

Taptué, André-Marie

**Working Paper**

## Comparing the size of the middle class using the alienation component of polarization

LIS Working Paper Series, No. 650

**Provided in Cooperation with:**

Luxembourg Income Study (LIS)

*Suggested Citation:* Taptué, André-Marie (2015) : Comparing the size of the middle class using the alienation component of polarization, LIS Working Paper Series, No. 650, Luxembourg Income Study (LIS), Luxembourg

This Version is available at:

<https://hdl.handle.net/10419/140652>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# LIS

## Working Paper Series

No. 650

### Comparing the size of the middle class using the alienation component of polarization

André-Marie Taptué

September 2015



CROSS-NATIONAL  
DATA CENTER  
*in Luxembourg*

---

**Luxembourg Income Study (LIS), asbl**

---

# Comparing the size of the middle class using the alienation component of polarization

André-Marie Taptué\*

revised version, August, 2015

## Abstract

This paper shows how to compare the size of the middle class in income distributions using a polarization index that do not account for identification. We derive a class of polarization indices where the antagonism function is constant in identification. The comparison of distributions using an index from this class motivates the introduction of alienation dominance surface, which is a function of an alienation threshold. We prove that a distribution has a large alienation component in polarization compared to another if the former always has a larger dominance surface than the latter regardless of the value of the alienation threshold. Then, we show that the distribution with large dominance surface is more concentrated in the tails and its middle class smaller than that of the other distribution. We implement statistical inference and test dominance between pairs of distributions using the asymptotic theory and Intersection Union tests. Our methodology is illustrated in comparing the declining of the middle class across pairwise distributions of twenty-two countries from the Luxembourg Income Study data base.

---

**Key words:** Alienation, identification, middle class, polarization.

---

\*Département D'Économique and CIRPÉE, Université Laval, Canada. email: andre-marie.taptue.1@ulaval.ca

# 1 Introduction

The importance of the middle class in society is no longer in question. [Thurow \(1984\)](#) put fear on the American government by writing that its middle class is disappearing. The fear was justified following what Karl Marx saw as an inevitable revolution based on the assumption that the disappearance of the middle class generates a bipolar income distribution composed of rich and poor. Once this bipolar situation existed, the poor would revolt, destroy the rich, protesting against the situation. Thus, the middle class is very important for the economic and political stability. In the economic aspect it is a source of the labour force and the key market for national product ([Foster and Wolfson, 2010/1992](#)). Moreover, the middle class is the one that provides directly or indirectly a large part of the country's tax revenue. A higher share of income for the middle class and lower ethnic divisions are empirically associated with higher income and higher growth ([Easterly, 2001](#)). What Easterly called the middle class consensus is also associated with more education, better health, better infrastructure, better economic policies, less political instability, less civil war and ethnic minorities at risk, more social modernization and more democracy.

In the political aspect a well administered government needs larger middle class for less factions and dissensions. It can be considered as an antidote to poverty, the sign of less income disparity and the goal many workers are trying to reach.

Although many studies such as [Blackburn and Bloom \(1985\)](#); [Foster and Wolfson \(2010/1992\)](#); [Leckie \(1988\)](#); [Levy \(1987\)](#); [Thurow \(1984\)](#) have attempted to give some approaches for the measurement of the middle class, there is no consensus among economists about the methodology to measure it. The middle class is mostly accepted to be composed of people whose income is in a given interval around the median. However, there is still a debate about the cut-offs of this interval since it varies from one author to another.

In this paper, we investigate the possibility of measuring and comparing the middle class using polarization indices of income. Polarization describes situations where the distribution of a variable of interest splits the population into two or more groups.

Studies in polarization traditionally use a framework based on alienation and identification, the basic components of polarization introduced by [Esteban and Ray \(1994\)](#). Alienation measures between-groups heterogeneity using distances between incomes and identification measures within-groups homogeneity using income density or group sizes. We analyze the possibility of abandoning the identification component and relying only on the alienation framework in a polarization index when comparing income distributions. For that purpose, we consider a polarization index developed by [Duclos et al. \(2004\)](#) and assume that identification is constant in the economy such that it does not interact with alienation and is not taken into account in the study of polarization.

The development leads to a restricted form of polarization indices that can be used to determine whether a distribution is more concentrated in the tails and has smaller middle class than another for a particular class of polarization indices and ranges of alienation thresholds. For instance, an income distribution in which incomes are far from the median means a hollowing out of its middle class.

We introduce alienation dominance surface defined as the proportion of individuals for whom alienation is greater than an alienation threshold. Alienation attributed to an individual is measured as the absolute relative deviation of his income from the median income, an approach that was used by [Foster and Wolfson \(2010/1992\)](#).

Whether incomes in a distribution are farther from their median than incomes in another distribution depends on the alienation threshold that has been fixed. But it is difficult in practice to agree on a non-arbitrary value of that threshold and different conclusions can be reached depending on thresholds. To address this difficulty, we adopt the methodology of stochastic dominance that allows us to determine a range of alienation thresholds, set robust comparisons and establish a partial ordering between two distributions.

Statistical inference is performed to guard against reaching conclusions that may be due to noise in samples. The comparison mentioned above leads to dominance curves that can intersect in the range of alienation thresholds. Also, distances between curves may be too tiny even when one curve always lies above another. One way to take this into account is to test statistically the significance of observed dominance. Thus, the second part of this paper develops statistical inference for stochastic dominance tests using an estimator of the alienation dominance surface. The statistical inference is implemented with asymptotic theory by deriving the asymptotic sampling distribution of the estimator and testing dominance using Intersection Union tests.

We begin in [Section 2](#) by first showing the decomposition of some polarization indices in order to isolate the components of alienation and identification. [Section 2](#) also addresses the measurement of alienation and establishes the relation between a polarization index and the alienation dominance surface. This relation characterizes the class of polarization indices compatible with comparison of income distributions in the alienation component of polarization.

Alienation dominance surface is developed in [Section 3](#) in the particular case of income distributions. This section identifies a distinction between an alienation dominance surface and an inequality measurement on the one hand, and similarities between alienation dominance and bi-polarization dominance on the other hand. [Section 3](#) also establishes the link between alienation dominance surface and middle class measurements. [Section 4](#) introduces an empirical illustration by comparing income distributions of 22 countries using 2004 data from the Luxembourg Income Study database. [Section 5](#) provides statistical inference by deriving the asymptotic properties of an estimator of the alienation dominance surface and testing statistically the results obtained in [Section 4](#). Finally, [Section 6](#) concludes.

## 2 Methodology

### 2.1 Decomposition of polarization indices

Our methodology is based on the fact that polarization indices can often be decomposed in a manner that isolates the components of alienation and identification. The decomposition gives a possibility of measuring each component alone without accounting for the other component. When one component is neglected, the antagonism function is written just as a function of the other component. We are examining the situation where the identification component is neglected, in which case the index captures the inequality (alienation) component of polarization. We first examine the decomposition of some polarization indices from the literature on polarization in order to define the measurement of alienation.

The first index we present is that axiomatized by [Esteban and Ray \(1994\)](#) who considered a discrete distribution of income. They measure polarization as the sum of all effective antagonisms in society where the vector of antagonism depends on the levels of alienation and identification for each value of income.

Consider a population of size  $N$  where the distribution of income bunches individuals in  $n$  groups of incomes  $y_1, y_2, \dots, y_i, \dots, y_n$  with sizes  $n_i$ . The polarization index is proportional to:

$$ER(\eta, \mathbf{F}) = \sum_{i=1}^n \sum_{j=1}^n n_i^{(1+\eta)} n_j |y_i - y_j| \quad (1)$$

where  $\mathbf{F}$  is the cumulative distribution function of income and  $\eta \in (0, 1.6]$  the relative importance of identification and alienation, commonly referred to as a parameter of polarization sensitivity. This polarization measure bears resemblance to the Gini coefficient: [Equation \(1\)](#) equals the Gini if  $\eta$  is set to zero. The fact that  $\eta$  can exceed zero distinguishes income polarization and inequality. The larger its value, the greater the departure from inequality measurement.

[Equation \(1\)](#) can be rewritten by isolating the two components of polarization:

$$ER(\eta, \mathbf{F}) = \sum_{i=1}^n n_i^\eta \sum_{j=1}^n n_i n_j |y_i - y_j|. \quad (2)$$

The identification felt by group  $i$  is  $\iota_i = n_i^\eta$ . This measure leads to the same level of identification in groups with the same sizes. However, identification does not only depend on the number of individuals with similar attributes, but also on the attributes themselves. For instance, the sense of identification must be lower in groups with small income and higher in groups with large income even if these all have the same masses. This view of identification as depending on the levels of attributes yields the following measure of identification:

$$\iota_i = n_i^\eta y_i^\beta \quad (3)$$

where  $\beta$  represents the sensitivity of group cohesion.

Alienation felt by group  $i$  towards groups  $j$  is  $|y_i - y_j|$ , and the overall alienation felt by group  $i$  towards all the other groups is  $a_i = \sum_{j=1}^n n_j |y_i - y_j|$ . Divided by the total size of the population, it is the sum of alienations between group  $i$  and groups  $j$  weighted by the population shares of groups  $j$ . The total alienation in society is  $a = \sum_i n_i a_i$ .

The second index is the index of [Duclos et al. \(2004\)](#) applicable to a continuous income distribution. It addresses two problems in the index of [Esteban and Ray \(1994\)](#), that of the discontinuity and the pre-existence of groups in which individuals are clustered. Their index is defined as follows:

$$DER(\eta) = \iint f^{(1+\eta)}(y) f(x) |y - x| dx dy \quad (4)$$

where  $f$  is the income density and  $\eta \in [0.25, 1]$  is still the polarization sensitivity.

[Equation \(4\)](#) is the continuous form of [Equation \(1\)](#). The identification generated in a group with income  $y$  is  $\iota(y) = f^\eta(y)$ . The alienation felt by a group whose income is  $y$  towards another group whose income is  $x$  is  $|y - x|$  such that  $a^*(y) = \int |y - x| dF(x)$  is the mean alienation felt by individuals whose income is  $y$  towards the rest of the population. A decomposition has been done in [Duclos et al. \(2004\)](#) where the index is rewritten isolating the components of identification and alienation:

$$DER(\eta) = \int f(y)^\eta \int f(y) f(x) |y - x| dx dy. \quad (5)$$

The average identification is defined by

$$\bar{\iota} \equiv \int f(y)^\eta dF(y) = \int f(y)^{1+\eta} dy \quad (6)$$

and the average alienation in society by

$$\bar{a} = \int a^*(y) dF(y) = \int \left( \int |y - x| dF(x) \right) dF(y). \quad (7)$$

Finally, it is proven that

$$DER(\eta) = \bar{a} \bar{\iota} [1 + \rho] \quad (8)$$

where  $\rho = \frac{cov(a^*(y), \iota(y))}{\bar{a} \bar{\iota}}$  is the normalized covariance between alienation and identification.

The alienation component of polarization can also be linked to the middle class and to bi-polarization, which measures the extent to which a population is divided into two separate groups. [Foster and Wolfson \(2010/1992\)](#) [FW] introduces two bi-polarization curves as well as a new Gini-like index of bi-polarization. The first curve says that bi-polarization is greater when the size of the middle class is smaller, *viz.*, when the distances to the median are larger. The second bi-polarization curve says that

bi-polarization is higher when the average distance from the median, on either side of the median, are larger. The index proposed by FW equals twice the area between the Lorenz curve and the tangent to the Lorenz curve at median income.

An income distance between two quantiles at percentiles  $q^-$  and  $q^+$  is defined by

$$S(q^-, q^+) = \tilde{Q}(q^+) - \tilde{Q}(q^-) \quad (9)$$

where  $\tilde{Q}(q) = \frac{1}{m}F^{-1}(q)$  is the median ( $m$ ) normalized income of the person at the  $q^{th}$  percentile.  $S(q^-, q^+)$  is the median-normalized income distance between the two quantiles  $q^-$  and  $q^+$ . If  $[q^-, q^+]$  overlaps 0.5, it can also be thought of as a measure of the size of the middle class. When  $S$  is large, there are fewer individuals near the middle, and consequently the middle class is deemed smaller and there is more bi-polarization. FW define a first-order polarization curve as:

$$S(q) = |S(q, 0.5)| \quad (10)$$

for  $0 \leq q \leq 1$ . For each  $q$ ,  $S(q)$  is the distance that separates median income from the income of the person situated at the  $q$ -th percentile. FW also define a second-order polarization curve as

$$B(p) = \left| \int_p^{0.5} S(q) dq \right| \quad (11)$$

for  $0 \leq p \leq 1$ .  $B(p)$  is the area under the  $S(q)$  curve between points  $p$  and 0.5. The bi-polarization index of Foster and Wolfson (2010/1992) is then defined as

$$FW = 2 \int_0^1 B(p) dp, \quad (12)$$

which is twice the area beneath the second-order bi-polarization curve  $B(p)$ .

The measurement of polarization based on the deviation from the median has several extensions. Wang and Tsui (2000) proposes two extensions that aggregate income distances from the median. The first one uses rank weighting and the corresponding measure is defined by:

$$WT_A = \int w(q) S(q) dq \quad (13)$$

where  $w(q)$  is positive and increasing for  $q \leq 0.5$ , and positive and decreasing for  $q \geq 0.5$ . The second one considers a polarization measure defined as a normalized transformation of distances from the median. A class of such indices is given by

$$WT_B = \int_0^1 \psi(S(q)) dq \quad (14)$$

where  $\psi(\cdot)$  is a continuous function.



## 2.2 Alienation measurement

At least in some of the indices presented above, alienation for a given income is assessed as the mean distance from some other incomes. The deviation may also be from the mean income, but we choose the median for two main reasons: The first one is the robustness of the median as a statistic compared to the mean income. The median income is not sensitive to the extreme values of a distribution and does not change if these values increase or decrease while the mean will be modified following any change in the extreme values. For the second reason, consider the expression of alienation drawn from the polarization index [Equation \(4\)](#) of [Duclos et al. \(2004\)](#) for a continuous distribution of income  $y$ .

$$a^*(y) = \int |y-x|dF(x). \quad (15)$$

We decompose  $a^*(y)$  in the sum of the distance when  $x < y$  and the distance when  $x > y$ :

$$a^*(y) = \int I(x < y)|y-x|dF(x) + \int (1-I(x < y))|y-x|dF(x). \quad (16)$$

Thus, alienation in  $y$  can exceed a threshold because of the gaps with incomes less than  $y$  or because of the gaps with incomes greater than  $y$ . The measurement in absolute value allows us to account for the two cases. Notice that alienation in  $y$  is the mean of absolute distances between  $y$  and all the other incomes. It is known that the median minimizes the expected absolute distance of any variable:

$$m = \arg \min_{\beta} E_X |X - \beta|. \quad (17)$$

Consequently, the average absolute deviation from the median is less than or equal to the average absolute deviation from the mean ( $E_X |X - m| \leq E_X |X - \mu|$ ) where  $\mu$  is the mean income. More generally, the average absolute deviation from the median is always less than or equal to the average absolute deviation from any income quartile.

