Measurement of Polarization

Rolf Aaberge

Research Department,Statistics Norway
and ESOP, University of Oslo

The Tenth Winter School of the University of Verona, Canazei, 11-16 January 2015

Motivation (1)

- *Polarization* is a concept which is intended to complement the concept of inequality rather than to substitute it
- Esteban and Ray (1991, 1994, 1999) referred to the alienation that individuals and groups feel from one another as motivation for introducing a *multi-group identification-alienation definition of polarization*
- Foster and Wolfson (1992/2010) was concerned about the sensitivity of conclusions to the – essentially arbitrary – definition of the middle class and defined *polarization in terms of a range around the median.*

Motivation (2)

- Aaberge and Atkinson (2013) is concerned with the definition and measurement of the concept *polarization* and its relationship to the concepts *dispersion, tail-heaviness and inequality*
- Since the median plays a key role in the definition of polarization they use the notion bi-polarization
- They introduce a bi-polarization curve capturing the distance from the median

Multi-group identification-alienation measures of polarization

- Esteban and Ray (1994, 1999) and Duclos et al. (2004)

adopt an identification-alienation framework by introducing the following family of polarization measures

$$P(F) = \int\int T(f(\mathrm{x}), |x-y|)f(x)f(y)dxdy$$

where T is a function increasing in its second argument.

Parametric subfamily

- Duclos, Esteban and Ray (2004) use an axiomatic approach to justify the following parametric subfamily of *P*(*F*)

$$P^{DER} = K \int \int |x - y| \, f(x)^{1+\alpha} \, f(y) dx dy$$

where $\alpha \in [0.25, 1]$ and *K*>0 is a constant.

Basic axiom

- **AXIOM 2** (*Duclos et al. (2002)*): If a symmetric distribution is composed of three basic densities with the same root and mutually disjoint supports, then a symmetric squeeze of the side densities cannot reduce polarization.

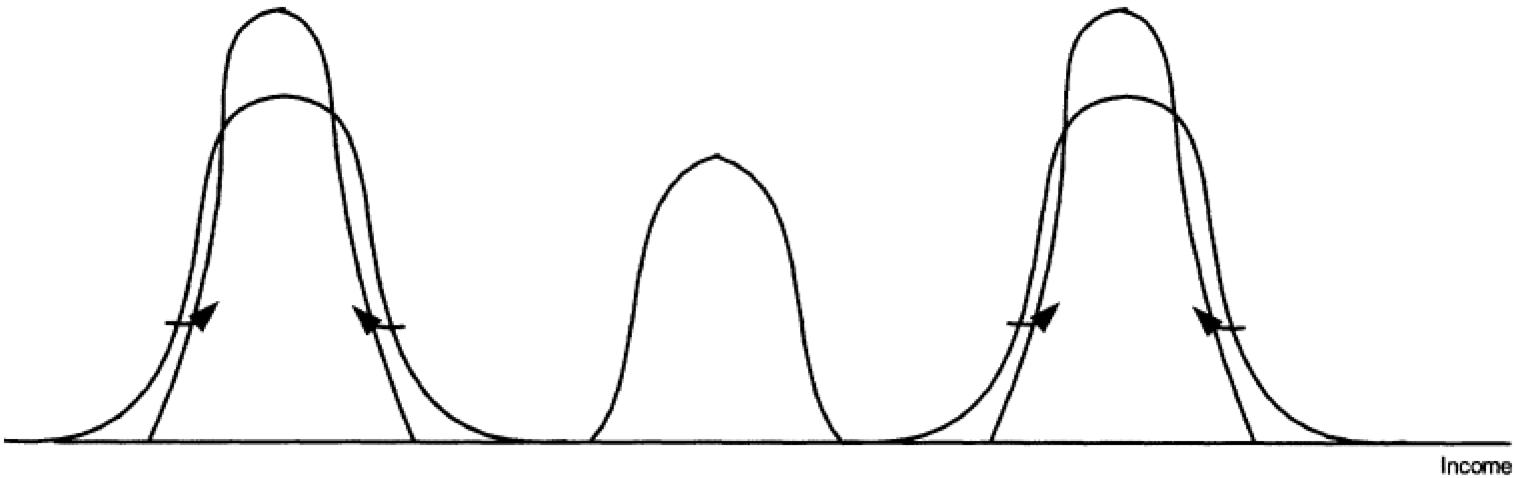


FIGURE 2.—A double squeeze cannot lower polarization.

