

**Working Paper No. 218**

**Extensions of a Measure of Polarization  
with an Application to the Income Distribution  
of Five OECD Countries**

**Joan Esteban  
Carlos Gradín  
Debraj Ray**

**November 1999**

# Extensions of a Measure of Polarization, with an application to the income distribution of five OECD countries<sup>S\*</sup>

Joan Esteban, Institut d'Anàlisi Econòmica, CSIC, Barcelona,  
Carlos Gradín, Universidad de Vigo

and

Debraj Ray, Boston University  
and Institut d'Anàlisi Econòmica, CSIC, Barcelona

**SUMMARY:** We introduce an extension of the Esteban and Ray [Econometrica, 1994] measure of polarization that can be applied to density functions. As a by-product we also derive the Wolfson [AER, 1994] measure as a special case. This derivation has the virtue of casting both measures in the context of a (statistically) unified framework. We study the polarization of the distribution of household income for five OECD countries (LIS database): US, UK, Canada, Germany and Sweden.

JEL classification: D31, D63, and I32.

Key words: polarization, income distribution, inequality.

November 1999

---

<sup>S</sup> Esteban and Gradín acknowledge the financial support of Fundació Barrié de la Maza and research grant DGICYT PB96--0897. Ray acknowledges financial support under National Science Foundation Grant SBR-9709254. We are grateful to R. Roy for his comments.

\* Corresponding author: Joan Esteban, Institut d'Anàlisi Econòmica, Campus de la UAB, 08193 Bellaterra-Barcelona, Spain. Tel. 349 3580 6612, Fax 349 3580 1452, mail:esteban@uab.es

# 1. Introduction

The concept and measurement of the polarization of a distribution has recently attracted some attention from economists. In independent work, Esteban and Ray [1994] (henceforth, ER) and Wolfson [1994] (henceforth, W) have developed measures of polarization which, while sharing similar motivations, turn out to have significant differences.<sup>1</sup> In this paper we introduce an extension of the ER measure of polarization which can be applied to data for which a density function for the income distribution has already been computed (either parametrically or non-parametrically). As a by-product of our investigation we also derive the W measure as a special case. This derivation has the virtue of casting both the ER and W measures in the context of a (statistically) unified framework.

The ER measure employs as its starting point notions of intra-group *identification* and inter-group *alienation*. The distribution is assumed to be pre-grouped into income classes in a way that individuals within each of the classes “identify” with their class members and feel “alienated” with members of other classes. Polarization is viewed as some increasing function of these two features. However, it is clear that the available data will, in general, be statistically grouped in a way that has little to do with such notions. Or (in the case of an estimated density) it may not be pre-grouped at all. To be sure, ER discusses these possibilities and suggests an extension of the measure. The purpose of this paper is to examine such extensions in more detail.

In the next section we develop our first extension of the E-R measure, derived from a behavioural approach. In section 3 we deal with our second extension, drawing from a "statistical" approach. We also examine the properties of this second type of extension for the special case of bi-polarization and derive Wolfson's measure as a particular case. In section 4,

---

<sup>1</sup> We are grateful to James Foster for bringing to our attention, after ER was published, the existence of Love and Wolfson [1976], in which similar concerns with the Pigou-Dalton principle of transfers were first raised

we study the polarization of the distribution of household income for five OECD countries (LIS database).

## 2. Conceptual Issues

Suppose that  $F$  is the (estimated or true) distribution for a particular income distribution, or simply the distribution of income data in pre-grouped form. As in ER, we will suppose that each individual is subject to two forces: he feels *identification* with those he considers to be members of his “own group”, and alienation from those he considers to be members of “other groups”. Thus, keeping matters deliberately abstract for the moment, suppose that an individual with income  $x$  feels group identification  $I(x, F)$  under the distribution  $F$ , and alienation  $r(x, y)$  with respect to some individual with income  $y$ . As in ER, we take the *effective antagonism* that individual  $x$  feels towards  $y$  as some increasing function  $T(I(x, F), r(x, y))$ . Effective antagonism increases with alienation, but this alienation is taken to be fueled by some sense of identification as well.

Polarization is the “sum” of all effective antagonisms:

$$P(F) = \int \int T(I(x, F), r(x, y)) dF(x) dF(y) \quad (1)$$

Described in this way, the measure is not very operational. Much is obviously left to the choice of identification and alienation functions, as well as the function  $T$  that combines them to precipitate a measure of polarization. The approach taken in ER is to combine this relatively broad starting point with a set of intuitive axioms that might compare polarization across distributions. These axioms yield restrictions on the functional forms that can be admitted into the general framework sketched in Esteban and Ray [1994]. The ER characterization is restricted, however, to distributions that are pre-arranged in groups so that for an individual with income  $x$  belonging to some group  $i$ ,  $I(x, F)$  simply equals  $p_i$ , the measure of individuals in that group (under the distribution  $F$ ).

But there is, of course, no reason to believe that the grouping of income distribution data will conveniently conform to the psychological demands of group identification. Consider, therefore, the following extension along the lines suggested by ER: 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  $F$ , define the identification felt by any individual with income  $x$  as

$$I(x, F) = \int_{\{y: |y-x| \leq D\}} w(|y-x|) dF(y) \quad (2)$$

Take alienation to be simply the linear distance between  $x$  and  $y$ , with the identification zone netted out:

$$r(x, y) = \max \{ |x-y| - D, 0 \}. \quad (3)$$

Then, a natural generalization of the ER measure is

$$P(\alpha, F) = \int_{\{y: |y-x| \leq D\}} \int_{\{x: |x-z| \leq D, 0\}} \alpha w(|y-x|) dF(y) dF(x) dF(z), \quad (4)$$

where  $\alpha$  is some positive constant capturing the importance of group identification.

Observe that if group identification is unimportant, then we can take  $D=0$  and  $\alpha=0$  as well, in which case the measure in (3) reduces to a measure of *inequality*, the Gini coefficient. Thus it is the presence of identification that makes a measure of polarization fundamentally different from one of inequality (for more on this, see ER).

We need to be aware of some features of (3). First, if the distribution is clustered entirely within a support of  $D$ , then polarization is zero. This is not a problem at all provided we do not insist that any difference in incomes, however slight, should result in some polarization. Second, and more problematic, is the fact that the measure is still not operational: how does one choose the weighting function, or indeed the domain of identification  $D$ ? It is clear that, beyond a point, it is difficult to nail these objects down: one can

only hope for robustness of the implied polarization ordering under various choices of the functional forms.

In the next section, we attempt a slightly different approach. To be sure, this cannot get around the problems raised here, but the measure that we obtain has the virtues that (a) it does not disagree with the overall conceptual scheme presented here, and (b) it is tractable and easy to implement.

### 3. A “Statistical” Approach

The view we develop in this extension of the ER measure of polarization can be summarised as follows. The ER polarization measure for discrete groups should be used only after the population has been regrouped in a way that captures the group identification structure of society. This regrouping or clustering will lose some of the initial information that concerns the dispersion of the population around the clusters that we are treating as single groups. Put another way, the artificial sharpness of identification induced by the ER measure needs to be corrected for. From a “statistical” perspective, the clustered data will contain some degree of error relative to the original information. The extended measure we propose is the polarization measure (as in Esteban and Ray [1994, equation 3]) on the clustered distribution minus a measure of the error made by clustering.

#### 3.1. The General Case

Suppose that income distribution data are provided by means of a density  $f$ . Let the support of the distribution be contained in some bounded interval  $[a,b]$ . Incomes are normalized to the expected income,  $\mu=1$ . An  $n$ -spike representation of  $f$  is a collection  $\rho$  of numbers

$$(y_0, y_1, \dots, y_n; \pi_1, \dots, \pi_n; \mu_1, \dots, \mu_n) \text{ such that}$$

$$a=y_0 < \dots < y_n=b, \text{ and}$$

$$\begin{aligned}\pi_i &= \int_{y_{i-1}}^{y_i} f(y)dy, \\ &\text{, for all } i=1, \dots, n. \\ \mu_i &= \frac{1}{\pi_i} \int_{y_{i-1}}^{y_i} yf(y)dy\end{aligned}\tag{5}$$

Each  $n$ -spike representation  $\rho$  of  $f$  induces an approximation error, which we denote by  $\varepsilon(f, \rho)$ . The error corresponds, as we have already discussed, to the implicit fuzziness of group identification: after all, the spikes are only a representation.

