**Appendix 2a: Comparisons of 30 indicators of inequality samples and a Principal Component Analysis PCA on 232 LIS**

We can compare the three ABG indices to other standardized inequality measures (Jenkins, 1999/2010, Abdelkrim and Duclos, 2013). These selected indicators are well-known or based on simple income ratios. We consider added ISO indicators at five different levels. In addition, the size (as a proportion in the total population) of five income classes are included: the poor (po), lower middle class (mcl), middle class (mc), upper middle class (mcu), and the rich (ri). Overall, our analysis covers a set of 30 variables and 232 data samples.

* **ABG class:**, , ; i.e., the three coefficients from the ABG method.
* **Atkinson class:**a2, a1, ahalf = Atkinson class of indices, respectively with parameters 2, 1, and ½ (Atkinson, 1970, also see Yitzhaki, 1983), the higher parameter (2) overweights the bottom of the distribution.
* **Generalized entropy class:**ge2, ge1, ge0, gem1 = Generalized entropy class of indices, respectively with parameters 2, 1, 0, -1 (Berry et al., 1983). The lower parameter (-1) implies a focus on the bottom of the distribution.
* **Gini inequality index:**The value of the standard Gini index (Gini, 1914).
* **Wolfson polarization index:**  
  The Wolfson index (Wolfson, 1986) of polarization.[[1]](#footnote-1)
* **Foster‐Greer‐Thorbecke poverty class:**fgt0, 1 and 2 show the Foster‐Greer‐Thorbecke (Foster et al., 1984) poverty index, with respectively parameters 0, 1, 2, and the poverty threshold of 60%. The higher the parameter, the greater the focus on very low incomes.
* **Income ratios:** 
  + p90p50 = decile9/median: this measures inequality at the top.
  + p50p10 = median /decile1: this measures inequality at the bottom.
  + pp907550 = (decile9/quartile3) / (quartile3/median): this measures the degree to which inequality accelerates near the top decile, compared to the degree of inequality between the median and the top quartile. This corresponds to over-elongation at the top.
  + pp251050 = (quartile1/decile1) / (median/quartile1): this measures the degree to which inequality accelerates at the lower decile, compared to the degree of inequality at the lower quartile. This corresponds to over-elongation at the bottom.
* **ISO(X) class of measure of inequality:**iso2, iso6, iso10, iso14, iso18 are respectively the values of ISO for the “vingtiles” (5% slices) 2, 6, 10, 14 and 18. These correspond to the values of X close to -3, -1, 0, +1 and +3, respectively.
* **Income class proportions:**po, mcl, mc, mcu, ri. These are respectively, the proportion of the poor (medi < .5), lower middle class (.5 <= medi < .75), middle class (.75 <= medi < 1.25), upper middle class (1.25 <= medi < 2) and rich (2 <= medi) in the total population.[[2]](#footnote-2)
* **Income class based indicator of polarization:**rpol = (mcl + mcu)/mc. This “polarization ratio” assesses the size of the lower and upper middle classes compared to the middle class, who are close to the median.

One important question is the relative position of the ABG parameters in the field of inequality measures. A first answer is given by an analysis of the correlation matrix of these indicators (Appendix3: the general correlation matrix of 30 inequality indicators): there is a very strong relation between  and the Gini index (R = +.95) thus confirming the relation of these two inequality measures when the CF approximation is acceptable. More generally, most of the measures correlate well with . This is good news for the ABG method, but then what is its intrinsic added value? A second answer is that we also see interesting correlations for the  and  coefficients, which thus provide complementary information to : the degree to which inequality moves at the top and at the bottom of the distribution. A third more systematic answer comes from the principal component analysis (PCA) of the whole table (Figure A2a1 depicts the correlation circle). The PCA is a type of factor analysis[[3]](#footnote-3) used for quantitative measures, and its application to our indicator set (in Table A2a1) helps us to understand the multidimensional relations between these indicators. The first axis of the PCA (69% of the total variance) reveals the similar nature of many inequality measures, including ; this axis picks up inequality intensity. The coefficient appears on the first axis of the PCA, along with the Atkinson parameters 1 (a1) and ½ (ahalf), the generalized entropy parameters 1 (ge1) and 0 (ge0), the Gini coefficient, a number of quantile ratios, as well as the Wolfson polarization index. This confirms that  is a new inequality parameter which is highly correlated with the main inequality measures, but is more sensitive (like the Wolfson index) to the median of the distribution.

