a set of axioms may be imposed to narrow down this class of measures. This set will be somewhat different from the axioms of that section, insofar as the latter has an implicit bias towards *symmetry* (for example, Axiom 2). This class of polarization measures will certainly *not* yield symmetric distributions as its maximal elements.

This last observation can be verified even without imposing any axiom at all. Consider the two point distribution in Section 3.3.1, with the mass points p and $(1-p)$ situated at a distance d. Any asymmetric measure in the general class yields polarization equal to

$$P \equiv p(1-p)\theta(p, d).$$

Suppose we were to maximize P with respect to a choice of $p \in [0,1]$. We observe that if θ is increasing and differentiable in p, with $\partial\theta/\partial p > 0$, then *the maximum of the polarization measure must be attained at values of p that lie strictly between $\frac{1}{2}$ and* 1.[34]

[34] Here is a quick proof. Note that any value of p that maximizes P must be strictly between 0 and 1. So the first-order necessary condition:

$$\theta(p^*, d)(1-2p^*) + p^*(1-p^*)\frac{\partial\theta(p^*, d)}{\partial p} = 0$$

must hold at any values of p^* that maximize P. But now observe that because $\partial\theta/\partial p > 0$, this condition *cannot* hold for any $p^* \leqslant \frac{1}{2}$.

This shows that asymmetric polarization models will not yield symmetric distributions as maximal elements.

5.3. *Weighted Identification and Alienation Functions*

We comment on two possible extensions. First, the sense of identification of an individual may depend not only on the *number* of other individuals with similar attributes, but also on the attributes themselves. In other words, the attributes may confer "power" on the group to effectively manifest its sense of identification. In the case where the space of attributes is income, this consideration may be of particular relevance. An equal number of people in two different income categories may possess an unequal degree of "effective identification." The resources of the richer group may contribute to the ability of that group to form lobbies, organize protests, push particular policies, and in general to manifest itself as a unified entity. This extension is easily incorporated into the model that we have set up here. The identification function of an individual with income y_i must now be extended to include not only the number π_i with that income but also y_i itself: $I = I(\pi_i, y_i)$. A typical member of the extended family of polarization measures would then be

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

which reduces to the current family in case $\beta = 0$. We have not axiomatized this extended family in the paper, though we conjecture that this would not pose serious difficulties.

Second, the alienation felt by an individual y towards another person y' may also depend upon the *number* of individuals with the latter income. In a certain sense, the notion of identification already incorporates this feature: the effective antagonism of the latter group towards y will have been weighted by *its* sense of identification. But it is certainly possible to create additional weighting. The point is that we have little to say, *a priori*, regarding the effect of such a weighting on effective antagonism. Does it increase or decrease the effective voicing of protest? In our mind, the answer is unclear, though such an extension merits inquiry.

5.4. *Methodological Remark*

Perhaps the most important drawback of our work is that it combines a behavioral story of individual attitudes with axiomatic restrictions placed not on the story itself but separately on distributions of characteristics. In this sense, there is perhaps a "deeper" axiomatization of the concept of polarization, one that is parallel to the theory of inequality measurement. Our efforts so far have not been successful in uncovering this more fundamental theory. While we believe that some progress has been achieved by motivating, defining, and characterizing a new concept and its measure, it should certainly be possible to do better.

Instituto de Análisis Económico, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

and

Department of Economics, Boston University, 270 Bay State Rd., Boston, MA 02215, U.S.A.

Manuscript received March, 1991; final revision received October, 1993.

REFERENCES

ACZÉL, J. (1966): *Lectures on Functional Equations and their Applications*. New York and London: Academic Press.

ADELMAN, I., AND C. T. MORRIS (1973): *Economic Growth and Social Equity in Developing Countries*. Stanford: Stanford University Press.

AHLUWALIA, M. (1976): "Inequality, Poverty and Development," *Journal of Development Economics*, 3, 307–432.

ALESINA, A., AND D. RODRIK (1991): "Distributive Politics and Economic Growth," NBER Working Paper No. 3668.

ALESINA, A., S. ÖZLER, N. ROUBINI, AND P. SWAGEL (1992): "Political Instability and Economic Growth," mimeo.

AMIEL, Y., AND F. A. COWELL (1992): "Measurement of Income Inequality: Experimental Test by Questionnaire," *Journal of Public Economics*, 47, 3–26.

ATKINSON, A. B. (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263.

BARRO, R. J. (1991): "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106, 407–443.

BARRO, R. J., AND X. SALA-I-MARTIN (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

BENHABIB, J., AND A. RUSTICHINI (1991): "Social Conflict, Growth and Income Distribution," mimeo.

BLACKORBY, C., AND D. DONALDSON (1978): "Measures of Relative Equality and their Meaning in Terms of Social Welfare," *Journal of Economic Theory*, 18, 59–80.

