the review of income and wealth

Review of Income and Wealth
Series 61, Number 3, September 2015
DOI: 10.1111/roiw.12107

A CLASS OF SOCIAL WELFARE FUNCTIONS THAT DEPEND ON MEAN INCOME AND INCOME POLARIZATION

BY JUAN GABRIEL RODRÍGUEZ*

Universidad Complutense de Madrid

A well-established strategy for evaluating alternative income distributions is based on the use of an abbreviated social welfare function that depends only on mean income and an inequality index. In keeping with this literature, we study the existence of social welfare functions that can be written as a trade-off between efficiency and income polarization. This paper proposes a class of social welfare functions consistent with the Esteban and Ray, and Duclos, Esteban and Ray income polarization indices. For this result, we expand the domain for personal preferences to incorporate not only own income but also the well-being of others. In addition, we link our proposal to the literature on relative satisfaction. The approach is illustrated by an empirical application using the CPS database for the United States in the period 1991–2010.

JEL Codes: D30, D63

Keywords: income polarization, relative deprivation, relative satisfaction, social welfare, utility function

1. Introduction

The evaluation of alternative income distributions in situations where unanimous preference is not attainable has been the focus of a large body of research on distribution. For this task, abbreviated social welfare functions that depend only on mean income and an inequality index have been proposed in the literature (see, for example, the discussion in Kondor (1975) for inequality indices that are differentiable). In this view, social welfare is the result of a trade-off between efficiency and equity, and the ranking of income distributions with equal means by social welfare functions and by inequality indices should be consistent. For example, the Kolm–Atkinson class of inequality indices (Kolm, 1969; Atkinson, 1970) is related to a consistent class of social welfare functions that can be written as a trade-off between efficiency and equity. In contrast, the use of a social welfare function (W) based on the Gini coefficient runs into considerable difficulties (Newbery, 1970; Dasgupta *et al.*, 1973; Lambert, 1985). In fact, a convincing rationale for it is not possible if individuals care only about their own income. Fortunately, Lambert (1985) provides a positive result for the Gini coefficient by expanding the domain for personal preferences to incorporate envy or altruism.

Note: I am grateful for helpful comments from the editor, Conchita D'Ambrosio, two anonymous referees, J.-Y. Duclos, A. Spadaro, J. M. Esteban, and M. Paniagua. I also acknowledge the comments and suggestions by the audience at XVI Encuentro de Economía Pública (Granada). This paper has benefited from the support of the Fundación Ramón Areces and Spanish Ministry of Education and Science (Project ECO2010-17590). The usual disclaimer applies.

*Correspondence to: Juan Gabriel Rodríguez, Departamento de Análisis Económico I, Universidad Complutense de Madrid, 28223 Madrid, Spain (juangabr@pdi.ucm.es).

In this paper, we study this issue for income polarization. For decades, inequality has been the summary concept with which the distributional effects of changes in the economic environment have been evaluated. Nowadays, however, a complementary summary concept of polarization is also used (Foster and Wolfson, 1992, 2010; Esteban and Ray, 1994; Wolfson, 1994, 1997; Gradín, 2000; Wang and Tsui, 2000; Chakravarty and Majumder, 2001; D'Ambrosio, 2001; Zhang and Kanbur, 2001; Prieto *et al.*, 2003; Rodríguez and Salas, 2003; Duclos *et al.*, 2004; Duclos and Échevin, 2005; Lasso de la Vega and Urrutia, 2006; Chakravarty *et al.*, 2007; Bossert and Schworm, 2008; Chakravarty and Maharaj, 2009). In fact, it is argued that polarization is a more appropriate criterion for explaining social conflict (see, for example, Esteban and Ray, 1999; Reynal-Querol, 2002). Accordingly, one could ask: can the welfare of society be related to the degree of income polarization? Does a trade-off between efficiency and polarization exist? A social welfare function that is consistent with a measure of income polarization could be used as a tool for evaluating alternative income distributions. In particular, we could use this welfare function to discriminate between policies with different effects on mean income and income polarization. However, there might be no way of aggregating the opposing interests to obtain a social valuation of the alternative income distributions in a polarized society.

Let P be the income polarization measure of either Esteban and Ray (1994) (ER henceforth) or Duclos *et al.* (2004) (DER henceforth). In this paper, we propose a class of social welfare functions that accords with P. To obtain these welfare functions, we make use of meaningful utility functions that depend not only on own income but also on the general well-being of society (represented by mean income), as long as it does not bring about any increase in individual relative deprivation.[1] Influential research on social preferences has recently pointed out that material self-interest is not the sole motivation of people (Kahana, 2005; Sobel, 2005; Fehr and Schmidt, 2006; Schwarze and Winkelmann, 2011). Evidence gathered by psychologists and experimental economists indicates that a substantial percentage of the population is strongly motivated by other-regarding preferences and that concerns for the well-being of others cannot be ignored in social interactions. We justify our proposal on the grounds of this literature.

As a by-product result, we obtain the average of individual satisfaction (Yitzhaki, 1979; Hey and Lambert, 1980; Chakravarty and Mukherjee, 1999) as a particular case of the proposed approach. Thus, the welfare analysis presented here bridges the gap between polarization measures and the literature on relative satisfaction.

We illustrate our approach by applying it to household data drawn from the Current Population Survey (CPS) dataset for the United States in the 1991–2010 period. After-tax and transfer household income and different parameterizations of personal utility are used.

In the next section, we introduce the Esteban and Ray (1994) polarization index. Then we define some general restrictions that the social welfare function should satisfy and present our main result. In Section 3, we apply our approach to the continuous case; that is, to the Duclos *et al.* (2004) polarization index. We

[1]According to Podder (1996), the concept that adequately represents the degree of discontent or injustice felt within a society is relative deprivation instead of inequality.

present an illustration of our proposal in Section 4. Finally, in Section 5, we close the paper with some concluding remarks.

2. The Discrete Case

2.1. *The Esteban and Ray (1994) Polarization Framework*

Let $X := (\pi_1, \ldots, \pi_n; x_1, \ldots, x_n)$ be a distribution for any positive integer n, where π_i is the percentage of population of group i, the vector of incomes is increasingly ordered, i.e. $0 < x_1 < x_2 < \ldots < x_n$ and the mean income is μ.

Esteban and Ray (1994) assume that each individual feeling of antagonism is subject to two forces: *identification* with members considered to belong to the same group, and *alienation* from those considered to belong to other groups. Effective antagonism increases in identification and alienation in such a way that increased intra-group identification reinforces the alienation effect. Polarization represents total effective antagonism. Accordingly, the absolute ER polarization index is proportional to:

$$P_\alpha^{ER} = \sum_{i=1}^{n}\sum_{j=1}^{n} \pi_i^{1+\alpha}\pi_j \left|x_i - x_j\right|, \tag{1}$$

where the parameter $\alpha \in [1, 1.6]$ represents the importance of group identification. The alienation term is $|x_i - x_j|$ and the identification term is π_i^{α}. Moreover, by dividing P_α^{ER} by the average μ, we obtain the relative ER polarization index, $RP_\alpha^{ER} = \frac{1}{\mu} P_\alpha^{ER}$.