Looking ahead, we are going to define our measure of extended polarization in the following way:

$$P(f; \alpha, \beta) = ER(\alpha, \rho) - \beta \varepsilon(f, \rho),\tag{6}$$

where  $ER(\alpha, \rho)$  is the ER measure of polarization with parameter  $\alpha$  applied to the  $n$ -spike representation  $\rho$ , given by

$$ER(\alpha, \rho) = \sum_{i < j} \pi_i^{1+\alpha} \pi_j |\mu_i - \mu_j|\tag{7}$$

and where  $\beta$  is a free parameter which measures the weight we attach to the “measurement error” (or lack of identification) in downscaling the ER polarization computed from the representation.<sup>2</sup> Since the ER measure is defined on the simplified representation of the distribution, we shall refer to it as “simple” polarization and use the term “extended” polarization for the measure on the complete distribution.

Now we turn to the question of an appropriate  $n$ -spike representation for the income distribution at hand. There are really two questions here: one has to do with the *number* of spikes involved in the representation, and the other has to do with their *locations*. We view the number as exogenous (for instance, standard economic categories may use the “poor”, the “middle class”, and the “rich”), but concentrate on the endogenous determination of

---

<sup>2</sup> Note that the measure  $P(F; \alpha, \beta)$  described in (6) is not a special case of the class of measures defined by (4), simply because we subtract any fuzziness in identification linearly and do not interact it with the alienation term. But it does respect the broad conceptual issues raised in that section.

the locations. To be sure, there is no single answer to this question, but it is clear that the locations should respect some notion of group identification: a *group*, represented by a typical interval of the form  $[y_{i-1}, y_i]$ , should not have a large dispersion in the characteristics of its members (relative to the dispersion in the overall distribution).

One way, then, to locate the spikes is to define the approximation error  $\varepsilon(f, \rho)$  as

$$\varepsilon(f, \rho) = \frac{1}{2} \sum_{i=1}^n \int_{y_{i-1}}^{y_i} \int_{y_{i-1}}^{y_i} |x - z| f(x) f(z) dx dz \quad (8)$$

and choose the approximation  $\rho$  (for given  $n$ ) that minimizes this error. In this way one minimizes the average difference of income pairs within the groups: thus, implicitly, the dispersion within each group is being measured by the Gini coefficient. This is precisely the approach taken by Aghevli and Mehran [1981] and Davies and Shorrocks [1989]. Let  $\rho^*$  be the  $n$ -spike representation that solves this problem. This solution is characterized by the condition that

$$\int_{y_{i-1}}^{y_i} f(x) dx = \int_{y_{i-1}}^{y_{i+1}} x f(x) dx, \text{ for } i=1, \dots, n-1. \quad (9)$$

Expression (9) has quite a simple interpretation. Using our notation it may be rewritten as

$$y_i = \frac{\pi_i \mu_i + \pi_{i+1} \mu_{i+1}}{\pi_i + \pi_{i+1}}. \quad (10)$$

That is, the dividing income between any two adjacent intervals has to be equal to the average income of these two intervals taken together.

Figure 1: An error-minimizing  $n$ -spike representation

Diagrammatically, an  $n$ -spike representation of is equivalent to transforming the original Lorenz curve into a piecewise linear Lorenz curve (with  $n$  pieces). Hence, the minimization of (8) is equivalent to minimizing



the area between the original Lorenz curve and the piecewise linear representation, as Figure 1 shows. It is therefore immediate that

$$\varepsilon(f, \rho^*) = G(f) - G(\rho^*), \quad (11)$$

where  $G(\cdot)$  assigns the Gini coefficient to the distribution variable in its argument. Combining equations (6) and (11), we see that

$$P(f, \alpha, \beta) = ER(\alpha, \beta) - \beta[G(f) - G(\rho^*)] . \quad (12)$$

This is the proposed extension of the extended polarization measure.

### 3.2. The case of bi-polarization

Of special interest is the question: how *bipolar* is society? How close is the distribution to the formation of two large groups, presumably identified within each and standing in antagonism to each other? This kind of question can be addressed by taking  $n=2$  in the exercise above. As a matter of fact, this might not be an accurate way of capturing the social groupings that actually do exist. Thus this question is different from the issue of how *polarized* a distribution actually is, but it is a perfectly legitimate handle on the related concept that we shall refer to as *bipolarization*.

It will be enough in this case to focus on  $y_1$ , the cut-off that divides the two presumed groups. We shall drop subscripts here and refer to this cut-off as  $y$ , while  $\pi$  will denote the value of the cumulative distribution up to  $y$ : that is,

$$\pi = \int_a^y f(x)dx. \quad (13)$$

Normalize so that mean income for the entire distribution equals unity. Let  $L(\pi)$  denote the ordinate of the Lorenz curve of  $f$  at the point  $\pi$ . Then it is easy to check that  $\mu_1 = L(\pi)/\pi$  and  $\mu_2 = [1-L(\pi)]/[1-\pi]$ . Consequently,

$$\begin{aligned} ER(\alpha, \rho) &= [\pi^{1+\alpha}(1-\pi) + (1-\pi)^{1+\alpha}\pi][\mu_2 - \mu_1] = \\ &= [\pi^\alpha + (1-\pi)^\alpha][\pi - L(\pi)] \end{aligned} \quad (14)$$

while

$$\varepsilon(f, \rho) = G - [\pi - L(\pi)]. \quad (15)$$

Combining (14) and (15) in the way indicated by (6), but not optimizing with respect to the error just yet, we see that

$$P(f; \alpha, \beta, y) = [\pi^\alpha + (1 - \pi)^\alpha][\pi - L(\pi)] - \beta\{G - [\pi - L(\pi)]\}, \quad (16)$$

where the inclusion of  $y$  inside the  $P$ -function reminds us that we have not yet optimized with respect to the cut-off income level.

Such optimization clearly means that we choose  $y$ , or equivalently  $\pi$ , to maximize the vertical difference between the Lorenz curve and the 45° line:

$$\max_{\pi} [\pi - L(\pi)].$$

If the Lorenz curve is strictly convex (as it typically will be if the data is presented in the form of a density), there is a unique solution to this problem,  $y = \mu$ . Taking into account that the mean deviation  $D$  is

$$D = \frac{1}{2\mu} \int_0^\mu (\mu - y)f(y)dy = \pi_\mu - L(\pi_\mu), \quad (17)$$

where  $\pi_\mu = F(\mu)$ . Bearing in mind that  $\mu = 1$ , we may rewrite (16) as

$$P(f; \alpha, \beta) = [\pi_\mu^\alpha + (1 - \pi_\mu)^\alpha]D - \beta(G - D) \quad (18)$$

Note finally that for  $\alpha = 1$  we have the following simple expression

$$P(f; 1, \beta) = (1 + \beta)D - \beta G \quad (19)$$

But (16) also holds a route to the measure of polarization proposed by Wolfson [1994]. To see this, suppose that we choose  $y$  to equal the median income, call it  $m$ . Then  $\pi = 1/2$ , so that (16) becomes

$$P(f; \alpha, \beta, m) = (2^{1-\alpha} + \beta) \frac{1}{2} - L \frac{1}{2} - \beta G. \quad (20)$$

If we specialize further to the case in which  $\alpha = \beta = 1$ , then we have

$$P(f; 1, 1, m) = 2 \frac{1}{2} - L \frac{1}{2} - G = \frac{m}{2} P^W(f), \quad (21)$$

where  $P^W(f)$  stands for the W measure of polarization.

Let us conclude this section with a word of caution on the use of measures of bipolarization. It should be clear that such measures do not provide a general representation of polarization. In order to see this clearly consider the following case. We start with a distribution on  $nk$  equally sized spikes. Suppose that we concentrate the population of each  $k$  adjacent spikes and generate a new distribution into  $n$  equidistant equally sized spikes. We now wish to compare the degree of bipolarization before and after the concentration has taken place. Then, this concentration of probability will imply an increase or decrease in bipolarization, as measured by (21), according to whether  $n$  is even or odd.<sup>3</sup> The reason is simply that, departing from the uniform distribution, an even number of spikes takes us closer to the symmetric bipolar distribution, while an odd number takes us farther. In contrast, the ER measure of polarization will always increase, as shown in Esteban and Ray [1994, section 3.4.3].

### 3.3. Concluding Remarks on the Extended Measure

In this paper, we show how the ER polarization measure can be extended to cover situations in which the primitive distributional data are captured by a density, or by some grouping that does not correspond naturally to the notion of group identification. A by-product of this extension is that it provides an independent route, via the ER measure, to the W measure of polarization. In this way, we attempt to provide both an extension of existing literature on polarization measurement and a synthesis of it.

---

<sup>3</sup> We make these computations by leaving a mass  $\frac{1}{2}$  at the median income and letting  $\frac{1}{2}$  go to zero.