Therefore, we measure alienation as the relative absolute deviation from the median income. If alienation measured like this is greater than a threshold, then it remains greater when measured as deviations from any other fixed income. This method of measuring alienation has already been used by [Foster and Wolfson \(2010/1992\)](#):

$$a(y) = \left| \frac{y}{m} - 1 \right|. \quad (18)$$

[Equation \(18\)](#) defines the degree of alienation attributed to individuals at each level of income. Using this measure in the average alienation expressed in [Equation \(7\)](#) leads to:

$$\bar{a} = \int \left( \int \left| \frac{y}{m} - 1 \right| dF(x) \right) dF(y) \quad (19)$$

$$= \int \left| \frac{y}{m} - 1 \right| dF(y). \quad (20)$$

### 2.3 Comparison principle

Consider two distributions from two distinct populations  $A$  and  $B$  that we want to compare in the alienation component of polarization in the sense described previously. Let  $P_A$  be the polarization index in population  $A$  as defined in [Duclos et al. \(2004\)](#):

$$P_A = \iint p(a, i) dF(a, i) \quad (21)$$

$$= \int_0^\infty \int_0^\infty p(a, i) f(a, i) da di \quad (22)$$

where  $p(., .)$  is the antagonism function taking alienation and identification as arguments,  $F(., .)$  is the joint cumulative distribution function of alienation and identification and  $f(., .)$  the corresponding density. Alienation and identification are two variables defined in  $\mathbb{R}^+$ .

When identification is neglected, the antagonism function does not depend on it. So,  $p(a, i) \propto p(a)$  and we can set

$$P_A = \int_0^\infty p(a) f(a) da \quad (23)$$

where  $f(a)$  is the density of alienation. To integrate  $P_A$ , we proceed first to the substitution  $\alpha = -a$ ; then

$$P_A = \int_{-\infty}^0 p(-\alpha) f(-\alpha) d\alpha \quad (24)$$

Next, integrating  $P_A$  by parts yields

$$P_A = p(-\alpha) \int_{-\infty}^\alpha f(-a) da \Big|_{\alpha=-\infty}^{\alpha=0} + \int_{-\infty}^0 p^1(-\alpha) \left( \int_{-\infty}^\alpha f(-a) da \right) d\alpha \quad (25)$$

$$= p(0) \int_{-\infty}^0 f(-a) da + \int_{-\infty}^0 p^1(-\alpha) \left( \int_{-\infty}^\alpha f(-a) da \right) d\alpha \quad (26)$$

$$= p(0) \int_0^\infty f(a) da + \int_0^\infty p^1(\alpha) \left( \int_\alpha^\infty f(a) da \right) d\alpha \quad (27)$$

where  $p^1$  is the first-order derivative of the antagonism function with respect to its first argument (alienation).

Let  $z_a \geq 0$  be a value of alienation and let  $H_A(z_a) = \int_{z_a}^\infty f_A(a) da$  denote the alienation dominance surface for population  $A$ . Hence,  $H_A(0) = 1$  and

$$P_A = p(0) + \int_0^\infty p^1(\alpha) H_A(\alpha) d\alpha. \quad (28)$$

The dominance surface is defined analogously for distribution  $B$  where [Equation \(28\)](#) is also applicable:

$$P_B = p(0) + \int_0^\infty p^1(\alpha) H_B(\alpha) d\alpha. \quad (29)$$

Let  $I^1$  denote the class of polarization indices in the form of Equation (23) where the antagonism function is increasing in alienation and constant in identification:

$$I^1 = \{P \text{ in the form of Equation (23)} \mid p^1 \geq 0\}. \quad (30)$$

We are searching conditions under which  $P_A - P_B \geq 0$  for any  $P \in I^1$ :

$$P_A - P_B = \int_0^\infty p^1(\alpha) (H_A(\alpha) - H_B(\alpha)) d\alpha. \quad (31)$$

Consider  $P \in I^1$  and suppose that  $H_A(\alpha) - H_B(\alpha) \geq 0$  whatever the value of alienation  $\alpha \geq 0$ . Since  $p^1 \geq 0$ ,  $\Delta P = P_A - P_B \geq 0$  for the integral of a positive function in a closed interval is positive. That is, the condition  $H_A(\alpha) - H_B(\alpha) \geq 0 \forall \alpha \geq 0$  is a sufficient condition to compare distributions  $A$  and  $B$  in the alienation component of polarization when polarization indices are taken in the class  $I^1$ .

For the necessity condition, suppose that  $P_A - P_B \geq 0$  for all  $P \in I^1$  and that on the other side,  $H_A(\alpha) - H_B(\alpha) < 0$  for  $\alpha \in [\bar{\alpha}, \bar{\alpha} + \varepsilon]$  for a small value of  $\varepsilon > 0$ . We define an antagonism function by:

$$p(\alpha, \mathfrak{t}) \propto p(\alpha) = \begin{cases} \bar{\alpha} & \text{if } \alpha \leq \bar{\alpha} \\ \alpha & \text{if } \bar{\alpha} < \alpha \leq \bar{\alpha} + \varepsilon \\ \bar{\alpha} + \varepsilon & \text{if } \alpha > \bar{\alpha} + \varepsilon \end{cases}$$

The function  $p(\alpha)$  is increasing in  $\alpha$  and  $\in I^1$ . The first derivative of that function is  $p^1(\alpha) = \frac{\partial p(\alpha)}{\partial \alpha} = 1$  in  $[\bar{\alpha}, \bar{\alpha} + \varepsilon]$  and 0 elsewhere. So, Equation (31) becomes:

$$P_A - P_B = \int_{\bar{\alpha}}^{\bar{\alpha} + \varepsilon} (H_A(\alpha) - H_B(\alpha)) d\alpha. \quad (32)$$

With the hypothesis that  $H_A(\alpha) - H_B(\alpha) < 0$ , we also have  $\int_{\bar{\alpha}}^{\bar{\alpha} + \varepsilon} (H_A(\alpha) - H_B(\alpha)) d\alpha < 0$  which implies that  $P_A - P_B < 0$  and contradicts the assumption of  $P_A - P_B \geq 0$ . The necessary condition is then verified.

From now on, we denote by  $z_a$  a threshold of alienation. When  $H_A(z_a) - H_B(z_a) \geq 0$  for all  $z_a \geq 0$ , we say that the inequality (alienation) component of polarization in distribution  $A$  is always greater or equal than that in distribution  $B$ . We note  $B \succcurlyeq A$  to express that distribution  $B$  stochastically dominates distribution  $A$  in the alienation component of polarization.

This relation reports that distribution  $B$  is more concentrated around its median and less in its tails than distribution  $A$ , or that groups in  $B$  are less heterogeneous than groups in  $A$ . According to the necessity condition, if there exists a level of alienation  $z_a$  for which  $H_A(z_a) - H_B(z_a) < 0$ , then we argue that the inequality (alienation) component of polarization in distribution  $A$  is not always greater or equal than that in distribution  $B$ .

**Proposition 2.1**  $I^1$  dominance: Given two distributions from two distinct populations  $A$  and  $B$ ,  $B \succcurlyeq A$  for all  $P \in I^1$  iff  $H_A(z_a) - H_B(z_a) \geq 0$  for all  $z_a \geq 0$ .

### 3 Case of income distributions

This section develops an alienation dominance surface for particular cases of income distributions. Income refers to any cardinal variable, such as wealth or wage. We intend to show in this section how the alienation dominance surface is distinct from inequality and how it is linked to a middle class measurement and bi-polarization.

#### 3.1 Alienation dominance surface

Let  $Y^A = y_1^A, y_2^A, \dots, y_{N_A}^A$  and  $Y^B = y_1^B, y_2^B, \dots, y_{N_B}^B$  be two vectors, of sizes  $N_A$  and  $N_B$ , of independent and identically distributed income observations randomly drawn from two populations  $A$  and  $B$ . The objective is to compare the two distributions in the alienation component of polarization. Alienation can then be measured for each value of income using [Equation \(18\)](#). For an alienation threshold, say  $z_a$ , consider the alienation dominance surface

$$H_A(z_a) = \int_{z_a}^{\infty} dF(a_A(y)) \quad (33)$$

$$= \int I(a_A(y) \geq z_a) dF_A(y) \quad (34)$$

where  $F_A(\cdot)$  is the cumulative distribution function of income in a population  $A$  and  $I(\cdot)$  the indicator function which is equal to 1 if its argument is true and 0 otherwise.

We can account now for the expression of alienation from [Equation \(18\)](#) and rewrite [Equation \(34\)](#) as follows:

$$H_A(z_a) = \int I\left(\left|\frac{y}{m_A} - 1\right| \geq z_a\right) dF_A(y) \quad (35)$$

$$= \int I(y \leq (1 - z_a)m_A) dF_A(y) + \int I(y \geq (1 + z_a)m_A) dF_A(y) \quad (36)$$

$$= \int_0^{(1-z_a)m_A} f_A(y) dy + \int_{(1+z_a)m_A}^{\infty} f_A(y) dy \quad (37)$$

$$= F_A((1 - z_a)m_A) + 1 - F_A((1 + z_a)m_A) \quad (38)$$

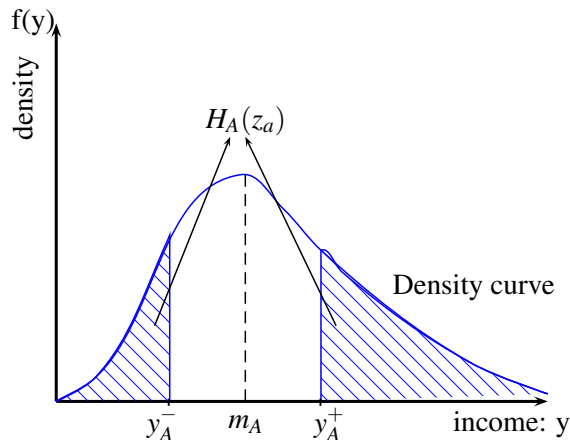
where  $m_A$  is the median of the income distribution in population  $A$ . [Proposition 2.1](#) becomes

$$B \succcurlyeq A \text{ for all } P \in I^1 \text{ iff } F_A((1 + z_a)m_A) - F_A((1 - z_a)m_A) \leq F_B((1 + z_a)m_B) - F_B((1 - z_a)m_B). \quad (39)$$

Thus, fixing a threshold  $z_a$  induces an interval which is a function of  $z_a$  and the median income, and within which each value of income must fall. The problem consists then to assess and compare the proportion of individuals for whom incomes are outside this interval  $[y_A^-, y_A^+]$  where  $y_A^- = (1 - z_a)m_A$  and  $y_A^+ = (1 + z_a)m_A$  for a population  $A$ . We conceptualize this interval graphically using the income density.

The interval  $[y_A^-, y_A^+]$  and the dominance surface  $H_A(z_a)$  are expressed in the space of income as shown on Figure 1.

Figure 1: Alienation dominance surface for a population A



Increasing a threshold ( $z_a$ ) widens the interval  $[y^-, y^+]$  and diminishes  $H(z_a)$ , which tends to zero for large values of thresholds. When the alienation threshold is greater than 1, the lower bound  $y^- = (1 - z_a)m$  is negative and can be set to zero if income cannot be negative. In the same manner, when  $z_a$  decreases, the interval  $[y^-, y^+]$  narrows and  $H(z_a)$  increases and equals 1 when  $z_a = 0$ . Remark that  $H(1)$  measures the proportion of individuals whose incomes are greater than twice the median.

### 3.2 Difference with inequality and (bi)polarization

The Pigou-Dalton principle of transfers stipulates that any rank preserving progressive transfer reduces inequality. Similarly, any rank preserving regressive transfer increases inequality. We are testing whether this principle is valid for the alienation dominance surface  $H(z_a)$ . In the income space (Figure 1), individuals are split in three groups of those below the lower bound ( $y^-$ ), those situated in the interval  $[y^-, y^+]$  and those above the upper bound ( $y^+$ ). Any rank preserving transfer within one of these groups will not affect  $H(z_a)$ , but will affect inequality. For instance, suppose a rank preserving progressive (regressive) transfer within one of these three groups, then inequality will decrease (increase), but the proportion of individuals outside the interval  $[y^-, y^+]$  will not change, that is,  $H(z_a)$  remains unchanged.

Suppose an economy with seven individuals with incomes 10\$, 30\$, 40\$, 50\$, 70\$, 90\$ and 100\$. The median of this distribution is 50\$. Let's set the alienation thresholds  $z_a = 0.5$ . Then  $y^- = 25$ ,  $y^+ = 75$ , and  $H(0.5) = 3/7$ . Suppose a transfer of 10\$ from the second individual to the fifth individual.

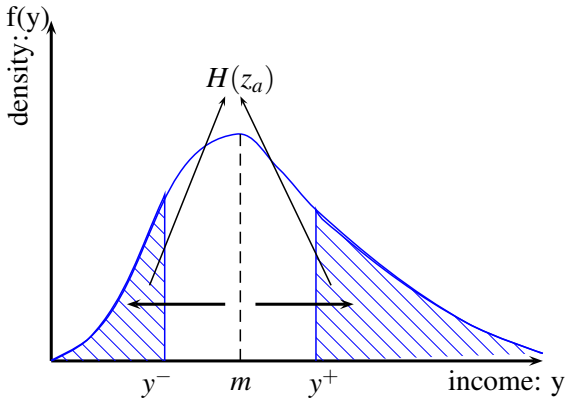
The new distribution is 10\$, 20\$, 40\$, 50\$, 80\$, 90\$ and 100\$. The median has not changed, the two individuals have moved outside the interval  $[y^-, y^+]$  and  $H(0.5) = 5/7$  that has increased. Inequality has also increased following the regressive transfer between the two individuals. Then suppose a transfer of ten dollars occurs between the sixth and the seventh individuals in any sense. Inequality increases if it is a progressive transfer and decreases if it is a regressive transfer, but the alienation dominance surface does not change.

Thus, though it is qualified as the inequality component of polarization, our measure of alienation dominance surface is different from inequality since it sometime fails to respect the Pigou-Dalton principle of transfers as shown above.

Bi-polarization is characterized by increased spread that is the movement away from the median and increased bi-polarity that is the clustering of individuals above and below the median each around their mean incomes. The method of FW measures bi-polarization in the dual space whereas an alienation dominance surface measures it in the primal space. An alienation dominance surface and bi-polarization are affected in the same manner by increased spread, but the impact of increased bi-polarity and that of transfers across the median are ambiguous on an alienation dominance surface.

