Keywords: Income Distribution, Polarization, Inequality. JEL Classification Code: D31, D63

Abstract

In this paper we analyze the trends in polarization in the US and other developed and transitional economies. Many studies have stressed the break in equalizing trends in income distributions in these countries. As polarization has emerged in the literature concerning the issue of income distribution, we here apply new contributions to the measurement of polarization proposed in Esteban, Ray and Gradín (1997) to describe their tendencies through time and across countries. Additionally, we investigate whether polarization was mainly due to increasing polarization between the groups or it was the result of their groups becoming internally more identified. This method allows us to analyze whether polarization between the groups has increased due to a larger income distance between the extreme groups or due to an increase in their sizes (shrinking the ámiddle class»).

* I acknowledge the provision of data necessary for this paper by both LIS and University of Michigan.

1. Introduction

The issue of the distribution of individual earnings and household’s incomes, has focused the attention of an increasing number of researchers during the last years in different countries. Levy and Murnane (1992) surveyed the main contributions to the study of increasing earnings inequality in the US while more recently Gottschalk and Smeeding (1997) have reviewed the distribution of both earnings and income inequality, analyzing the way by which the break in the equalizing trend that had characterized the post-war period has been shared by other countries with varying intensity.

Several empirical analyses have left the standard methodology of studying inequality, focusing instead on alternative tools to describe the shape of distributions in order to investigate whether the population has moved away from the middle to the extremes of the income space, and whether the extremes moved away from each other, the distribution thus becoming more bipolarized.

The áquantile/share approach» has been used to study the structural changes experienced in the labor market in the US and Canada towards grea-
ter dualization, testing the hypothesis of the disappearance of the ámiddle class». This class was defined as those workers or categories of workers whose wages fell within an interval with bounds defined as proportions of either the mean or the median.

A different approach was undertaken to study the increasing dualization of incomes in the UK through the inspection of non-parametric density estimates. These studies\(^3\) have shown that during the 1980s part of the population shifted from the middle to both extremes of the distribution of absolute income, but with more intensity towards the top. Alternative tools such as those based on árelative distributions» have also focused on similar issues\(^4\).

Inequality measures based on the Pigou-Dalton Principle of Transfers fail to distinguish whether the distribution exhibits convergence to the global mean or is clustering around a few local poles. This failure motivated the distinction between polarization and inequality, recently conceptualized by Wolfson (1994) and Esteban and Ray (1994) - ER hereafter-, opening a new branch for empirical research. Distributive analyses based on inequality measures must be completed with analyses based on measures of polarization.

In Esteban, Ray and Gradín (1997) - ERG for short- a new methodology which extends the ER approach was introduced to study the degree of polarization displayed by continuous distributions (or with a large number of observations). This generalization of the measure of polarization has the index proposed by Wolfson as a particular case. It is based on the ER model of attitudes stating that an individual feels some identification with members of his own group, but alienation from members of other groups, and that polarization is defined as the addition of the effective antagonism that results from the interaction between both feelings. Taking the number of groups as exogenous, the method builds a k-spike representation of the distribution which minimizes the error of approximation when it fits the original one. Then overall polarization is defined as the level of polarization displayed by this k-spike representation, as measured by the ER index, minus the effect of the lack of identification within the groups, accounted for by the error that the simplified representation induces.

In this paper we use this new methodology in order to analyze the polarization trends in the US, as well as in other developed and transitional countries. This empirical work makes use of the large data sets provided by the Luxembourg Income Study (LIS) and the cross-section of the panel data of the University of Michigan.

We conduct the analysis first for the US and then for other countries. In each case, this approach can be structured into three stages. The first stage analyzes the levels and trends of polarization and inequality. The second stage deals with the question of whether polarization is produced either because groups have become more


polarized or because they are internally more homogenous. The third stage undertakes the analysis of the $k$-spike representation, trying to identify whether polarization is the result of an increasing gap between extreme groups or of these groups becoming larger.

The paper is organized as follows. The next two sections present the methodology proposed in ERG. Section 4 describes the data and variables used for the analysis, and section 5 describes the trends in polarization and inequality in the US. Section 6 establishes whether polarization is the result of heterogeneity between groups or homogeneity within them, and section 7 deals with the analysis of optimal partitions in order to identify whether the source of polarization has been the gap between the extreme groups or the concentration of population on the extremes. Section 8 extends the analysis to other LIS countries with long series of data. Section 9 inspects their optimal partitions and section 10 extends the approach to the rest of LIS countries. The final section provides the concluding remarks.

2. The measurement of Polarization

Let us consider a particular distribution defined by a density $f$ or its respective cumulative density $F$ over the bounded support $[a, b]$ with mean income normalized to $\mu=1$. As in the ER approach, we assume that each member of the distribution with income $x$ feels some degree of group identification $I(x,F)$ with members of his own group and alienation $r(x,y)$ from those with income $y$ belonging to a different group. The interaction between both feelings gives rise to the effective antagonism $T(I(x,F),r(x,y))$ individual $y$ feels towards $x$, increasing in both terms such that a higher degree of intra-group identification reinforces the effect of alienation. Polarization is defined as the addition of all effective antagonism in the distribution:

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

The approach in ER states furthermore a set of axioms restricting the functional forms of the identification and the alienation functions, leading to their measure of polarization which is described below.

The ER characterization needs the distribution to be previously pre-arranged in groups, so that identification for any individual with his group simply depends on the relative size of the group. There is no reason to expect that these arbitrarily given groups will capture in an accurate way the feelings of identification and alienation in that distribution. Otherwise, if $F$ is ungrouped then we need some reasonable criterion to define the groups to implement the ER approach. In both cases, ungrouped and pre-grouped $F$, the specific way identification is defined raises the question of the substantial loss of information on the intragroup distribution of income.

The approach proposed in ERG is an extension of ER which preserves the same notion of polarization as in (1) but makes it more operational regardless of how data are organized, and
incorporates in any case the information on the intragroup distribution of income.