Polarization in society is, therefore, the sum of all the effective antagonisms. The additive postulate is justified (following Harsanyi, 1953, among others) in terms of an impartial individual who might use the expected value of his effective antagonism to judge polarization in society.

Now, let W be a social welfare function defined on utilities that are represented by the function U. The end-product we seek is a social welfare function of the form

$$W(U_1(X), U_2(X), \ldots, U_n(X)) = V\left(\mu, P_\alpha^{ER}\right), \tag{2}$$

where $V : S \subseteq \mathbb{R}_+^2 \to \mathbb{R}$ is a function such that $\frac{\partial V}{\partial \mu} > 0$ and $\frac{\partial V}{\partial P_\alpha^{ER}} < 0$. In expression (2), the trade-off between income polarization and efficiency is explicit. Provided it has a plausible rationale, such a social welfare function W would allow us to make meaningful comparisons in situations where unanimous preference is not attainable, because efficiency can be traded for income polarization. Moreover, polarization indices do not fulfill the principle of progressive transfers so the social welfare function will turn out not to be Schur-concave.[2] The question is whether this

[2]A function g: $\mathbb{R}^n \to \mathbb{R}$ is said to be Schur-concave if *Lorenz domination* of distribution X over distribution Y implies that $g(X) \geq g(Y)$. Thus, all Schur-concave functions satisfy the principle of progressive transfers.

function W exists or not. A negative answer would make a plausible case for arguing that income polarization and welfare cannot be meaningfully combined.

2.2. *The Result*

To study whether polarization and welfare can be meaningfully combined, we initially impose three restrictions.

First, all individuals have the same utility function $U(\cdot)$ for symmetry of the social welfare function. In this manner, social preferences are impersonal or disinterested.

Second, U is not an individualistic function in the sense that people care not only about their own income but also about the distribution they inhabit. At this point, it is important to recall that initial configurations of the whole distribution are relevant for the measurement of polarization (Esteban and Ray, 1994; Wolfson, 1994). Accordingly, preferences should not be individualistic if they are to take into account the whole distribution; that is, the utility function should be $U(x_i, X)$ for all $i = 1, \ldots, n$. In this manner, we expand the domain for personal preferences to incorporate envy or altruism.

Third, social welfare is utilitarian; that is, $W(X) = \sum_{i=1}^{n} U(x_i, X) \cdot \pi_i$. From d'Aspremont and Gevers (1977, 2002) we know that utilitarianism is characterized by the *weak Pareto* principle, the *anonymity* axiom, and the axiom of *invariance with respect to common rescaling and individual change of origin*. The first two axioms are well-known principles, commonly used in the literature. However, the last axiom needs some explanation. Under this axiom of invariance each individual evaluation is measured on a cardinal scale. Furthermore, interpersonal comparisons of welfare gains are permitted, while interpersonal comparisons of welfare levels are prohibited. However, we have assumed above that all individuals have the same utility function (see the first restriction). Therefore, we are implicitly assuming that utility functions are co-cardinal, i.e. cardinal and fully comparable. This is a strong assumption, albeit it is in agreement with the existing literature on polarization (see, for example, Esteban and Ray, 1994; Duclos *et al.*, 2004). In this manner, impartial observers who evaluate overall polarization according to the expected value would apply the same criterion to welfare in society.

Once we assume these restrictions, we propose a plausible social welfare function W that satisfies (2).

Let $D(x_i; x_j)$ be the relative deprivation felt by an individual with income x_i in relation to an individual with income x_j, as follows (Runciman, 1966; Yitzhaki, 1979; Hey and Lambert, 1980; Chakravarty, 1997):

(3)
$$D(x_i; x_j) = x_j - x_i \quad \text{if } x_i < x_j,$$
$$D(x_i; x_j) = 0 \quad \text{if } x_i \geq x_j.$$

Then, the relative deprivation felt by an individual with income x_i is:

(4)
$$D(x_i) = \sum_{j=1}^{n} D(x_i; x_j)\pi_j = \sum_{j=i+1}^{n} (x_j - x_i)\pi_j.$$

We can rewrite $D(x_i)$ as $\mu - \sum_{j=1}^{i} x_j \pi_j - x_i \sum_{j=i+1}^{n} \pi_j$. Following Chakravarty and Mukherjee (1999) we regard the complement $S(x_i) = \sum_{j=1}^{i} x_j \pi_j + x_i \sum_{j=i+1}^{n} \pi_j$ of $D(x_i)$ to the mean income μ as the satisfaction function of the person with income x_i. That is, $S(x_i) = \mu - D(x_i)$. Given the position of a person in an income distribution, she may be regarded as being either frustrated or satisfied. Therefore, since $S(x_i)$ is based on the censored income distribution $(x_1, \ldots, x_{i-1}, x_i, \ldots, x_i)$ in which individual i has no feeling of frustration, it can be considered as his satisfaction function.[3]

Given the restriction that U cannot be individualistic and expression (4), we define the utility function of an individual with income x_i as follows:

$$U(x_i, X) = (1-\theta_i)x_i + \theta_i[\mu - 2D(x_i)] \quad 0 \le \theta_i = k\pi_i^{\beta} \le 1, \tag{5}$$

where $k, \beta \ge 0$. The individual cares not only about his own income but also about the distribution to which he belongs. Utilities are, therefore, not individualistic. In fact, personal utility is a convex combination of two arguments: own income and income distribution in terms of mean income and relative deprivation.

Most economists routinely assume that material self-interest is the sole motivation of people. However, this practice contrasts with a large body of evidence gathered by psychologists and experimental economists over the last two decades (see Sobel, 2005; Fehr and Schmidt, 2006; and the references therein). In particular, this evidence indicates that a substantial percentage of the population is strongly motivated by other-regarding preferences and that concerns for the well-being of others cannot be ignored in social interactions. In line with this literature, we can interpret expression (5) as follows: individuals care about their own income and, in addition, would like to increase the mean income for the society they inhabit as long as it does not bring about any increase in their own relative deprivation. Here, we find two opposite sentiments. On the one hand, individuals feel altruism (Kahana, 2005) because they care for the general well-being of society (represented by μ). On the other hand, individuals feel envy because they compare their positions to those of others and realize that they have less than them (represented by $D(x_i)$).[4] Looking to reality, we find this interpretation particularly appealing.

Now, we provide three alternative interpretations of the utility function in (5). First, we note the following:

$$\begin{aligned} \mu - 2D(x_i) &= \sum_{j=1}^{n} x_j \pi_j - \sum_{j=i+1}^{n} x_j \pi_j + \sum_{j=i+1}^{n} x_i \pi_j - D(x_i) \\ &= \sum_{j=1}^{n} \min\{x_j, x_i\}\pi_j - D(x_i). \end{aligned} \tag{6}$$

[3]The generalized satisfaction quasi-ordering introduced by Chakravarty (1997) is denoted non-deprivation quasi-ordering by Magdalou and Moyes (2009).