In the case of increased spread, individuals are moving towards the tails. Bi-polarization as well as the alienation dominance surface  $H(z_a)$  increase regardless of the value of the alienation threshold  $z_a$  (see Figure 2). The above example with seven persons illustrates the case. The second and the fifth individuals are moving from the median to the tails of the distribution.

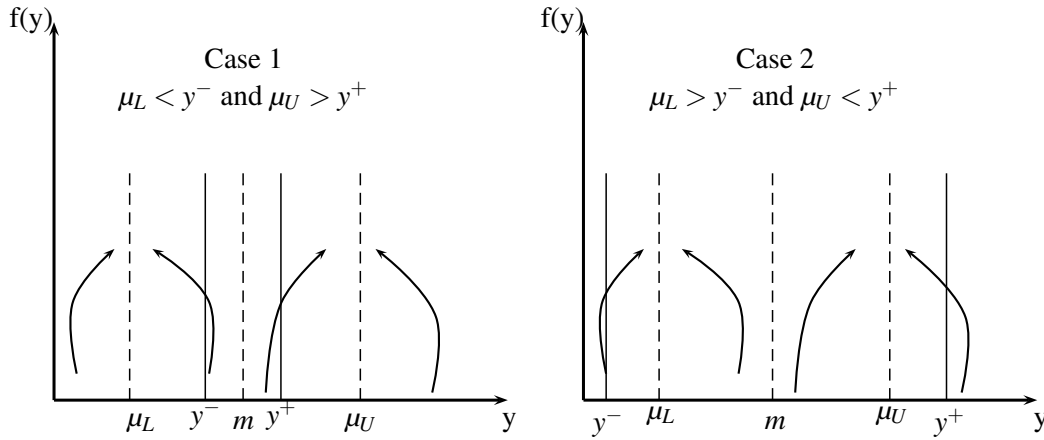
Figure 2: Increased spread increases bi-polarization and alienation dominance surface regardless of alienation threshold



Then consider increased bi-polarity where individuals above and below the median each spread around their mean incomes. Its impact on an alienation dominance surface depends on the size of the alienation threshold and is ambiguous. For small values of  $z_a$ ,  $y^-$  and  $y^+$  are close to the median, the mean of incomes below the median ( $\mu_L$ ) may be lower than  $y^-$ , and the mean of those above the median

( $\mu_U$ ) may be higher than  $y^+$ . This is the situation of case 1 in Figure 3 where increased bi-polarity can increase the alienation dominance surface. As  $z_a$  increases, the interval  $[y^-, y^+]$  enlarges,  $\mu_L$  may be higher than  $y^-$ , and  $\mu_U$  may be lower than  $y^+$ . The case 2 in Figure 3 describes this situation in which increased bi-polarity decreases the alienation dominance surface. However, another situation may occur where nothing can be said about the dominance surface. Take the situation of Figure 4 where  $\mu_L < y^-$  and  $\mu_U < y^+$  as  $z_a$  moves. In increased bi-polarity, the increased spread on the left side of the median increases the alienation dominance surface while the decreased spread on its right induces the opposite effect. The overall effect on the alienation dominance surface is undetermined. An ambiguity is also obtained when  $\mu_L > y^-$  and  $\mu_U > y^+$  as  $z_a$  changes.

Figure 3: Increased bi-polarity can increase (case 1) or decrease (case 2) the alienation dominance surface



Case 1: individuals are moving outside  $[y^-, y^+]$  and the alienation dominance surface increases.

Case 2: individuals are moving inside  $[y^-, y^+]$  and the alienation dominance surface decreases.

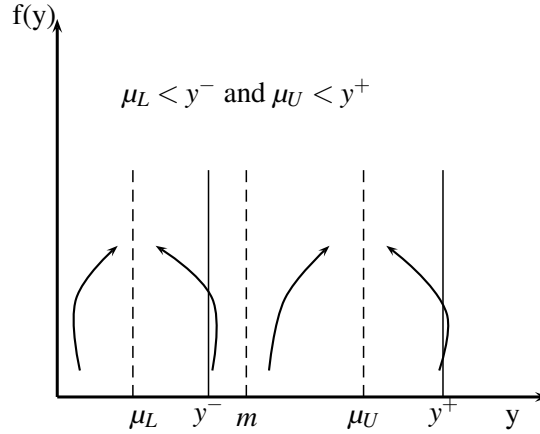
Likewise, transfers that displace individuals from one side of the median to another can change the median and can increase or decrease the alienation dominance surface. When the median increases,  $y^-$  and  $y^+$  also increase, the interval  $[y^-, y^+]$  widens and the alienation dominance surface decreases. But when the median decreases,  $y^-$  and  $y^+$  also decrease, the interval  $[y^-, y^+]$  narrows in and the alienation dominance surface increases.

In the dual space, the first order dominance curve in bi-polarization is defined with median-normalized distances between the median and incomes of population percentiles:

$$S(p) = \frac{1}{m} |F^{-1}(p) - m| \quad (40)$$

for  $0 \leq p \leq 1$ . In the primal space, the alienation dominance surface defines a first order dominance curve in alienation with the sum of the proportion of individuals whose incomes are lower than  $1 - z_a$

Figure 4: Increased bi-polarity can have an ambiguous effect on the alienation dominance surface



*On the left side of the median some individuals are moving outside  $[y^-, y^+]$  while on its right, some individuals are moving inside. The impact on the alienation dominance surface is undetermined.*

times the median and the proportion of those whose incomes are greater than  $1 + z_a$  times the median across alienation thresholds  $z_a$ :

$$H(z_a) = F((1 - z_a)m) + 1 - F((1 + z_a)m) \quad (41)$$

for  $z_a \geq 0$ . A comparison between the two curves is possible only for values of alienation thresholds less than or equal to 1 since  $S(z_a)$  is not defined for  $z_a > 1$ . A higher value of  $S$  means that there are fewer incomes near the median. This remark is true for a higher value of  $H$ . So,  $H$  can be considered as a measure of bi-polarization like  $S$ .

### 3.3 Link with middle class measurements

The alienation dominance surface for a given alienation threshold also relates to the size of the middle class. The most influential measures of the middle class have relied on the position of two income cutoffs around the median and have defined the middle class as the share of the population with income within these cutoffs. The middle class is hence defined as the share of the population with incomes between 75% and 125% of the median income by [Thurow \(1984\)](#); [Blackburn and Bloom \(1985\)](#) broaden that middle income range to 60% to 225% of the median income; in [Leckie \(1988\)](#), the middle class is defined on the basis of an income range of 85% to 115% of the median wage. The resulting middle class index, denoted by  $M$ , is then the share of the population found within these cutoff points. This index is



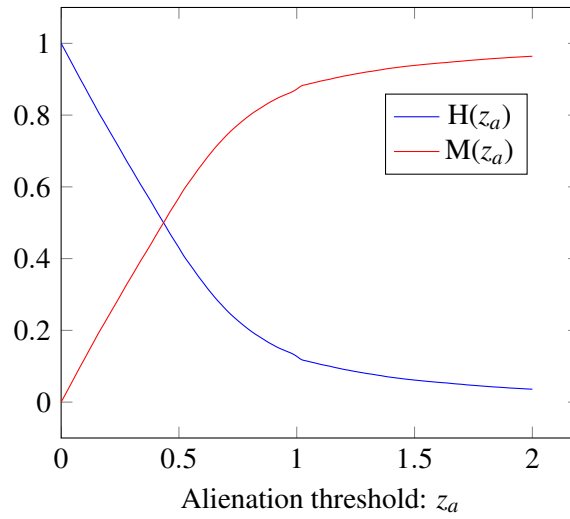
related to the alienation dominance surface upon the definition of the alienation threshold:

$$M(z_a) = 1 - H(z_a) \quad (42)$$

$$= F((1 + z_a)m) - F((1 - z_a)m). \quad (43)$$

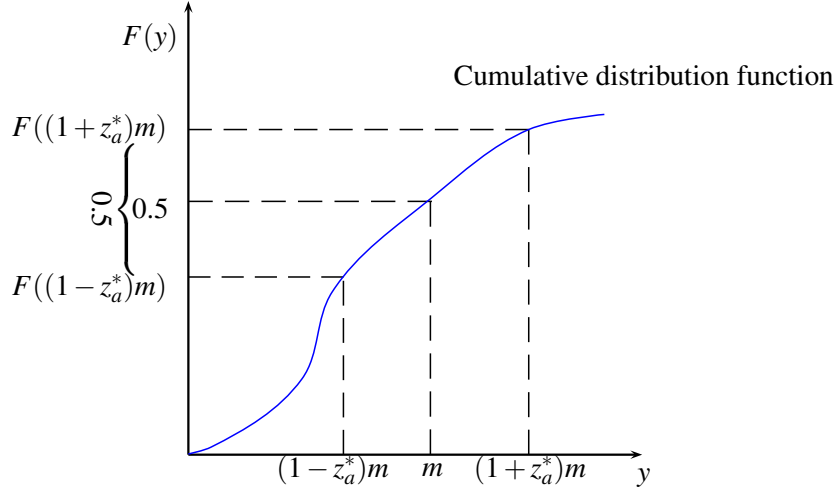
Graphically, the two measures move in opposite directions (Figure 5). Thus, actions which decrease  $H(z_a)$  simultaneously increase the middle class ( $M(z_a)$ ). Likewise, actions which increase  $H(z_a)$  simultaneously decrease the middle class ( $M(z_a)$ ).

Figure 5: Curves of the alienation dominance surface and the size of the middle class as a function of alienation thresholds



The two curves cross at the alienation threshold  $z_a^*$  defined by  $(0.5 - F((1 - z_a^*)m)) + (F((1 + z_a^*)m) - 0.5) = 0.5$ . This value of  $z_a^*$  can be determined graphically as in Figure 6 using the cumulative distribution function of income.

Figure 6: Finding the value  $z_a^*$  of the alienation threshold for which  $H(z_a^*) = M(z_a^*)$



## 4 Application

### 4.1 Data and statistics

We apply the above methodology to 2004 data of 22 countries from the Luxembourg Income Study (LIS) database (see [Table 1](#)). The LIS database contains harmonized microdata from countries which are exclusively middle and high-income countries. It also contains variables on household income, taxes and transfers, households and individual characteristics, labor market outcomes, and expenditures. We use disposable household income, defined as post tax and post transfer income of the household.

The household income is adjusted through the division by an adult equivalent scale  $s^{0.5}$  where  $s$  is the household size. This operation is introduced to provide comparable incomes for individuals living in households of different sizes. We also weighted all household observations by the household weight "hweight normalized" multiplied by the number of people in the household unit. Some (less than 0.5%) incomes are negative in each country. We could have set these values to zero but we did not assuming their impact is not significant on results. Income is also normalized by the mean to eliminate the problem of measurement units.

To compute some statistics like the mean and median incomes, the national currencies are converted into US \$ using the purchasing power parities (PPP) in 2005 from the 2005 International Comparison Program. Sample sizes for each country are shown in the first column of [Table 1](#). This table also summarizes the mean and the median of the adjusted household net income of each country set in US \$. In addition to [Table 1](#), we plot [Figure 7](#) which shows that Luxembourg is far from the others in terms of

these two statistics. Its mean income is \$39,657.23, the highest among the 22 countries. It is followed by the USA where the mean income is \$31,616.15 and then Switzerland with mean of \$29,205.55. The lowest mean income is that of Peru (\$4,761.52). For the countries like Mexico, Estonia, Poland and Slovak Republic, the mean incomes are less than 10,000 dollars, but are greater than that of Peru. The last column of [Table 1](#) presents the median income of the set of the 22 countries. The Luxembourg median income found to be \$34,997.17 is still the highest and is far from the rest. It is followed by Switzerland and the USA with the median incomes found to be respectively \$26,652 and \$25,696.50. The lowest values of the median income are those of Peru (\$2,964.09) and Mexico (\$4,425.04). These two countries are situated at the bottom South West in [Figure 7](#). In this last figure, countries are all situated close to 45 degree line meaning that their median incomes are almost equal to their mean incomes.

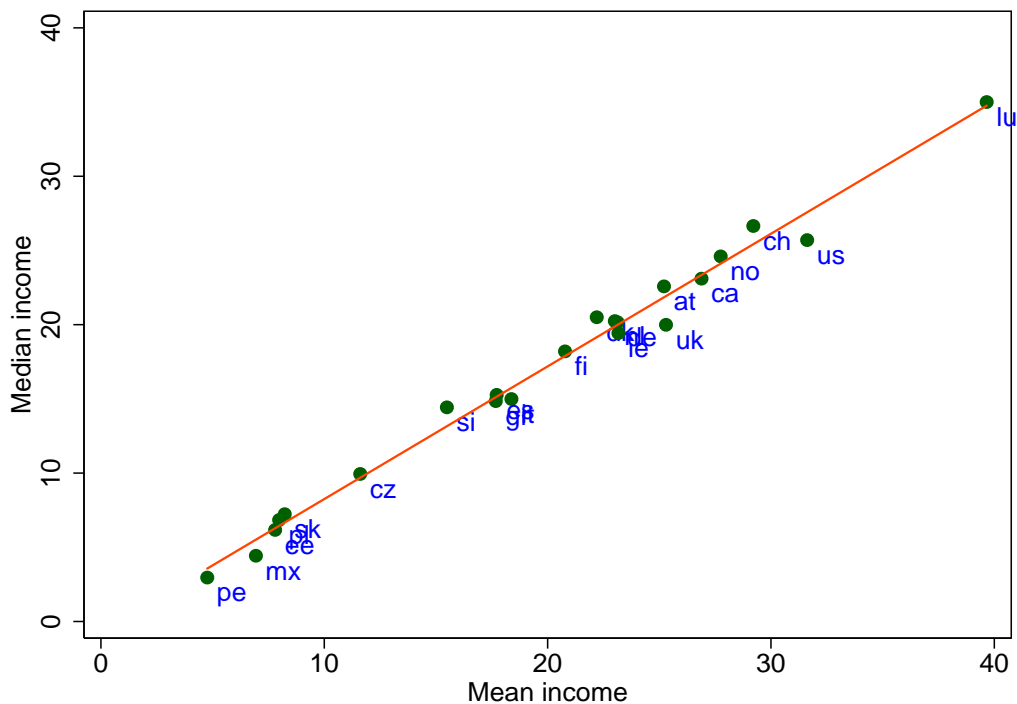
We set the interval of alienation thresholds to be between 0 and 2 for all countries. We first compute the alienation dominance surface for the USA, the country to which we want to compare all others. There is no loss to set the maximum alienation to 2 since few people have income more than three times greater than the median.

Table 1: Sample sizes, mean and median income for 22 countries. Year=2004.

