Measuring Social Polarization with Ordinal and Categorical data: an application to the missing dimensions of poverty in Chile

Conchita D'Ambrosio
conchita.dambrosio@unibocconi.it
Iñaki Permanyer
inaki.permanyer@uab.es

CPRC International Conference 2010

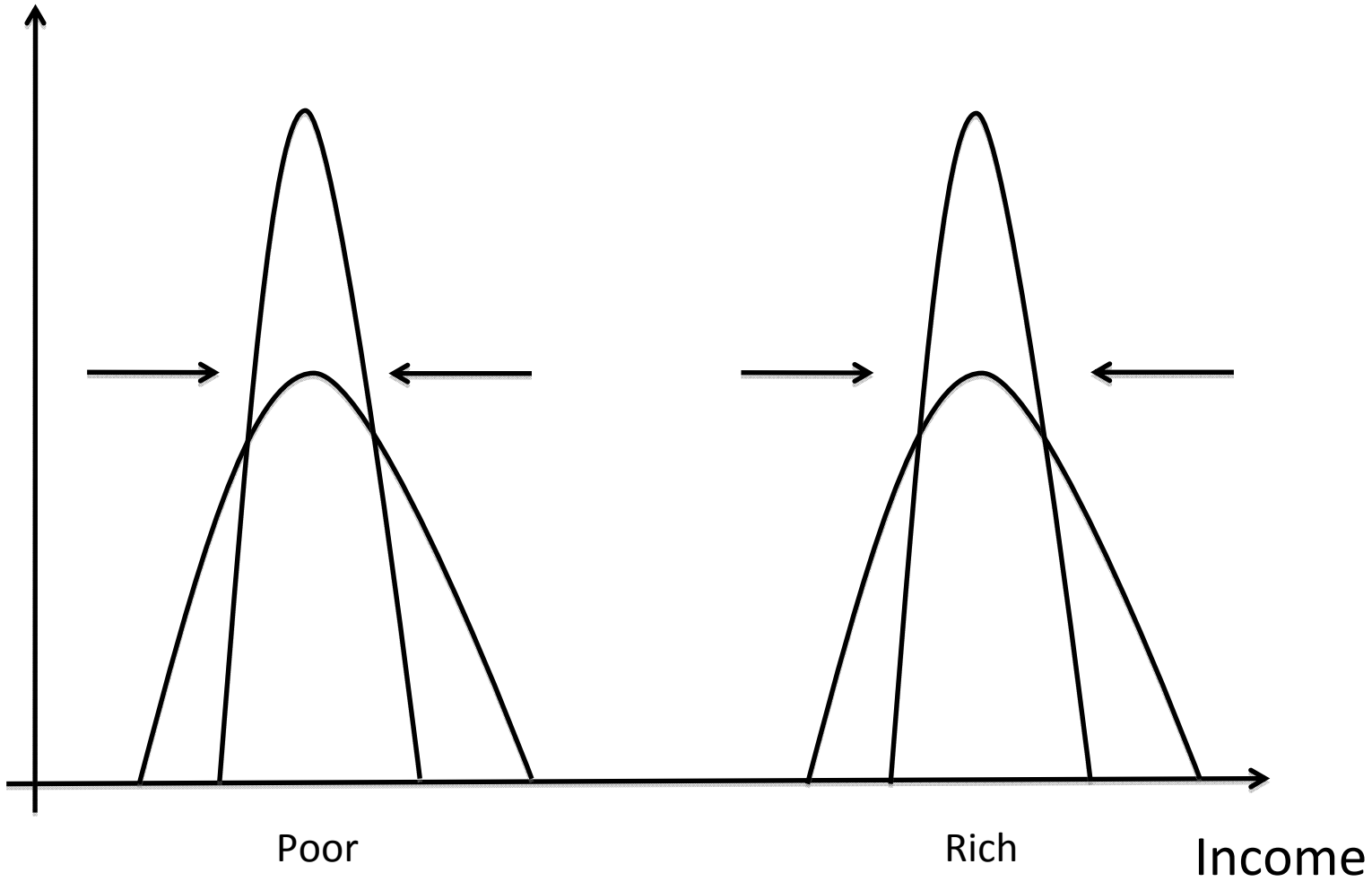
Summary of the talk

- Introduction/motivation: the measurement of polarization
  - Income polarization
  - Polarization with ordinal data
  - Social Polarization
- Some new proposals
  - Axiomatic characterization
  - New polarization indices
- Empirical application: the case of Chile.