**Table A2a1: PCA Scores: Correlation between the Principal Components and 30 Indicators of Inequality**

|  |  |  |  |
| --- | --- | --- | --- |
| Indicator | v1 | v2 | v3 |
|  | 0.2143 | -0.1114 | 0.0138 |
|  | -0.0548 | 0.4194 | -0.244 |
|  | -0.0565 | 0.3174 | 0.4321 |
| a2 | 0.1166 | 0.0958 | 0.3377 |
| a1 | 0.2171 | 0.0666 | -0.0154 |
| Ahalf | 0.2145 | 0.0875 | -0.0738 |
| ge2 | 0.1373 | 0.1738 | -0.1968 |
| ge1 | 0.2092 | 0.1123 | -0.1184 |
| ge0 | 0.2154 | 0.0879 | -0.0416 |
| gem1 | -0.0002 | 0.131 | 0.2562 |
| Gini | 0.2174 | 0.0239 | -0.0381 |
| Wolfson | 0.2183 | -0.0179 | -0.0404 |
| fgt0 | 0.2183 | 0.0094 | 0.0044 |
| fgt1 | 0.215 | 0.0889 | 0.0055 |
| fgt2 | 0.2096 | 0.123 | 0.0089 |
| p90p50 | 0.2108 | 0.0632 | -0.1424 |
| p50p10 | 0.2098 | 0.0245 | 0.1522 |
| pp907550 | -0.0266 | 0.3587 | -0.355 |
| pp251050 | -0.0451 | 0.3495 | 0.3557 |
| iso2 | 0.2096 | -0.0046 | 0.1789 |
| iso6 | 0.2118 | -0.103 | 0.0745 |
| iso10 | 0.2115 | -0.1105 | 0.0226 |
| iso14 | 0.2148 | -0.0662 | -0.0305 |
| iso18 | 0.2147 | 0.0142 | -0.1062 |
| Po | 0.2085 | -0.0219 | 0.1833 |
| Mc | -0.2015 | 0.1644 | -0.0753 |
| Mcl | -0.0646 | -0.3677 | -0.2298 |
| Mcu | -0.1162 | -0.298 | 0.2478 |
| Ric | 0.2146 | -0.0182 | -0.0723 |
| Rpol | 0.1859 | -0.2517 | 0.0745 |

The role of and  becomes apparent on axes 2 and 3 (12% and 7% of the variance, respectively), which reveal the shape of inequality but not its intensity.

* On the second PCA axis,  and  are strongly correlated in the same direction as pp251050 and pp907550, the two measures of the over-elongation of the extreme deciles compared to the quartiles. The correlation with mcu and mcl (respectively, the upper and lower middle classes) is negative: the elongation at the top (resp. bottom) implies a smaller upper (resp. the lower) middle class that is stretched out. Positive values on the second axis reflect greater inequality at the extremes. Here, the generalized entropy index with parameter 2 is more strongly correlated on axis 2 than are the other traditional measures.
* Axis 3 shows the difference between  and , along with the contrast between pp251050 and pp907550. On this axis, the generalized entropy index with parameter 1 (gem1) and the Atkinson index with parameter 2 (a2) are located on the same side of axis3 as . All of these indicators are relatively more sensitive to inequality at the bottom. Conversely, the generalized entropy index with parameter 2 (ge2), located on the same side of axis3 as , is sensitive to inequality at the top. Therefore,  and  pick up salient features of the distribution that are less-well detected by other measures.

**Figure A2a1: The unrotated PCA components of the 30 indicators of inequality (PCA scores)**



Axis 2

Axis 3

**Table A2a2: OLS Coefficients: Regression of the Gini, Atkinson 2 and Generalized Entropy 2 indices on the ABG Coefficients**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coefficient | S.E. | T | P > t | 95% C.I. min | 95% C.I. max |
| Gini index r2 = 0.9849 | | | | | | |
|  | 0.8978 | 0.0076 | 117.9 | 0.0000 | 0.8828 | 0.9128 |
|  | 0.4839 | 0.0160 | 30.2 | 0.0000 | 0.4523 | 0.5154 |
|  | 0.1192 | 0.0141 | 8.4 | 0.0000 | 0.0913 | 0.1471 |
| Cons. | 0.0277 | 0.0024 | 11.3 | 0.0000 | 0.0228 | 0.0325 |
| Atkinson 2 r2 = 0.3964 | | | | | | |
|  | 1.4425 | 0.1322 | 10.9 | 0.0000 | 1.1820 | 1.7030 |
|  | 0.0293 | 0.2778 | 0.1 | 0.9160 | -0.5182 | 0.5767 |
|  | 1.9340 | 0.2457 | 7.9 | 0.0000 | 1.4500 | 2.4181 |
| Cons. | -0.0506 | 0.0425 | -1.2 | 0.2350 | -0.1344 | 0.0331 |
| Generalized entropy 2 r2 = 0.4131 | | | | | | |
|  | 3.0979 | 0.2582 | 12.0 | 0.0000 | 2.5891 | 3.6067 |
|  | 3.8710 | 0.5426 | 7.1 | 0.0000 | 2.8018 | 4.9403 |
|  | -0.0388 | 0.4798 | -0.1 | 0.9360 | -0.9842 | 0.9067 |
| Cons. | -0.5315 | 0.0830 | -6.4 | 0.0000 | -0.6951 | -0.3680 |

Note: VIF < 1.28; N = 232

The results here confirm that the estimated ABG parameters reflect central features of empirical distributions, and help us to understand the role played by other indicators. **TABLE A2a2** uses the results from our 232 samples to shed light on the relation between the Gini index, the Atkinson 2 index, the generalized entropy 2 index and the ABG coefficients.