Let the collection of numbers $D=(z_0, z_1, ..., z_k; y_1, ..., y_k; p_1, ..., p_k)$ be any non-intersecting partition of $F$ with $a=z_0<...<z_k=b$ and:

\[
(2) \\
p_i = \int_{z_{i-1}}^{z_i} f(y)dy \\
y_i = \frac{1}{p_i} \int_{z_{i-1}}^{z_i} f(y)ydy
\]

indicating respectively population shares of the groups and conditional means. This partition defines groups as intervals of incomes $[z_{i-1}, z_i]$ for $i=1, ..., k$.

From a statistical point of view, the $k$-spike distribution $D$ is a representation of $F$ which induces an error of approximation $(F,D)$, standing for the lack of identification within the groups as there exists some internal dispersion. The simplest way to incorporate the lack of identification into the level of polarization given by $F$ is by defining a general index which subtracts the error from the polarization in the representation $D$, as measured by index $ER(\hat{\alpha}, D)$ in $ER$:

\[
(3) \\
P(F; \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = ER(\hat{\alpha}, \hat{\beta}) - \hat{\alpha} \hat{\beta}(F; \hat{\gamma})
\]

where $\hat{\alpha} > 0$ indicates the weight assigned to the error in the representation. If $\hat{\alpha} = 0$, expression (3) just represents the measure $ER$, expressed as:

\[
(4) \\
ER(\hat{\alpha}, \hat{\beta}) = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{1}{p_i} p_j |y_i - y_j|
\]

where $\hat{\alpha}$ represents the sensitivity to polarization, and belongs to the interval $[1, 1.6]$.

The implementation of (3) involves two different choices: the number of groups and their locations.

One might infer from the inspection of the distribution which is the reasonable representation, or the nature of the attribute we deal with might involve such a partition. But where there is no such information $ERG$ propose a method which, treating $k$ as an exogenous number, provides the partition which minimizes aggregate error. For instance, we can be concerned with the extent to which the distribution displays a bottom, a poor and a middle group in order to study what happened with the middle class. Given $k$, we endogenously determine their locations. The optimal partition satisfies:

\[
(5) \\
\hat{\beta}(F; \hat{\gamma}) = \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{z_{i-1}}^{z_i} \int_{z_{j-1}}^{z_j} \left|x-y\right|dF(x)dF(y)
\]

This definition considers the error as the level of inequality within groups measured by the Gini coefficient $G$; that is, the inequality in the original distribution $F$ minus the inequality in the representation $D$: 
This approach has already been employed by Aghevli and Mehran (1981) and Davies and Shorrocks (1989) to obtain optimal partitions of grouped data minimizing the loss of information.

The representation $D^*$ which solves this problem is characterized by the condition:

That is, the cut-off between two adjacent groups is the combined average income of both groups.

Expressing (3) accordingly with this optimal collection $D^*=(z_0^*, z_1^*, ..., z_k^*; y_1^*, ..., y_k^*; p_1^*, ..., p_k^*)$ leads, finally, to the ERG measure of polarization:

This specific way to obtain $D^*$ can be represented in the Lorenz framework as figure 1 shows. Let $L_F$ and $L_D$ denote the Lorenz curves respectively for the original distribution $F$ and for its representation $D$. The latter consists of $k$ linear segments. The value of $(F; D')$ is twice the area between both curves, so that minimizing $\phi$ is equivalent to minimize this area.
Thus far, we have assumed the number of groups as given. If we have no information about the \( k \) which corresponds to \( F \), we can choose between the analysis of polarization for different possible \( k \)s or set some fix value. In the later case, one reasonable procedure is choosing \( k \) so as to maximize polarization.

3. The special case of two groups: bipolarization

One case that can be considered of special interest is to ask the extent to which the distribution is becoming more bimodal. This question restricts the notion of polarization imposing the existence of exactly two poles, with the advantage of an easier interpretation and the cost depending on the closeness of the distribution to its bimodal representation. This case will be referred to as bipolarization and can be addressed using the previous approach fixing \( k=2 \).

According to the statistical approach in ERG, we fit a degenerate bimodal distribution to the original \( F \), revealing two distinct elements: the severity of the cleavage between both groups and the proximity of the original distribution to this simplified representation. In this case, there is one unique cut-off income \( z/z_1 \), \( y_1 \) and \( y_2 \) are the respective conditional mean incomes for both groups, and \( p/F(z) \) is the relative size of population in the poor group. Let \( L(p) \) indicate the value of the Lorenz curve at the population share \( p \). Then, noting that \( y_1=L(p)/p \) and \( y_2=[1-L(p)]/(1-p) \), the measure \( ER \) is, in this special case:

\[
ER(\hat{a},\hat{\alpha}) = \frac{[p^{\hat{a}}(1-p)+(1-p)^{\hat{a}}p][y_2-y_1]}{[p^{\hat{a}}+(1-p)^{\hat{a}}][p-L(p)]},
\]

where the partition is not necessarily the optimal one. Finally, the measure of polarization is given by:

\[
P(F;\hat{a},\hat{\alpha},z) = [p^{\hat{a}}+(1-p)^{\hat{a}}][p-L(p)] - \hat{\alpha}G(F)-(p-L(p))
\]

In the simplest case of \( \hat{a}=1 \), we obtain:

\[
P(F;\hat{a}=1,\hat{\alpha},z) = (1+\hat{\alpha})[p-L(p)] - \hat{\alpha}G(F)
\]

Additionally, let us consider the case where we want to know the degree to which the distribution is divided into two equally sized, homogenous groups. Then, the reasonable partition would consist of setting the unique cut-off income to be the median, \( me \) and then \( p=\frac{1}{2} \). Furthermore, specializing on the case of equal weights \( \hat{\alpha}=1 \), we obtain the measure \( W \) proposed by Wolfson (1994) as a particular case of the ERG approach:

\[
P(F;\hat{a}=1,\hat{\alpha}=1, z=me) = 2[p-L(\frac{1}{2})] - G(F) = \frac{me}{2}
\]