## 4. Income Polarization in the US and some Selected OECD Countries

Using the approach developed above we shall now analyze the level of polarization in the size distribution of household income in the US and in four selected OECD countries. We shall use the LIS (Luxembourg Income Study) database, which furnishes standardized household income distributions for a number of countries, permitting meaningful international comparisons. We shall thus work with household disposable income as defined in the LIS (household yearly earnings + cash property income + social insurance + social transfers - tax - mandatory contributions) adjusted accordingly with the OECD equivalence scales. The cost of this approach is that we are faced with severe limitations on the countries and years covered. We shall have to work with the five countries for which we have data for five/six years. These countries are: the United States, the United Kingdom, Canada, Germany and Sweden. However, there is no one single year for which we have information for five countries at the same year.

We have computed the extended polarization measure for two, three and four groups and for different values of the sensitivity parameter ( $\alpha = 1, 1.3, \text{ and } 1.6$ ). Furthermore, we have taken  $\beta = 1$  throughout.

### 4.1. The Polarization of the US distribution of personal income

#### 4.1.1. Trends in polarization

In Table 1 we report the level of extended polarization  $P$  exhibited by the distribution of household income in the years 1974, 1979, 1986, 1991, 1994 and 1997. We can observe that for two, three and four groups, and for all the values of the sensitivity parameter  $\alpha$ , polarization has increased by 10/20% over these twenty years. Most of this increase has taken place between 1979-1986 and 1991-1994. Figure 2 depicts these changes in polarization for  $\alpha = 1.6$ . It is worth noticing here that the rapid increase recorded in 1979-1986 is

consistent with the evidence of a “disappearing” middle class which has motivated a particular strand of literature since the second half of the eighties.<sup>4</sup> This stylized fact also seems to have played a major role in inspiring Wolfson’s (1994) way of thinking about polarization.

Figure 2

From Table 2 we can observe that the simplified polarization follows a similar pattern: it increases in steps. Interestingly, the within-group heterogeneity (the “error” term) increases when polarization increases. Therefore, the rising polarization displayed by the simplified representation of the distribution goes together with an increase in the fuzziness of the distribution of the population within the groups. The intra-group dispersion is depicted in Figure 3.

Figure 3

In order to understand the processes behind the increases in polarization we need to examine the changes experienced by the simplified representation of the income distribution and their impact on the simplified polarization and the dispersion within the groups. To do so we shall separately examine the bi-polar and multi-polar representations.

#### 4.1.2. Polarization in the bi-polar representation.

The optimal bi-polar representation of the income distribution takes the mean income as the optimal cut-off level between the two groups. In Table 3 we have that the idealized two-group representation of the income distribution accounts for more than 70% of the Gini coefficient of the original distribution, the rest being within group dispersion.

The within group dispersion has been rising steadily since 1979. This increase, however, has essentially been driven by the dispersion within the top income group with nearly a 40% increase in the same period. In contrast, the low income group has kept a constant level of internal dispersion. The

---

<sup>4</sup> See for instance Duncan et al. [1991], Horrigan and Haugen [1988], Kosters and Ross [1988], MacMahon and Tschetter [1986] and Rosenthal [1985].

recorded higher intra-group dispersion has thus diluted the changes in simple polarization.

The level of simple polarization displayed by this simplified bi-polar representation depends on the relative size of the two groups and on the distance between the two representative incomes. In Tables 6 and 7 we have the evolution of these two figures. It can be readily seen that the share of the households below the mean income fell in 1974-1979 from 59.2 to 57.5% and has continuously risen thereafter up to 63.2% in 1997. Since the sizes of the two groups have become increasingly uneven, this process by itself reduces the level of bi-polarization.

At the same time, the income distance between the two groups became smaller in the first period and has increased since, most specially in 1979-1986 and 1991-1994. The average income of the upper group was 2.62 times larger than the average income of the lower group in 1979 and 3.10 in 1997. As the two groups get farther apart from each other polarization increases *caeteris paribus*.

We thus have that the level of simplified polarization in the simplified bi-polar representation results from two processes moving in opposite directions. Hence, the net overall effect might critically depend on the value of  $\alpha$ . In our case, however, the evolution of the simplified polarization has a similar pattern for the three values of  $\alpha$ , which cover the entire range of acceptable values: it decreases first and rises after, most significantly in 1979-1986 and 1991-1994. In sum, the increase in the distance between the two groups has been so important relative to the higher unevenness in their sizes that simplified polarization has resulted in an increase for all acceptable values of the sensitivity parameter .

Taking together the measure of simplified polarization and the representation error, we have that extended bi-polarization has been increasing throughout the entire period 1974-1994 for all the values of and for =1. For the period 1994-1997, the increase in simplified bi-polarization has been outweighed by the increase in within-group dispersion, yielding a stabilization/slight decline in extended polarization (see Figure 2).

### 4.1.3. Polarization in the multi-polar representation.

Let us now examine the case of three and four groups. Representing the income distribution by three or four, rather than two, groups has a dramatic effect on its explanatory power: it explains over 86% and 92% of the recorded dispersion, respectively, to be contrasted with the 70% with two groups. The remaining part corresponds to within-group dispersion, i.e. the error made by this simplified representation.

In Table 3 we can observe that, for both three- and four-group representations, the intra-group dispersion fell in 1974-1979 and increased thereafter, except for 1986-1991 in which it remained constant. Therefore, on the average, income groups have become more dispersed. This increase in dispersion has not been homogeneous across groups in the period 1979-1997. For the three group case, the dispersion within the middle and upper income groups grew by 20 and 30%, respectively, while the low income group saw its dispersion slightly reduced. For the four group case, we have that the bottom group is the only one that reduces its internal dispersion in the period 1979-1997 (by nearly 10%). The other three groups increase their internal dispersion in increasingly proportional amounts: 10, 25 and 80%, respectively. This process of the groups becoming internally more diluted has contributed *per se* to the reduction of extended polarization.

#### Figure 4

Let us now turn to the characteristics of the simplified representations. Here again there is a clear change in pattern in 1979. Therefore, we shall concentrate on the period 1979-1997. Starting with the three-group case, we have in Table 6 that the population share of the low-income group has increased by nearly 15%, at the entire expense of the high-income group. This change has made group sizes more uneven, reducing *caeteris paribus* simple polarization. At the same time, income distances have spread out. The mean income of the low group has decreased by 5%, while the mean income of the top group increased by 25%. The net result of these two forces working in opposite directions has been to increase simple polarization for all values of  $\alpha$ . As for the four-group case, we observe a decrease of the share of the higher income group by nearly 50% and of the middle-high by 5%. Both the

middle-low and the lower income group shares increase by 6 and by 17%, respectively. Here again population shares become more uneven, with a shift towards the most populated side of the distribution, thus reducing polarization. As for income distances, we can see in Table 7 that in the period 1979-1997 the representative incomes of the different groups have spread out from the mean, with a significant shift in the mean income of the higher-income group.<sup>5</sup> The joint changes in group size and location are represented in Figure 4. All taken together, these changes result in an increase in simple polarization for all values of  $\alpha$ . Once again the increase in inter-group distances has been so significant that it has overridden the effect of an increased unevenness in the group population shares.

To close our analysis, and jointly considering the evolution of simplified polarization and intra-group dispersion, we conclude that extended polarization has increased over the period 1979-1997 for all values of  $\alpha$  and for three and four income groups.

## 4.2. An International Comparison

We shall be taking here the US income distribution as a benchmark against which we shall contrast the performance of a number of OECD countries. In the LIS database there are only the UK, Canada, Germany and Sweden with data covering a similar period. Unfortunately, neither the initial nor the final year of the period covered can be made to coincide for the five countries under consideration.

### 4.2.1. General Trends

---

<sup>5</sup> Our analysis of the group-simplified representation of the distribution reveals a pattern that appears to fit with casual evidence, but that does not quite conform to the stereotype disappearance of the middle class. In essence we have a low-income group which has become poorer, larger and with higher internal cohesion, together with a middle-high and a high income group (notice that we have 10% of households in this group) which is becoming thinner and substantially more disperse internally. This view can be supplemented with our results on the evolution of the cut-off incomes. The spread of the support of the centre groups has become larger.