[4]Relative deprivation is considered to be caused by individual envy throughout this paper. Alternatively, relative deprivation might be also caused by individual injustice because people could believe that higher incomes are due to unambiguous unfair allocations of income. In any case, relative deprivation will cause social tension or unrest.

For a given personal income x_i, the term $\sum_{j=1}^{n}\min\{x_j, x_i\}\pi_j$ increases when income differences between x_i and those who are worse off decrease. However, income deprivation increases when income differences between x_i and those who are better off increase. Therefore, individual utility for a given personal income x_i will increase when the gap with lower incomes and individual deprivation decrease. Again, we find two opposite sentiments: people are willing to give up some material payoff to move in the direction of more equitable outcomes (altruism), but only if such an equalization of incomes does not increase individual deprivation (envy).

Second, when we substitute equation (4) in the utility function definition, we obtain the following expression:

$$U(x_i, X) = (1-\theta_i)x_i + \theta_i\left[\mu - 2\sum_{j=i+1}^{n}(x_j - x_i)\pi_j\right]. \tag{7}$$

We know that $\mu = \sum_{j=1}^{n} x_j\pi_j$, therefore we obtain the following:

$$\begin{aligned} U(x_i, X) &= x_i + \theta_i\left[\sum_{j=1}^{n} x_j\pi_j - 2\sum_{j=i+1}^{n} x_j\pi_j + 2\sum_{j=i+1}^{n} x_i\pi_j - x_i\right] \\ &= x_i + \theta_i\left[2\sum_{j=1}^{i} x_j\pi_j - \sum_{j=1}^{n} x_j\pi_j + 2\sum_{j=i+1}^{n} x_i\pi_j - \sum_{j=1}^{n} x_i\pi_j\right] \\ &= x_i - \theta_i\left[\sum_{j=1}^{n}\left(x_i + x_j - 2\min\{x_i, x_j\}\right)\pi_j\right] \end{aligned} \tag{8}$$

Since $|x_i - x_j| = x_i + x_j - 2\min\{x_i, x_j\}$, we have:

$$U(x_i, X) = x_i - \theta_i\sum_{j=1}^{n}|x_i - x_j|\pi_j. \tag{9}$$

Individuals care not only about their own income but also about the incomes of those who are worse off and better off in the income distribution. On the one hand, individuals are altruistic because they have an unselfish preference for those who are poorer. On the other hand, individuals are envious because they prefer the impoverishment of those who are richer. Once again, individuals feel altruism or envy depending on the income under consideration. In this respect, it is worth noting that Bolton and Ockenfels (2000) have proposed a model where agents' preferences are an increasing function of their own payoff and their relative payoff. Fehr and Schmidt (1999) have introduced a model with a similar motivation but assumed an individual utility function under which an agent cares about his own monetary payoff and, in addition, would like to reduce the inequality in payoffs across all agents. In line with these studies, we propose a utility function that depends on own income and on the degree of income differences.

Third, considering the definition of the satisfaction function, we can rewrite (5) as follows:

$$U(x_i, X) = (1-\theta_i)x_i + \theta_i[S(x_i) - D(x_i)]. \tag{10}$$

In addition to personal income, individual utility increases with personal relative satisfaction, and decreases with personal relative deprivation. That is, people care not only for their own income, but also for the sentiments of satisfaction and deprivation that the distribution they face brings about.[5]

Utility, therefore, is a balance between an individual's income and the distribution to which he belongs in terms of mean income and relative deprivation. In the most general case, this balance will depend on individual characteristics like income, size of his group, socioeconomic and cultural background, and so on. In this case, every individual will have a particular weight. In contrast, the most restrictive case will consist of a society where all individuals present the same weight. We focus on an intermediate case where individuals balance both arguments (own income and $\mu - 2D(\cdot)$) according to the size of their group. In this manner, every income group has a different weight factor θ_i. In this paper, we assume a positive relationship between the weight factor θ_i and the size of group *i*.

According to the "group-size paradox" (Olson, 1965), larger groups may be less successful than smaller groups in furthering their interests. First, individuals bear the adverse consequences of reducing their contributed effort only partially and, consequently, collective effort typically falls below the group-optimal level. This is the well-known free-rider or collective action problem. Second, if the collective good is purely private, so that it can be divided up among the group members, the larger the group, the smaller the individual reward. However, Chamberlin (1974), McGuire (1974), Marwell and Oliver (1993), and Esteban and Ray (2001), among others, have pointed out that collective action critically depends on the possibility of distributing the benefits from cooperation in ways that pay all potential partners to cooperate. They argue that when the collective good is public—so that individual reward is non-excludable—Olson's result is reversed: larger groups are more successful than smaller groups in producing the collective good.

In our framework, individuals value not only their own income but also the general well-being of society (as long as their relative deprivation is not affected). Therefore, there are two types of individual reward: own income which is private in nature because it is excludable; and moreover the general well-being of society which is public in nature because it is non-excludable. From the discussion above, it is clear that Olson's thesis can be applied in the first case but not in the second. Accordingly, we weigh the two arguments of the utility function as follows: own income is weighted by the factor $1 - \theta_i$, which depends negatively on the size of

[5]Following the proposal in Fehr and Schmidt (1999), D'Ambrosio and Frick (2012) comment on this utility function:

$$U(x_i, X) = x_i + \alpha D(x_i) + \beta S(x_i),$$

where $\alpha \leq \beta \leq 0$. In this case, individual well-being depends negatively not only on deprivation, but also on satisfaction.

group i, while the general well-being of society is weighted by the factor θ_i, which depends positively on the size of group i.

Relative weights depend also on parameters β and k. The percentage of population, π, is by definition lower than one for all income groups. Therefore, the lower the value of parameter β, the larger the weight θ_i. In fact, all individuals will weigh their income distribution by the parameter k if $\beta = 0$, while individuals will not care about the distribution they inhabit if the value of β is sufficiently large. Accordingly, we assume that $\beta \in (0, \bar{\beta})$, where $\bar{\beta}$ is an upper bound.

In contrast, the higher the value of parameter k, the larger the weight given to the income distribution. However, the constant k should have an upper bound. In principle, parameter k should be at most equal to $\frac{1}{\sup_i \pi_i^\beta}$ to guarantee that $\theta_i \leq 1$ for all $i = 1, \ldots, n$. By definition, we know that $\sup_i \pi_i^\beta \leq 1$ for all i, hence $\frac{1}{\sup_i \pi_i^\beta} \geq 1$. Thus, we can assume that $k \in [0, 1]$.

After justifying the meaningfulness of the utility function in (5), the following result proposes an additively separable function W that satisfies (2) for the absolute and relative cases.