Sigle	Country	Sample size	\$US, ppp=2005	
			Mean	Median
at	Austria	5,147	25,206.96	22,583.33
ca	Canada	27,820	26,886.37	23,103.57
cz	Czech Rep	4,351	11,611.72	9,938.78
dk	Denmark	83,349	22,197.02	20,504.93
ee	Estonia	4,155	7,802.82	6,165.41
fi	Finland	11,229	20,784.08	18,199.34
de	Germany	11,294	23,125.56	20,162.36
gr	Greece	5,568	17,677.8	14,846.15
ie	Ireland	6,080	23,171.57	19,431.43
it	Italy	8,012	18,384.46	14,989.77
lu	Luxembourg	3,622	39,657.23	34,997.17
mx	Mexico	22,595	6,938.19	4,425.04
nl	Netherlands	9,356	23,008.95	20,237.78
no	Norway	13,131	27,754.11	24,599.64
pe	Peru	18,383	4,761.52	2,964.09
pl	Poland	32,214	7,975.03	6,837.35
sk	Slovak Rep	5,147	8,233.35	7,231.4
si	Slovenia	3,725	15,494.17	14,433.74
es	Spain	12,996	17,724.50	15,269.55
ch	Switzerland	3,270	29,205.17	26,652.3
uk	United Kingdom	27,753	25,304.55	19,989.23
us	United States	76,447	31,616.15	25,696.5

*The 2005 ppp figures are from the 2005 International Comparison Program*

Figure 7: Mean and median income (thousands of \$ US) of 22 countries



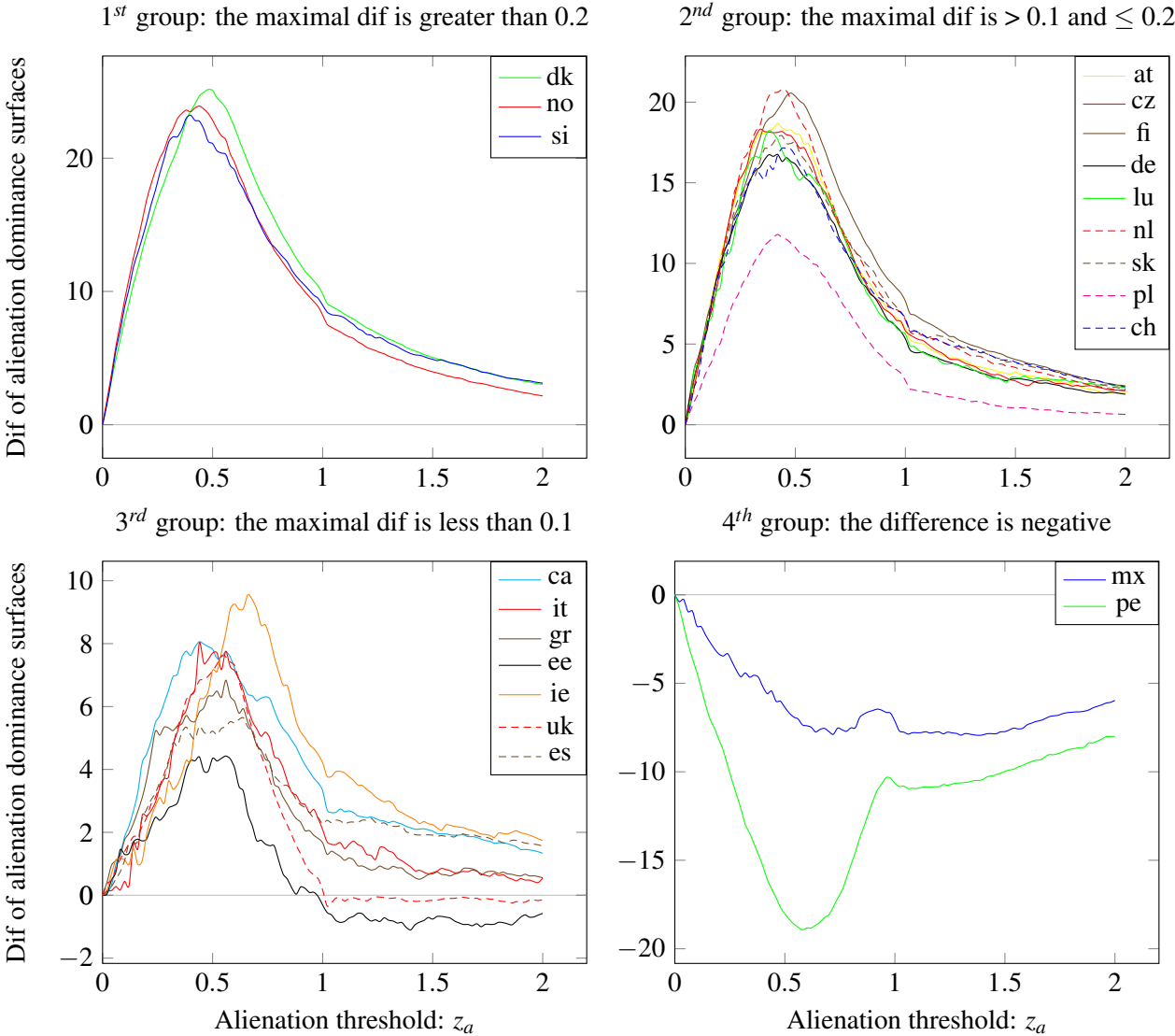
## 4.2 Results

To present results, consider the following question: Among the 22 countries, which ones are better in terms of income concentration around the median, or equivalently which ones are worse in terms of income concentration in the tails of the distribution? We start the application by comparing the 21 other countries to the USA taken as a benchmark. All the countries, except Mexico and Peru, are better than the USA. Whatever the alienation threshold, the alienation dominance surfaces in these countries are lower than that in the USA and their curves always lie under the one of the USA. Thus, these countries stochastically dominate the USA in the alienation component of polarization, but the USA stochastically dominates Mexico and Peru. These results are illustrated both in Figure 8 below and in the Figures in the section on statistical tests between the USA and each country. Figure 8 shows the difference between the alienation dominance surface of the USA and that of the other countries for each value of the alienation threshold. The difference is negative between the USA and Mexico and between the USA and Peru. This confirms that the alienation dominance surface in these two countries is always greater than that in the USA. Income in the USA is more concentrated around its median (with a large middle class) and less in its tails than income in Mexico and Peru.

Except Mexico and Peru, the curves of the difference between the USA alienation dominance surface and the other countries all lie in the positive area of the plan. Incomes in these countries are more concentrated around their medians and less in their tails than income in the USA regardless of the alienation threshold. In [Figure 8](#) the curves of the difference between the USA and Estonia curves and between the USA and UK curves are positive for some alienation thresholds before falling under zero. Before that frontier from which the differences become negative, Estonia and UK stochastically dominate the USA in the alienation component in polarization (see the third panel of [Figure 8](#)).

Then we compare the countries in pairs as shown in [Figure 8](#) and [Table 2](#). Mexico and Peru, whose curves stand out from other countries, are dominated by all the other countries. Above all the curves are located the curve between the USA and Norway, that between the USA and Denmark and the one between the USA and Slovenia. These countries are expected to stochastically dominate all the other countries in a restricted interval of alienation threshold. As  $z_a$  increases the difference between the USA alienation dominance surface and those of the countries of the first group, namely Denmark, Norway and Slovenia exceeds 0.2 for some values of  $z_a$ . For the second group of countries, the difference between their alienation dominance surfaces and that of the USA exceeds 0.1, but is lower than 0.2. The third group of the countries are those for which the difference between their alienation dominance surfaces and the one of the USA never exceeds 0.1 and the fourth group composed of Mexico and Peru are the countries for which the difference between their alienation dominance surfaces and the one of the USA is always negative.

Figure 8: Curves of the difference between the alienation dominance surface of the USA and those of other countries multiplied by 100



In [Table 2](#), we identify three groups of countries which are similar in terms of income concentration in the tails. The first group is composed of Greece, Ireland, Italy and Estonia where the concentration of income in the tails is almost the same. The second group is composed of some European Northern countries like Norway, Finland and Denmark in addition to Slovenia. Finally the third and last group is composed of the countries of Central Europe like Austria, Czech Republic, Germany and Switzerland.

Table 2: Result of the alienation dominance between 22 countries

	at	ca	cz	dk	ee	fi	de	gr	ie	it	lu	mx	nl	no	pe	pl	es	ch	si	sk	uk	us
at		⋈	≈	≈	⋈	≈	≈	⋈	⋈	⋈	≈	⋈	≈	⋈	⋈	⋈	⋈	≈	⋈	≈	≈	⋈
ca			⋈	⋈	⋈	⋈	⋈	≈	≈	≈	⋈	⋈	⋈	⋈	⋈	≈	≈	⋈	⋈	⋈	≈	⋈
cz				≈	⋈	≈	≈	⋈	⋈	⋈	≈	⋈	≈	⋈	⋈	⋈	⋈	≈	⋈	≈	≈	⋈
dk					⋈	≈	⋈	⋈	⋈	⋈	≈	⋈	≈	≈	⋈	⋈	⋈	≈	≈	⋈	⋈	⋈
ee						⋈	⋈	≈	≈	≈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	≈
fi							≈	⋈	⋈	⋈	≈	⋈	≈	≈	⋈	⋈	⋈	≈	≈	≈	≈	⋈
de								⋈	⋈	⋈	≈	⋈	≈	≈	⋈	⋈	⋈	≈	≈	≈	≈	⋈
gr									≈	≈	⋈	⋈	⋈	⋈	⋈	⋈	≈	⋈	⋈	⋈	⋈	⋈
ie										≈	⋈	⋈	⋈	⋈	⋈	≈	≈	≈	⋈	⋈	⋈	⋈
it											⋈	⋈	⋈	⋈	⋈	≈	≈	≈	⋈	⋈	⋈	⋈
lu												⋈	≈	⋈	⋈	⋈	⋈	≈	⋈	⋈	⋈	⋈
mx													⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈
nl														⋈	⋈	⋈	⋈	≈	≈	≈	⋈	⋈
no															⋈	⋈	⋈	⋈	≈	≈	⋈	⋈
pe																	⋈	⋈	⋈	⋈	⋈	⋈
pl																		≈	⋈	⋈	⋈	⋈
es																			⋈	⋈	⋈	≈
ch																				⋈	≈	⋈
si																					⋈	⋈
sk																						⋈
uk																						
us	⋈	⋈	⋈	⋈	≈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	⋈	≈

*Note:* In a cell, the  $\succ$  sign means that the country in line stochastically dominates the country in column; the  $\preceq$  sign means that the country in line is stochastically dominated by the country in column; the  $\approx$  sign means that the curve of the country in line is close to the one of the country in column.

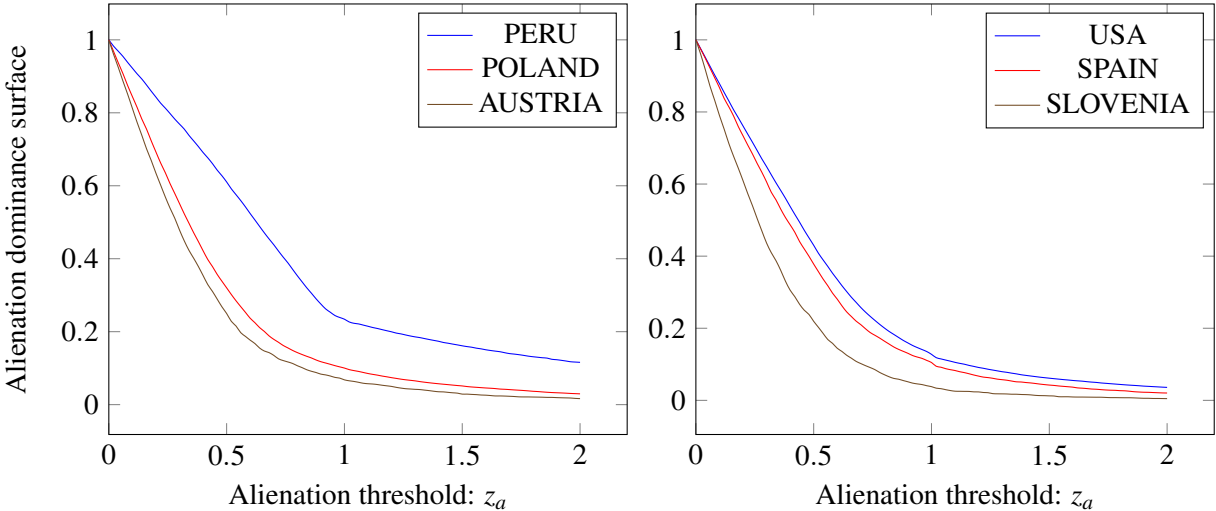
## 5 Estimation and inference

This section is devoted to the development of methods to statistically infer the dominance relation described above. Sometimes, the bulk of observed dominance between two distributions may come from sample noise. For example, suppose Austria is a reference to which we are comparing Poland and Peru.



In Figure 9, we observe that the curve of Austria always lies below those of Poland and Peru, but it is close to that of Poland. Although the curve of Poland is always above the one of Austria, the distance between them is small such that it is difficult to conclude that the two distributions are significantly different in terms of concentration in the tails and that Austria dominates Poland without a statistical test. The two countries might be similar in terms of concentration of income around the median or in the tails of the distribution and the dominance observed in curves could come from noise in samples. The same pattern is observed when the USA is taken as a reference in comparison with Spain and Slovenia. The distance between the USA and Slovenia curves is sufficiently large to indicate the dominance of the USA by Slovenia and that Slovenia is better than the USA. However a conclusion cannot be drawn easily between the USA and Spain. Though the USA curve is always above that of Spain, the distance between the two curves is small and reduce the chance of true dominance of the USA by Spain. It would be more accurate to say that the two distributions are similar in terms of concentration in the queues.

Figure 9: Alienation dominance curves between Peru, Poland and Austria, and between USA, Spain and Slovenia



So, statistical inference allows to guard against wrong conclusions about dominance induced by sample noise. We propose an estimator of the alienation dominance surface and we verify the convergence. Then we test dominance using an Intersection-Union test, an approach introduced by Berger (1982) and used by Berger and Sinclair (1984) and Berger (1996) that consists of choosing the null hypothesis as a union of simple hypotheses and the alternative hypothesis as an intersection of simple hypotheses.