From Table 1 we can see that the US and the UK have the highest levels of polarization over this set of countries. The differences in polarization with respect to the other three countries have increased over the period. These results hold qualitatively true for every number of groups and all values of  $\alpha$ . However, the evolution within the period is different in the UK where, in contrast to the continued increase recorded in the US, there has been only one –and dramatic- episode of increase in polarization: 1986-1991. Furthermore, whereas in the US increases in polarization went hand in hand with increases in inequality, for the case of the UK inequality already experienced an important increase in the previous period, 1986-1991, without bringing about any significant increase in polarization. Canada and Germany have a moderate –but steady- decrease in polarization, except for the last period. In Sweden, polarization experienced a significant decrease in 1975-1981 and increased thereafter. In Figure 5 we plot the evolution of polarization in three groups end for  $\alpha=1.6$ .

Figure 5

From Table 2 we can see that for all the countries and groupings the behaviour of the measure of extended polarization has essentially followed that of the simple polarization. Changes in intra-group dispersion have not been strong enough to counter the changes in simple polarization, except for the case of bi-polarization in the UK and Canada in the periods 1974-1979 and 1975-1981, respectively.

#### 4.2.2. Bi-polarization

The bi-polar representation of the income distribution turns out to be for all the countries considered as accurate as for the US. We can observe in Table 3 some differences. The unexplained dispersion (i.e. the intra-group dispersion) for the UK is around 28% of the value of the Gini coefficient, whilst this is over 30% for Sweden. Table 4 gives the internal dispersion of each group. Canada and most specially Sweden have less intra-group dispersion at the top income group than at the bottom group. The other three

countries have very similar degrees of dispersion in each of the two income groups.

Simple polarization combines the effects of evenness in group size with their distance –their alienation. The share of the low-income group varies quite significantly across countries (see Table 6). If we take the last year for every country, the bottom share varies from 63% in the US and the UK, to 60% in Canada and Germany, and to 56% in Sweden. Notice that from the viewpoint of polarization the more even the share of the two groups the higher will be the recorded polarization. As for the relative income distance between the two groups, in Table 7 we see that the ratio of the mean incomes of the two groups in the last year ranges from 3.1 (US), to 2.7 (UK), 2.4 (Canada), 2.2 (Germany), and to 2.0 (Sweden). We thus have that countries with more even groups also have them closer together. Therefore, bearing in mind the ranking in terms of simple polarization, we can conclude that for the set of countries under consideration, inter-group distance always dominates over the evenness of their respective share.

#### 4.2.3. Multi-group polarization

Simplified representations of distributions over more than two groups permit substantial improvements in the fit. The unexplained dispersion (the remaining dispersion within groups) drops to 14/15% and to 8/9% of the value of the global Gini coefficient, as it can be seen in Table 3. These values are quite similar among the different countries, with the exception of Sweden, who has higher residual intra-group dispersion for all groupings (two, three and four).

This intra-group dispersion varies across groups and countries. In Table 4 we can observe the different patterns. The first observation is that the US presents the highest dispersion group by group, over all countries in the last year. For obvious reasons, all countries show higher dispersion in the extreme groups than in the center group(s). The US, the UK and Germany have a very similar dispersion at the two extreme groups, while Canada and Sweden have a more concentrated top income group, relative to the low

income group. As for the evolution of the intra-group dispersion over the entire period, we have that the dispersion has increased in all groups in the UK and Sweden, decreased in the bottom group in the US and Canada, and remained about the same in Germany.

Let us now examine the two components explaining simple polarization. We shall consider separately the case of three and four groups. We start with the population shares for the three group case. We observe that in the last year the UK and the US have the largest low-income group, with 43/44% of the population, followed by Germany and Canada with 39/40%, and Sweden with 35% has the smallest such group. Again, the US and the UK have the smallest upper-income group, with 18/19% of the population, Germany has 21% in this group, and Canada and Sweden have 23% and 25%, respectively. In sum, the middle-income group has a similar size in all countries, but we observe significant shifts of the population from high- to low-income groups in the US and the UK. It follows that these two countries present the highest degree of unevenness in group size.

#### Figure 6

Let us now turn to the four group case. Here again there appear not to be major differences across countries with respect to the population share of the two middle groups –the low-middle with 32% and the high-middle with 26/29% of the population. The most notorious differences are in the share of the bottom and top-income groups. The US and the UK have 11/12% of the population in the top group, Germany 13%, and Canada and Sweden 15%, respectively (see Figure 6).

The second component of simple polarization is inter-group distance.<sup>6</sup> For the three group case, the country with highest alienation is the US, with a mean income for the top group 5.4 time the mean income of the bottom group. This ratio is 4.3 in the UK, 3.4 in Canada, 3.2 in Germany and 2.7 in Sweden. For the four group case, the ratio of mean incomes of the top and bottom groups is 7.9 (US), 5.8 (UK), 4.5 (Canada), 4.3 (Germany), and 3.6 (Sweden). These relative income distances have significantly increased during the period in the US and the UK, remained stable in Germany and

Sweden, and decreased in Canada. The relative income distances between the two extremes in the four-group case is represented in Figure 7.

Figure 7

The overall picture can be summarized as follows. There are essentially two types of patterns. On the one side, the US and the UK have a more uneven structure of group sizes, with large low-income groups and small high-income groups. On the other side, these countries have large (and increasing) income distances between high- and low-income groups. While the first characteristic tends to reduce polarization, the second tends to increase it. The net effect of these two opposite forces is driven by the significant size of the differences in income distances across countries. Both the US and the UK have increases in simple polarization, while the others have stability or decreases.

### 4.3. Summary of empirical findings

We observe two basic patterns for income distributions, each showing a distinct polarization level and structure as well as a distinct evolution through the period covered by our data. One pattern corresponds to the US and the UK and the other to Germany, Canada and Sweden.

In the first set of countries we observe high polarization and high inequality. During the period, the US polarization and inequality have followed a sustained increase, with two periods of rapid increases (1979-1986 and 1991-1994). This behaviour is common to all groupings and values of  $\alpha$ . In the UK we observe that inequality has had substantial increases in 1979-1986 and 1986-1991, while polarization has been decreasing in every period, except for a dramatic increase in 1986-1991. The second set of countries has lower polarization, lower inequality, and both have tended to decrease over the period studied.

The level and evolution of intra-group dispersion has played no critical role in modifying the qualitative behaviour of simple polarization.

---

<sup>6</sup> In Table 5 we give the endogenous cut-off incomes.

This might be due to the fact that we have been using a moderate value for the parameter  $\alpha$  ( $\alpha = 1$ ) giving weight to the error term in the definition of extended polarization.

In order to study the forces driving simple polarization we have analyzed the main features of the simplified representations of the original distributions. We have found that here again there are fundamental differences between the two sets of countries. The US and the UK have smaller sized upper-income groups and larger sized lower-income groups. This creates a distribution in group sizes that is more uneven than for the second set of countries. At the same time, inter-group income distances are larger for the US and the UK than for the other countries. The first factor tends to diminish polarization whilst the second tends to increase it. The net effect is driven by the significant differences in inter-group distances recorder in the first set of countries. Consequently, simple polarization is larger and increases for these countries and decreases for Canada, Germany and Sweden.

Let us finish with a remark on the number of groups to be used. In section 3 we have left open the issue of how many groups there should be. One possible criterion is to select the number of groups for which extended polarization is maximal. In view of Table 1, this depends on the value of parameter  $\alpha$ . If we have high sensitivity towards polarization –and hence use a high value for  $\alpha$  – we should use a three-group representation of the income distribution for *all* OECD countries under consideration. However, for low values of  $\alpha$ , extended polarization is maximal when we represent the income distributions by two groups. Interestingly enough, in this case, having a high sensitivity towards polarization drives to using a richer representation than the crude bi-modal.

## References

- Aghevli, B.B. and F. Mehran (1981), "Optimal Grouping of Income Distribution Data", *Journal of the American Statistical Association*, **76**, 22--26.
- Davies, J.B. and A.F. Shorrocks (1989), "Optimal Grouping of Income and Wealth Data", *Journal of Econometrics*, **42**, 97--108.
- Duncan, G., T.M. Smeeding and W. Rodgers (1991), "W(h)ither the Middle Class?. A Dynamic View", the Levy Institute Conference on Income Inequality, Bard College, June.
- Esteban, J. and D. Ray (1994), "On the Measurement of Polarization", *Econometrica*, **62**, 819--851.
- Esteban, J. and D. Ray (1997), "Conflict and Distribution", mimeo, Boston University and Institut d'Anàlisi Econòmica.
- Horrigan, M.W. and S.E. Haugen (1988), "The Declining Middle-Class Thesis: A Sensitivity Analysis", *Monthly Labour Review* **111**, 3-13.
- Kosters, M.H. and M.N. Ross (1988), "A Shrinking Middle Class?", *The Public Interest* **90**, 3-27.
- Love, R. and M.C. Wolfson (1976), *Income Inequality: Statistical Methodology and Canadian Illustrations*, Catalogue 13-559 Occasional. Ottawa: Statistics Canada.
- MacMahon, P.J. and J.H. Tschetter (1986), "The Declining Middle: a Further Analysis", *Monthly Labour Review* **109**,
- Rosenthal, N.H. (1985), "The Shrinking Middle Class: Myth or Reality?", *Monthly Labour Review* **108**,
- Wolfson, M.C. (1994), "When Inequalities Diverge", *American Economic Review* **84**, Papers and Proceedings, 353--358.