Instead, if we now wish to chose the optimal cut-off point, rather than an arbitrary one, then accordingly with (8) \( z=\mu=1 \). In this case \( p/\mu/F(\mu) \) represents the size of the relatively poor group and the measure becomes a function of two well-known indices of inequality: the Gini coefficient and the Relative Mean Deviation, \( D(F)=[p_F-L(p_F)] \). Indeed:
Concentrating again on the case where $\hat{a}=1$, we get:

$$P(F; \hat{a}, \hat{a}, z^* = \mu) = [p_n^\hat{a} + (1-p_n)^\hat{a}]D(F) - \hat{a}[G(F) - D(F)]$$

Finally, and in contrast with (13), when $\hat{a}=1$ we obtain:

$$P(F; \hat{a} = \hat{a} = z = 1) = 2D(F) - G(F)$$

4. The data and methodology

This study uses data from the Luxembourg Income Study (LIS) database because it provides the best available information for international comparisons. In order to facilitate international comparisons, this data set uses the "LIS Disposable Income" as the standardized unit of account. It includes yearly earnings, cash property incomes, social insurance and social transfers, net of income tax and mandatory contributions. The number of years available varies across countries, always within the period between 1967 and 1995. The data set covers twenty-five countries in total: most of the Eastern and Western European countries and the US, Canada, Australia, Israel and Taiwan.

For a more detailed analysis of the US case we use cross-sectional data from the Panel Study of Income Dynamics (PSID) of the University of Michigan. Our variable here is also some definition of disposable income: the sum of all household members taxable income and transfers, net of federal taxes, for each year of the period 1968-1991.

In both cases, LIS and PSID data, the unit of analysis is the household. To deal with heterogeneity in household size and composition we use the standard OECD equivalence scales. Hence our variable will be *Household Equivalent Disposable Income* at current prices expressed in respective national currency.

We have organized the empirical study so that we first focus on the US case in detail, then we extend the same analysis to the UK, Canada, Germany and Sweden and finally to other developed and transitional economies. We try to answer three main questions: (a) how were the trends in polarization and inequality in each country?, (b) was increasing polarization due to higher polarization between the groups or increasing homogeneity within them? and (c) was increasing polarization the

---


6 Some data sets were not yet available for different reasons: Luxembourg 1991, Italy 1995, and Austria 1989. In the US case, the data set corresponding to 1969 was not included because it does not report "LIS Disposable Income". In the case of France there are two different data sets: an income survey for 1979 and 1984 and an expenditure survey for 1984 and 1989. In the case of Germany, information for 1973, 1978 and 1983 comes from an income and expenditure survey, for 1981 from a transfer income survey and for 1984 and 1989 from a cross section of the Socioeconomic Panel Study.

7 The fact that this source is a panel allows for future research about household mobility across groups in the distribution.

8 They assign a weight 1 to the first adult, 0.7 to each additional adult, and 0.5 to each child.
result of an increasing gap between the extreme groups or the result of extreme groups becoming larger?

For the first question we compute the polarization index $P(F;\hat{\alpha},\hat{\alpha}=1,\hat{n}^*)$ in (9) defined for the optimal partition with two and three groups, where we compute the polarization index $ER$ for the representation, expressed in terms of the natural logarithm of incomes:

\begin{equation}
ER(\hat{\alpha};\hat{n}) = \sum_{i=1}^{k} \sum_{j=1}^{k} p_i^{\hat{\alpha}} p_j^{\hat{\alpha}} |\ln(y_i) - \ln(y_j)|
\end{equation}

To address the second question, we analyze separately both terms in (9) while for the last question we inspect the optimal partition $\hat{n}^*$.

5. Polarization in the US

The US is the case which has attracted most attention because it has experienced a dramatic increase in inequality since the end of the seventies, as it has been stressed by many analysts. In order to investigate whether this trend has also created higher polarization, we compute the index of polarization $P(F;\hat{\alpha},\hat{\alpha}=1,\hat{n}^*)$ between 1968 and 1991 together with the Gini index of inequality using PSID data, reported in table 1. Figure 2 shows the evolution of polarization in the case of intermediate $\hat{\alpha}$ and three groups, together with inequality, where both indices have been normalized to the base year 1968.

The results reveal that polarization follows a U-shaped trend regardless of the number of groups and the weight we put on the error term. For the first half of the 1970s, a similar decreasing tendency is observed for inequality and polarization, which seems to level off for the second half of that decade, with the exception of 1976-77 where polarization increases despite inequality declines. For the 1980s, both polarization and inequality increase, showing the highest rates for the first half of the decade. However, polarization increased more persistently than inequality in this decade. Polarization increases despite the declines in inequality in several periods: 1981-82, 1983-84 and 1986-87, for all sensitivity to polarization $\hat{\alpha}$ and for two and three groups. In the case of three groups the same occurs also between 1989 and 1990.

The results also suggest that the distributional change is better described by the characterization of the distribution by three groups, as polarization for $\hat{\alpha}=1$ was maximum in this case regardless of $\hat{\alpha}$, and it increased more persistently than in the case of two groups. Bipolarization declines for

---

9 The reason for using logs is that, in our view, in the two-spike distribution, for a given distance $d=y_2-y_1$, polarization should be maximum when both groups have the same size, $p=1-p=1/2$. This is satisfied by log-incomes but not by the ratio of incomes to the global mean.

10 We use the Gini index in order to be consistent with the way we measured the dispersion when choosing the optimal partition.
1978-79, 1980-81, 1984-85 and 1989-90 despite tri-polarization goes up. We obtain similar results when we use instead the LIS data set between 1974 and 1994. These results are reported in table 2. The advantage of the LIS data set is that it furnishes standardized information for a substantial number of developed countries allowing for a more meaningful international comparison. For the USA, polarization strongly increases for the first half of both 1980s and 1990s, remaining stable for the rest of periods. Between 1974 and 1979, we find increasing bipolarization while inequality substantially declines and polarization slightly decreases in the case of three groups. The value of the polarization index is maximized for the case of three groups regardless of the value of the parameter $\alpha$.