## 5.1 Estimator and critical frontiers

Consider a random sample of  $N$  independently and identically distributed observations of income  $y_1, \dots, y_N$  drawn from a distribution with cumulative distribution function  $F$ . An estimator of  $H(z_a)$  is then:

$$\begin{aligned}\hat{H}(z_a) &= \int I(\hat{a}(y) \geq z_a) d\hat{F}(y) \\ &= N^{-1} \sum_{i=1}^N I(y_i \leq (1 - z_a)\hat{m}) + N^{-1} \sum_{i=1}^N I(y_i \geq (1 + z_a)\hat{m}).\end{aligned}\quad (44)$$

where  $\hat{a}(y) = |\frac{y}{\hat{m}} - 1|$ . The estimator of  $m$  is defined by  $\hat{m} = \inf\{y | \hat{F}(y) \geq \frac{1}{2}\}$  and  $\hat{F}(y) = N^{-1} \sum_{i=1}^N I(y_i \leq y)$  denotes the empirical distribution function of the sample.

The existence of the population moment of order 1 allows us to apply the law of large numbers to Equation (44) from which we deduce that  $\hat{H}(z_a)$  is a consistent estimator of  $H(z_a)$ . The central limit theorem allows us to assert that this estimator is root- $N$  consistent and asymptotically normal. Finally, Theorem 5.1 that follows defines the asymptotic distribution of the estimator of the alienation dominance surface.

**Theorem 5.1**  $N^{\frac{1}{2}}(\hat{H}(z_a) - H(z_a))$  is asymptotically normal with mean zero, and the structure of the variance is given by:

$$\begin{aligned}\lim_{N \rightarrow \infty} N \text{Var}(\hat{H}(z_a) - H(z_a)) &= \frac{1}{4f(m)^2} [-y^+ f(y^+) + y^- f(y^-)]^2 \\ &+ (1 - F(y^+))F(y^+) + (1 - F(y^-))F(y^-) \\ &+ \frac{1}{f(m)} [-y^+ f(y^+) + y^- f(y^-)](1 - F(y^+) - F(y^-)).\end{aligned}\quad (45)$$

*Proof: See the Appendix.*

The dominance surface  $H(z_a)$  equals 1 when  $z_a = 0$  and 0 for large values of  $z_a$ . Hence, dominance generally exists in an interval of thresholds excluding 0 and large values of alienation. Consider two income distributions from two populations  $A$  and  $B$  where the interest is about the dominance of  $A$  by  $B$ . For distribution  $B$  to dominate distribution  $A$ , there must exist a threshold  $z_a^-$  from which  $H_A(z_a)$  begins to be greater than  $H_B(z_a)$ :

$$z_a^- = \inf\{z_a | H_B(z_a) < H_A(z_a)\}.\quad (46)$$

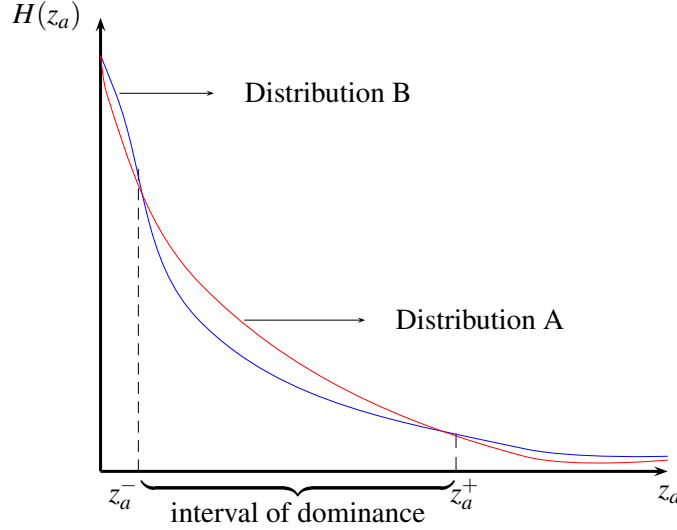
From the lower critical frontier  $z_a^-$ , either  $H_B(z_a) \leq H_A(z_a)$  for all  $z_a \geq z_a^-$  and distribution  $B$  stochastically dominates distribution  $A$  everywhere and the maximum alienation threshold can be set to a large value, or else there exists a reversal threshold up to which dominance applies. The upper critical frontier  $z_a^+$  is the

maximum common alienation threshold up to which distribution  $B$  stochastically dominates distribution  $A$ . It is defined by:

$$z_a^+ = \inf\{z_a > z_a^- | H_A(z_a) < H_B(z_a)\}. \quad (47)$$

Therefore,  $z_a^-$  and  $z_a^+$  are the critical frontiers of the dominance interval as shown in Figure 10.

Figure 10: Critical frontiers and the interval of dominance of distribution A by distribution B



In this section, we also study the convergence in probability of the estimator of the alienation dominance surface through a Monte Carlo simulation. We first take an alienation threshold  $z_a = 0.5$ . Next we draw three sets of 1,000 samples from a log-normal distribution of mean zero and standard deviation 1 with sample sizes set respectively to 1,000, 5,000 and 10,000 observations. We draw the density of the estimators obtained from the 1000 simulated samples. In Figure 11, the density curves narrow in as the sample size increases. This behaviour supports the proof that the estimator  $\hat{H}(z_a)$  is convergent. The same exercise is conducted for different values of alienation threshold,  $z_a = 0.2, 1$  and  $1.5$  and the results shown in Figure 12 lead to the same conclusion of convergence.

Figure 11: Density curves showing the convergence of  $\hat{H}(z_a)$  when  $z_a = 0.5$ . For each size  $N$ , 1,000 samples are drawn from a log-normal distribution of mean zero and standard deviation 1

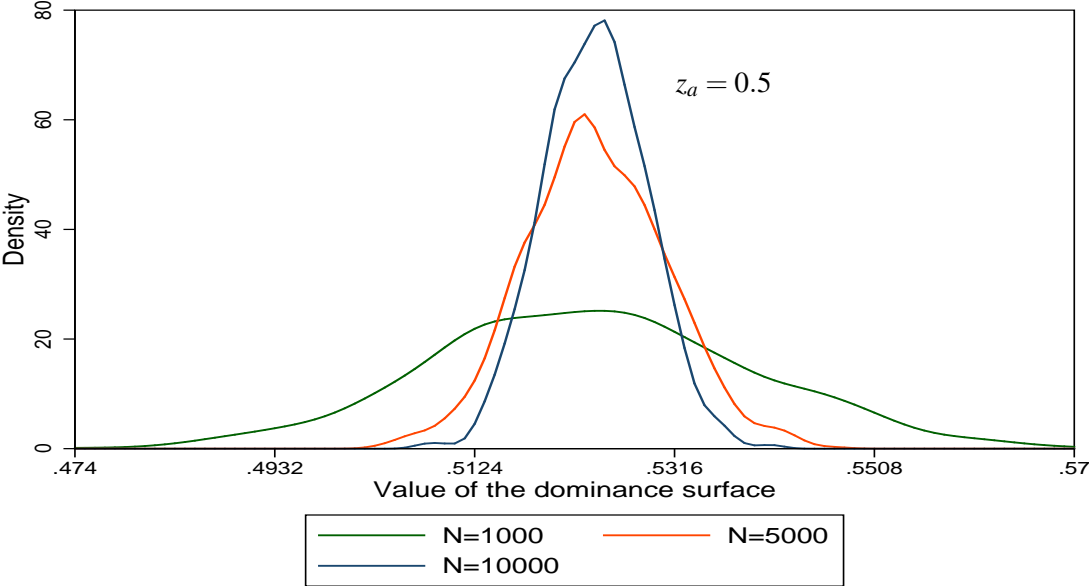
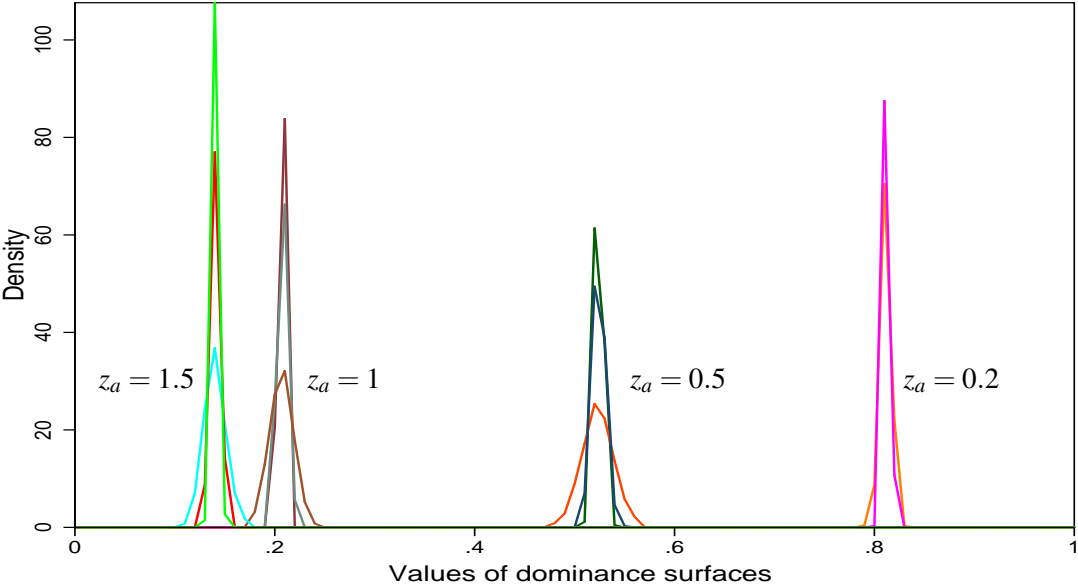


Figure 12: Density curves of  $\hat{H}(z_a)$  when  $z_a = 0.2, 0.5, 1$ , and  $1.5$ . For sizes  $N = 1,000; 5,000$  and  $10,000$ ; 1,000 samples are drawn from a log-normal distribution of mean zero and standard deviation 1



## 5.2 Testing dominance

Testing for dominance implies addressing two issues, that of the choice of the null and alternative hypotheses and that of the test statistic. Davidson and Duclos (2000) present three nested hypotheses that could be used either as the null or the alternative hypothesis in dominance tests. The first one is  $H_0 : H_A(z_a) - H_B(z_a) = 0$  for all  $z_a \leq z_a^+$ . The second one is  $H_1 : H_A(z_a) - H_B(z_a) \geq 0$  for all  $z_a \leq z_a^+$  and the third and last one is  $H_2$  without any restriction on  $H_A(z_a) - H_B(z_a)$ . The three hypotheses are nested in the sense that  $H_0 \subset H_1 \subset H_2$ .

McFadden (1989) proposes to test  $H_0$  against  $H_1$  with the test statistic of the supremum. It is a form of the Kolmogorov-Smirnov test used to test the identity of two distributions. The main difficulty in the test of McFadden is to trace the asymptotic properties of the test statistic under the null hypothesis. Also, it supposes that the distributions have the same sizes, which is not always the case.

The approach proposed by Kaur et al. (1994) [KPS] seems more appropriate. The minimum value of test statistics at different alienation thresholds is used as the test statistic for the null of nondominance against the alternative of dominance. One advantage of this formulation is that if we reject the null hypothesis, all that remains is the alternative of dominance (Davidson and Duclos, 2013). This test is interpreted as the Intersection-Union test (IU) and the probability of rejection of the null when it is true is shown to be asymptotically bounded by the nominal level of a test based on the standard normal distribution (Davidson and Duclos, 2000).

So, we propose the following test of  $H_0^a$  against  $H_1^a$  to test the dominance of  $A$  by  $B$ :

$$\begin{cases} H_0^a : H_A(z_a) \leq H_B(z_a) \text{ for some } z_a \in [z_a^-, z_a^+] \\ H_1^a : H_A(z_a) > H_B(z_a) \text{ for all } z_a \in [z_a^-, z_a^+]. \end{cases}$$

A procedure to perform this test consists of using a predetermined grid of points corresponding in our case to alienation thresholds (Howes, 1993). The only precaution is to be sure to entirely cover the interval of interest. The properties of the IU test using a grid of points are similar to those of the KPS test.

At the tails of the distribution of alienation, the two alienation dominance surfaces are identical. That is  $H_A(0) = H_B(0) = 1$  and  $H_A(z_a) = H_B(z_a) = 0$  for large values of  $z_a$ . Accordingly, it is impossible to reject the null of nondominance at the tails of the distributions.

The null hypothesis is a union of simple hypotheses and the alternative hypothesis is an intersection of simple hypotheses. In this case, the evidence against the null hypothesis exists when it also exists for every simple null hypotheses. If each simple test is of level  $\alpha$ , then the IU test is also of level  $\alpha$  following theorems 1 and 2 of Berger (1982). Furthermore, Berger and Sinclair (1984) show that not only the IU test is of level  $\alpha$ , but it is exactly of size  $\alpha$ .

Consider two distributions drawn from two distinct populations  $A$  and  $B$ . We suppose that the two distributions are independent. The KPS statistic is defined as the minimum over alienation thresholds of simple test statistics. Consider the following simple test of  $H_0^{z_a}$  against  $H_1^{z_a}$  for the dominance of  $A$  by  $B$  when the alienation threshold is  $z_a$ .

$$\begin{cases} H_0^{z_a} : H_A(z_a) \leq H_B(z_a) \\ H_1^{z_a} : H_A(z_a) > H_B(z_a) \end{cases}$$

The test statistic associated to this simple test is

$$t_{z_a} = \frac{\hat{H}_A(z_a) - \hat{H}_B(z_a)}{\sqrt{\text{Var}(\hat{H}_A(z_a)) + \text{Var}(\hat{H}_B(z_a))}}$$

Because of the assumption on the independence of the two distributions, the variance of  $\hat{H}_A(z_a) - \hat{H}_B(z_a)$  is the sum of their variances. Under the null hypothesis, there is no loss to take  $H_A(z_a) = H_B(z_a)$ , a condition under which  $t_{z_a}$  tends asymptotically to the standard normal distribution  $\mathcal{N}(0, 1)$ . Therefore, the test statistic  $t_{z_a}$  is pivotal. At a nominal level  $\alpha$ , there is a strong piece of evidence against the null hypothesis of this simple test if the test statistic  $t_{z_a}$  is greater than the critical value of the standard normal distribution at the level  $\alpha$ .

The test statistic of the IU test is defined by:  $t = \min\{t_{z_a}\}$  for  $z_a$  in the grid of thresholds .