Duclos et al. (2004) "In some sense, this is the defining axiom of polarization, and may be used to motivate the concept".

Discrete versions and an extension of *P*(*F*) as measures of conflict potential

- Consider a population composed of *m* groups, where p_i is the proportion belonging to group *i* and δ_{ij} is the "distance" between groups *i* and *j*. *Then*

$$P_{\alpha}^{DER} = \sum_{i=1}^{m}\sum_{j=1}^{m}\delta_{ij}\, p_i^{1+\alpha}\, p_j \qquad for\ \alpha \in [0.25, 1]$$

It is common to assume that $\delta_{ij} = 1\ \ for\ all\ i \neq j$ and $\delta_{ii} = 0$, which yields

$$P_{\alpha}^{DER} = K\sum_{i=1}^{m} p_i^{1+\alpha}(1 - p_i) \qquad for\ \alpha \in [0.25, 1]$$

Extended *P*(*F*) as measure of conflict potential

- Kovacic and Zoli (2012) introduce the following extension of P_{α}^{DER}

$$P_{\alpha}^{KZ} = \sum_{i=1}^{m}\sum_{j=1}^{m} \varphi(p_i, p, \delta_{ij})\, p_i\, p_j$$

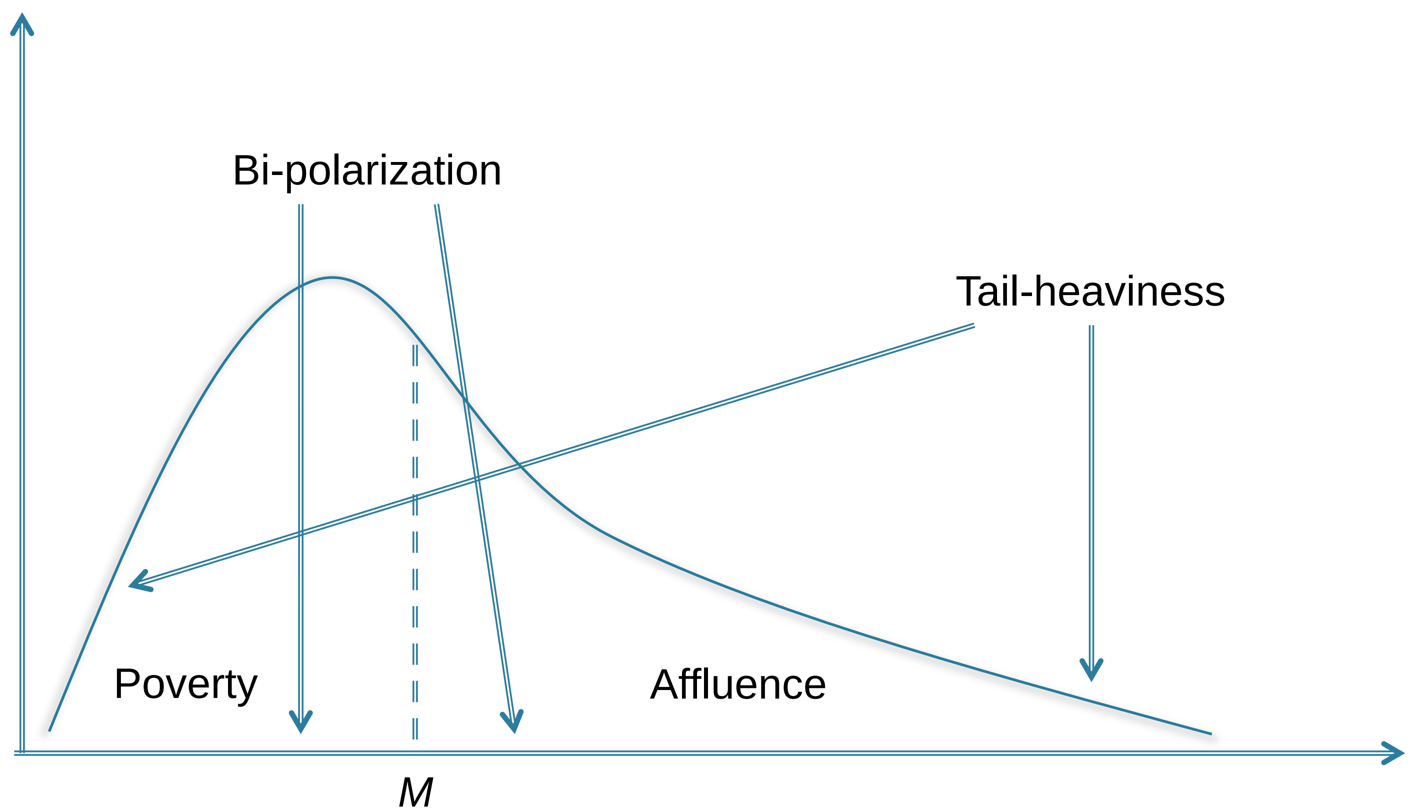
For $\delta_{ij} = 1 \ \ for\ all\ i \neq j$ we get