Introduction

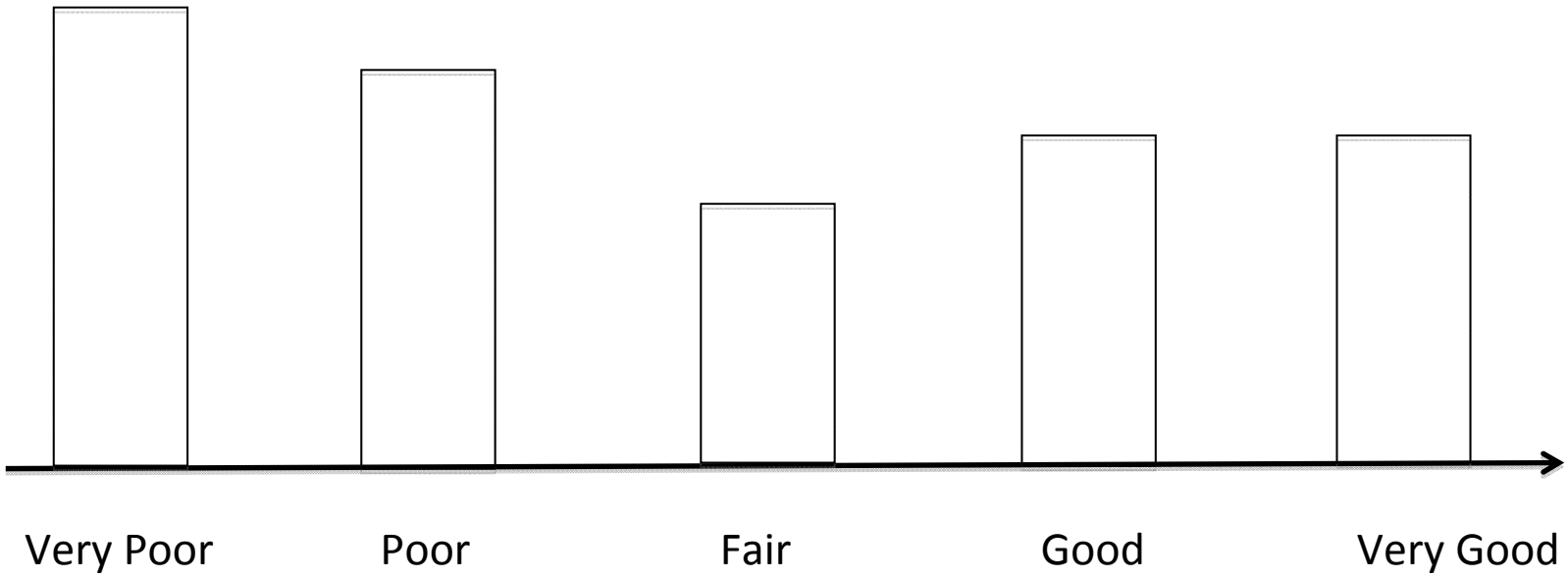
- Polarization measures try to capture the extent to which individuals are clustered around certain “poles”, thus forming cohesive groups that might potentially express their unrest into social action or revolt.
- Polarization is very important for its relationship with the occurrence of conflict.
- **Key question**: What is the most relevant variable/dimension in generating alienated groups?

Income polarization



Polarization with ordinal data

- Apouey (2007) proposes a measure of polarization for ordinal data (useful for self-assessments of health, well-being, and so on).