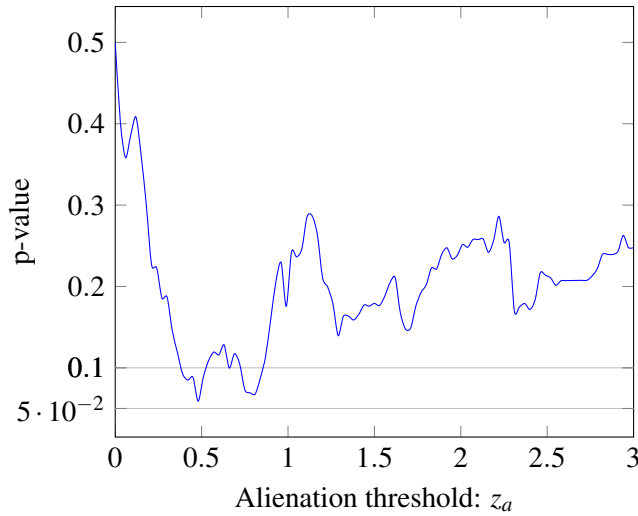
Thus, there is a strong evidence against the null hypothesis of the IU test at the nominal level  $\alpha$  if there is a strong evidence against all simple tests. In that case, the test statistic  $t$  is greater than the critical value of the standard normal distribution at the level  $\alpha$ .

The evidence against the null hypothesis can also be tested using a p-value. The p-value is defined as a measure of the strength of the evidence against the null hypothesis and is determined as the probability of getting a value greater than or equal to the test statistic. Under the null hypothesis, the test statistic above follows the standard normal distribution using the central limit theorem. In consequence,  $\text{p-value} = \int_t^\infty \phi(u) du = 1 - \Phi(t)$  where  $t$  is the test statistic,  $\phi$  and  $\Phi$  are respectively the density and the distribution function of the standard normal distribution. If  $\alpha$  is a given significance level, then reject the null hypothesis if  $\text{p-value} \leq \alpha$ .

A decision can be taken graphically for each significance level of the test. For instance, consider [Figure 13](#) below of the dominance test between two hypothetical distributions  $A$  and  $B$ , on which we have p-values on the vertical axis and alienation thresholds on the horizontal axis. For this illustrative example and for the traditional significance level of 5%, the null hypothesis is never rejected since p-values are always greater than 0.05 regardless of the threshold. However, the null hypothesis is sometimes rejected when the significance level is set to 10%. In this last case, the null hypothesis cannot be rejected until the threshold equals 0.39. From this threshold, p-values remain less than 10% until the threshold equals 0.51. The critical frontiers can then be fixed respectively to  $z_a^- = 0.39$  and  $z_a^+ = 0.51$ , values between

which distribution  $B$  stochastically dominates distribution  $A$ .

Figure 13: Graphical illustration of a statistical test with p-values at different alienation thresholds



### 5.3 Application with p-value using LIS data

We now apply this methodology to real data of the 22 countries that we have used for dominance curves in the previous section. We intend to test dominance which has been observed between the 22 countries. We graph the p-values of tests with the null  $H_A(z_a) \leq H_B(z_a)$  against the alternative that  $H_A(z_a) > H_B(z_a)$  at various values of alienation thresholds  $z_a$  over a range of 0 to 3. The threshold of 3 is large enough for a dominance interval to encompass potential upper critical frontiers. When the threshold is set to 3,  $H(3)$  measures the proportion of individuals whose income is four times greater than the median, a proportion which is very low in most societies.

Graphs are plotted for various pairs of countries. These graphs have an advantage that for different values of significance levels, a conclusion can be drawn on the rejection of the null hypothesis and the critical frontiers can be deduced from the graph.

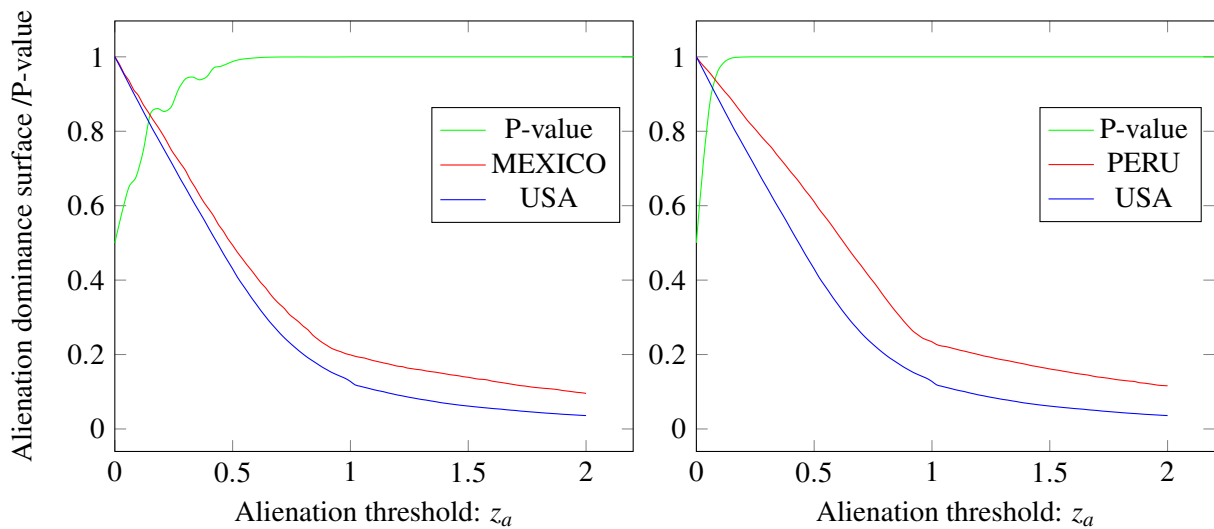
When the threshold is zero, the p-value always equals 0.5, and no dominance can be expected. So the lower critical frontier must be greater than zero in any case. The critical frontiers also depend on the intervals of the grid such that in cases where the lower critical frontier might be close to zero, it will be set to the first point of the grid.

Statistical tests allow to confirm or not the conclusion drawn from dominance curves. Comparing other countries to the USA, the null hypothesis is the non-dominance of the USA by the other coun-

try ( $H_{USA}(z_a) \leq H_{Country}(z_a)$ ) and the alternative is that the USA is dominated by that other country ( $H_{USA}(z_a) > H_{Country}(z_a)$ ).

First consider Mexico and Peru which were shown not to dominate the USA. We are expecting the p-value to be always greater than the significance level of 5% or even 10%. This is exactly what is obtained in Figure 14. The p-values of the test between the USA and Mexico and those between the USA and Peru are never lower than 0.5. This confirms the result obtained with the dominance curves that neither Mexico nor Peru stochastically dominates the USA in the alienation component of polarization. Given that neither the USA and Mexico curves nor the USA and Peru curves were confounded, we run the test the other way around taking the null hypothesis as the alternative and the alternative hypothesis as the null for the three countries. At the significance level of 5%, the test reveals the dominance of Mexico by the USA when the lower critical frontier is  $z_a^- = 0.06$  and the upper critical frontier,  $z_a^+$ , greater than 3. It also shows the dominance of Peru by the USA with the lower critical frontier  $z_a^- = 0.03$  and the upper critical frontier greater than 3.

Figure 14: Alienation dominance curves and p-values at different alienation thresholds for the USA and Mexico, and the USA and Peru



Statistical tests confirm the observed dominance in a restricted alienation interval between the USA and Greece and between the USA and Italy. The lower critical frontiers are far from zero in these countries as in Ireland and Spain and shown in Figure 15 and Table 3. Despite the relative closeness of the dominance curves of these countries with that of the USA, there exist intervals of alienation within which the USA is stochastically dominated by the two other countries. For Greece, the interval of dominance is  $[0.24, 0.83]$  in which the p-value is always lower than 5% and ranges between 0.021 and 0.038. The



dominance interval for Italy is translated on the right compared to that of Greece. Its lower bound is 0.33 and its upper bound 0.96 with p-values of 0.020 for the lower bound and 0.040 for the upper bound. The lower bound of the interval of dominance is 0.42 for Ireland and 0.21 for Spain with p-values set respectively to 0.040 and 0.047. The upper bound is greater than 3 for the two countries.

Figure 15: Alienation dominance curves and p-values at different alienation thresholds for the USA and: Greece, Ireland, Italy and Spain

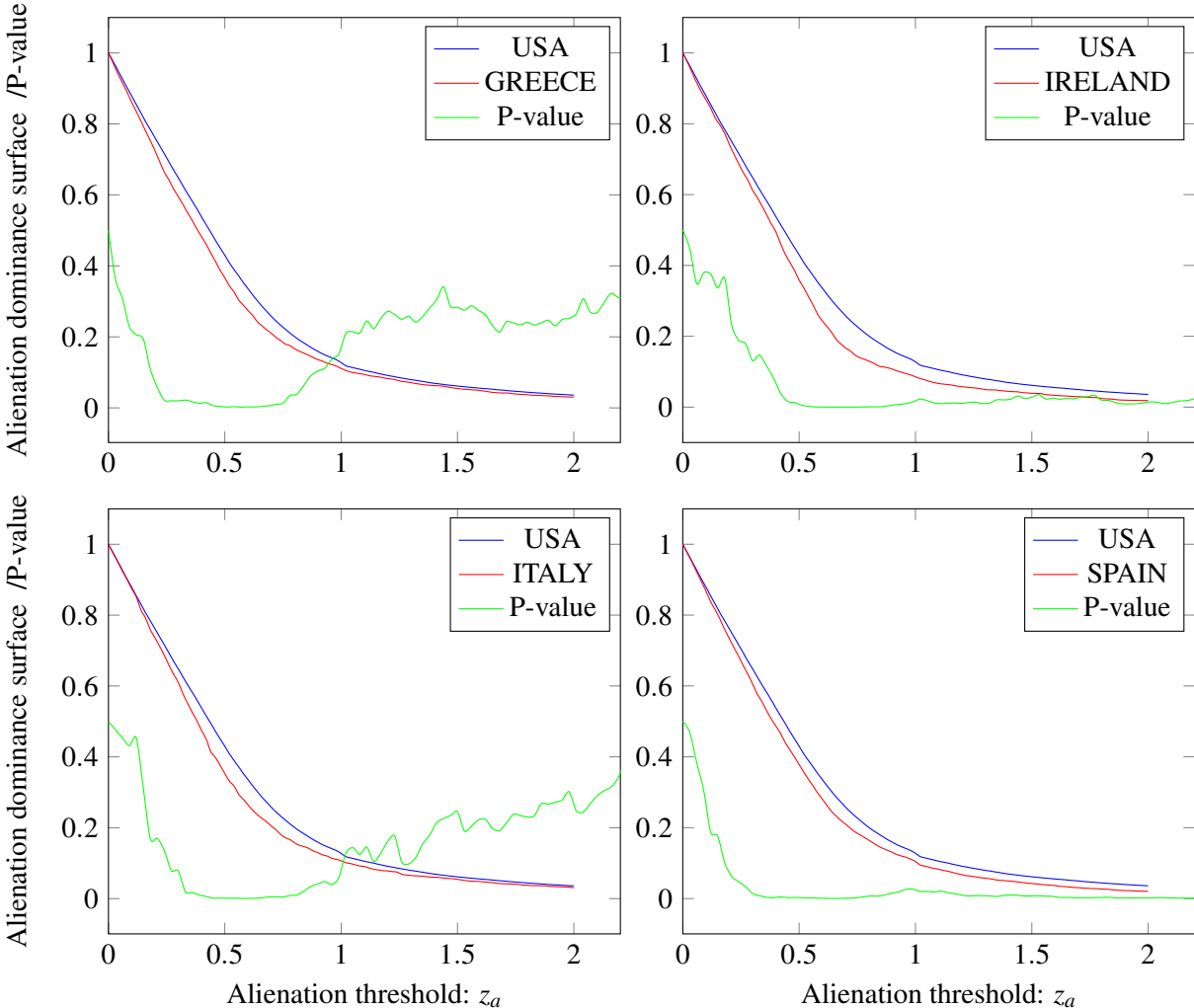


Figure 16 presents the results of tests where the alienation thresholds are set over an interval of the form  $[0, z_a^+]$  for various values of  $z_a^+ \leq 2$ . The null hypothesis is therefore  $H_{USA}(z_a) \leq H_{Country}(z_a)$  for at least some  $z_a$  in  $[0, z_a^+]$  against the alternative that  $H_{USA}(z_a) > H_{Country}(z_a)$  over the entire range

of  $[0, z_a^+]$ . We notice that the estimator of the alienation dominance surface  $\hat{H}(z_a)$  in the USA is always higher than that in other countries, but the distance between the two empirical distribution functions diminishes as  $z_a^+$  increases and tends to 2. At the significance level of 5%, it is not difficult to see that the null hypothesis of non-dominance is rejected in every threshold range in favour of the alternative of dominance for each  $z_a^+$  outside a close neighbourhood of zero, because the p-value is always lower than 0.05. The p-values move rapidly from 0.5 for thresholds very close to zero and then stay at zero for every  $z_a^+$ .

The USA is then stochastically dominated by the other countries for any alienation threshold interval in  $]0, 2]$ . The lower critical frontier is found to be 0.06 for Austria, 0.15 for Canada, 0.09 for the Czech Republic, 0.12 for Luxembourg, 0.06 for Poland, 0.09 for Switzerland, 0.06 for Slovenia and 0.06 for Slovak Republic (See [Table 3](#)). These lower critical frontiers of the dominance interval are greater than the minimal point of the grid (0.03) and represent a clear estimation of  $z_a^-$  for the tests between the USA and other countries. The p-values associated to these lower critical bounds ranges from 0.004 in Germany to 0.048 in Poland.

The null hypothesis of non-dominance is rejected until  $z_a^+ = 2$ , leading to an impossibility to define the upper critical frontiers for all these countries. We perform the test once more setting the alienation thresholds over an interval of the form  $[0, z_a^+]$ , but for various  $z_a^+ \leq 3$ . We verify the results with the above countries for which we first set dominance interval in  $]0, 2]$ . The pattern does not change for  $z_a^+ > 2$ , and dominance still exists until  $z_a^+ = 3$ . The lower critical frontiers did not change, but still the upper critical frontiers appeared to be greater than 3 for all the countries. Thus, the USA is stochastically dominated for  $z_a^+$  up to certain point greater than 3. Situations where  $z_a^+$  are extended to 3 is illustrated in [Figure 17](#) where the p-values are always close to zero until  $z_a^+ = 3$ . This confirms the dominance in the entire interval  $]0, 3]$ .