$$P_{\alpha}^{KZ} = \sum_{i=1}^{m} \varphi(p_i, p)\, p_i\, (1 - p_i)$$

Polarization and tail-heaviness as complementary concepts of dispersion

- The aim of Aaaberge and Atkinson (2013) is to place these concepts within a common framework and to identify the way in which different classes of income transfers contribute to different objectives
- In particular, they examine the role of transfers that preserve both the mean and the median, and the importance of distinguishing between transfers across the median and transfers on one side of the median

Poverty, Affluence, Bi-polarization and Tail-heaviness



Dispersion

A general definition of dispersion is given by Bickel and Lehmann (1979, page 34) as follows: the distribution F is less dispersed than the distribution G if for all

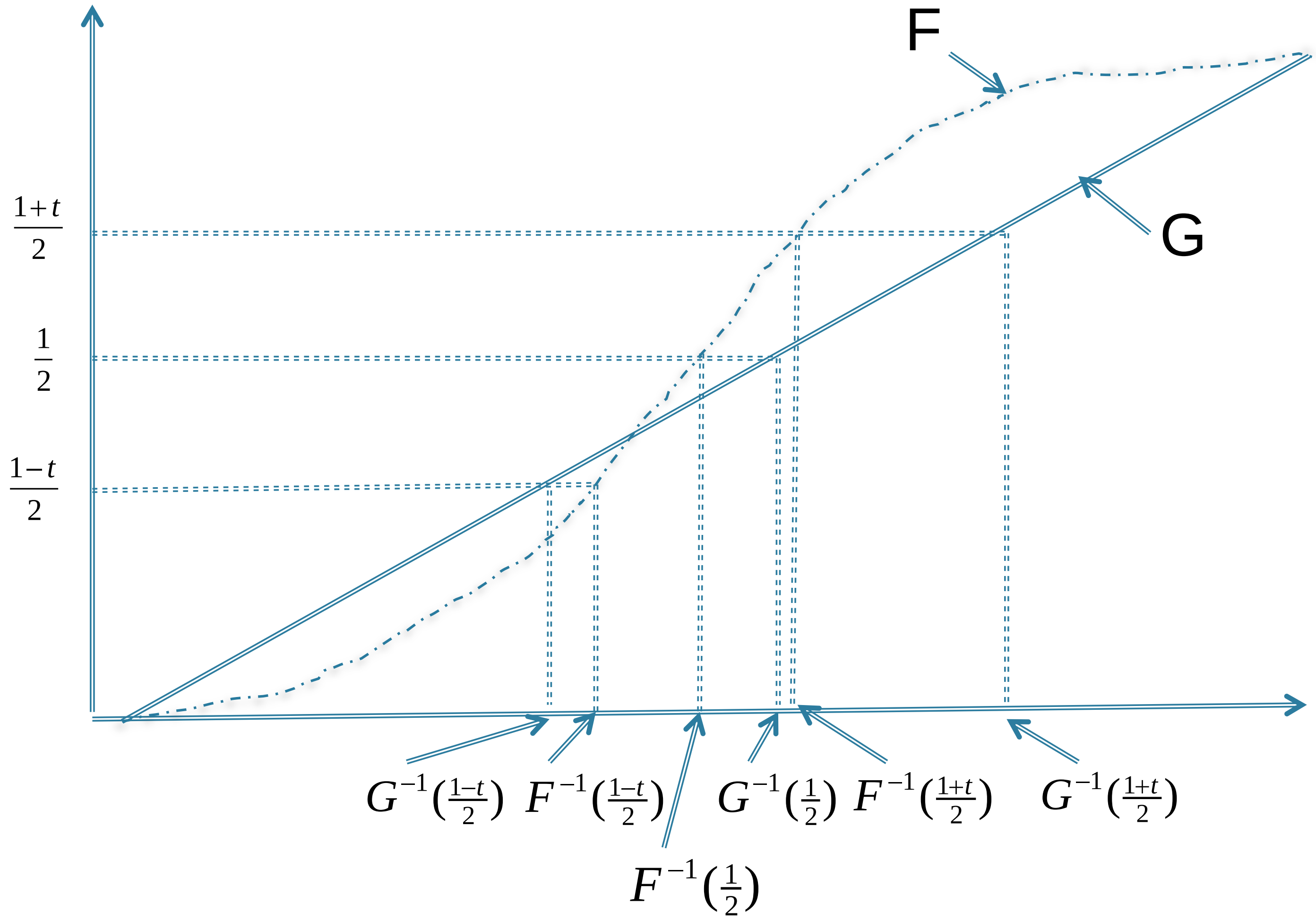
$$0 < u < v < 1,\ \ F^{-1}(v) - F^{-1}(u) \le G^{-1}(v) - G^{-1}(u)\,^{1}.$$