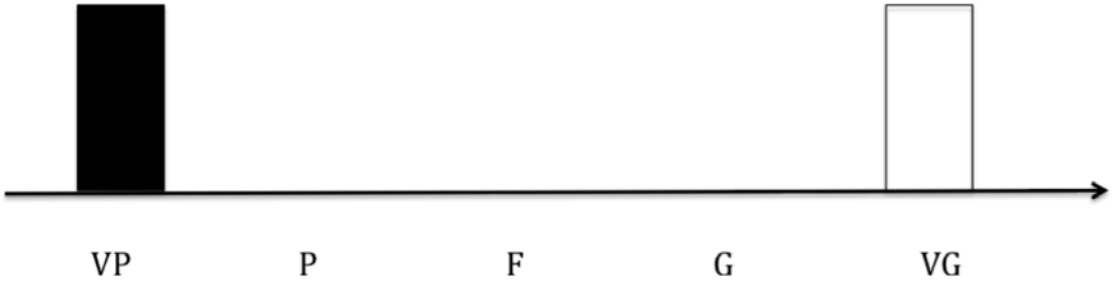
Social Polarization

- Individuals are no longer classified according to their income levels but by another salient characteristic (e.g: race, ethnicity, religion,…). This generates *k* groups with population shares $\pi_1,\ldots,\pi_k$.
- A well-known social polarization index is the Reynal-Querol index:

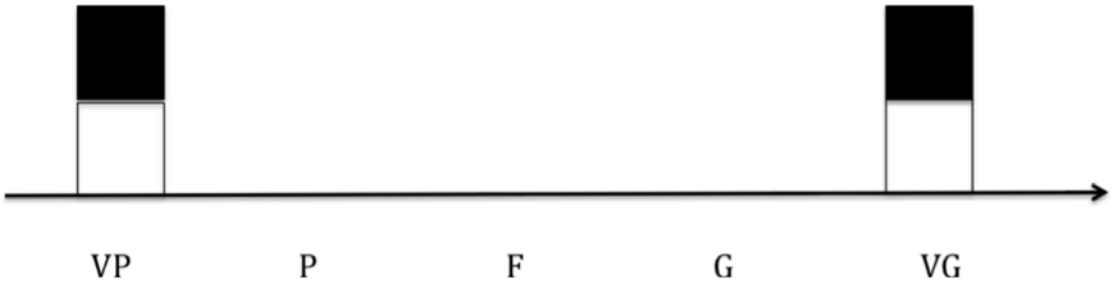
$$RQ = 4\sum_{s=1}^{k}\sum_{t=1}^{k}\pi_s^2\pi_t$$

A disturbing example

Scenario 1



Scenario 2



Income polarization, ordinal polarization and social polarization indices are unable to distinguish between both scenarios!