### Table 1. Polarization and Inequality in the US, 1968-1991, $\beta=1$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha=1.0$</td>
<td>2</td>
<td>0.246</td>
<td>0.249</td>
<td>0.238</td>
<td>0.242</td>
<td>0.237</td>
<td>0.237</td>
<td>0.232</td>
<td>0.236</td>
<td>0.235</td>
<td>0.237</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.211</td>
<td>0.214</td>
<td>0.203</td>
<td>0.207</td>
<td>0.204</td>
<td>0.203</td>
<td>0.199</td>
<td>0.201</td>
<td>0.201</td>
<td>0.203</td>
<td>0.199</td>
</tr>
<tr>
<td>$\alpha=1.3$</td>
<td>2</td>
<td>0.201</td>
<td>0.205</td>
<td>0.195</td>
<td>0.198</td>
<td>0.195</td>
<td>0.194</td>
<td>0.190</td>
<td>0.193</td>
<td>0.193</td>
<td>0.194</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.155</td>
<td>0.157</td>
<td>0.148</td>
<td>0.152</td>
<td>0.149</td>
<td>0.148</td>
<td>0.145</td>
<td>0.147</td>
<td>0.147</td>
<td>0.148</td>
<td>0.145</td>
</tr>
<tr>
<td>$\alpha=1.6$</td>
<td>2</td>
<td>0.166</td>
<td>0.168</td>
<td>0.160</td>
<td>0.163</td>
<td>0.160</td>
<td>0.159</td>
<td>0.156</td>
<td>0.158</td>
<td>0.159</td>
<td>0.159</td>
<td>0.155</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.114</td>
<td>0.116</td>
<td>0.109</td>
<td>0.112</td>
<td>0.110</td>
<td>0.109</td>
<td>0.106</td>
<td>0.108</td>
<td>0.108</td>
<td>0.108</td>
<td>0.106</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha=1.0$</td>
<td>2</td>
<td>0.143</td>
<td>0.147</td>
<td>0.141</td>
<td>0.143</td>
<td>0.139</td>
<td>0.137</td>
<td>0.135</td>
<td>0.139</td>
<td>0.138</td>
<td>0.141</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.163</td>
<td>0.167</td>
<td>0.156</td>
<td>0.161</td>
<td>0.158</td>
<td>0.157</td>
<td>0.154</td>
<td>0.156</td>
<td>0.156</td>
<td>0.158</td>
<td>0.155</td>
</tr>
<tr>
<td>$\alpha=1.3$</td>
<td>2</td>
<td>0.099</td>
<td>0.102</td>
<td>0.098</td>
<td>0.099</td>
<td>0.096</td>
<td>0.094</td>
<td>0.093</td>
<td>0.096</td>
<td>0.096</td>
<td>0.097</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.107</td>
<td>0.109</td>
<td>0.102</td>
<td>0.106</td>
<td>0.103</td>
<td>0.103</td>
<td>0.100</td>
<td>0.101</td>
<td>0.101</td>
<td>0.103</td>
<td>0.101</td>
</tr>
<tr>
<td>$\alpha=1.6$</td>
<td>2</td>
<td>0.063</td>
<td>0.066</td>
<td>0.063</td>
<td>0.063</td>
<td>0.062</td>
<td>0.059</td>
<td>0.059</td>
<td>0.061</td>
<td>0.061</td>
<td>0.063</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.066</td>
<td>0.068</td>
<td>0.063</td>
<td>0.066</td>
<td>0.063</td>
<td>0.066</td>
<td>0.061</td>
<td>0.062</td>
<td>0.062</td>
<td>0.063</td>
<td>0.062</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gini</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha=1.0$</td>
<td>2</td>
<td>0.354</td>
<td>0.359</td>
<td>0.339</td>
<td>0.345</td>
<td>0.341</td>
<td>0.340</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
<td>0.337</td>
<td>0.337</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.102</td>
<td>0.103</td>
<td>0.097</td>
<td>0.099</td>
<td>0.098</td>
<td>0.100</td>
<td>0.097</td>
<td>0.097</td>
<td>0.097</td>
<td>0.096</td>
<td>0.096</td>
</tr>
<tr>
<td>$\alpha=1.3$</td>
<td>2</td>
<td>0.048</td>
<td>0.048</td>
<td>0.046</td>
<td>0.046</td>
<td>0.046</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.111</td>
<td>0.118</td>
<td>0.118</td>
<td>0.121</td>
<td>0.121</td>
<td>0.128</td>
<td>0.127</td>
<td>0.127</td>
<td>0.128</td>
<td>0.131</td>
<td>0.132</td>
</tr>
</tbody>
</table>

### 6. Homogeneity within groups or Heterogeneity between groups?

Polarization in a country may increase either because the population has
inequality has also increased.

The trend in polarization between groups is shown to exhibit a less erratic and more pronounced U shape. It has been strongly increasing since 1978, though it seems to decelerate at the end. From the figure it can be appreciated that it better fits the trend in $P$. The latter is in general the result of increasing polarization between the groups, partially offset by increasing lack of identification since 1978.

Indeed, increasing polarization in the US almost always corresponds to periods of increasing polarization between the groups, but there alternate periods in which these groups become more identified, with periods of decreasing identification. In the reference case with the PSID data the former occurs for 1976-77, 1983-84, 1989-90, while the latter is true for 1974-75, 1978-79, 1980-81, 1982-83 and 1984-85. Identification remains stable for the rest of the periods. Note that only in the first case can we assert that polarization has increased regardless of the weight $\hat{\alpha}$.

The lack of identification remained steady until 1978, showing a clear increasing trend, though erratic, ever since. It represents on average 29.1% and 13.7% of global Gini for two and three groups respectively. These shares correspond to the loss of information when we use the simplified representation rather than the original. As figure 4 presents, the evolution through time of these shares is also very erratic since 1978, with only a slightly increasing trend because global inequality has also increased.

been clustering around local poles or because these poles become more outstanding and with a larger distance between them. Both elements may be reinforcing each other. The analysis of the decomposition of the index $P(F;\hat{\alpha},\hat{n}=1,\bar{n}^*)$ into polarization between the groups $ER(\hat{\alpha},\bar{n}^*)$ and the error or lack of group identification $(F;\bar{n}^*)$ will allow us to identify the source of polarization. Furthermore, higher identification leads to increasing inequality but decreasing polarization, so this term explains many of the discrepancies between these two different ways to approach the distribution of income. Figure 3 shows this decomposition for the case of $\hat{\alpha}=1$, intermediate $\hat{\alpha}$ and three groups.