CHAYANOV, A. V. (1923): *The Theory of Peasant Economy* (1966 edition edited by D. Thorner, B. Kerblay, and R. Smith). Homewood, IL: R.D. Irwin.

COSER, L. A. (1956): *The Functions of Social Conflict*. Glencoe IL: The Free Press.

DAHRENDORF, R. (1959): *Class and Class Conflict in Industrial Society*. Stanford, CA: Stanford University Press.

DASGUPTA, P., A. SEN, AND D. STARRETT (1973): "Notes on the Measurement of Inequality," *Journal of Economic Theory*, 6, 180–187.

DEUTSCH, M. (1971): "Conflict and its Resolution," in *Conflict Resolution: Contributions of the Behavioral Sciences*, ed. by C. G. Smith. Notre Dame: University of Notre Dame Press.

DUNCAN, G., T. M. SMEEDING, AND W. RODGERS (1991): "W(h)ither the Middle Class? A Dynamic View," Levy Institute Conference on Income Inequality, Bard College.

DUTTA, B., AND J. ESTEBAN (1992): "Social Welfare and Equality," *Social Choice and Welfare*, 9, 267–276.

EBERT, U. (1987): "Size and Distribution of Incomes as Determinants of Social Welfare," *Journal of Economic Theory*, 41, 23–33.

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

FIELDS, G. S. (1980): *Poverty, Inequality and Development*. Cambridge: Cambridge University Press.

FIELDS, G. S., AND J. C. H. FEI (1978): "On Inequality Comparisons," *Econometrica*, 46, 303–316.

FOSTER, J. (1985): "Inequality Measurement," in *Fair Allocation*, ed. by H. Peyton Young. Providence, R.I.: American Mathematical Society.

GURR, T. R. (1970): *Why Men Rebel*. Princeton: Princeton University Press.

——— (ED.) (1980): *Handbook of Political Conflict: Theory and Research*. New York: The Free Press.

HARSANYI, J. C. (1953): "Cardinality Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, 61, 434–435.

HORRIGAN, M. W., AND S. E. HAUGEN (1988): "The Declining Middle Class Thesis: A Sensitivity Analysis," *Monthly Labour Review*, 111, 3–13.

KERNER COMMISSION (UNITED STATES) (1968): *Report of the National Advisory Commission on Civil Disorders*. Washington, D.C.: Gov't. Printing Office.

KOLM, S-CH. (1969): "The Optimal Production of Social Justice," in *Public Economics*, ed. by J. Margolis and H. Guitton. London: Macmillan; New York: St. Martin's Press.

KOSTERS, M. H., AND M. N. ROSS (1988): "A Shrinking Middle Class?" *The Public Interest*, 90, 3–27.

KRAVIS, I. (1960): "International Differences in the Distribution of Income," *Review of Economics and Statistics*, 42, 408–416.

KUZNETS, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, 45, 1–28.

LANDSBERGER, M., AND I. MEILIJSON (1987): "A Tale of Two Tails: Playing Bold and Some New Definitions of Risk," mimeo, Haifa University.

LENIN, V. I. (1899): *The Development of Capitalism in Russia*, 1956 edition. Moscow: Foreign Languages Publishing House.

MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

MIDLARSKI, M. I. (1988): "Rulers and the Ruled: Patterned Inequality and the Onset of Mass Political Violence," *American Political Science Review*, 82, 491–509.

MULLER, E. N., M. A. SELIGSON, AND H. FU (1989): "Land Inequality and Political Violence," *American Political Science Review*, 83, 577–586.

NAGEL, J. (1974): "Inequality and Discontent: a Non-linear Hypothesis," *World Politics*, 26, 453–472.

PERSSON, T., AND G. TABELLINI (1991): "Is Inequality Harmful for Growth?: Theory and Evidence," mimeo, Institute for International Economic Studies, University of Sweden.

QUAH, D. (1992): "Empirical Cross-section Dynamics in Economic Growth," mimeo, London School of Economics.

SEN, A. K. (1973): *On Economic Inequality*. Oxford: Clarendon Press.

SHANIN, T. (1966): "The Peasantry as a Political Factor," *Sociological Review*, 14, 5–27.

SHESHINSKI, E. (1972): "Relation between a Social Welfare Function and the Gini Index of Income Inequality," *Journal of Economic Theory*, 4, 98–100.

SIMMEL, G. (1955): *Conflict*. Glencoe, IL: The Free Press.

TILLY, C. (1978): *From Mobilization to Revolution*. New York: Random House.

WEISSKOFF, R. (1970): "Income Distribution and Economic Growth in Puerto Rico, Argentina and Mexico," *Review of Income and Wealth*, 16, 303–332.