## Tables and Figures

**Table 1. Extended Polarization  $P$  ( $\beta=1$ )**

N <sup>er</sup> . of Groups	$\alpha=1$			$\alpha =1.3$			$\alpha =1.6$			Gini
	2	3	4	2	3	4	2	3	4	
<b>U. S.</b>										
<b>1974</b>	0.1410	0.1485	0.1289	0.0965	0.0945	0.0767	0.0605	0.0558	0.0422	0.3478
<b>1979</b>	0.1381	0.1439	0.1242	0.0949	0.0917	0.0736	0.0599	0.0544	0.0404	0.3311
<b>1986</b>	0.1507	0.1553	0.1342	0.1044	0.0994	0.0801	0.0670	0.0593	0.0444	0.3563
<b>1991</b>	0.1524	0.1570	0.1355	0.1059	0.1005	0.0811	0.0684	0.0601	0.0452	0.3590
<b>1994</b>	0.1619	0.1670	0.1440	0.1126	0.1074	0.0863	0.0730	0.0646	0.0482	0.3841
<b>1997</b>	0.1607	0.1685	0.1466	0.1120	0.1093	0.0891	0.0730	0.0669	0.0509	0.3903
<b>U. K.</b>										
<b>1974</b>	0.1226	0.1252	0.1090	0.0857	0.0806	0.0659	0.0560	0.0487	0.0374	0.2871
<b>1979</b>	0.1227	0.1231	0.1053	0.0863	0.0794	0.0633	0.0570	0.0481	0.0355	0.2769
<b>1986</b>	0.1276	0.1296	0.1118	0.0885	0.0823	0.0660	0.0571	0.0485	0.0358	0.3078
<b>1991</b>	0.1526	0.1544	0.1335	0.1079	0.1007	0.0816	0.0720	0.0622	0.0471	0.3495
<b>1995</b>	0.1471	0.1503	0.1282	0.1033	0.0976	0.0767	0.0682	0.0597	0.0427	0.3462
<b>Sweden</b>										
<b>1975</b>	0.1000	0.0996	0.0862	0.0689	0.0627	0.0505	0.0437	0.0361	0.0270	0.2331
<b>1981</b>	0.0803	0.0835	0.0733	0.0540	0.0517	0.0427	0.0327	0.0290	0.0225	0.2011
<b>1987</b>	0.0903	0.0955	0.0849	0.0597	0.0586	0.0493	0.0350	0.0322	0.0260	0.2359
<b>1992</b>	0.0925	0.0987	0.0868	0.0616	0.0611	0.0505	0.0364	0.0343	0.0267	0.2413
<b>Germany</b>										
<b>1973</b>	0.1170	0.1220	0.1063	0.0805	0.0776	0.0633	0.0512	0.0457	0.0349	0.2877
<b>1978</b>	0.1134	0.1184	0.1030	0.0781	0.0756	0.0615	0.0498	0.0450	0.0341	0.2780
<b>1983</b>	0.1099	0.1157	0.1000	0.0759	0.0742	0.0598	0.0486	0.0446	0.0333	0.2690
<b>1984</b>	0.1072	0.1119	0.0976	0.0735	0.0712	0.0582	0.0463	0.0420	0.0321	0.2640
<b>1989</b>	0.1024	0.1107	0.0962	0.0689	0.0699	0.0567	0.0418	0.0408	0.0307	0.2648
<b>1994</b>	0.1086	0.1163	0.1020	0.0736	0.0736	0.0607	0.0454	0.0430	0.0334	0.2778
<b>Canada</b>										
<b>1971</b>	0.1544	0.1596	0.1379	0.1067	0.1020	0.0823	0.0682	0.0607	0.0456	0.3673
<b>1975</b>	0.1348	0.1404	0.1213	0.0924	0.0893	0.0719	0.0581	0.0526	0.0393	0.3250
<b>1981</b>	0.1324	0.1350	0.1168	0.0917	0.0859	0.0694	0.0588	0.0508	0.0382	0.3122
<b>1987</b>	0.1299	0.1323	0.1141	0.0901	0.0843	0.0678	0.0580	0.0499	0.0372	0.3066
<b>1991</b>	0.1253	0.1278	0.1105	0.0865	0.0810	0.0653	0.0552	0.0475	0.0355	0.2992
<b>1994</b>	0.1279	0.1287	0.1113	0.0891	0.0818	0.0660	0.0578	0.0482	0.0362	0.2993

**Table 2. Extended Polarization by Components: ER**

N <sup>er</sup> . of Groups	$\alpha = 1$			$\alpha = 1.3$			$\alpha = 1.6$		
	2	3	4	2	3	4	2	3	4
<b>U. S.</b>									
<b>1974</b>	0.2444	0.1971	0.1571	0.1998	0.1431	0.1048	0.1639	0.1044	0.0703
<b>1979</b>	0.2346	0.1887	0.1500	0.1914	0.1365	0.0995	0.1565	0.0992	0.0663
<b>1986</b>	0.2535	0.2034	0.1619	0.2072	0.1474	0.1078	0.1698	0.1073	0.0722
<b>1991</b>	0.2557	0.2049	0.1634	0.2092	0.1484	0.1089	0.1717	0.1080	0.0730
<b>1994</b>	0.2730	0.2189	0.1743	0.2238	0.1593	0.1166	0.1842	0.1166	0.0785
<b>1997</b>	0.2755	0.2221	0.1775	0.2268	0.1630	0.1199	0.1879	0.1206	0.0818
<b>U. K.</b>									
<b>1974</b>	0.2049	0.1639	0.1313	0.1680	0.1194	0.0883	0.1383	0.0875	0.0598
<b>1979</b>	0.1998	0.1594	0.1269	0.1634	0.1157	0.0849	0.1341	0.0844	0.0572
<b>1986</b>	0.2177	0.1738	0.1377	0.1786	0.1266	0.0919	0.1472	0.0927	0.0617
<b>1991</b>	0.2510	0.2009	0.1610	0.2063	0.1472	0.1091	0.1705	0.1087	0.0746
<b>1995</b>	0.2467	0.1978	0.1560	0.2029	0.1451	0.1045	0.1678	0.1073	0.0705
<b>Sweden</b>									
<b>1975</b>	0.1666	0.1325	0.1055	0.1355	0.0955	0.0698	0.1103	0.0690	0.0463
<b>1981</b>	0.1407	0.1128	0.0904	0.1144	0.0811	0.0598	0.0931	0.0584	0.0397
<b>1987</b>	0.1631	0.1314	0.1058	0.1325	0.0945	0.0703	0.1078	0.0682	0.0469
<b>1992</b>	0.1669	0.1348	0.1080	0.1359	0.0973	0.0717	0.1108	0.0704	0.0478
<b>Germany</b>									
<b>1973</b>	0.2024	0.1624	0.1298	0.1659	0.1180	0.0867	0.1365	0.0862	0.0584
<b>1978</b>	0.1957	0.1574	0.1256	0.1604	0.1147	0.0841	0.1321	0.0840	0.0567
<b>1983</b>	0.1894	0.1526	0.1215	0.1555	0.1112	0.0813	0.1281	0.0815	0.0547
<b>1984</b>	0.1856	0.1494	0.1194	0.1519	0.1086	0.0800	0.1247	0.0795	0.0539
<b>1989</b>	0.1836	0.1490	0.1189	0.1501	0.1083	0.0794	0.1231	0.0792	0.0533
<b>1994</b>	0.1932	0.1562	0.1249	0.1582	0.1134	0.0836	0.1301	0.0829	0.0564
<b>Canada</b>									
<b>1971</b>	0.2609	0.2094	0.1672	0.2132	0.1517	0.1116	0.1747	0.1104	0.0748
<b>1975</b>	0.2299	0.1849	0.1472	0.1875	0.1337	0.0978	0.1532	0.0971	0.0652
<b>1981</b>	0.2223	0.1777	0.1417	0.1816	0.1286	0.0943	0.1487	0.0935	0.0631
<b>1987</b>	0.2183	0.1745	0.1387	0.1785	0.1265	0.0924	0.1463	0.0921	0.0618
<b>1991</b>	0.2122	0.1696	0.1349	0.1734	0.1228	0.0896	0.1421	0.0893	0.0599
<b>1994</b>	0.2136	0.1700	0.1352	0.1748	0.1232	0.0899	0.1436	0.0896	0.0601