Proposition 1. Given the social welfare function $W(X) = \sum_{i=1}^{n} U(x_i, X) \cdot \pi_i$, for every income distribution X:

$$U(x_i, X) = (1 - \theta_i) x_i + \theta_i [\mu - 2D(x_i)],\ \theta_i = k\pi_i^\beta \text{ and } \beta = \alpha \in [1, 1.6] \Rightarrow$$
$$W_{k,\alpha}(X) = \mu - kP_\alpha^{ER} = \mu\left(1 - kRP_\alpha^{ER}\right).$$

Proof. When we substitute equation (9) in the social welfare function $W(X)$, we obtain the following expression:

$$W(X) = \sum_{i=1}^{n} \left(x_i - k\pi_i^\alpha \sum_{j=1}^{n} |x_i - x_j| \pi_j \right) \pi_i.$$

We need only consider the definitions of absolute and relative income polarization to complete the proof.

This result shows that an abbreviated social welfare function that depends only on mean income and a polarization index (absolute or relative) can be used as a tool for choosing between alternative distributions of income when the range of β is restricted to the interval [1, 1.6].

It is interesting to note that the non-decreasingness of the social welfare function, while capturing the desire for higher incomes, may come into conflict with the desire for lower polarization. For example, an increase in the income of the richest person implies an increase not only in efficiency, but also in the relative deprivations of all the other persons. For this reason, we consider an alternative notion of efficiency proposed by Shorrocks (1983). In particular, we adopt the *scale improvement condition*, which demands that welfare improves if all the incomes are increased equiproportionally. The social welfare function W satisfies scale improvement if for all X,

$$W(\lambda X) \geq W(X) \tag{11}$$

where $\lambda \geq 1$. This condition indicates preference for higher incomes, while keeping relative deprivations constant. It is clear from above that the social welfare function in Proposition 1 verifies the scale improvement condition.

Some further observations about Proposition 1 are in order.

First, the rate of substitution between income polarization and efficiency at a constant welfare level in the absolute case depends on parameter k as follows:

$$\left.\frac{\Delta P_\alpha^{ER}}{\Delta \mu}\right|_W = \frac{1}{k}. \tag{12}$$

The higher the value of parameter k, the more sensitive social welfare is to income polarization. For the relative case, the rate of substitution between income polarization and efficiency is

$$\left.\frac{\Delta RP_\alpha^{ER}}{\Delta \mu}\right|_W = \frac{1-kRP_\alpha^{ER}}{k\mu}, \tag{13}$$

which closely resembles the rate of substitution for the Gini coefficient (G).[6]

Second, the social welfare function is $W_{k,\alpha}(X) = \mu(1 - 2kG(X))$ when $\alpha = 0$ because in this case the relative ER polarization index boils down to twice the relative Gini coefficient. Likewise, social welfare is equal to the average of individual satisfaction, i.e. $W_{k,\alpha}(X) = \sum_{i=1}^{n} S(x_i)\pi_i$ when $k = 0.5$ and $\alpha = 0$. Therefore, the average of individual satisfaction is equal to the Gini social welfare function, i.e. $\sum_{i=1}^{n} S(x_i)\pi_i = \mu(1-G(X))$ when $k = 0.5$ and $\alpha = 0$.

Third, the welfare index $W_{k,\alpha}(X)$ cannot reach negative values. We know that the distribution that assigns half the population to the lowest income class and the remainder to the highest income class is more polarized than any other distribution (see ER, p. 837). Therefore, the most adverse case is given by the bimodal distribution $\left(\frac{1}{2}, 0, \ldots, 0, \frac{1}{2}; x_1, \ldots, x_n\right)$ and $\alpha = 1$. For this case, it can be shown that $RP_1^{ER} = G(X) = \frac{1}{2}\frac{x_n - x_1}{x_n + x_1}$ which is always less than or equal to 0.5.

Fourth, consider the absolute ER polarization index with asymmetric alienation (AER):[7]

[6]Recall that the rate of substitution between inequality (measured by the Gini coefficient) and efficiency is $\left.\frac{\partial G}{\partial \mu}\right|_W = \frac{1-kG}{k\mu}$ under the welfare function $W = \mu(1 - kG)$ (see Lambert, 2001).

[7]This index is commented on by ER, though its axiomatic foundation remains undone.

$$P_\alpha^{AER} = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \pi_i^{1+\alpha} \pi_j (x_j - x_i) \tag{14}$$

where the poor feel alienated from the rich but the rich do not feel alienated from the poor. Likewise, let the relative ER polarization index with asymmetric alienation be $RP_\alpha^{AER} = \frac{1}{\mu} P_\alpha^{AER}$. Then, if we adopt the utility function $U(x_i, X) = x_i - k\pi_i^\beta D(x_i)$ and $\beta = \alpha \in [1, 1.6]$, our results for social welfare and income polarization, $W_{k,\alpha}(X) = \mu - kP_\alpha^{AER}$ and $W_{k,\alpha}(X) = \mu(1 - kRP_\alpha^{AER})$, will again be obtained. Now, utility depends only on own income and the deprivation felt by the individual. In this respect, it is worth noting that D'Ambrosio and Frick (2007) have found that subjective well-being depends not only on absolute levels of income, but also on relative deprivation.[8] In the same vein, Clark and Oswald (1996) have found that workers' reported levels of well-being are inversely related to their comparison wage rates and Cabrales *et al.* (2007) have recently proposed a model where workers, in addition to the utility they obtain from their own wage, experience disutility from the wage of fellow workers who enjoyed similar circumstances in the near past and have a higher wage than their own.

Finally, the result in Proposition 1 contradicts the idea that welfare and inequality are inversely related. It is easy to see that any change in within-group inequality always provokes an opposite effect on income polarization through the variation in the identification component. Consequently, the well-established notion of abbreviate social welfare functions that are inequality averse must be abandoned in our case. However, far from limiting the arsenal at hand for researchers and policymakers, our proposal enlarges it by focusing on a different dimension of the income distribution that has not been considered yet, income polarization.

In the next section we briefly extend our results to the Duclos *et al.* (2004) polarization index for continuous distributions.

3. The Continuous Case

3.1. *The Duclos et al. (2004) Polarization Framework*

Let $f(x)$ be a frequency density function for $x \in [a, b]$, where x is income and $[a, b]$ is a positive bounded interval that contains the support of the distribution. We assume that $f(x)$ is differentiable on the open interval (a, b) and $\mu = \int_a^b xf(x)\,dx$ is the mean income. The corresponding distribution function is $F(x) \in \Phi$, where Φ is the class of income distributions. Then, the DER polarization index is proportional to:

$$P_\alpha^{DER}(F) = \int_a^b \int_a^b f(x)^{1+\alpha} f(y)|y - x|\,dy\,dx \tag{15}$$

for $\alpha \in [0.25, 1]$, which represents the importance of identification. Again the additive postulate is assumed; that is, polarization is the sum of all antagonisms.