The lack of identification remained steady until 1978, showing a clear increasing trend, though erratic, ever since. It represents on average 29.1% and 13.7% of global Gini for two and three groups respectively. These shares correspond to the loss of information when we use the simplified representation rather than the original. As figure 4 presents, the evolution through time of these shares is also very erratic since 1978, with only a slightly increasing trend because global inequality has also increased.

The trend in polarization between groups is shown to exhibit a less erratic and more pronounced U shape. It has been strongly increasing since 1978, though it seems to decelerate at the end. From the figure it can be appreciated that it better fits the trend in $P$. The latter is in general the result of increasing polarization between the groups, partially offset by increasing lack of identification since 1978.

Indeed, increasing polarization in the US almost always corresponds to periods of increasing polarization between the groups, but there alternate periods in which these groups become more identified, with periods of decreasing identification. In the reference case with the PSID data the former occurs for 1976-77, 1983-84, 1989-90, while the latter is true for 1974-75, 1978-79, 1980-81, 1982-83 and 1984-85. Identification remains stable for the rest of the periods. Note that only in the first case can we assert that polarization has increased regardless of the weight $\hat{\alpha}$.
In the case of LIS data, we find higher polarization for the 1979-86 and 1991-94 periods because of increasing polarization between the groups despite lower identification. However increasing bipolarization during 1974-79 was due to groups becoming far more identified despite the decline in polarization between them. In the case of three groups this increasing identification was not enough to offset the decline in the between-groups term.

7. Disappearing middle class or increasing cleavage in the US?

It has been a standard practice in economic analysis to simplify distributions by constructing a small number of adjacent intervals, such as the use of different quantiles, to study distributions of income. Of special interest was the quantile/share approach to study the size and economic position of the middle class in salary distribution during the 1980s in the US. To this effect,
The problem of the "disappearance of the middle class" cannot be addressed only by measuring whether there have been population shifts from the middle to the extremes of the distribution. It also involves determining whether the extreme groups have moved far from each other. For this effect, we need endogenously to determine the location of groups. Thus, a proper analysis must combine the relative size of the groups with the distance between their "gravity centers".

Figure 4: The Gini ratio as a share of global Gini, US, 1968-91

Source: Own construction using PSID data

It was also customary, for instance, Rosenzweig (1985) or McPherson and Schuessler (1986), to divide all categories of workers into three groups, each with the same number of occupations. From a conceptual point of view, speaking of lower, middle and top groups implies that we are implicitly simplifying the distribution. It seems reasonable to minimize the induced loss of information. That is precisely what we do in our approach minimizing the intra-group dispersion.
Focusing on the most representative case of three groups, figure 5a shows the trends for the group sizes and figure 5b presents the trends for the relative income means together with the corresponding cut-off incomes.

It can be observed that the relative size of the middle group has diminished progressively between 1974 and 1991, but not by a significant amount, from 39.1% of all households to 37.9%. We observe a continuous increase in the size of the bottom group, from 38% of households in 1974 to 41% by 1994. Indeed there was a shift of population towards the bottom group: from the top and middle groups between 1979 and 1986, and only from the top of the distribution between 1991 and 1994.

More substantial changes are observed in the increasing distance between the extreme poles since 1979. We find an increase in the top group average income (from 188% to 214% of the overall mean) together with the decline in the mean of the bottom group (from 45% to 41% of the overall average). However, the income of the middle group has remained steady around the global mean.

Our approach endogenously determines the cut-off incomes $z_1$ and $z_2$, that play the role of a poverty and a wealth line respectively. This partially explains the limited change in the relative size of the middle group. The interval of incomes defining the middle group has extended since 1979, mostly in the top cut-off. While in that year the middle consisted of those households whose incomes fell in the interval between 73% and 134% of the overall mean, by 1994 the interval was defined by 70%-141%.

We think that an appropriate analysis must endogenously determine the location of groups in the simplified representation, and then the cut-offs should vary
through time. Despite this, it still seems reasonable to ask what would have happened had the cut-offs remained fix since 1974. This exercise is done in figure 6 which shows the relative sizes of groups. A substantial share of households moved away from the middle towards the bottom group since 1979. The reduction in the size of the middle group was equivalent to 5% of all households (a 13% of its group members).

Inequality in the middle group has increased substantially since 1979. This has lead their limits to expand in order to compensate for that increase and obtain a better defined group -with lower inequality.

In conclusion, the increasing separation between extreme groups, rather than the declining size of the endogenous middle group, explains the increasing polarization found between the groups in the USA since 1979.

---

12 Using the optimal partition to compute polarization, rather than the one with fix cut-offs, induces no clear effect on the level of polarization. Among all LIS countries only Canada and particularly Sweden would increase polarization by more than 1% if the latter partition were employed. The UK and Germany would display more polarization, and the US, less, but all of them by small percentages.

13 The middle was the group with the largest increase in inequality. Inequality, as measured by the Gini index, was 0.143 in 1979 in this group. By 1994 it was 0.197 and 0.188 for the cases of constant and endogenous cut-off incomes respectively.
8. Polarization trends in other developed countries: UK, Sweden, Canada and Germany

It is known that increasing inequality was a problem not only in the US, but other countries have also shared the same tendency with different intensity and at different times. This raises the question of whether they have also shared the polarizing trends found in the US.

The results for UK, Sweden, Canada and Germany are reported in table 3 for polarization and table 4 for inequality. Figure 7 illustrates the case of maximum polarization including the US. Polarization and inequality trends in these countries are plotted in figures 8a to 9b together with the decomposition of polarization into a between-groups and an intragroup term.

Polarization has been increasing since 1974 in the UK with the strongest intensity found during the second half of 1980s according to figure 8a. During the second half of 1980s we find a large increase in polarization bet-
Polarization starts to increase in 1974, while inequality begins in 1979. The discrepancy occurs because during the 1974-1979 period the higher intragroup identification more than compensated for smaller polarization between the groups.

The case of Sweden, shown in figure 8b, is similar to the UK, but with the major increase in polarization during the first half of the 1980s, while for the last period 1987-92 polarization increase was very small with three groups and changed to a decline in the case of two groups. The strongest increase in polarization was due to a large increase the $ER$ term, despite the lower identification.