**Table 3. Extended Polarization by Components: Intra-group Dispersion**

N <sup>er</sup> . of Groups	2		3		4	
	Absolute Level	Relative (%Gini)	Absolute Level	Relative (%Gini)	Absolute Level	Relative (%Gini)
<b>U. S.</b>						
1974	0.1034	29.73	0.0486	13.98	0.0282	8.10
1979	0.0965	29.15	0.0448	13.52	0.0259	7.81
1986	0.1028	28.85	0.0480	13.47	0.0278	7.79
1991	0.1033	28.78	0.0479	13.34	0.0278	7.75
1994	0.1111	28.93	0.0519	13.52	0.0302	7.87
1997	0.1148	29.42	0.0537	13.75	0.0308	7.90
<b>U. K.</b>						
1974	0.0822	28.65	0.0388	13.51	0.0223	7.78
1979	0.0771	27.85	0.0363	13.11	0.0216	7.81
1986	0.0901	29.27	0.0443	14.38	0.0259	8.42
1991	0.0985	28.18	0.0465	13.30	0.0274	7.85
1995	0.0996	28.76	0.0475	13.73	0.0278	8.03
<b>Sweden</b>						
1975	0.0665	28.54	0.0328	14.09	0.0193	8.27
1981	0.0604	30.03	0.0294	14.60	0.0172	8.53
1987	0.0728	30.87	0.0359	15.23	0.0209	8.88
1992	0.0744	30.82	0.0361	14.97	0.0212	8.77
<b>Germany</b>						
1973	0.0853	29.66	0.0404	14.05	0.0235	8.15
1978	0.0823	29.60	0.0390	14.04	0.0226	8.13
1983	0.0795	29.57	0.0369	13.74	0.0214	7.97
1984	0.0784	29.70	0.0375	14.19	0.0218	8.25
1989	0.0812	30.68	0.0384	14.49	0.0227	8.55
1994	0.0846	30.46	0.0398	14.33	0.0229	8.25
<b>Canada</b>						
1971	0.1065	28.98	0.0497	13.54	0.0293	7.97
1975	0.0951	29.26	0.0444	13.67	0.0259	7.97
1981	0.0899	28.79	0.0427	13.68	0.0249	7.97
1987	0.0884	28.82	0.0422	13.75	0.0246	8.02
1991	0.0869	29.05	0.0418	13.97	0.0243	8.14
1994	0.0857	28.64	0.0414	13.82	0.0239	7.98

Table 4. Intra-group-Dispersion by Income Groups

N <sup>er</sup> . of Groups	2		3			4			
	1	2	1	2	3	1	2	3	4
<b>U. S.</b>									
<b>1974</b>	0.248	0.197	0.245	0.100	0.175	0.258	0.087	0.076	0.166
<b>1979</b>	0.244	0.174	0.239	0.097	0.143	0.245	0.085	0.071	0.129
<b>1986</b>	0.259	0.189	0.249	0.106	0.158	0.251	0.093	0.078	0.145
<b>1991</b>	0.255	0.193	0.242	0.107	0.161	0.241	0.092	0.081	0.146
<b>1994</b>	0.269	0.216	0.257	0.114	0.187	0.261	0.097	0.086	0.178
<b>1997</b>	0.250	0.247	0.230	0.114	0.233	0.225	0.094	0.089	0.233
<b>U. K.</b>									
<b>1974</b>	0.162	0.182	0.131	0.087	0.168	0.116	0.069	0.067	0.169
<b>1979</b>	0.161	0.160	0.131	0.085	0.136	0.124	0.070	0.066	0.127
<b>1986</b>	0.189	0.190	0.182	0.092	0.171	0.210	0.071	0.074	0.165
<b>1991</b>	0.204	0.217	0.168	0.107	0.199	0.156	0.088	0.082	0.204
<b>1995</b>	0.207	0.217	0.185	0.105	0.197	0.189	0.083	0.083	0.191
<b>Sweden</b>									
<b>1975</b>	0.163	0.117	0.154	0.073	0.098	0.164	0.062	0.049	0.093
<b>1981</b>	0.148	0.103	0.159	0.060	0.084	0.191	0.048	0.045	0.079
<b>1987</b>	0.189	0.121	0.214	0.067	0.108	0.279	0.060	0.049	0.106
<b>1992</b>	0.173	0.136	0.185	0.068	0.122	0.222	0.057	0.053	0.118
<b>Germany</b>									
<b>1973</b>	0.175	0.181	0.161	0.084	0.164	0.161	0.066	0.068	0.158
<b>1978</b>	0.162	0.181	0.143	0.081	0.170	0.137	0.064	0.066	0.169
<b>1983</b>	0.152	0.179	0.129	0.078	0.167	0.118	0.061	0.065	0.163
<b>1984</b>	0.154	0.171	0.134	0.077	0.165	0.126	0.061	0.063	0.173
<b>1989</b>	0.164	0.172	0.151	0.074	0.168	0.150	0.061	0.061	0.176
<b>1994</b>	0.171	0.181	0.161	0.078	0.170	0.161	0.062	0.066	0.167
<b>Canada</b>									
<b>1971</b>	0.275	0.193	0.273	0.108	0.161	0.289	0.094	0.080	0.148
<b>1975</b>	0.237	0.173	0.229	0.095	0.146	0.234	0.083	0.070	0.135
<b>1981</b>	0.211	0.171	0.196	0.094	0.145	0.199	0.079	0.070	0.134
<b>1987</b>	0.199	0.173	0.185	0.092	0.148	0.186	0.076	0.072	0.139
<b>1991</b>	0.196	0.170	0.185	0.090	0.147	0.190	0.073	0.069	0.141
<b>1994</b>	0.190	0.170	0.176	0.091	0.146	0.178	0.072	0.071	0.137

(\*) groups are ordered by (ascending) average income.

**Table 5. Inter-Group Cut-off Incomes (relative to the mean income)**

N <sup>er</sup> . of Groups	2	3		4		
	1 and 2	1 and 2	2 and 3	1 and 2	2 and 3	3 and 4
<b>U. S.</b>						
1974	1	0.728	1.357	0.589	1.004	1.603
1979	1	0.733	1.335	0.593	1.001	1.553
1986	1	0.713	1.370	0.574	1.009	1.626
1991	1	0.707	1.372	0.574	1.009	1.640
1994	1	0.696	1.408	0.553	1.006	1.702
1997	1	0.706	1.429	0.569	1.015	1.764
<b>U. K.</b>						
1974	1	0.773	1.310	0.672	1.020	1.545
1979	1	0.770	1.302	0.672	1.019	1.511
1986	1	0.758	1.328	0.635	0.977	1.540
1991	1	0.727	1.400	0.610	1.033	1.703
1995	1	0.735	1.391	0.602	0.990	1.648
<b>Sweden</b>						
1975	1	0.805	1.235	0.698	1.010	1.371
1981	1	0.827	1.181	0.726	0.981	1.288
1987	1	0.803	1.212	0.688	0.991	1.341
1992	1	0.805	1.225	0.695	0.990	1.369
<b>Germany</b>						
1973	1	0.773	1.292	0.667	0.999	1.516
1978	1	0.785	1.293	0.681	1.003	1.506
1983	1	0.790	1.281	0.691	0.999	1.489
1984	1	0.795	1.275	0.693	1.009	1.469
1989	1	0.796	1.265	0.690	1.001	1.452
1994	1	0.783	1.275	0.678	0.996	1.483
<b>Canada</b>						
1971	1	0.704	1.381	0.565	1.020	1.652
1975	1	0.737	1.328	0.603	1.008	1.546
1981	1	0.745	1.322	0.625	1.009	1.545
1987	1	0.751	1.321	0.631	1.001	1.535
1991	1	0.758	1.308	0.637	0.994	1.510
1994	1	0.755	1.311	0.639	0.991	1.522

(\*) groups are ordered by (ascending) average income.