[8]These authors have found that the correlation between subjective well-being and income is 0.36, while the correlation with relative deprivation is –0.44.

Also, note that homogeneity of degree one and zero is achieved by multiplying the expression in (15) by μ^{α} and $\mu^{\alpha-1}$, respectively.

Then, we may ask whether there exists a social welfare function W that satisfies the following:

(16) $$W(\cdot) = V(\mu, P_{\alpha}^{DER})$$

for some function $V: Z \subseteq \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}$, where $\frac{\partial V}{\partial \mu} > 0$ and $\frac{\partial V}{\partial P_{\alpha}^{DER}} < 0$.[9] To answer this question, we impose the restrictions specified in Section 2 in continuous terms. First, we assume that all individuals have the same utility function $U(\cdot)$ for symmetry of the social welfare function. Second, utility is non-individualistic; that is, the utility of an individual with income x in distribution $F(x)$ is $U(x, F)$. Finally, social welfare is an additive separable function; that is, $W(F) = \int_{a}^{b} U(x, F) f(x) dx$. Given our problem and these restrictions, we replicate the result in Section 2 for continuous distributions.

3.2. *The Result*

Let $D(x)$ be the relative deprivation felt by an individual with income x (Runciman, 1966; Yitzhaki, 1979; Hey and Lambert, 1980; Chakravarty, 1997):

(17) $$D(x) = \int_{x}^{b} (y - x) f(y) dy.$$

Having more people with higher income increases personal deprivation.

Moreover, let $U(x, F)$ be the utility function of an individual with income x, as follows:

(18) $$U(x, F) = (1 - \theta_x) x + \theta_x [\mu - 2D(x)] \quad 0 \leq \theta_x = k_x (f(x)\mu)^{\beta} \leq 1.$$

As in the discrete case, individuals care about their own income and, in addition, the general well-being of society (represented by μ) as long as it does not bring about any increase in their relative deprivation. However, in contrast with the discrete case, we multiply the term $f(x)^{\beta}$ by μ^{β} in the weight factor θ_x to make the latter invariant to changes in the scale of incomes.

Unfortunately, we cannot assume that $k \in [0, 1]$ in the continuous case, because θ_x may be larger than 1 since $f(x)$ has no upper bound. For this reason, we have to assume that $k_x \in \left[0, \frac{1}{\sup_x (f(x)\mu)^{\beta}}\right]$. We denote the parameter k with the subscript x to make explicit that the upper bound of k depends on the distribution considered, though it is invariant to changes in the scale of incomes. Despite this, we can safely use the interval [0, 1] in empirical studies because the term $\sup_x (f(x)\mu)^{\beta}$ converges to 1. In real data the value of $f(x)\mu$ is below 1 since $f(x)$

[9]For the same reason as in the discrete case, the social welfare function W will turn out not to be Schur-concave.

tends to be very small. Then, the smaller the parameter β, the larger the term $\sup_x (f(x)\mu)^{\beta}$. It is straightforward to see that the latter converges to 1 when $\beta \to 0$. We check for the irrelevance of the term $\sup_x (f(x)\mu)^{\beta}$ with real data in the illustration below.

The following result proposes an additively separable function W that satisfies (16).

Proposition 2. Given the social welfare function $W(F) = \int_a^b U(x, F) f(x) dx$ for every income distribution $F(\cdot) \in \Phi$:

$$U(x, F) = (1-\theta_x)x + \theta_x[\mu - 2D(x)], \theta_x = k_x \cdot (f(x)\mu)^{\beta} \text{ and } \beta = \alpha \in [0.25, 1] \Rightarrow$$
$$W_{k,\alpha}(F) = \mu - k_x P_{\alpha}^{DER(1)} = \mu\left(1 - k_x RP_{\alpha}^{DER}\right),$$

where $P_{\alpha}^{DER(1)}$ denotes the DER polarization index that is homogeneous of degree 1 and RP_{α}^{DER} is the relative (homogeneous of degree zero) DER polarization index.

Proof. When we substitute equation (17) in the utility function definition we obtain the following expression:

$$U(x, F) = x - \theta_x x + \theta_x\left[\mu - 2\int_x^b (y - x) f(y) dy\right].$$

We know that $\int_a^b f(y) dy = 1$, therefore we can rewrite the above expression as follows:

$$U(x, F) = x - \theta_x \int_a^b xf(y) dy + \theta_x\left[2\left(\mu - \int_x^b yf(y) dy\right) + 2\int_x^b xf(y) dy - \mu\right].$$

Furthermore, $\mu = \int_a^b yf(y) dy$ so we derive the following:

$$\begin{aligned} U(x, F) &= x + \theta_x\left[2\int_a^x yf(y) dy + 2\int_x^b xf(y) dy - \int_a^b yf(y) dy - \int_a^b xf(y) dy\right] \\ &= x + \theta_x\left[2\int_a^b \min\{y, x\} f(y) dy - \int_a^b yf(y) dy - \int_a^b xf(y) dy\right] \\ &= x - \theta_x \int_a^b (y + x - 2\min\{y, x\}) f(y) dy. \end{aligned}$$

Since $|y - x| = y + x - 2\min\{y, x\}$, we have the following:

$$U(x, F) = x - \theta_x \int_a^b |y - x| f(y) dy.$$

If $W(F) = \int_a^b U(x, F) f(x) dx$, $\theta_x = k_x \cdot (f(x)\mu)^{\beta}$ and $\beta = \alpha \in [0.25, 1]$, then

$$W(F) = \int_a^b \left(x - k_x \cdot f(x)^{\alpha} \mu^{\alpha} \int_a^b |y - x| f(y) dy\right) f(x) dx$$

We need only consider expression (15) to complete the proof.

An abbreviated social welfare function that depends only on mean income and a polarization index can be used as a tool for choosing between alternative distributions of income when the range of β is restricted to the interval [0.25, 1]. Notice that the social welfare function in Proposition 2 verifies the scale improvement condition.