Motivation

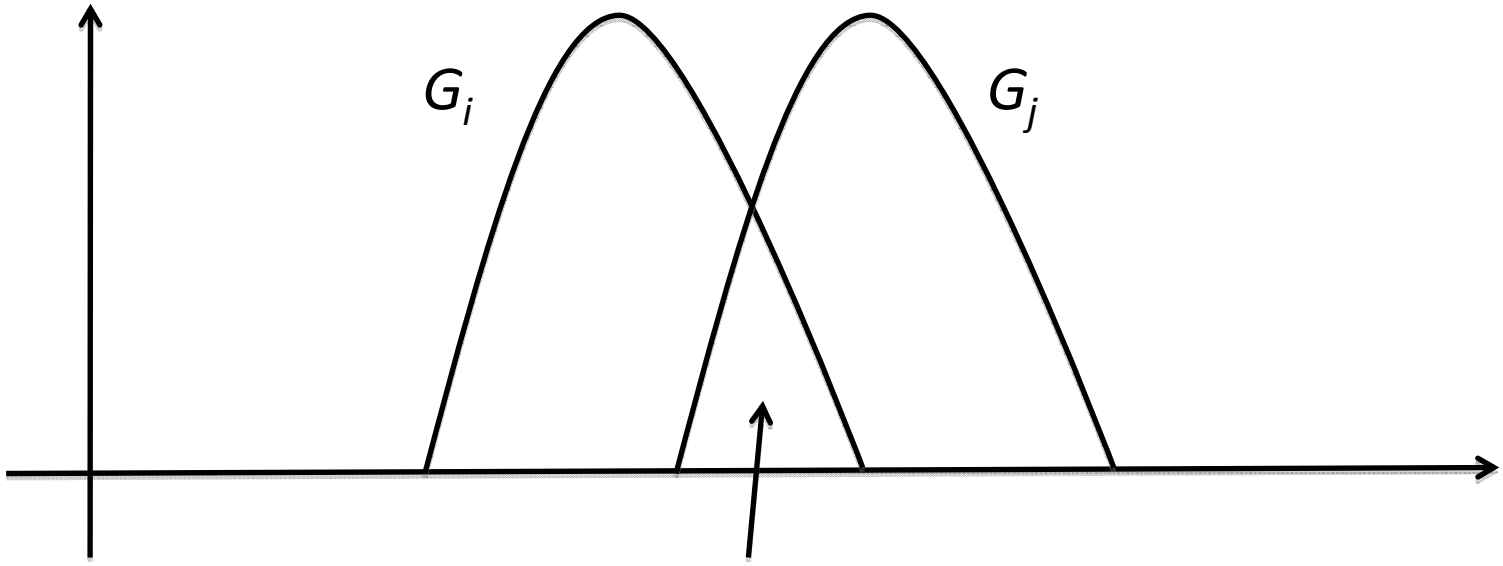
- We present new polarization indices that incorporate both intuitions *at the same time*, that is: individuals can be identified according to categorical (ethnicity, religiosity) *and* ordinal/cardinal characteristics (e.g:income, health, well-being).
- These indices are very useful in OPHI's "Missing Dimensions framework", where many variables are categorical or ordinal.

Some notation

- We consider k exogenously given groups $\boldsymbol{G}:=\{G_1,\ldots,G_k\}$ with population shares $\pi_1,\ldots,\pi_k$ and absolute sizes $N_1,\ldots,N_k$.
- We consider variables with C different categories (could be ordinal or categorical).
- We denote by $p_{i,c}$ the share of group G_i that belongs to category c.

Symmetric distance between groups

- Given any two groups G_i and G_j we will assume that the alienation felt between them is given by a function of the overlap coefficient (Anderson 2010).