Figure 16: Alienation dominance curves and p-values at different alienation thresholds for the USA and: Austria, Canada, Czech Rep, Luxembourg, Poland, Switzerland, Slovenia and Slovak Rep

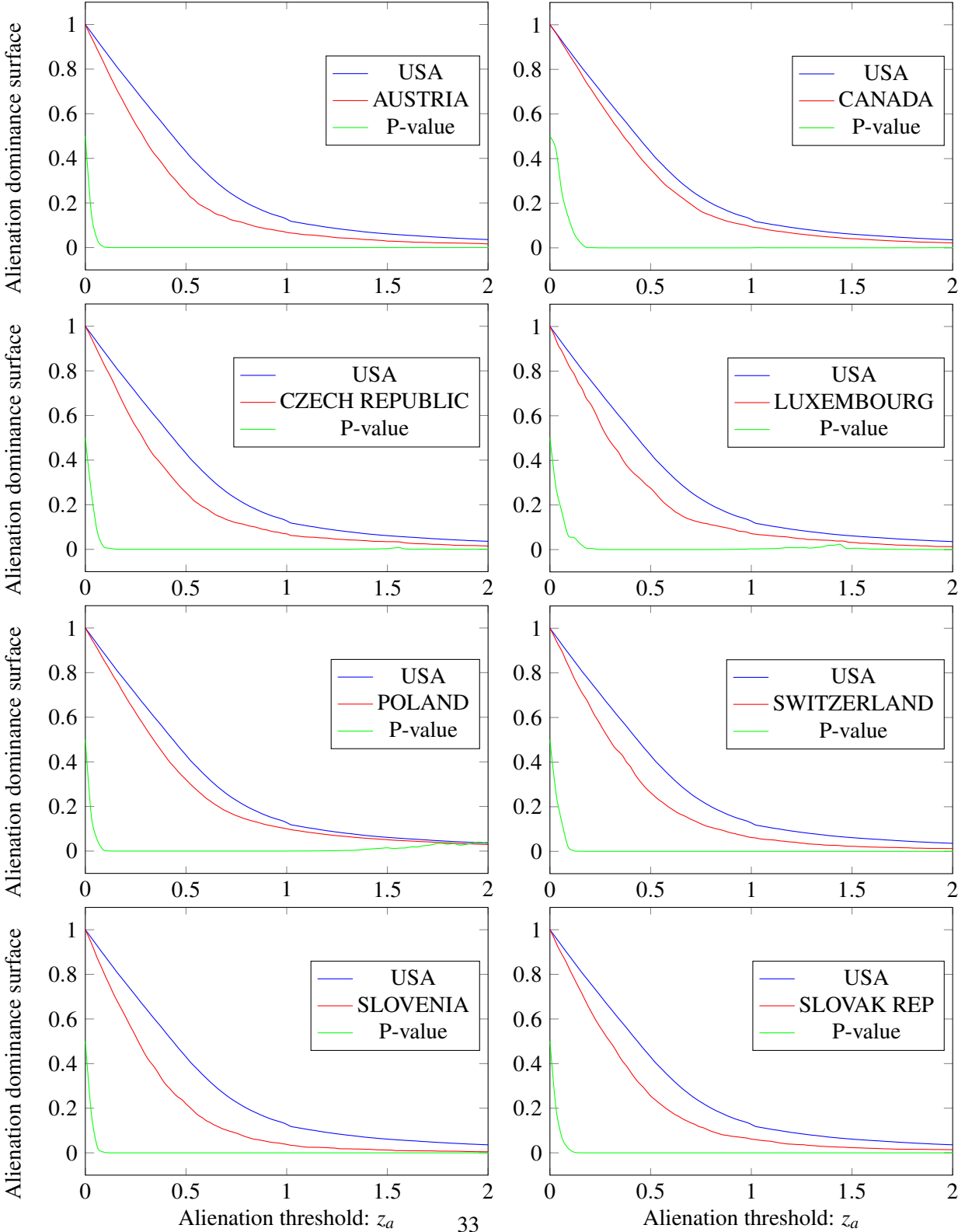
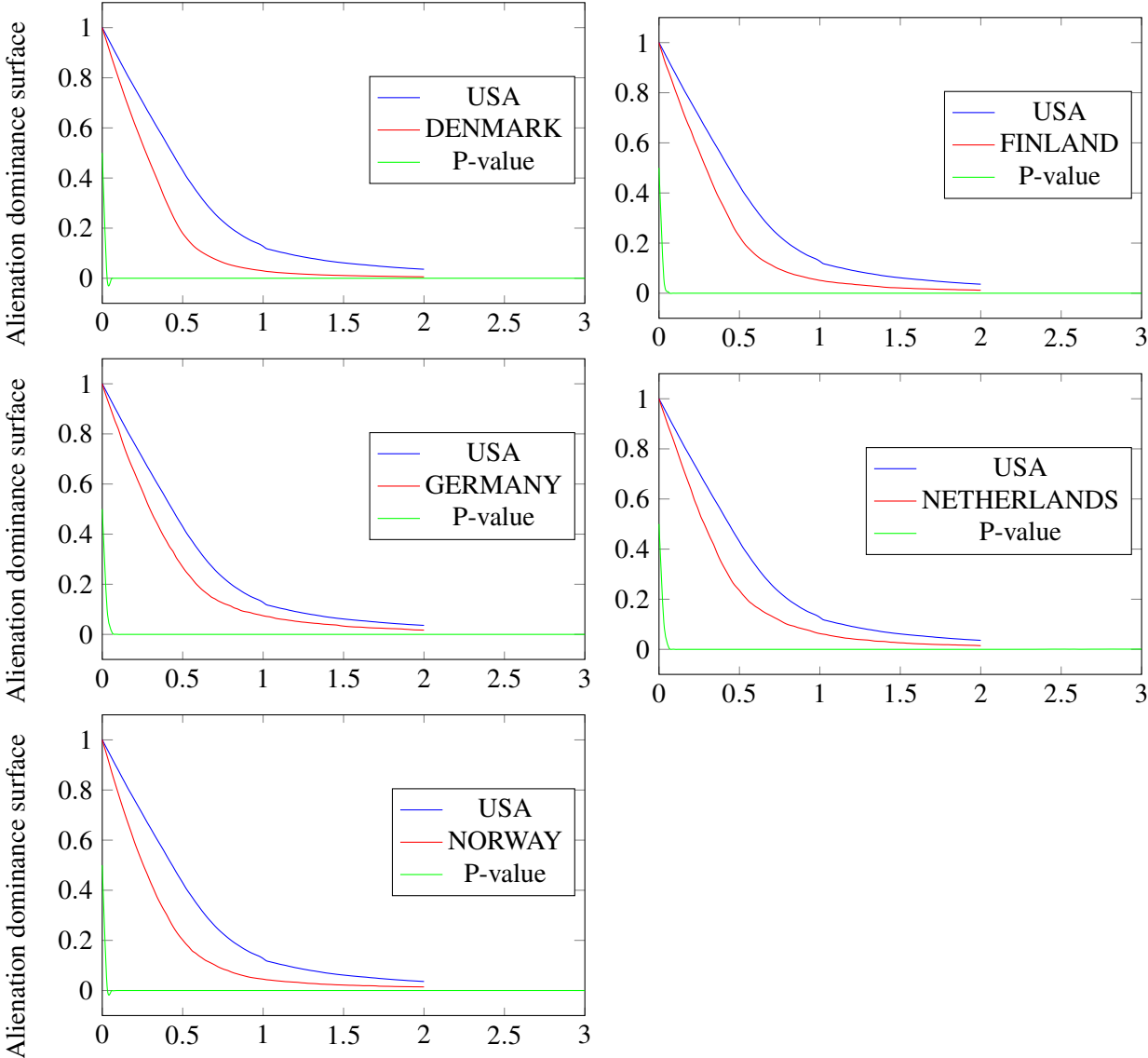


Figure 17: Alienation dominance curves and p-values at different alienation thresholds for the USA and: Denmark, Finland, Germany, Netherlands and Norway

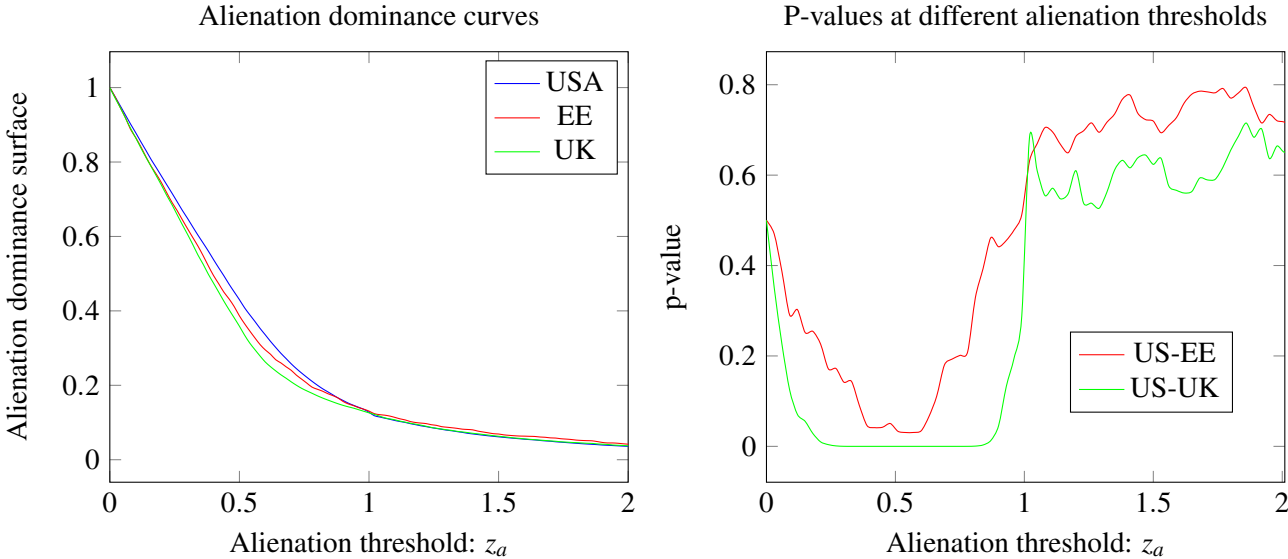


The analysis comparing the USA and other countries shows that dominance always exists when curves are not very close, suggesting dominance when curves are very distant. Consequently, there is no value added to perform the test in pairs between countries especially when the distance between curves is large enough to show clear distinction.

However, when curves are close, we cannot draw strong conclusions with respect to strict domi-

nance according to the alienation dominance curves as we have done previously. Dominance tests reveal that dominance can exist even when distance between curves is small. The examples between the USA and Estonia and between the USA and UK are illustrative of these cases. Using the dominance curves, we have concluded that the USA and UK as well as the USA and Estonia cannot be distinguished in terms of alienation dominance and income concentration in the tails or around the median. On the contrary, statistical tests show dominance between the USA and Estonia on the one hand and between the USA and UK on the other hand. The p-values of the corresponding tests in Figure 18 fall below 0.1 and even below 0.05 for some values of alienation thresholds. The lower and upper critical frontiers between the USA and Estonia are 0.39 with a p-value of 0.046 and 0.60 with a p-value of 0.034. Inside these frontiers, the p-value is always less than 0.05 showing the dominance of the USA by Estonia. The critical frontiers between the USA and UK are 0.18 and 0.9 with corresponding p-values of 0.028 and 0.045. Inside these frontiers, the p-value is lower than 0.05 and is sometimes close to zero, proving that the USA is stochastically dominated by the UK.

Figure 18: Alienation dominance curves and p-values between the USA and Estonia, and the USA and UK at different alienation thresholds



We now apply dominance tests only between countries whose dominance curves are close suggesting they are similar in terms of income concentration in the tails.

Within the first group composed of Greece, Estonia, Ireland and Italy, we have presumed that these countries are similar based on their dominance curves. With statistical tests, this is exactly the case between Greece and Italy and between Greece and Estonia where the absence of dominance is

confirmed. However, statistical tests show dominance of Greece by Ireland in a restricted interval of alienation thresholds with the lower critical frontier equal to 0.09 and the upper critical frontier greater than 3. Also, we notice the dominance of Italy and Estonia by Ireland, and that of Estonia by Italy, contrary to what we have assumed with the dominance curves. The lower critical values are respectively 0.48 for the test between Ireland and Italy, 0.39 for the test between Ireland and Estonia and 0.21 for the test between Italy and Estonia. The upper critical frontiers are greater than 3 in all the three cases.

In the second group of Norway, Denmark, Finland and Slovenia, statistical tests confirm that there is no dominance between Norway and the other three countries. Neither does Norway dominate these countries, nor do these countries dominate Norway. Also, the absence of dominance is confirmed between Finland and Slovenia. Furthermore, statistical tests show that Denmark dominates Finland and Slovenia respectively in the intervals of [0.36, 1.49] and [0.45, 0.99] of alienation thresholds.

In the third group composed of Austria, Czech Republic, Germany and Switzerland, statistical tests confirm that Austria does not dominate the three other countries, nor do these three countries dominate Austria. Tests also confirm that there is no dominance neither between Czech Republic and Germany, nor between Czech Republic and Switzerland or between Germany and Switzerland, exactly what we have presumed with dominance curves.

Table 3: Critical frontiers and corresponding p-values of the tests between the USA and other countries at the significance level of 5%

Country	$z_a^-$ (p-value)	$z_a^+$ (p-value)	Country	$z_a^-$ (p-value)	$z_a^+$ (p-value)
Austria	0.06 (0.034)	> 3 (-)	Mexico	- (-)	- (-)
Canada	0.15 (0.024)	> 3 (-)	Netherlands	0.06 (0.008)	> 3 (-)
Czech Rep	0.09 (0.013)	> 3 (-)	Norway	< 0.03 (-)	> 3 (-)
Denmark	< 0.03 (-)	> 3 (-)	Peru	- (-)	- (-)
Estonia	0.39 (0.046)	0.6 (0.034)	Poland	0.06 (0.048)	2.19 (0.048)
Finland	0.06 (0.004)	> 3 (-)	Slovak Rep	0.09 (0.021)	> 3 (-)
Germany	0.06 (0.009)	> 3 (-)	Slovenia	0.06 (0.026)	> 3 (-)
Greece	0.24 (0.021)	0.83 (0.038)	Spain	0.24 (0.047)	> 3 (-)
Ireland	0.42 (0.040)	> 3 (-)	Switzerland	0.09 (0.021)	> 3 (-)
Italy	0.33 (0.020)	0.96 (0.040)	United Kingdom	0.18 (0.028)	0.9 (0.045)
Luxembourg	0.15 (0.026)	> 3 (-)			

*Numbers in parentheses are the p-values associated to critical frontiers.*

*The "-" sign means that there is no interval for dominance like in Mexico and Peru, the lower frontier is less than 0.0.3, or the upper frontier is higher than 3.*

## 6 Conclusion

This paper examines how to compare the size of the middle class in income distributions using a polarization index that does not account for identification. For that purpose, we define a class of polarization indices where the antagonism function is constant in identification and depends only on alienation. The possibility of comparing income distributions with polarization indices from this class leads to an alienation dominance surface defined as the proportion of individuals for whom the relative deviation of incomes with respect to the median is greater than an alienation threshold. This proportion captures the extent to which an income distribution is far away from the median, is concentrated in the tails and is losing its middle class.