The rate of substitution between polarization and efficiency at a constant welfare level is $\frac{1}{k_x}$ for $P_\alpha^{DER(1)}$ and $\frac{1-k_x RP_\alpha^{DER}}{k_x\mu}$ for RP_α^{DER}. As in the discrete case, social welfare is a function of the Gini coefficient, i.e. $W_{k,\alpha}(F) = \mu(1 - 2k_x G(F))$ when $\alpha = 0$; it is equal to the expected value of individual satisfaction, i.e. $W_{k,\alpha}(F) = \int_a^b S(x) f(x) dx$ when $k_x = 0.5$ and $\alpha = 0$; and the Gini social welfare function and the expected value of individual satisfaction are the same, i.e. $\int_a^b S(x) f(x) dx = \mu(1 - G(F))$ when $k_x = 0.5$ and $\alpha = 0$. Likewise, if we assume that the limit distribution $F^*(x)$, that assigns half the population to the lowest income value a and the remainder to the highest income value b, is more polarized than any distribution in Φ, it can be shown that $\lim_{F\to F^*} W_{k,\alpha}(F) \geq 0$ for all k_x and α. Finally, if we consider the Duclos *et al*. (2004) polarization index with asymmetric alienation (ADER):

$$P_\alpha^{ADER}(F) = \int_a^b \int_x^b f(x)^{1+\alpha} f(y)(y-x)\,dy\,dx \tag{19}$$

and its relative version $RP_\alpha^{ADER}(F) = \frac{1}{\mu} P_\alpha^{ADER}(F)$, and adopt the utility function $U(x, F) = x - k_x f(x)^\beta D(x)$ and $\beta = \alpha \in [0.25, 1]$, our results for social welfare and income polarization, $W_{k,\alpha}(F) = \mu - kP_\alpha^{ADER}$ and $W_{k,\alpha}(F) = \mu\left(1 - kRP_\alpha^{ADER}\right)$, will again be obtained.[10]

4. Empirical Application: The Case of the United States (1991–2010)

For our empirical exercise we use data drawn from the Current Population Survey (CPS) dataset for the United States in the period 1991–2010.[11] The income definition considered is disposable household income calculated as pre-tax income plus cash transfers from the government, minus income tax payments and social security contributions. We deflate nominal values to real values by using the consumer price index (CPI) with base period 2000. Moreover, incomes are normalized by an adult-equivalence scale defined as $s^{0.5}$, where s is household size.

Observations with non-positive incomes are removed and household observations are weighted by the CPS sample weights times the number of persons in the household. We estimate income polarization following the proposal in Duclos *et al*. (2004). In particular, we use the Rosenblatt–Parzen kernel density estimator (Rosenblatt, 1956; Parzen, 1962), the Gaussian kernel, and the homogeneity-of-degree-zero normalization, i.e. the indices P_α^{DER} are multiplied by $\mu^{\alpha-1}$. Moreover,

[10]In this case, there is no problem in assuming that $k \in [0, 1]$.

[11]That is, we use the CPS waves of the period 1992–2011. See http://www.census.gov/cps/ for detailed information on the structure of this data.

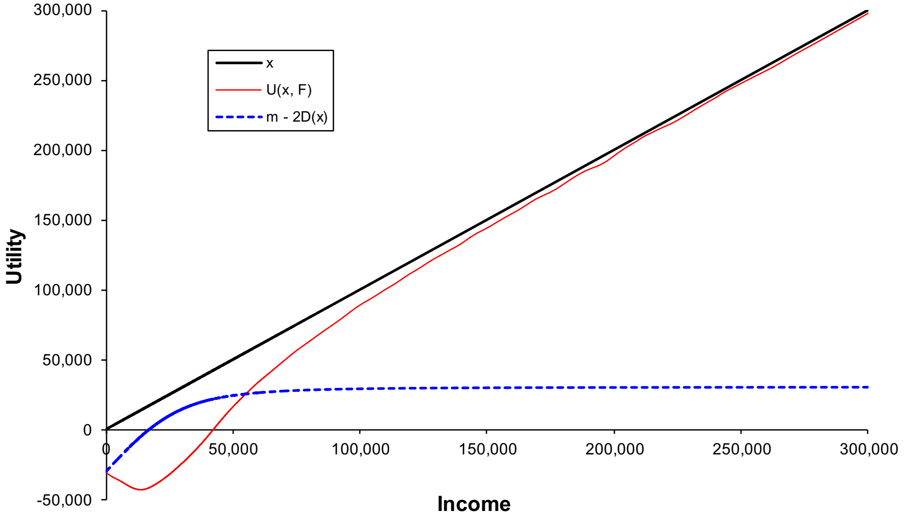

Figure 1. Individual Utility in the United States (2010) ($\alpha = 0.5$ and $k = 0.5$)

we provide the standard error estimates calculated by bootstrapping.[12] The results are shown in the Appendix.

First, we focus on the utility function described in (18). As commented above, the estimates of $f(x)$ are very small[13] so that the term $f(x)\mu$ is always below 1. As a result, the smaller the parameter α, the larger the term $\sup(f(x)\mu)^{\alpha}$. For this reason we compute only the values of the upper-bound of k for $\alpha = 0.25$ (the most adverse case). As expected, this upper-bound is always higher than 1 because it converges to 1 when $\alpha = 0$. Thus, it is shown that the theoretical upper-bound of k, $\frac{1}{\sup_x (f(x)\mu)^{\alpha}}$, is irrelevant when working with real income data and, consequently, that the interval [0, 1] for k can be safely used in empirical studies. From now onwards we consider $k = 0.5$ in order to compare welfare based on polarization indices ($\alpha = 0.25$, 0.5, and 1) with welfare based on relative satisfaction ($\alpha = 0$).

In Figure 1, we represent the term $\mu - 2D(x)$ and individual utility for people who only care about their own income ($k = 0$) and people who also care about the distribution they inhabit in the United States (2010).[14] For the latter case, we assume the values $\alpha = 0.5$ and $k = 0.5$ only for illustrative purposes. We note that

[12]In our calculations, we have assumed the formula (Davison and Hinkley, 2005):

$$\hat{\sigma}(\hat{I}) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}(I^* - \bar{I}^*)^2}$$

where I is the corresponding index and $R = 1000$ is the number of replications.

[13]This result is usually found in empirical studies based on real data (see, for example, Jenkins and van Kerm, 2005).

[14]To make more evident the shape of these functions in Figure 1, we have disregarded those observations with incomes higher than \$300,000. They are only 54 observations and represent less than 0.1 percent of the total sample.

the term $\mu - 2D(x)$ is an increasing and concave function of income. It is negative for low incomes because their deprivation is large, while it is almost constant (close to μ) for high incomes because their deprivation is very small. In contrast with utility for $k = 0$ which is a linear function of income, utility for $k = 0.5$ is an increasing and convex function of income. The presence of deprivation turns individual utility for low incomes into negative values, while the lack of deprivation and the small value of $f(x)$ for high incomes make $U(x, F)$ to converge to x.

Next, we comment on the estimated values of polarization. In Figure 2 (Panel 1) we show the polarization trend for disposable income in the United States assuming $\alpha = 0.5$. We see that polarization significantly increased over the 1994–2006 period. In particular, it experienced significant growth during the 1995–97 period, a large decrease in 1998 and 1999, and finally, a new significant increase from 2000 to 2006. With the advent of the Great Recession, polarization decreased in 2007 and, despite its recovery in 2009, it decreased in 2010 again. These results are, in general, robust to different values of parameter α.