$$\theta_{ij} = \sum_{c=1}^{C} \min\{p_{i,c}, p_{j,c}\} = \theta_{ji}$$

Asymmetric distance between groups

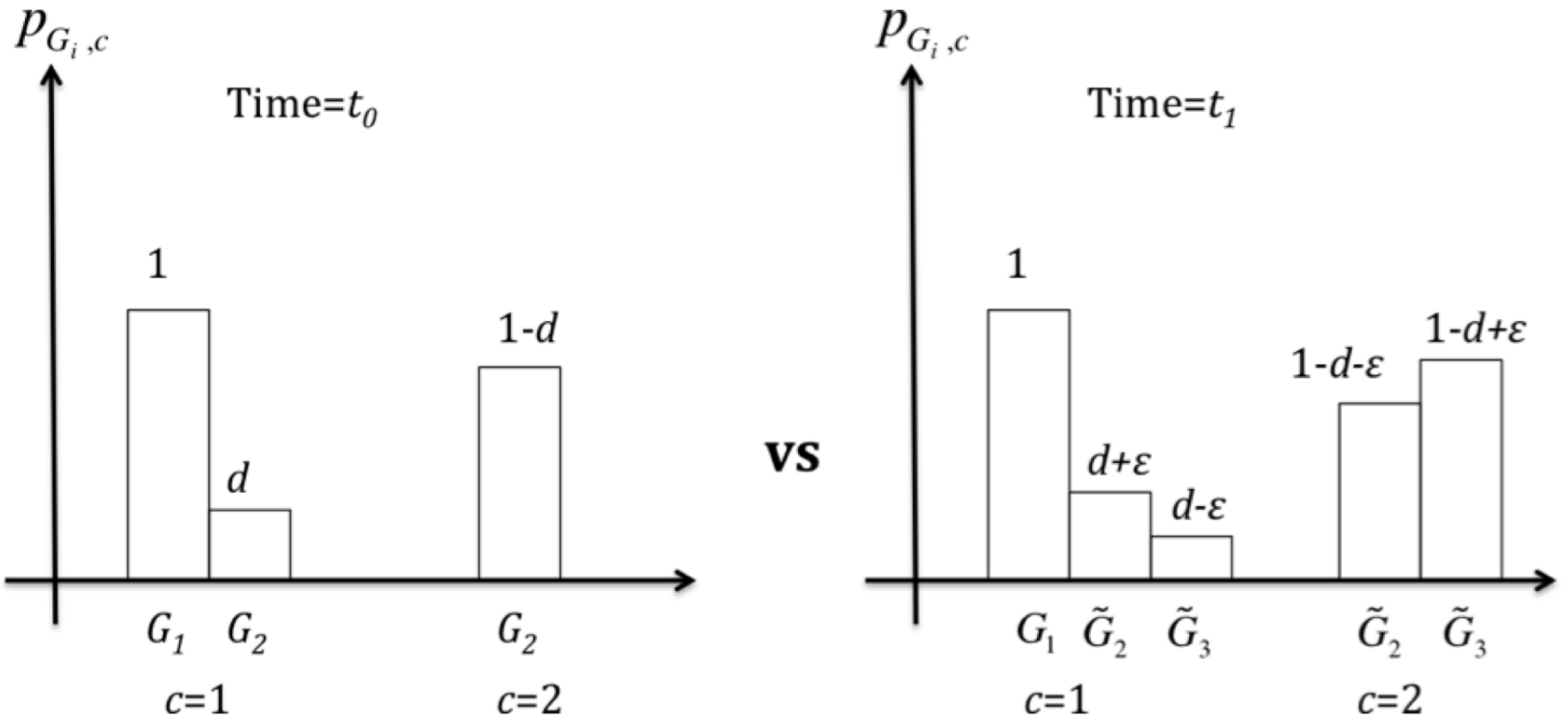
- Given any two groups G_i and G_j we will assume that the alienation felt between them is given by a function of the following coefficient:

$$A_{ij} = \frac{\sum_{s=1}^{N_i}\sum_{t=1}^{N_j}\delta_{st}}{N_i N_j}$$

  where δ_{st} is 1 if individual '*s*' from group '*i*' is **ranked below** individual '*t*' from group '*j*' and 0 otherwise.