Aaberge and Atkinson (2013) apply a weaker version where u = (1-t)/2 and v = (1+t)/2. In other words, we use the following curve, denoted the dispersion curve,

$$D(t) = \frac{1}{M}\left(F^{-1}\left(\frac{1+t}{2}\right) - F^{-1}\left(\frac{1-t}{2}\right) \right), \quad t \in [0,1].$$

As t approaches 1, the distance becomes $D(1) = \left(F^{-1}(1) - F^{-1}(+0)\right)\big/M$.

Dispersion curve



Summary measures of dispersion

$$\Delta_c = \int_0^1 c(t)D(t)dt$$

Δ_c is completely characterized by a bi-polarization curve ordering that is continuous, transitive and complete and satisfies the dual independence axiom and first-degree bi-polarization dominance

Axiom (Dual independence): *Let* D_1*,* D_2 *and* D_3 *be members of* ***D*** *and let* $\alpha \in [0,1]$*. Then* $D_1 \succeq D_2$ *implies* $\left(\alpha D_1^{-1} + (1-\alpha)D_3^{-1}\right)^{-1} \succeq \left(\alpha D_2^{-1} + (1-\alpha)D_3^{-1}\right)^{-1}$*.*

Polarization curve

Theorem 4.1 justifies the function P defined by

$$P(u) = \int_0^u D(t)\,dt, \;\; u \in [0,1],$$

as a device for comparing polarization between distribution functions. Accordingly,
P is denoted *the polarization curve*. The following alternative expression for P

$$P(u) = \frac{u}{M}\left[E\left(X \,\middle|\, M \le X \le F^{-1}\left(\frac{1+u}{2}\right)\right) - E\left(X \,\middle|\, F^{-1}\left(\frac{1-u}{2}\right) \le X \le M \right)\right], u \in [0,$$

,

provides an intuitive justification for why it makes sense to consider P as a polarization curve.