Now, we focus on the evolution of welfare in the United States for five different valuations: mean income, which represents the welfare of people who only care about their own income ($k = 0$), and welfare assuming $k = 0.5$ for $\alpha = 0$, 0.25, 0.5, and 1. Note that the second valuation corresponds to the case when welfare is equal to the average of individual satisfaction. We chose these values only for illustrative purposes. In principle, there are two possibilities. A first possibility is that income polarization and mean income experience variations with opposite signs. In this case, a change in income polarization reinforces a change in mean income. Examples of this are the changes in disposable income between 1997 and 1998, and 2003 and 2004 for every value of α (see Appendix). A second possibility is that income polarization and mean income experience variations with the same sign. In this case, the change in welfare will depend on the relative magnitude of both variations and on the value of parameters α and k. For example, the variation in income polarization dominates the variation in mean income between 2000 and 2001 for every value of α and k (see Appendix).

Figure 2 (Panel 2) shows the evolution of welfare in the United States for the five cases mentioned above. The five series exhibit a similar pattern. Thus, welfare according to all valuations remains constant during the first three years of the period, then increases significantly from 1994 to 2008, and finally, initiates a descent with the deterioration of the Great Recession. It is interesting to note that the series of welfare based on polarization always move between two extremes, average income and average satisfaction.

5. Concluding Remarks

This paper constitutes a first attempt to relate the level of welfare in society consistently to the degree of income polarization. We have shown that such a relationship may be established if we expand the domain for individual preferences to include the sentiments of altruism and envy. Policy makers and researchers who have usually justified their policy analyses on the trade-off between efficiency and income inequality can now also implement their analyses of such policies on the basis of the trade-off between efficiency and income polarization.

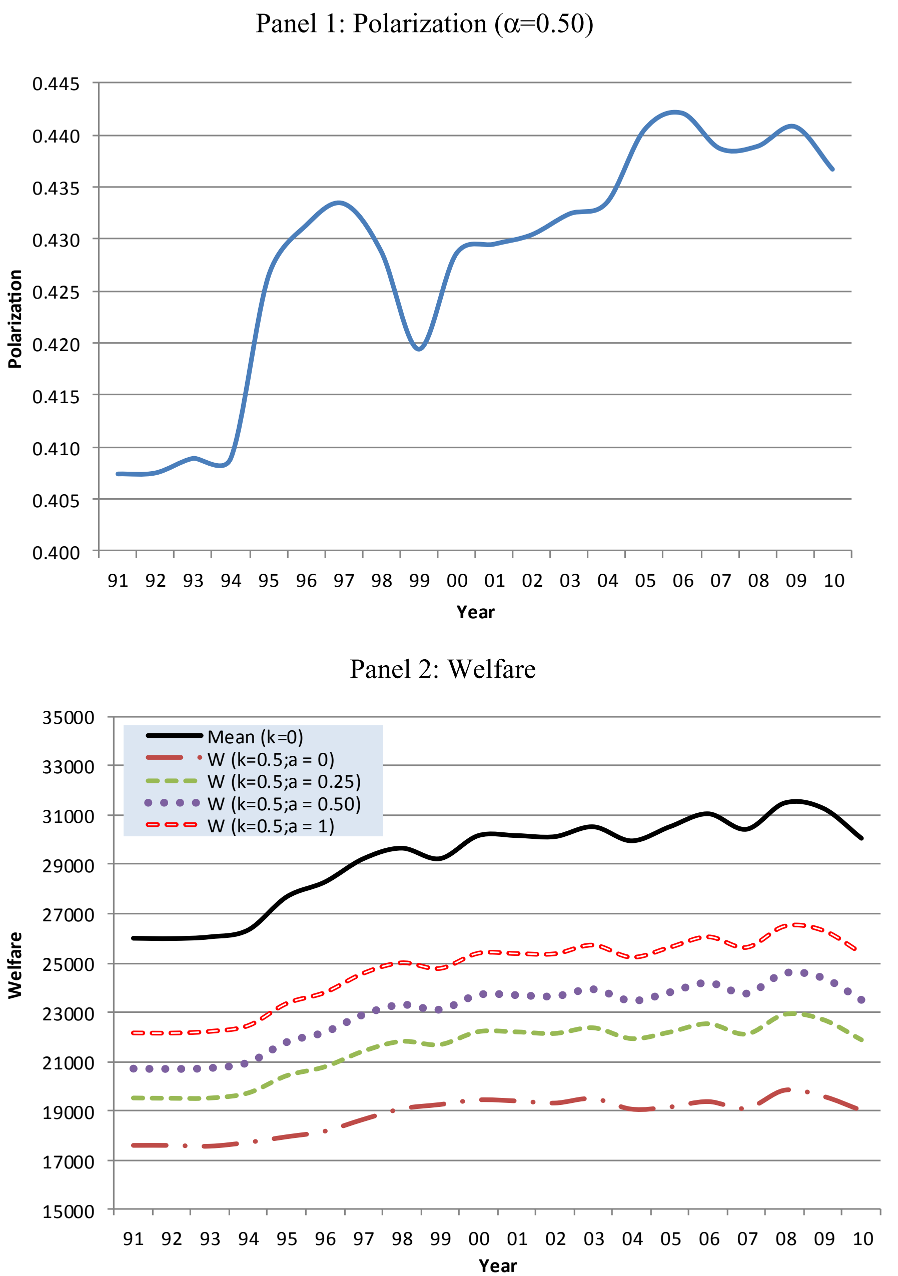

Figure 2. Evolution of Income polarization and Welfare in the United States (1991–2010)

It is important to acknowledge that any proposal of this kind has limitations. For example, the additive separable hypothesis for the social welfare function might not be a reasonable assumption. Likewise, the behavioral assumptions of the proposed utility function might be not so real after all. Nevertheless, our proposal does not need to be ironclad in order for it to constitute a useful tool for the applied researcher who, until now, has lacked a convincing way of evaluating income distributions in terms of mean income and income polarization.

References

Atkinson, A. B., "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–63, 1970.

Bolton, G. and A. Ockenfels, "ERC: A Theory of Equity, Reciprocity and Competition," *American Economic Review*, 90, 166–93, 2000.

Bossert, W. and W. Schworm, "A Class of Two-Group Polarization Measures," *Journal of Public Economic Theory*, 10, 1169–87, 2008.

Cabrales, A., A. Calvó-Armengol, and N. Pavoni, "Social Preferences, Skill Segregation and Wage Dynamics," *Review of Economic Studies*, 75, 65–98, 2007.

Chakravarty, S. R., "Relative Deprivation and Satisfaction Orderings," *Keio Economic Studies*, 34, 17–31, 1997.

Chakravarty, S. R. and B. Maharaj, "Measuring Polarization by Reduced-Form Indices," ECINEQ Working Paper No. 135, 2009.

Chakravarty, S. R. and A. Majumder, "Inequality, Polarization and Welfare: Theory and Applications," *Australian Economic Papers*, 40, 1–13, 2001.

Chakravarty, S. R. and D. Mukherjee, "Measures of Deprivation and Their Meaning in Terms of Social Satisfaction," *Theory and Decision*, 47, 89–100, 1999.

Chakravarty, S. R., A. Majumder, and S. Roy, "A Treatment of Absolute Indices of Polarization," *Japanese Economic Review*, 58, 273–93, 2007.