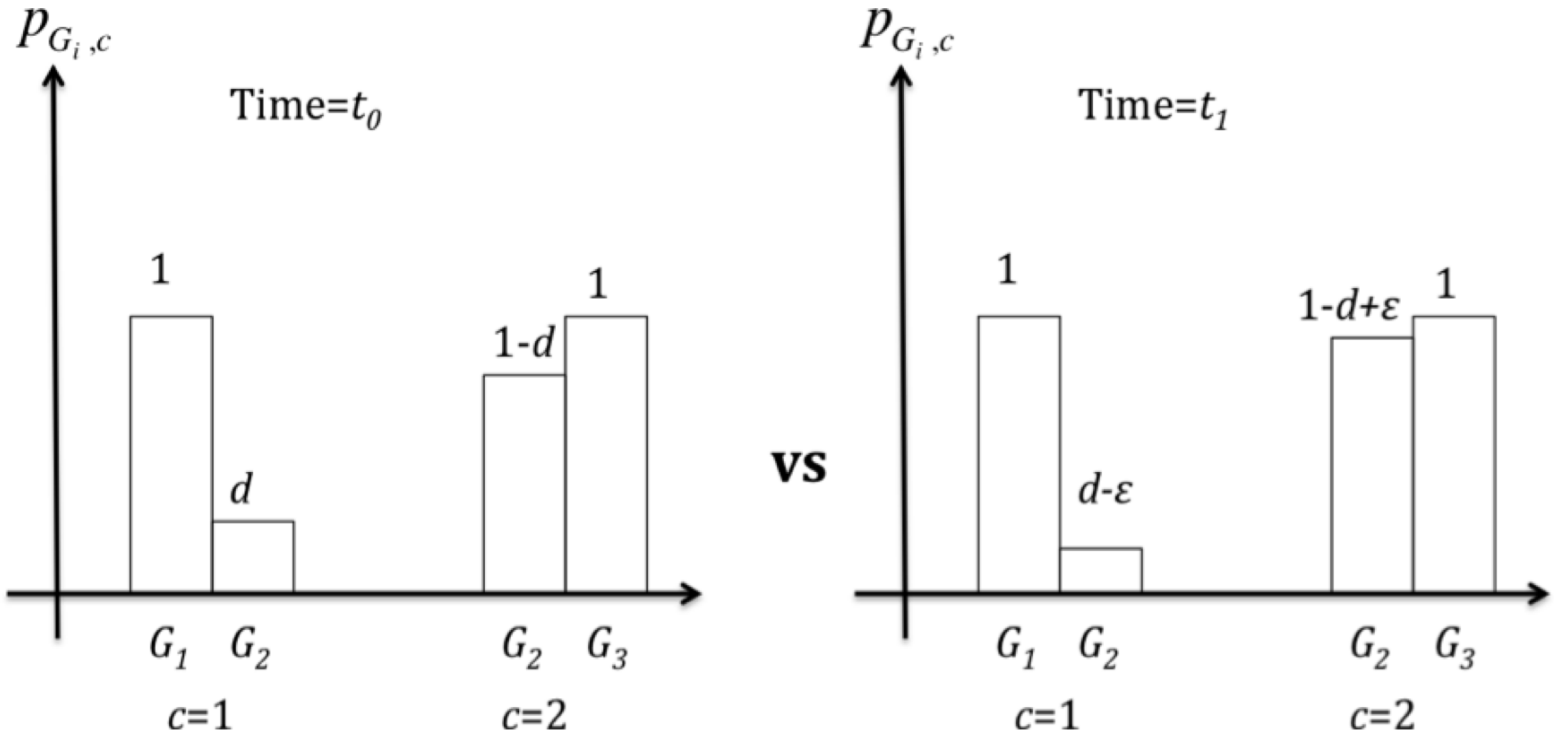
- Recall that $A_{ij} \neq A_{ji}$.

Axiomatic characterization (I)



Axiom 1: Polarization should decrease in that case.

Axiomatic characterization (II)



Axiom 2: Polarization should increase in that case.

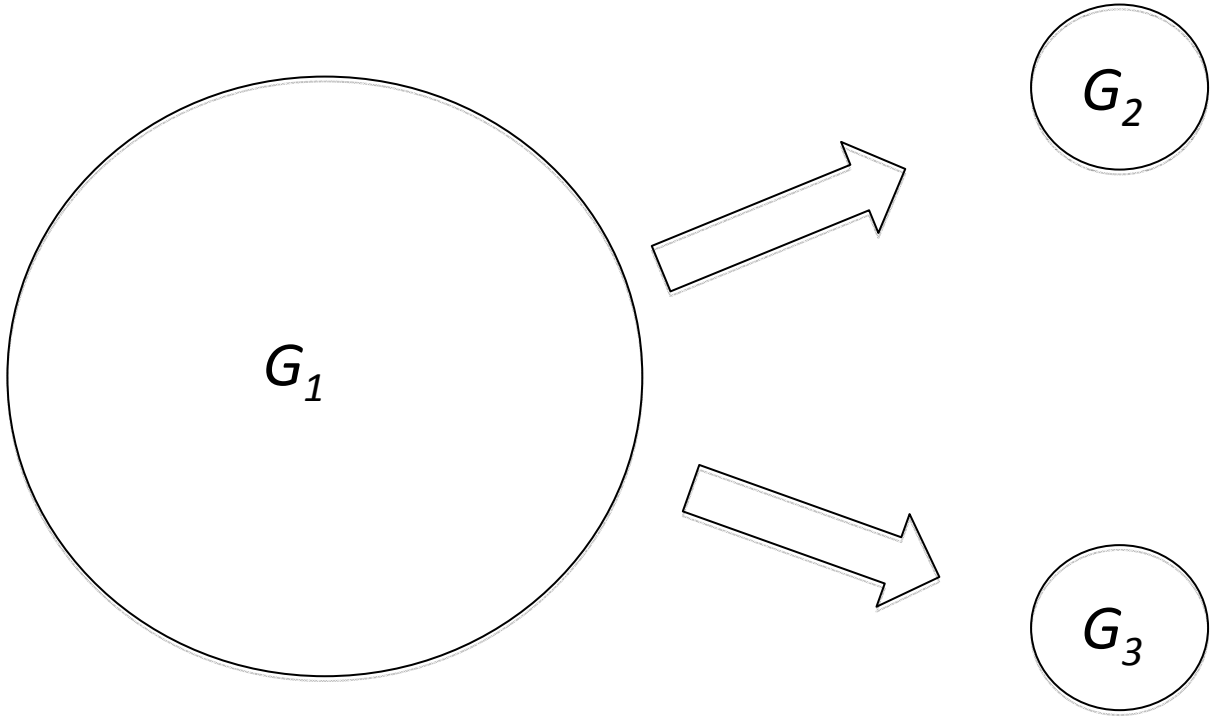
Axiomatic characterization (III)