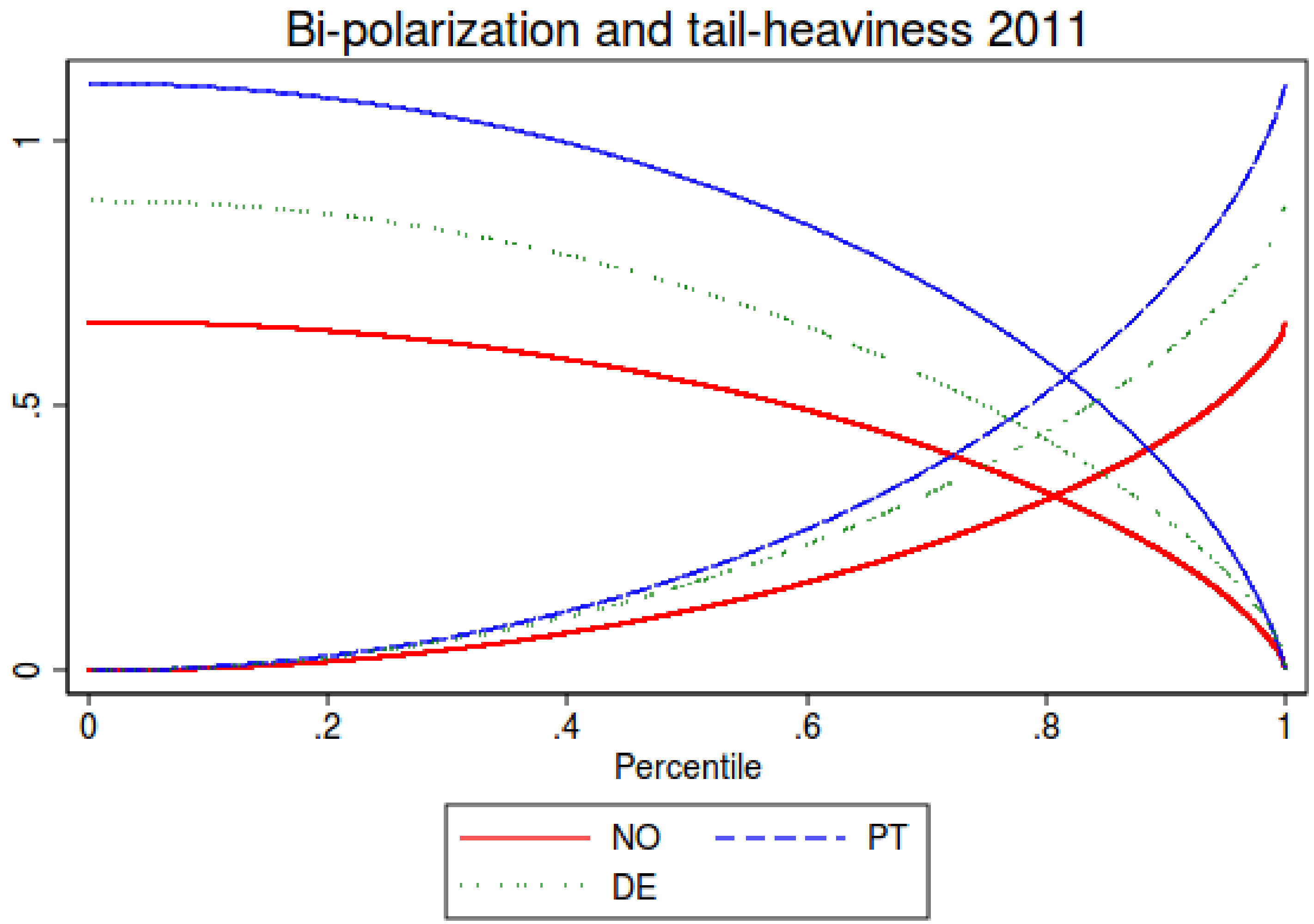
,

Tail-heaviness curve

Dispersion curve: $$D(t)=\frac{1}{M}\left(F^{-1}(\frac{1+t}{2})-F^{-1}(\frac{1-t}{2})\right), \quad t\in[0,1],$$

Tail-heaviness curve: $$T(u)=\int_u^1 D(t)dt,\ u\in[0,1],$$

$$T(u)=\frac{(1-u)}{M}\left[E\left(X\,\middle|\,X\geq F^{-1}\left(\frac{1+u}{2}\right)\right)-E\left(X\,\middle|\,X\leq F^{-1}\left(\frac{1-u}{2}\right)\right)\right], u\in[0,1]$$



Bi-polarization and tail-heaviness 2011
0
.5
1
0
.2
.4
.6
.8
1
Percentile
NO
PT
DE

Summary measures of bi-polarization and tail-heaviness

Δ_c can be given the following alternative expression in terms of the bi-polarization curve P,

(4.8) $$\Delta_c = -\int_0^1 c'(u)P(u)du \,,$$

where $c(t)$ is decreasing.

Moreover, Δ_c can be expressed in terms of the tail-heaviness curve T,

(4.17) $$\Delta_c = \int_0^1 c'(u)T(u)du \,.$$

where $c(u)$ is an increasing function of u.