The alienation dominance surface can be used to compare the disappearance of the middle class in a distribution across a range of alienation thresholds. The alienation component of polarization is greater in one distribution than in another regardless of the polarization index in a well-defined class of polarization indices if and only if the former distribution always has a larger alienation dominance surface than the latter regardless of the alienation threshold. The development induces a partial ordering between distributions through alienation dominance curves that can be used to determine whether a distribution has smaller middle class than another.

A greater value of an alienation dominance surface means fewer individuals near the median and many individuals in the tails. Thus, it is the other side of the middle class: when the alienation dominance surface is high, the middle class is low and reciprocally, and when the alienation dominance surface increases, the middle class decreases. A distribution always has smaller middle class than another if its dominance curve is always above that of the other distribution. The alienation dominance surface is different from the measure of inequality since it does not respect the Pigou-Dalton transfer principles, but it bears similarity with bi-polarization dominance.

We derive an estimator of the alienation dominance surface as well as its asymptotic sampling distribution and prove its convergence. Statistical tests are performed with the approach of Intersection Union tests and help confirm the dominance observed with dominance curves. The results of statistical tests are interpreted with p-values which indicate for each value of alienation thresholds whether the dominance holds.

The methodology is then applied to compare pairwise the shrinking of the middle class in income distributions of 22 countries. Incomes distributions in all the countries, except Mexico and Peru, are less concentrated in their tails and have larger middle class than the income distribution in the USA. The analysis leads to restricted dominance in the sense that a country always has smaller middle class than another only in a restricted interval of alienation thresholds. For each pair of countries, we estimated the lower and upper bounds along with corresponding p-values of the interval of dominance.

## APPENDIX

THEOREM 5.1: Distribution of the estimator of the alienation dominance surface.

The alienation dominance surface is defined by:

$$H(z_a) = \int I\left(\left|\frac{y}{m} - 1\right| \geq z_a\right) dF(y) \quad (48)$$

where  $F$  is the cdf of income. It can be decomposed as:

$$H(z_a) = \underbrace{\int_0^{(1-z_a)m} dF(y)}_{y \leq m} + \underbrace{\int_{(1+z_a)m}^{\infty} dF(y)}_{y > m}. \quad (49)$$

An estimator of  $H(z_a)$  is

$$\hat{H}(z_a) = \underbrace{\int_0^{(1-z_a)\hat{m}} d\hat{F}(y)}_{y \leq \hat{m}} + \underbrace{\int_{(1+z_a)\hat{m}}^{\infty} d\hat{F}(y)}_{y > \hat{m}}. \quad (50)$$

We have:

$$\hat{H}(z_a) - H(z_a) = \int_0^{(1-z_a)\hat{m}} d\hat{F}(y) + \int_{(1+z_a)\hat{m}}^{\infty} d\hat{F}(y) - \int_0^{(1-z_a)m} dF(y) - \int_{(1+z_a)m}^{\infty} dF(y). \quad (51)$$

Let  $y^+ = (1+z_a)m$  and  $y^- = (1-z_a)m$ . Then  $\hat{y}^+ = (1+z_a)\hat{m}$  and  $\hat{y}^- = (1-z_a)\hat{m}$ . Let  $\xi^+(y_i) = -(1+z_a) \left( \frac{I(y_i < m) - 0.5}{F'(m)} \right)$ <sup>1</sup> and  $\xi^-(y_i) = -(1-z_a) \left( \frac{I(y_i < m) - 0.5}{F'(m)} \right)$ <sup>2</sup>. Then  $\hat{y}^+ = N^{-1} \sum_{i=1}^N \xi^+(y_i) + o(1)$  and  $\hat{y}^- = N^{-1} \sum_{i=1}^N \xi^-(y_i) + o(1)$ . It follows that  $\hat{y}^+ - y^+ = O(N^{-\frac{1}{2}})$  and  $\hat{y}^- - y^- = O(N^{-\frac{1}{2}})$ .

$$\text{Consider } \hat{H}_1(z_a) = \int_{\hat{y}^+}^{\infty} d\hat{F}(y) - \int_{y^+}^{\infty} dF(y) \quad (52)$$

$$= \int_{\hat{y}^+}^{y^+} d\hat{F}(y) + \int_{y^+}^{\infty} d\hat{F}(y) - \int_{y^+}^{\infty} dF(y) \quad (53)$$

$$= \int_{\hat{y}^+}^{y^+} d(\hat{F} - F)(y) + \int_{\hat{y}^+}^{y^+} dF(y) + \int_{y^+}^{\infty} d\hat{F}(y) - \int_{y^+}^{\infty} dF(y) \quad (54)$$

---

<sup>1</sup>Davidson and Duclos (2000)

<sup>2</sup>Davidson and Duclos (2000)



By properties of the empirical distribution,  $\hat{F} - F = O(N^{-\frac{1}{2}})$  and  $\hat{y}^+ - y^+ = O(N^{-\frac{1}{2}})$  such that the first term of  $\hat{H}_1(z_a)$  is  $O(N^{-1})$  is of negligible order. The second and third terms are of order  $N^{-\frac{1}{2}}$ .

$$\text{Hence } \hat{H}_1(z_a) = O(N^{-1}) + \int_{\hat{y}^+}^{y^+} dF(y) + \int_{y^+}^{\infty} d\hat{F}(y) - \int_{y^+}^{\infty} dF(y) \quad (55)$$

$$= O(N^{-1}) - (F(\hat{y}^+) - F(y^+)) + N^{-1} \sum_{i=1}^N I_{(y_i \geq y^+)} - (1 - F(y^+)) \quad (56)$$

$$\cong -(\hat{y}^+ - y^+)F'(y^+) + N^{-1} \sum_{i=1}^N I_{(y_i \geq y^+)} - (1 - F(y^+)). \quad (57)$$

Using a Taylor expansion at first order, we have  $F(\hat{y}^+) - F(y^+) = (\hat{y}^+ - y^+)F'(y^+)$ .

$$\text{Also consider } \hat{H}_2(z_a) = \int_0^{\hat{y}^-} d\hat{F}(y) - \int_0^{y^-} dF(y) \quad (58)$$

$$= \int_0^{y^-} d\hat{F}(y) + \int_{y^-}^{\hat{y}^-} d\hat{F}(y) - \int_0^{y^-} dF(y) \quad (59)$$

$$= \int_{y^-}^{\hat{y}^-} d(\hat{F} - F)(y) + \int_{y^-}^{\hat{y}^-} dF(y) + \int_0^{y^-} d(\hat{F} - F)(y). \quad (60)$$

Using the properties of the empirical distribution,  $\hat{F} - F = O(N^{-\frac{1}{2}})$  and  $\hat{y}^- - y^- = O(N^{-\frac{1}{2}})$  such that the first term of  $\hat{H}_2(z_a)$  is  $O(N^{-1})$  is of negligible order. The second and third terms are of order  $N^{-\frac{1}{2}}$ .

$$\hat{H}_2(z_a) = \int_{y^-}^{\hat{y}^-} d(\hat{F} - F)(y) + \int_{y^-}^{\hat{y}^-} dF(y) + \int_0^{y^-} d(\hat{F} - F)(y) \quad (61)$$

$$= O(N^{-1}) + F(\hat{y}^-) - F(y^-) + \int_0^{y^-} d\hat{F}(y) - \int_0^{y^-} dF(y) \quad (62)$$

$$\cong (\hat{y}^- - y^-)F'(y^-) + N^{-1} \sum_{i=1}^N I_{(y_i \leq y^-)} - F(y^-). \quad (63)$$

Using a Taylor expansion at first order,  $F(\hat{y}^-) - F(y^-) = (\hat{y}^- - y^-)F'(y^-)$ .

Considering the fact that  $E(\hat{y}^+ - y^+) = E(\hat{y}^- - y^-) = 0$ ,  $E(N^{-1} \sum_{i=1}^N I_{(y_i \geq y^+)}) = 1 - F(y^+)$  and  $E(N^{-1} \sum_{i=1}^N I_{(y_i \leq y^-)}) = F(y^-)$ , it follows that  $E(\hat{H}_1(z_a)) = E(\hat{H}_2(z_a)) = 0$ .

*Asymptotic distribution* : Following the development above, we can state that  $N^{\frac{1}{2}}(\hat{H}(z_a) - H(z_a)) = N^{\frac{1}{2}}(\hat{H}_1(z_a) + \hat{H}_2(z_a))$  is asymptotically normal with mean zero.

*Development of the variance*

We have:

$$\hat{H}(z_a) = -(\hat{y}^+ - y^+)F'(y^+) + N^{-1} \sum_{i=1}^N I_{(y_i \geq y^+)} + (\hat{y}^- - y^-)F'(y^-) + N^{-1} \sum_{i=1}^N I_{(y_i \leq y^-)} \quad (64)$$

The variance of  $\hat{H}(z_a)$  can be defined as follows:

$$\begin{aligned} \text{Var}(\hat{H}(z_a) - H(z_a)) &= [-(1+z_a)F'(y^+) + (1-z_a)F'(y^-)]^2 \text{var}(\hat{m} - m) \\ &+ N^{-1}[(1-F(y^+))F(y^+) + (1-F(y^-))F(y^-)] \\ &+ 2[-(1+z_a)F'(y^+) + (1-y_a)F'(y^-)] \text{cov}(\hat{m} - m, N^{-1} \sum_{i=1}^N (I_{(y_i \geq y^+)} + I_{(y_i \leq y^-)})) \end{aligned} \quad (65)$$

$$\hat{m} - m = -\frac{1}{NF'(m)} \sum_{i=1}^N (I_{(y_i < m)} - 0.5) + O(N^{-3/4}(\log N)^{3/4})^3$$

$$\text{var}(\hat{m} - m) = \frac{1}{NF'^2(m)} \text{var}(I_{(y < m)}) \quad (66)$$

$$= \frac{N^{-1}}{F'(m)^2} F(m)(1-F(m)) = \frac{N^{-1}}{4f(m)^2} \quad (67)$$

$$\text{cov}(\hat{m} - m, N^{-1} \sum_{i=1}^N (I_{(y_i \geq y^+)} + I_{(y_i \leq y^-)})) = \text{cov}\left(-\frac{1}{NF'(m)} \sum_{i=1}^N (I_{(y_i < m)} - 0.5), N^{-1} \sum_{i=1}^N (I_{(y_i \geq y^+)} + I_{(y_i \leq y^-)})\right) \quad (68)$$

$$= N^{-2} \sum_{i=1}^N \text{cov}\left(-\frac{1}{F'(m)} I_{(y_i < m)}, I_{(y_i \geq y^+)} + I_{(y_i \leq y^-)}\right) \quad (69)$$

$$= -\frac{N^{-1}}{F'(m)} \text{cov}(I_{(y < m)}, I_{(y \geq y^+)}) - \frac{N^{-1}}{F'(m)} \text{cov}(I_{(y < m)}, I_{(y \leq y^-)}) \quad (70)$$

$$= \frac{N^{-1}}{F'(m)} F(m)(1-F(y^+)) - \frac{N^{-1}}{F'(m)} [F(\min(m, y^-)) - F(m)F(y^-)] \quad (71)$$

$$= \frac{N^{-1}}{F'(m)} [F(m)(1-F(y^+)) - F(y^-)(1-F(m))] \quad (72)$$

$$= \frac{N^{-1}}{2f(m)} (1-F(y^+) - F(y^-)) \quad (73)$$

$$\begin{aligned} \text{As a result, } \lim_{N \rightarrow \infty} N \text{Var}(\hat{H}(z_a) - H(z_a)) &= \frac{1}{4f(m)^2} [-y^+ F'(y^+) + y^- F'(y^-)]^2 \\ &+ (1-F(y^+))F(y^+) + (1-F(y^-))F(y^-) \\ &+ \frac{1}{f(m)} [-y^+ F'(y^+) + y^- F'(y^-)] (1-F(y^+) - F(y^-)) \end{aligned} \quad (74)$$

---

<sup>3</sup>Davidson and Duclos (2000)

## References

- BERGER, R. L. (1982): “[Multiparameter Hypothesis Testing and Acceptance Sampling](#),” *Technometrics*, 24, 295–300.
- (1996): “[Likelihood Ratio Tests and Intersection-Union Tests](#),” *Institute of Statistics Mimeo Series*.
- BERGER, R. L. AND D. F. SINCLAIR (1984): “[Testing Hypotheses concerning Unions of Linear Subspaces](#),” *Journal of the American Statistical Association*, 79, 158–163.
- BLACKBURN, M. AND D. BLOOM (1985): “[What is happening to the middle class?](#)” *American Demographics/Advertising Age*, 7, 19–25.
- DAVIDSON, R. AND J.-Y. DUCLOS (2000): “[Statistical Inference for stochastic dominance and for the measurement of poverty and inequality](#),” *Econometrica*, 68, 1435–1464.
- (2013): “[Testing for Restricted Stochastic Dominance](#),” *Econometric Reviews*, 32, 84–125.
- DUCLOS, J.-Y., J.-M. ESTEBAN, AND D. RAY (2004): “[Polarization: Concepts, Measurement, Estimation](#),” *Econometrica*, 72, 1737–1772.
- EASTERLY, W. (2001): “[The Middle Class Consensus and Economic Development](#),” *Journal of Economic Growth*, 6, 317–335.
- ESTEBAN, J.-M. AND D. RAY (1994): “[On the Measurement of Polarization](#),” *Econometrica*, 62, 819–851.
- FOSTER, J. E. AND M. C. WOLFSON (2010/1992): “[Polarization and the decline of the middle class: Canada and the U.S.](#)” *Journal of Economic Inequality*, 8, 247–273.
- HOWES, S. (1993): “Asymptotic Properties of Four Fundamental Curves of Distributional Analysis,” *Unpublished Paper, STICERD, London School of Economics*.
- KAUR, A., B. PRAKASA RAO, AND H. SINGH (1994): “[Testing for Second-Order Stochastic Dominance of Two Distributions](#),” *Econometric Theory*, 10, 849–866.
- LECKIE, N. (1988): *The declining middle and technological change: Trends in the distribution of employment income in Canada, 1971-84*, Economic Council of Canada, Ottawa.
- LEVY, F. (1987): “[The Middle Class: Is It Really Vanishing?](#)” *The Brookings Review*, 5, 17–21.

LUXEMBOURG INCOME STUDY (LIS) DATABASE, <http://www.lisdatacenter.org> (multiple countries; October 2014). Luxembourg: LIS.

MCFADDEN, D. (1989): “[Testing for Stochastic Dominance](#),” in *Studies in the Economics of Uncertainty*, ed. by T. Fomby and T. Seo, Springer New York, 113–134.

THUROW, L. C. (1984): “[The disappearance of the middle class](#),” *New York Times*, Feb-5th, 2.

WANG, Y.-Q. AND K.-Y. TSUI (2000): “[Polarization Ordering and New Classes of Polarization Indices](#),” *Journal of Public Economic Theory*, 2, 349–363.