- **Axiom 3.** If $P(\boldsymbol{G}_1) \geq P(\boldsymbol{G}_2)$ and $q>0$, then $P(q\boldsymbol{G}_1) \geq P(q\boldsymbol{G}_2)$, where $q\boldsymbol{G}_1$, $q\boldsymbol{G}_2$ represent population scalings of $\boldsymbol{G}_1, \boldsymbol{G}_2$ respectively.

- **Axiom 4.** Keeping all else equal, an increase in the number of groups (k) leads to a decrease of polarization.

- **Axiom 6.** For any grouping of the population and any distribution, $0 \leq P(\boldsymbol{G}) \leq 1$.

Axiomatic characterization (IV)



Axiom 5: Polarization should increase in that case.

New polarization indices (I)

- **Theorem 1.** The only **symmetric** polarization index satisfying the previous axioms is given by

$$P_S(G) = 4\sum_{s=1}^{k}\sum_{t=1}^{k} \pi_s^{1+\alpha}\,\pi_t(1-\theta_{st})$$

$$where\ \alpha \in [\alpha^*,1],\ with\ \alpha^* = \frac{2-\log_2 3}{\log_2 3 - 1} \approx 0.71$$

New polarization indices (II)

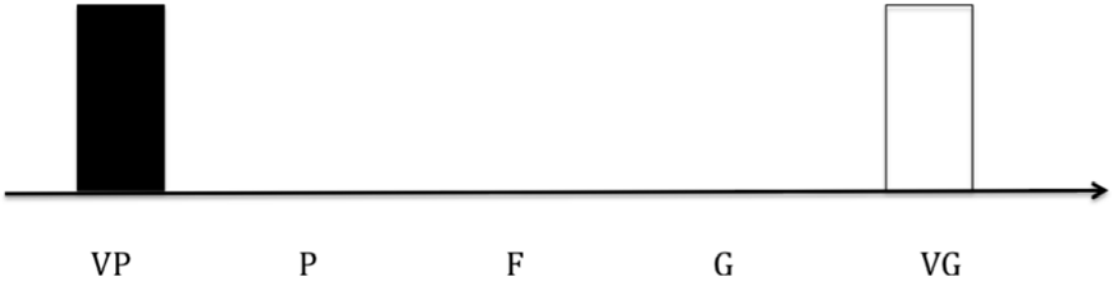
- **Theorem 2.** The only **asymmetric** polarization index satisfying the previous axioms is given by

$$P_A(G) = \frac{27}{4} \sum_{s=1}^{k} \sum_{t=1}^{k} \pi_s^{1+\alpha} \pi_t A_{st}$$