**Table 6. Group Population Shares (% Households)**

N <sup>er</sup> . of Groups	2		3			4			
	1	2	1	2	3	1	2	3	4
<b>U. S.</b>									
<b>1974</b>	59.19	40.81	39.24	39.06	21.70	28.13	31.32	26.86	13.69
<b>1979</b>	57.48	42.52	37.92	38.74	23.34	27.27	30.28	26.58	15.87
<b>1986</b>	58.95	41.05	39.28	37.90	22.83	28.72	30.74	25.81	14.73
<b>1991</b>	59.63	40.37	39.31	37.86	22.83	29.39	30.77	25.39	14.46
<b>1994</b>	60.82	39.18	40.89	38.01	21.10	30.15	31.02	25.50	13.33
<b>1997</b>	63.19	36.81	42.94	38.81	18.25	31.97	32.10	25.22	10.71
<b>U. K.</b>									
<b>1974</b>	60.93	39.07	41.37	37.60	21.03	31.55	30.87	25.03	12.55
<b>1979</b>	59.55	40.45	41.11	35.66	23.23	31.46	29.45	24.77	14.31
<b>1986</b>	61.28	38.72	41.60	37.39	21.01	26.73	32.96	26.60	13.71
<b>1991</b>	62.33	37.67	43.84	36.40	19.76	34.13	30.26	24.18	11.43
<b>1995</b>	62.66	37.34	43.76	37.16	19.08	30.20	31.80	25.61	12.40
<b>Sweden</b>									
<b>1975</b>	54.11	45.89	36.58	37.04	26.37	25.73	29.31	27.38	17.58
<b>1981</b>	53.58	46.42	32.54	39.52	27.94	20.19	31.11	28.46	20.23
<b>1987</b>	53.03	46.97	32.82	39.89	27.29	20.36	31.85	29.92	17.86
<b>1992</b>	55.73	44.27	34.94	40.33	24.74	22.77	32.06	28.69	16.48
<b>Germany</b>									
<b>1973</b>	60.79	39.21	39.54	39.16	21.30	28.48	32.29	26.00	13.23
<b>1978</b>	60.96	39.04	40.75	38.74	20.51	29.19	32.09	26.15	12.57
<b>1983</b>	61.43	38.57	40.60	39.08	20.33	29.41	31.88	25.82	12.89
<b>1984</b>	59.81	40.19	40.19	38.70	21.11	28.84	31.71	26.89	12.56
<b>1989</b>	59.15	40.85	38.73	40.18	21.09	27.61	31.78	27.37	13.23
<b>1994</b>	60.30	39.70	38.75	40.20	21.05	26.93	33.06	27.35	12.65
<b>Canada</b>									
<b>1971</b>	58.83	41.17	39.18	38.01	22.81	29.79	30.20	25.57	14.45
<b>1975</b>	57.27	42.73	37.91	38.54	23.55	27.83	30.02	26.70	15.46
<b>1981</b>	58.53	41.47	38.99	37.62	23.39	28.44	30.79	26.01	14.77
<b>1987</b>	59.20	40.80	39.80	37.34	22.85	27.97	31.25	26.24	14.55
<b>1991</b>	58.84	41.16	39.16	37.78	23.06	27.10	31.36	26.49	15.05
<b>1994</b>	59.86	40.14	39.74	37.15	23.11	27.38	31.78	25.83	15.02

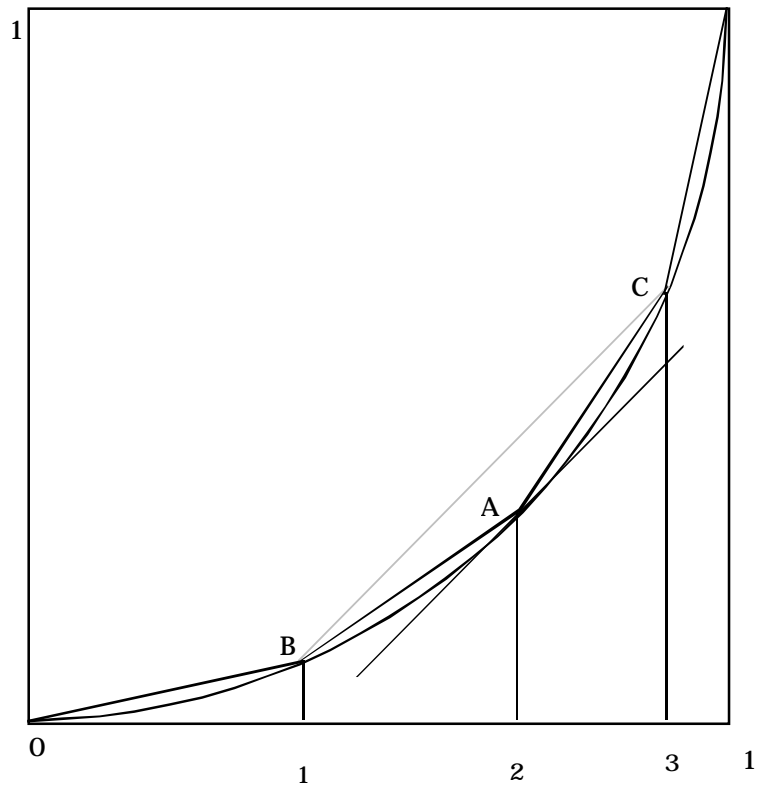
(\*) groups are ordered by (ascending) average income.

Table 7. Mean Incomes by Groups (relative to the mean)

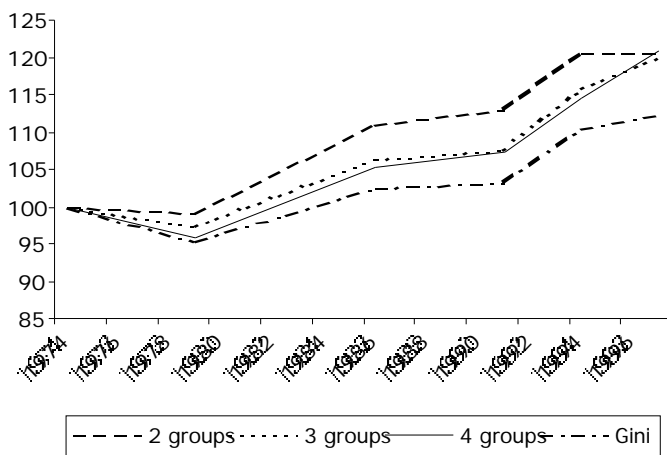
<b>N<sup>er</sup>. of Groups</b>	<b>2</b>		<b>3</b>			<b>4</b>			
<b>group (*)</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b>U. S.</b>									
<b>1974</b>	0.587	1.599	0.448	1.009	1.983	0.365	0.790	1.255	2.284
<b>1979</b>	0.592	1.552	0.451	1.009	1.876	0.369	0.794	1.237	2.080
<b>1986</b>	0.570	1.618	0.427	1.008	1.972	0.348	0.785	1.276	2.237
<b>1991</b>	0.571	1.633	0.427	0.998	1.990	0.354	0.786	1.281	2.276
<b>1994</b>	0.551	1.697	0.410	1.004	2.136	0.333	0.768	1.296	2.481
<b>1997</b>	0.564	1.748	0.431	1.012	2.314	0.360	0.779	1.314	2.833
<b>U. K.</b>									
<b>1974</b>	0.664	1.524	0.561	1.007	1.850	0.511	0.837	1.244	2.145
<b>1979</b>	0.665	1.494	0.566	1.006	1.758	0.520	0.835	1.239	1.982
<b>1986</b>	0.645	1.562	0.538	1.002	1.911	0.451	0.786	1.215	2.170
<b>1991</b>	0.597	1.666	0.488	1.016	2.107	0.437	0.807	1.319	2.518
<b>1995</b>	0.606	1.661	0.497	1.016	2.123	0.421	0.774	1.260	2.453
<b>Sweden</b>									
<b>1975</b>	0.692	1.363	0.592	1.018	1.541	0.525	0.849	1.184	1.661
<b>1981</b>	0.737	1.303	0.625	0.994	1.445	0.531	0.853	1.121	1.525
<b>1987</b>	0.693	1.347	0.564	1.001	1.523	0.452	0.839	1.154	1.655
<b>1992</b>	0.700	1.377	0.582	0.999	1.592	0.490	0.841	1.156	1.741
<b>Germany</b>									
<b>1973</b>	0.667	1.516	0.552	0.996	1.838	0.487	0.826	1.214	2.108
<b>1978</b>	0.679	1.501	0.575	1.005	1.834	0.513	0.833	1.212	2.116
<b>1983</b>	0.692	1.491	0.589	1.001	1.820	0.531	0.838	1.197	2.076
<b>1984</b>	0.690	1.462	0.591	1.007	1.766	0.530	0.842	1.206	2.037
<b>1989</b>	0.690	1.449	0.580	1.006	1.760	0.514	0.844	1.186	2.003
<b>1994</b>	0.680	1.487	0.564	0.995	1.812	0.491	0.830	1.195	2.106
<b>Canada</b>									
<b>1971</b>	0.557	1.634	0.410	1.007	2.003	0.339	0.789	1.292	2.288
<b>1975</b>	0.599	1.538	0.462	1.008	1.853	0.387	0.803	1.239	2.074
<b>1981</b>	0.620	1.536	0.496	1.004	1.833	0.426	0.808	1.247	2.070
<b>1987</b>	0.631	1.535	0.515	1.004	1.839	0.440	0.803	1.236	2.075
<b>1991</b>	0.639	1.516	0.522	1.002	1.809	0.444	0.803	1.220	2.024
<b>1994</b>	0.643	1.532	0.528	0.997	1.816	0.451	0.801	1.225	2.036

(\*) groups are ordered by (ascending) average income.