Chamberlin, J., "Provision of Collective Goods as a Function of Group Size," *American Political Science Review*, 68, 707–16, 1974.

Clark, A. E. and A. J. Oswald, "Satisfaction and Comparison Income," *Journal of Public Economics*, 61, 359–81, 1996.

D'Ambrosio, C., "Household Characteristics and the Distribution of Income in Italy: An Application of Social Distance Measures," *Review of Income and Wealth*, 47, 43–64, 2001.

D'Ambrosio, C. and J. R. Frick, "Income Satisfaction and Relative Deprivation: An Empirical Link," *Social Indicators Research*, 81, 497–519, 2007.

———, "Individual Well-Being in a Dynamic Perspective," *Economica*, 79, 284–302, 2012.

Dasgupta, A., A. Sen, and D. Starrett, "Notes on the Measurement of Inequality," *Journal of Economic Theory*, 6, 180–87, 1973.

d'Aspremont, C. and L. Gevers, "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies*, 44, 199–209, 1977.

———, "Social Welfare Functional and Interpersonal Comparability," in K. J. Arrow, A. K. Sen, and K. Suzumura (eds), *Handbook of Social Choice and Welfare*, Elsevier, North-Holland, Amsterdam, 2002.

Davison, A. C. and D. V. Hinkley, *Bootstrap Methods and Their Application*, 7th edition, Cambridge University Press, New York, 2005.

Duclos, J.-Y. and D. Échevin, "Bi-Polarization Comparisons," *Economic Letters*, 87, 249–58, 2005.

Duclos, J.-Y., J. Esteban, and D. Ray, "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 72, 1737–72, 2004.

Esteban, J. and D. Ray, "On the Measurement of Polarization," *Econometrica*, 62, 819–51, 1994.

———, "Conflict and Distribution," *Journal of Economic Theory*, 87, 379–415, 1999.

———, "Collective Action and the Group Size Paradox," *American Political Science Association*, 95, 663–72, 2001.

Fehr, E. and K. Schmidt, "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics*, 114, 817–68, 1999.

———, "The Economics of Fairness, Reciprocity and Altruism—Experimental Evidence and New Theories," in S. Kolm and J. M. Ythier (eds), *Handbook on the Economics of Giving, Reciprocity and Altruism*, Elsevier, North-Holland, Amsterdam, 2006.

Foster, J. E. and M. Wolfson, "Polarization and the Decline of the Middle Class: Canada and the U.S.," Mimeo, Vanderbilt University, 1992.

———, "Polarization and the Decline of the Middle Class: Canada and the U.S.," *Journal of Economic Inequality*, 8, 247–73, 2010.

Gradín C., "Polarization by Sub-Populations in Spain, 1973–91," *Review of Income and Wealth*, 46, 457–74, 2000.

Harsanyi, J. C., "Cardinality Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, 61, 434–35, 1953.
Hey, J. D. and P. J. Lambert, "Relative Deprivation and the Gini Coefficient: Comment," *Quarterly Journal of Economics*, 95, 567–73, 1980.
Jenkins, S. P. and P. van Kerm, "Accounting for Income Distribution Trends: A Density Function Decomposition Approach," *Journal of Economic Inequality*, 3, 43–61, 2005.
Kahana, N., "On the Surge of Altruism," *Journal of Population Economics*, 18, 261–6, 2005.
Kolm, S., "The Optimal Production of Social Justice," in J. Margolis and H. Guitton (eds), *Public Economics: An Analysis of Public Production and Consumption and their Relations to the Private Sectors*, Macmillan, London, 1969.
Kondor, Y., "Value Judgements Implied by the Use of Various Measures of Income Inequality," *Review of Income and Wealth*, 21, 309–21, 1975.
Lambert, P. J., "Social Welfare and the Gini coefficient Revisited," *Mathematical Social Sciences*, 9, 19–26, 1985.
———, *The Distribution and Redistribution of Income*, Manchester University Press, Manchester, 2001.
Lasso de la Vega, M. C. and A. M. Urrutia, "An Alternative Formulation of the Esteban–Gradín–Ray Extended Measure of Polarization," *Journal of Income Distribution*, 15, 10–42, 2006.
Magdalou, B. and P. Moyes, "Deprivation, Welfare and Inequality," *Social Choice and Welfare*, 32, 253–73, 2009.
Marwell, G. and P. Oliver, *The Critical Mass in Collective Action. A Micro-Social Theory*, Cambridge University Press, Cambridge, 1993.
McGuire, M. C., "Group Size, Group Homogeneity and the Aggregate Provision of a Pure Public Good under Cournot Behavior," *Public Choice*, 18, 107–26, 1974.
Newbery, D., "A Theorem on the Measurement of Inequality," *Journal of Economic Theory*, 2, 264–66, 1970.
Olson, M., *The Logic of Collective Action*, Harvard University Press, Cambridge, MA, 1965.
Parzen, E., "On Estimation of a Probability Density Function and Mode," *Annals of Mathematical Statistics*, 33, 1065–76, 1962.
Podder, N., "Relative Deprivation, Envy and Economic Inequality," *Kyklos*, 49, 353–56, 1996.
Prieto, J., J. G. Rodríguez, and R. Salas, "Polarization Characterization of Inequality-Neutral Tax Reforms," *Economics Bulletin*, 4, 1–7, 2003.
Reynal-Querol, M., "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution*, 46, 29–54, 2002.
Rodríguez, J. G. and R. Salas, "Extended Bi-Polarization and Inequality Measures," *Research on Economic Inequality*, 9, 69–83, 2003.
Rosenblatt, M., "Remarks on Some Nonparametric Estimates of a Density Function," *Annals of Mathematical Statistics*, 27, 832–37, 1956.
Runciman, W. G., *Relative Deprivation and Social Justice*, Routledge and Kegan Paul/Penguin Books, London, 1966.
Schwarze, J. and R. Winkelmann, "Happiness and Altruism within the Extended Family," *Journal of Population Economics*, 24, 1033–51, 2011.
Shorrocks, A. F., "Ranking Income Distributions," *Economica*, 50, 3–17, 1983.
Sobel, J., "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43, 392–436, 2005.
Wang, Y. Q. and K. Y. Tsui, "Polarization Orderings and New Classes of Polarization Indices," *Journal of Public Economic Theory*, 2, 349–63, 2000.
Wolfson, M. C., "When Inequalities Diverge," *American Economic Review*, 84, 353–58, 1994.
———, "Divergent Inequalities: Theory and Empirical Results," *Review of Income and Wealth*, 4, 401–22, 1997.
Yitzhaki, S., "Relative Deprivation and the Gini Coefficient," *Quarterly Journal of Economics*, 93, 321–24, 1979.
Zhang, X. and R. Kanbur, "What Difference do Polarization Measures Make? An Application to China," *Journal of Development Studies*, 37, 85–98, 2001.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix: Relative polarization after taxes and transfers in the USA (1991–2010)