Specific measures of bi-polarization and tail-heaviness

Let $c(t) = 2(1-t)$. Then

(4.9) $$\Delta_c = \Delta_2^* \equiv \Delta_1 - \frac{\mu_l}{M} G_{l,2} - \frac{\mu_u}{M} G_{u,2} = 2\Delta_1 - 4\frac{\mu}{M} G$$

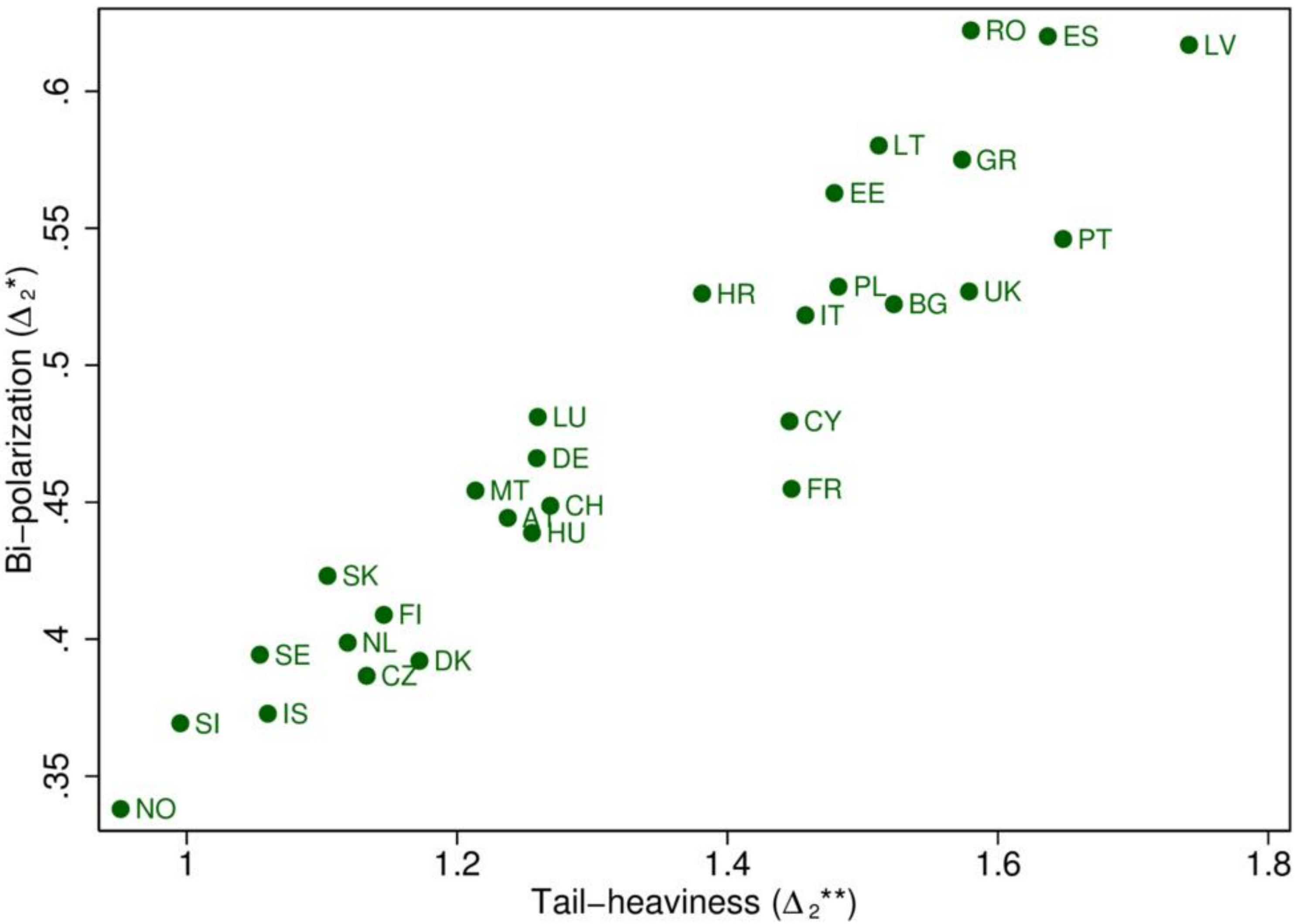
where μ_l *and* $G_{l,2}$ are the mean and the Gini coefficient of the distribution of incomes below the median and μ_u *and* $G_{u,2}$ are the mean and the Gini coefficient of the distribution of incomes above the median.