**Figure 1: Error-minimizing  $n$ -spike representation ( $n=4$ )**

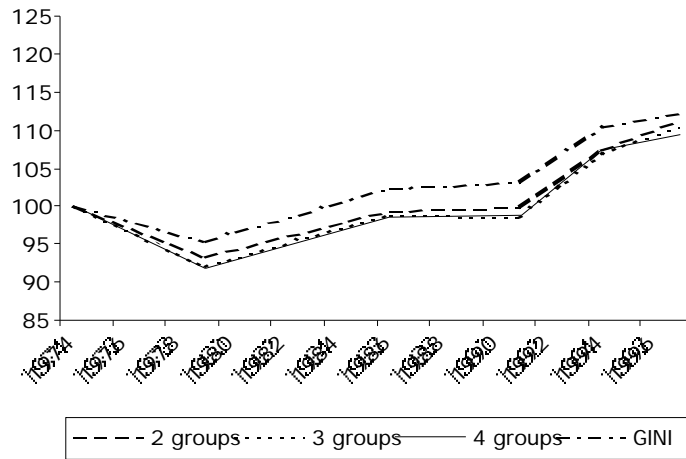


**Figure 2: Polarization and Inequality in the US Income Distribution. ( $\alpha=1.6$ )**



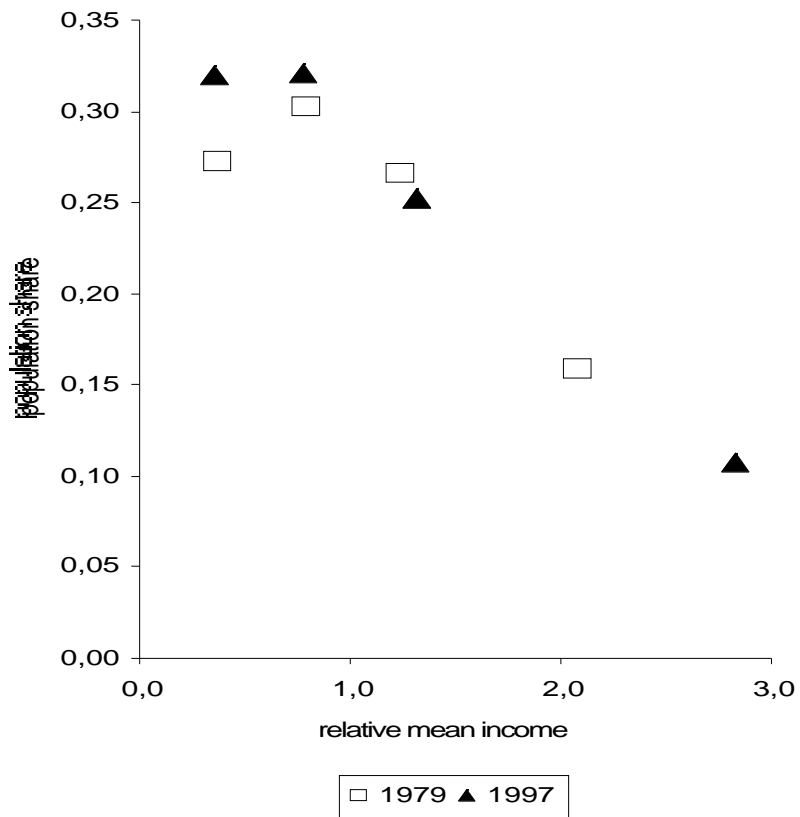
**Note: indices with 1974=100.**

**Figure 3: US Income Distribution: Intra-group Dispersion**

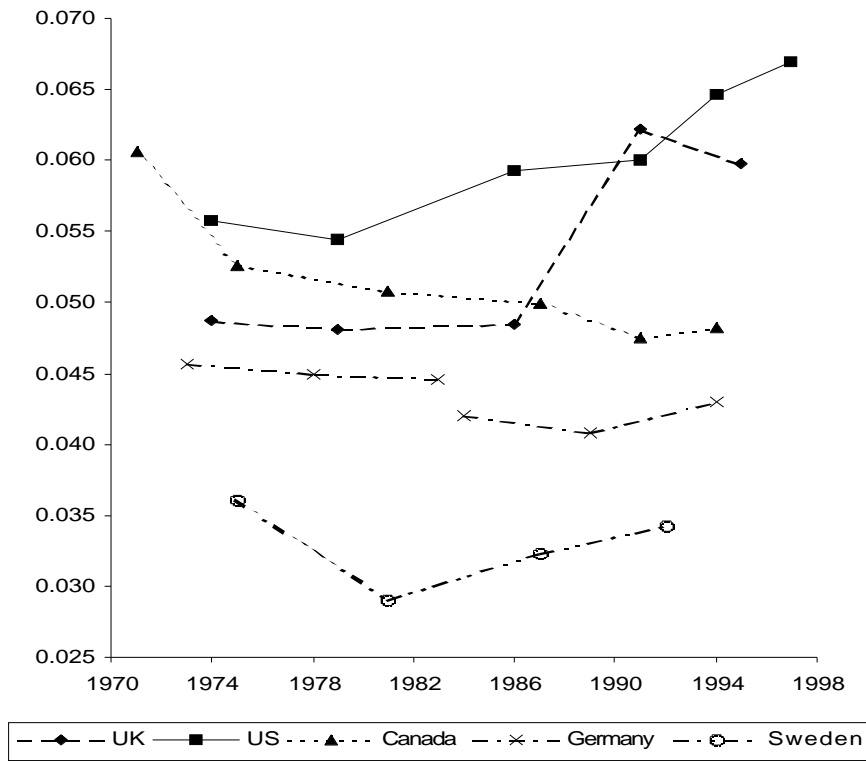


Note: indices, 1974=100.

**Figure 4: US Income Distribution: the 4-Group Representation**



**Figure 5: Polarization in Five OECD Countries: 3 Groups and ( $\alpha=1.6$ )**



**Figure 6: Share of the Low-Income Group in a 4-Group Partition**

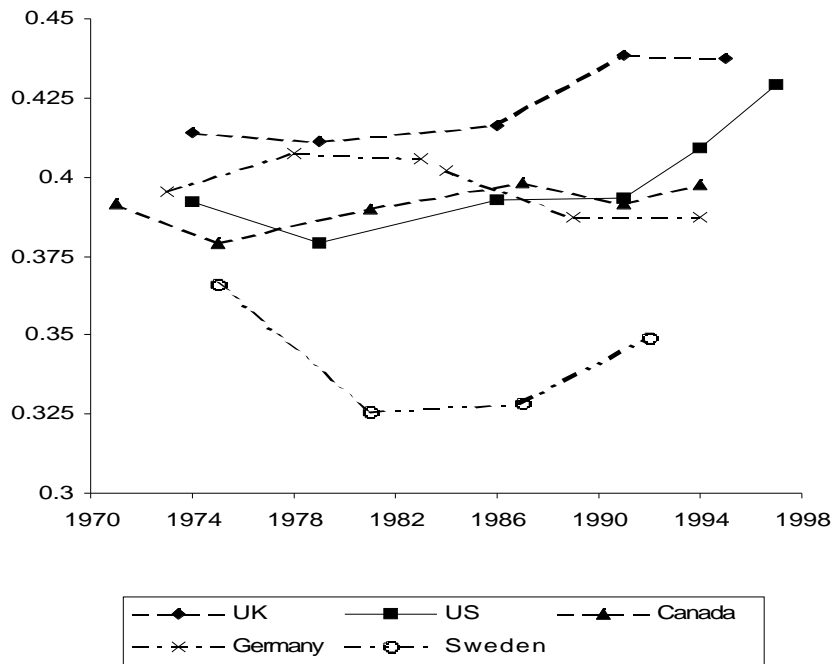




Figure 7: Relative Income Distance High/Low Income Groups in 4-Group Partition