In Germany polarization was declining but progressively stabilizing because both $ER$ and $\epsilon$ were decreasing at a decelerating rate as figure 9a shows. For the last period a slight increase in polarization between groups can be observed which did not lead to increasing polarization because the groups became also less identified. In contrast, inequality increased in this period.

In Canada, a long depolarizing trend is also found, due especially to decreasing polarization between the groups at a decelerating rate, with identification remaining almost constant. This is illustrated in figure 9b.

In conclusion, for this selected group of countries it is evident that Sweden and the UK share a similar pattern to the US, while Canada and Germany display continuous, though decelerating, declines in polarization.
9. The groups in the UK, Sweden, Canada and Germany

As in the US case, the inspection of what happened to the size and average income of the bottom, middle and top groups using the optimal partitions helps to understand how polarization between groups increased in UK and Sweden, and decreased in Canada and Germany during the 1980s. Figure 10a shows the evolution of the respective middle group sizes and figure 10b the trends for the ratio of extreme groups average incomes (the top to the bottom).

In the case of UK we can appreciate why polarization increased so much during the second half of 1980s. The reason is that during that period the ratio of extreme incomes increased from 3.5 to 4.3, at the same time that the size of the middle group diminished by 1% of all households. During that period population shifted from the middle and top towards the bottom, which increased from 41.6% to 43.8% of households. During the first half, however, polarization increase was less signifi-
distribution because they focused on absolute income, while our focus is relative income.

The distribution in Sweden exhibits a different pattern from the US and the UK, because relative sizes of groups have been quite stable during the episode of high increase in polarization (1981-87), with the middle group accounting for almost 40% of households. Between 1987 and 1992 a transfer of households from the top and middle towards the bottom is observed. Polarization is mostly explained by the ascending ratio of extreme incomes that for the 1981-87 period rose from 3.2 to 3.6. For constant cut-off incomes, the middle has strongly declined in size since 1981 from 65.8% to 59.1% of households.

Both Canada and Germany display decreasing polarization between the groups due to decreasing distances between extreme groups, but with several oscillations in their sizes. This approximation between extreme poles occurs with more intensity in Canada, where the income ratio diminished con-
tinuously from 4.9 in 1971 to 3.8 by 1994. These two countries exhibit increasing middle-group household shares for constant relative cut-off incomes. For Canada, from 38% to 46.5% between 1971 and 1991, with a slight decrease ever since. In the case of Germany, from 39.2% by 1967 to 44.1% in 1989.

Among this set of four countries, changes in the distance between extreme groups explain most changes observed in polarization either increasing or decreasing. Additionally in the case of the UK, these changes have been reinforced by a shrinking middle group since 1986. In Sweden, the size of the groups has been quite stable, while in Germany and Canada they exhibit different oscillations. As in the US case, changes in the size of the middle groups are more outstanding if these groups are treated as exogenous.

10. Polarization in other developed countries: the rest of the LIS countries
In other developed countries, we can observe different tendencies. Table 5 shows the results for inequality and polarization in the case of intermediate sensitivity to polarization and with two and three groups. Regardless of the number of groups, Australia during 1981-89, Finland (1987-91), Hungary (1991-95), Poland (1986-92) and Russia (1992-95) are the countries with polarizing income distributions, while Denmark (1987-92), Israel (1979-1992), Italy (1986-91), Luxembourg (1985-1991), Norway (1979-91) and Spain (1980-90) show a depolarizing tendency. France (1979-81) displays declining polarization for three groups -which is maximum- and Belgium a constant one, while bipolarization is increasing in both cases. Netherlands increases its level of polarization for 1983-87 but decreases it for 1987-91.

Increasing polarization is always related to increasing polarization between groups and increasing lack of identification, with the exceptions of France (1979-81) and Belgium (1988-92) for two groups, where both terms decrease, but the latter with more intensity.

For depolarization, in most cases we face decreasing polarization between groups and decreasing lack of identification, but with several exceptions. It is the case of France (1981-84), Israel (1986-92) and Belgium (1985-88 only for 3 groups) for which both terms increase, but the greater intensity in the lack of identification causes a decline in polarization despite higher inequality. Other exceptions are Norway (1986-91) and Taiwan (1981-86 for 2 groups) for which the polarization declines regardless of the weight \( \tilde{a} \), due to decreasing polarization between groups becoming less identified. This leads to a decline in polarization which is more intense than the decline in inequality in Norway, and to increasing inequality in Taiwan.

For a better understanding of trends in polarization between groups, we report the optimal partition in table 6. We can observe significant stability of the size of the middle groups within countries, with the largest increases being Norway (1979-91) from 37.3% to 39.5% of households, France (1981-84) from 37.7% to 40.1% and Luxembourg (1987-91), 37.8% to 39.8% and the largest declines, Belgium (1988-92), from 40.5% to 38.9% of households, and Netherlands (1987-91) from 40.1% to 37.4%. In the case of fix cut-offs, not reported here, the variability would be larger, but without the intensity of the US or the UK.\(^{15}\)

In general, polarization trends are determined by the distance between extreme incomes.

\(^{14}\) For the Spanish case see Gradín (1998) who analyzes the period 1973-80 and 1980-90 in more details, directly using data from Spanish official statistics. For the Italian case, see D’Ambrosio (1998) who finds increasing polarization for the whole period 1986-95, with different methodology.