Let $c(t) = 2t$. Then

(4.18) $$\Delta_c = \Delta_2^{**} \equiv \Delta_1 + \frac{\mu_l}{M} G_l + \frac{\mu_u}{M} G_u = \frac{4\mu}{M} G,$$

where G_l is G_u the Gini coefficients of the conditional income distributions given that incomes takes values below and above the median, respectively.

Ranking countries by measures of tail-heaviness and bi-polarization



Association between bi-polarization and tail-heaviness

Bi-polarization	Tail-heaviness	Correlation coeffcient	Spearman coeffcicient
Δ_2^*	Δ_2^{**}	.94	.94
Δ_3^*	Δ_3^{**}	.89	.90
Δ_4^*	Δ_4^{**}	.85	.85

Conclusions

- *Polarisation* and *tail-heaviness* can be considered as complementary measures of dispersion, but with the crucial difference that *polarization* cumulates from the median and gives more weight to transfers near the middle, whereas *tail-heaviness* cumulates from the tails and gives more weight to transfers far removed from the median.
- The concept of tail-heaviness fits more naturally with measures of inequality.

Con....

Effect of progressive mean- median-preserving transfers

	1.	4.	5.
	Inequality	Bi-Polarization	Tail-heaviness
Across median	Fall	Fall	Fall
Below median	Fall	Rise	Fall
Above median	Fall	Rise	Fall

- The empirical results show strong association between bi-polarization and tail-heaviness, which means that most countries are similar for bi-polarization and tail-heaviness.
- France stands out as an exception with a more favourable rank in bi-polarization than in tail-heaviness

Asymptotic estimation theory

Let the empirical processes $P_n(x)$ and $Q_n(t)$ be defined by

$$P_n(x) = n^{\frac{1}{2}}\left(F_n(x) - F(x)\right) \text{ and } Q_n(t) = n^{\frac{1}{2}}\left(F_n^{-1}(t) - F^{-1}(t)\right)$$

and let $W_0(t)$ denote a Brownian Bridge on [0,1], that is, a Gaussian process with mean zero and covariance function $s(1-t)$, $0 \le s \le t \le 1$.

$P_n(x)$ and $Q_n(t)$ converge in distribution to the Gaussian processes

$$W_0(t) \text{ and } W_0(t) \big/ f\left(F^{-1}(t)\right).$$

See Billingsley (1968) and Doksum (1974)

Dual measures

Let $w(t)$ be a positive non-increasing or non-decreasing function of x and let $0 \leq m < r \leq 1$. In order to study the asymptotic behaviour of the empirical counterparts of Π_p, Γ_q and Δ_c it is convenient to consider the empirical process

$$Y_n = \int_m^r w(t) Q_n(t) dt ,$$

Proposition A.1b. Suppose that F has a continuous nonzero derivative f on $[a,b]$. Then $Y_n(u)$ converges in distribution to the process

$$Y = \int_m^r w(t) \frac{W_0(t)}{f(F^{-1}(t))} dt .$$

Normally distributed

$$V_N(t) = \frac{2^{\frac{1}{2}}}{f\left(F^{-1}(t)\right)} \sum_{j=1}^{N} \frac{\sin(j\pi t)}{j\pi} Z_j$$ where $Z_1, Z_2, \ldots$ are independent $N(0,1)$ variables

and note that

(3.5) $$2\sum_{j=1}^{\infty} \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} = s(1-t), \quad 0 \le s \le t \le 1 .$$

Thus, the process $V_N(t)$ is Gaussian with mean zero and covariance function

$$\operatorname{cov}\left(V_N(s), V_N(t)\right) = \frac{2}{f\left(F^{-1}(s)\right) f\left(F^{-1}(t)\right)} \sum_{j=1}^{N} \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} \to \operatorname{cov}\left(V(s), V(t)\right),$$

where

$$V(t) = \frac{W_0(t)}{f\left(F^{-1}(t)\right)} .$$

Aaberge, R. (2006): "Asymptotic Distribution Theory of Empirical Rank-dependent Measures of Inequality".
In V. Nair (Ed.): Advances in Statistical Modeling and Inference - Essays in Honor of Kjell A. Doksum, World Scientific.