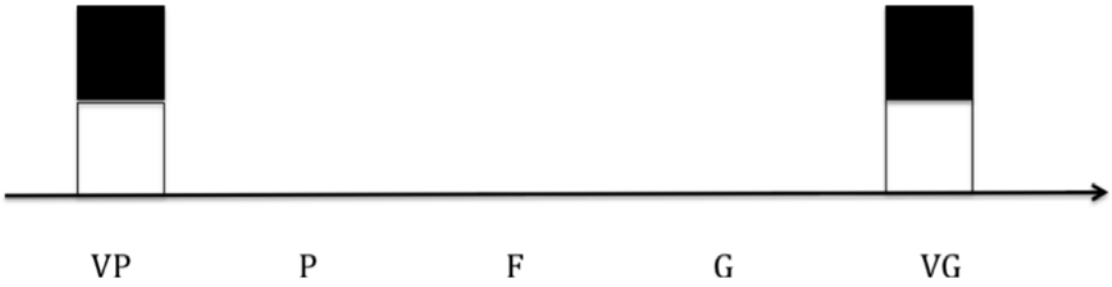
$$where\ \alpha \in [\alpha^*, 1],\ with\ \alpha^* = \frac{2 - \log_2 3}{\log_2 3 - 1} \approx 0.71$$

The previous example revisited

Scenario 1



Scenario 2



Now P_S =1 , P_A =0.84375 in scenario 1 and
P_S =0.5 , P_A =0.421875 in scenario 2, as expected.

Empirical illustration: the Chilean case (I)

- Ordinal variable: “Self-assessed health” (Very Poor, Poor, Fair, Good, Very Good).
- Groups classified by “ethnicity” (the Chilean *Fundación para la Superación de la Pobreza* highlights the underprivileged situation of those individuals with partly indigenous descent in their report “Social Thresholds for Chile: Towards Future Social Policy”).

Empirical illustration: the Chilean case (II)

Region name	DER	GINI	RQ	Apouey (2007)	PS	PA
Tarapacá	0,370 (9)	0,647 (1)	0,876 (1)	0,249 (9)	0,131 (4)	0,458 (1)
Antofagasta	0,329 (12)	0,530 (12)	0,571 (3)	0,273 (6)	0,193 (2)	0,242 (5)
Atacama	0,392 (6)	0,589 (4)	0,525 (4)	0,271 (7)	0,197 (1)	0,118 (6)
Coquimbo	0,414 (4)	0,632 (2)	0,268 (8)	0,234 (11)	0,104 (5)	0,045 (10)
Valparaíso	0,367 (10)	0,585 (5)	0,111 (10)	0,298 (4)	0,035 (9)	0,009 (12)
O'Higgins Rancagua	0,395 (5)	0,585 (6)	0,015 (12)	0,281 (5)	0,017 (12)	0,018 (11)
Maule Talca	0,486 (1)	0,548 (10)	0,101 (11)	0,232 (12)	0,030 (10)	0,048 (9)
Biobío Concepción	0,389 (7)	0,597 (3)	0,299 (7)	0,311 (2)	0,055 (8)	0,114 (7)
Araucanía Temuco	0,428 (3)	0,535 (11)	0,736 (2)	0,268 (8)	0,082 (6)	0,355 (2)
Los Lagos Puerto Montt	0,373 (8)	0,563 (9)	0,505 (5)	0,243 (10)	0,067 (7)	0,285 (4)
Aysén Coyhaique	0,453 (2)	0,566 (8)	0,447 (6)	0,328 (1)	0,156 (3)	0,300 (3)
Metropolitana	0,354 (11)	0,577 (7)	0,173 (9)	0,301 (3)	0,022 (11)	0,095 (8)

Authors' calculations using OPHI dataset for Chile. Region Rankings in parentheses.

Future Research

- Extend the previous empirical analysis to find out the dimensions that divide the Chilean society more deeply.
- Identify the couples groups that feel more alienated vis-à-vis each other.

Measuring Social Polarization with Ordinal and Categorical data: an application to the missing dimensions of poverty in Chile

Conchita D'Ambrosio
conchita.dambrosio@unibocconi.it
Iñaki Permanyer
inaki.permanyer@uab.es

THANK YOU!