\(^{15}\) The largest increase is found in Norway from 37.3% of households in the middle group in 1979 to 42.8% by 1991, and the largest decline in Hungary from 40.4% in 1991 to 35.9% by 1995.
Table 5. Polarization and Inequality for other LIS countries, [α=1.3]

<table>
<thead>
<tr>
<th>Country year</th>
<th>2 groups</th>
<th></th>
<th></th>
<th></th>
<th>3 groups</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Gini</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ER</td>
<td>ε</td>
<td>P</td>
<td></td>
<td></td>
<td>ER</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>0.129</td>
<td>0.083</td>
<td>0.006</td>
<td>0.033</td>
<td>0.040</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.182</td>
<td>0.085</td>
<td>0.008</td>
<td>0.136</td>
<td>0.040</td>
<td>0.016</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>0.119</td>
<td>0.089</td>
<td>0.009</td>
<td>0.142</td>
<td>0.043</td>
<td>0.008</td>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>0.126</td>
<td>0.056</td>
<td>0.006</td>
<td>0.050</td>
<td>0.031</td>
<td>0.062</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.153</td>
<td>0.082</td>
<td>0.002</td>
<td>0.091</td>
<td>0.034</td>
<td>0.006</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0.137</td>
<td>0.075</td>
<td>0.002</td>
<td>0.101</td>
<td>0.036</td>
<td>0.005</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.134</td>
<td>0.081</td>
<td>0.003</td>
<td>0.099</td>
<td>0.031</td>
<td>0.005</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>0.148</td>
<td>0.019</td>
<td>0.019</td>
<td>0.078</td>
<td>0.028</td>
<td>0.010</td>
<td>0.106</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.153</td>
<td>0.067</td>
<td>0.006</td>
<td>0.137</td>
<td>0.041</td>
<td>0.006</td>
<td>0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>0.129</td>
<td>0.070</td>
<td>0.001</td>
<td>0.104</td>
<td>0.029</td>
<td>0.006</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>0.134</td>
<td>0.073</td>
<td>0.001</td>
<td>0.130</td>
<td>0.034</td>
<td>0.006</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.165</td>
<td>0.075</td>
<td>0.003</td>
<td>0.103</td>
<td>0.035</td>
<td>0.008</td>
<td>0.118</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>0.171</td>
<td>0.093</td>
<td>0.001</td>
<td>0.136</td>
<td>0.044</td>
<td>0.006</td>
<td>0.030</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>0.108</td>
<td>0.086</td>
<td>0.002</td>
<td>0.126</td>
<td>0.031</td>
<td>0.005</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income survey</td>
<td>0.169</td>
<td>0.094</td>
<td>0.075</td>
<td>0.128</td>
<td>0.045</td>
<td>0.003</td>
<td>0.068</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>0.184</td>
<td>0.103</td>
<td>0.083</td>
<td>0.142</td>
<td>0.049</td>
<td>0.003</td>
<td>0.036</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>0.173</td>
<td>0.096</td>
<td>0.007</td>
<td>0.132</td>
<td>0.046</td>
<td>0.007</td>
<td>0.062</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>0.157</td>
<td>0.088</td>
<td>0.008</td>
<td>0.118</td>
<td>0.041</td>
<td>0.007</td>
<td>0.180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.167</td>
<td>0.092</td>
<td>0.004</td>
<td>0.126</td>
<td>0.041</td>
<td>0.008</td>
<td>0.314</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>0.191</td>
<td>0.100</td>
<td>0.001</td>
<td>0.145</td>
<td>0.049</td>
<td>0.007</td>
<td>0.316</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>0.104</td>
<td>0.055</td>
<td>0.009</td>
<td>0.147</td>
<td>0.041</td>
<td>0.163</td>
<td>0.358</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td>0.110</td>
<td>0.019</td>
<td>0.001</td>
<td>0.141</td>
<td>0.041</td>
<td>0.101</td>
<td>0.332</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>0.185</td>
<td>0.087</td>
<td>0.006</td>
<td>0.145</td>
<td>0.044</td>
<td>0.006</td>
<td>0.320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>0.173</td>
<td>0.093</td>
<td>0.004</td>
<td>0.133</td>
<td>0.042</td>
<td>0.009</td>
<td>0.207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>0.130</td>
<td>0.088</td>
<td>0.002</td>
<td>0.127</td>
<td>0.041</td>
<td>0.008</td>
<td>0.300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0.139</td>
<td>0.072</td>
<td>0.008</td>
<td>0.103</td>
<td>0.034</td>
<td>0.006</td>
<td>0.217</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>0.137</td>
<td>0.073</td>
<td>0.005</td>
<td>0.102</td>
<td>0.033</td>
<td>0.009</td>
<td>0.247</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.166</td>
<td>0.087</td>
<td>0.007</td>
<td>0.123</td>
<td>0.041</td>
<td>0.009</td>
<td>0.295</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>0.160</td>
<td>0.082</td>
<td>0.007</td>
<td>0.116</td>
<td>0.041</td>
<td>0.007</td>
<td>0.282</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>0.145</td>
<td>0.073</td>
<td>0.002</td>
<td>0.130</td>
<td>0.039</td>
<td>0.007</td>
<td>0.233</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>0.158</td>
<td>0.072</td>
<td>0.006</td>
<td>0.102</td>
<td>0.031</td>
<td>0.006</td>
<td>0.193</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>0.136</td>
<td>0.073</td>
<td>0.002</td>
<td>0.101</td>
<td>0.033</td>
<td>0.006</td>
<td>0.212</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>0.149</td>
<td>0.076</td>
<td>0.004</td>
<td>0.131</td>
<td>0.035</td>
<td>0.005</td>
<td>0.262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>0.164</td>
<td>0.047</td>
<td>0.007</td>
<td>0.233</td>
<td>0.044</td>
<td>0.203</td>
<td>0.261</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>0.123</td>
<td>0.133</td>
<td>0.109</td>
<td>0.098</td>
<td>0.004</td>
<td>0.136</td>
<td>0.417</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.252</td>
<td>0.155</td>
<td>0.117</td>
<td>0.206</td>
<td>0.064</td>
<td>0.143</td>
<td>0.470</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.252</td>
<td>0.135</td>
<td>0.117</td>
<td>0.206</td>
<td>0.064</td>
<td>0.143</td>
<td>0.470</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.009</td>
<td>0.000</td>
<td>0.000</td>
<td>0.073</td>
<td>0.082</td>
<td>0.006</td>
<td>0.180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.164</td>
<td>0.090</td>
<td>0.005</td>
<td>0.126</td>
<td>0.044</td>
<td>0.009</td>
<td>0.336</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.175</td>
<td>0.053</td>
<td>0.003</td>
<td>0.132</td>
<td>0.044</td>
<td>0.009</td>
<td>0.377</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.184</td>
<td>0.107</td>
<td>0.076</td>
<td>0.144</td>
<td>0.051</td>
<td>0.004</td>
<td>0.311</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source own construction using LIS Dataset. The cut-off for 2 groups is always ε. Income expressed as the ratio to the global mean.
Table 7. Correlation between polarization and ER and for all 69 LIS observations

<table>
<thead>
<tr>
<th></th>
<th>ER(Æw)</th>
<th>(F,Æ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>á=1  á=1.  á=1.</td>
<td></td>
</tr>
<tr>
<td>P(F,α;β,ρ)</td>
<td>0.98  0.965</td>
<td>0.87  0.853</td>
</tr>
<tr>
<td>2 groups</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>P(F,α;β,ρ')</td>
<td>0.99  0.995</td>
<td>0.92  0.919</td>
</tr>
<tr>
<td>3 groups</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: own construction using LIS Dataset

Table 8. Correlation between polarization and group components for all 69 LIS observations

<table>
<thead>
<tr>
<th>Income</th>
<th>households shares</th>
<th>cut-offs</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(F,α=1.3;β=1,ρ)</td>
<td>0.91  0.3</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: own construction using LIS Dataset

in LIS—all years and countries—we can give a global characterization of polarization which coincides in its basic features with that described for the US. For that we compute in table 8 the correlation between \( P(F;â;â',n') \), and each element in its decomposition: polarization between the optimal groups \( ER (n';â) \) and the lack of identification \( (F'n') \). Additionally, table 7 reports the coefficient of correlation between the polarization index \( P(F;â=1.3;â=1,n') \) and each of the elements of \( D' \), i.e. group population sizes and incomes, and cut-off incomes.

We can observe that \( P(F;â;â=1,n') \) is increasing with both \( ER (n';â) \) and \( (F'n') \), this correlation decreasing with sensitivity to polarization \( â \) and increasing with the number of groups. We find that the term between groups is more determinant than the term within groups.

According to the two-spike optimal representation \( n' \), a relatively bipolarized distribution is determined by a large distance between the rich and the poor groups, and a large poor group. Ac-
According to the three-spike optimal representation, polarization is determined by shifting households from the middle and top to the bottom of the distribution, but especially by increasing distance between mean incomes at the extreme groups, as well as widening out the income range defining the middle group. For both representations, of two and three groups, polarization is more correlated with distances between incomes than with households shares\textsuperscript{16}.

Despite the fact that it is well known that international comparisons with heterogeneous information based on trends in income distributions are far more reliable than those based on their levels, it is worthwhile to look at the rankings among countries on the basis of our results. The countries with the highest levels of polarization regardless of the specification we use are Russia, the US, the UK, Israel and Australia, while those with the lowest le-

\textsuperscript{16} The correlation between the polarization and the size of the middle would be -0.41, larger than the current -0.29, if we use the sizes obtained with fix cut-offs rather than variables.
vels are Slovakia, the Czech Republic, Austria, Belgium and Scandinavia. Figure 12 provides the ranking for maximum polarization with intermediate sensitivity to polarization $\alpha$ and weight $\alpha=1$ for the observations around 1990. Inequality levels are also reported for comparison. For instance, Australia exhibits more polarization than Ireland, Switzerland or Taiwan while it displays less inequality.

The ranking in figure 12 corresponds to $\alpha=1$. How far can we rank countries according to polarization without any explicit weighting for the identification term? In order to answer this, we need to consider both elements in polarization separately: polarization between group and identification. For a given number of groups and a given sensitivity to polarization, one country shows unambiguously more polarization than another when it exhibits groups that are both more polarized and internally more identified. Otherwise, the ranking will depend on $\alpha$. Figure 13 summarizes the results for one particular case for the middle 1980s: three groups and intermediate $\alpha$. It shows that we can rank several countries. For instance we can say that Australia displays more polarization than Canada and Italy, and that these two countries have greater polarization than the UK, which shows higher polarization than France.

Figure 13
Polarization ranking for 1984-87 regardless of $\beta$, 3 groups and $\alpha=1.3$
11. Conclusions

Polarization has increased since the end of the 1970s in a large number of countries. This has been the case of the US with more intensity over the first halves of the 1980s and the 1990s. Sweden and the UK, among others, have also experienced episodes of a strong increase in polarization during the first and second half respectively of the 1980s. In the UK and the US polarization increases have been more persistent than in the case of inequality.

Other group of countries have shown decreasing trends in polarization but with a tendency to level off this process. This has been the case of Canada and Germany, also Norway, Denmark and Spain.

The nature of the index of polarization used here allows us to go further and characterize which changes in the shape of distributions lie behind increasing polarization. The results emphasize that increases in polarization are mainly due to changes in polarization between the groups rather than changes in their internal identification.

These changes in distribution between the groups entail enlarging the gap between extreme groups, and increasing the population size at the bottom end of the distribution. The size and average income of the middle groups appear to display an important element of stability over time and across countries, though the size has diminished in the UK and the US.

The size of the middle did not shrink substantially in polarizing countries because the notion of who belonged to the middle was widening, basically due to higher dispersion within that group. If the middle is a fix interval in terms of relative incomes then polarization entails a substantial share of households moving away from the middle of the distribution.

If we assume that the groups underlying the distribution as an endogenous element, then we should say that what has led to higher polarization in the US and the UK has been the increasing income distance between the extreme poles rather than the shrinking middle class. For asserting the opposite we need to assume the middle group as exogenous.

Despite the fact that it is not the crucial element in explaining polarization trends, the lack of identification also plays a role because there is evidence of years for which, in the US, increasing identification within groups reinforces increasing polarization. In France, Belgium or Israel, the lack of identification is crucial in asserting whether polarization increases or not, while in Norway and Taiwan it reinforces decreasing polarization. Some countries can be ranked regardless of the weight of the identification term because they show, at the same time, more polarization and more identification than others.

It has been shown also that countries such as Russia, the US, the UK, Israel or Australia constitute the group of those with the highest levels of polarization while the lowest levels are found in Slovakia, the Czech Republic, Austria, Belgium and the Scandinavian countries.
References