* The Gini index is very similar to  and is also correlated with the values of  (showing inequality at the top), but has almost no relation to . As a measure of inequality, the Gini index is (1) sensitive to the median (as is ), and (2) rich-oriented (like ).
* The Atkinson 2 index is more sensitive to lower-tail inequality. In the Atkinson 2 regressions  and  are very significant, but  is not: the Atkinson 2 index is sensitive to both poverty and general inequality (the Gini coefficient).
* The generalized entropy 2 index is correlated with both  and .

This analysis of correlations then suggests that the triple ABG parameters can be seen as contenders for the three coefficients of the Gini, Atkinson 2 and Generalized entropy 2 indices (GA2GE2). To see which triple performs best, we consider nested models of the five income-class proportions (po, mcl, mc, mcu, ri). Table A2a3 compares the goodness of fit (in terms of delta r2) when ABG is first and GA2GE2 second, and vice versa. This comparison shows that the ABG triple always outperforms the GA2GE2 triple, with the advantage of ABG being particularly striking for the explanation of mcl and mcu, the lower and upper middle class respectively. In these 232 cases, ABG generally outperforms the GA2GE2 triple in terms of the prediction of income-class size.

We can also ask whether the ABG method provides a better assessment of polarization than the Wolfson index (Wolfson, 1986). The Wolfson index was developed from the Gini index, improving its sensitivity to median stretches when the other indices remain almost unchanged. Here, the ratio rpol = (mcl + mcu)/mc, as defined earlier, should rise with polarization. The linear correlation matrix in Table A2a4 shows that, in terms of r2, the Wolfson index is indeed better than the Gini in predicting the rpol ratio, but  is even better.

**TABLE A2a3 : R2 Added Value in Nested Models of Income-Class Proportions of the ABG Coefficients and the GA2GE2 Triple Coefficients** **(Gini Index, Atkinson 2, Generalized Entropy 2)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ABG first | GA2GE2 delta r2 | Improvement sig. p |  | GA2GE2 first | ABG delta r2 | Improvement sig. p |
| Po | 0.9775 | 0.0011 | 0.0110 |  | 0.8855 | 0.0931 | 0.0000 |
| Mcl | 0.3642 | 0.0460 | 0.0007 |  | 0.1139 | 0.2963 | 0.0000 |
| Mc | 0.9231 | 0.0091 | 0.0000 |  | 0.8668 | 0.0654 | 0.0000 |
| Mcu | 0.4566 | 0.0490 | 0.0001 |  | 0.3753 | 0.1302 | 0.0000 |
| Ri | 0.9837 | 0.0006 | 0.0400 |  | 0.9655 | 0.0187 | 0.0000 |

**Table** A2a4**: Correlation between the Ratio of Polarization, Gini, Wolfson Index and ABG coefficients**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Var | Rpol | Gini | Wolfson |  |  |  |
| Rpol | 1 |  |  |  |  |  |
| Gini | 0.8263 | 1 |  |  |  |  |
| Wolfson | 0.8499 | 0.9831 | 1 |  |  |  |
|  | 0.9197 | 0.9524 | 0.9778 | 1 |  |  |
|  | -0.5652 | -0.1603 | -0.2373 | -0.4252 | 1 |  |
|  | -0.3861 | -0.2392 | -0.3229 | -0.3784 | 0.3605 | 1 |

We now consider a nested model comparison of the 232 datasets with respect to middle-class polarization (rpol). When entered first, the Gini coefficient explains more than half of the variation in rpol, with the Wolfson adding a further 4.2%, which is significant; when the Wolfson index is entered first, the r2 is 72.2%, with the Gini adding no further significant explanatory power. The Wolfson index does therefore act as a good measure of polarization, although to this extent  performs better (see Table A2a5). In general, for the different aspects of inequality measurement, the ABG method offers interpretable parameters that generally outperform the other methods in terms of the description of the distribution and the size of income classes.

**Table** A2a5**: R2 Added Value in Nested Models of Rpol (Middle Class Polarization) of the Gini Coefficient and Wolfson Index**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Gini first | Wolfson delta r2 | Improvement sig. p |  | Wolfson first | Gini delta r2 | Improvement sig. p |
| Rpol | 0.6828 | 0.0422 | 0.0000 |  | 0.7224 | 0.0026 | 0.1430 |
|  | first | Wolfson delta r2 | Improvement sig. p |  | Wolfson first | delta r2 | Improvement sig. p |
| Rpol | 0.8458 | 0.0556 | 0.0000 |  | 0.7224 | 0.1789 | 0.0000 |

**References of this appendix**

Abdelkrim, A. and J.-Y. Duclos *User Manual for Stata Package DASP: Version 2.3*, PEP, World Bank, UNDP and Université Laval, 2013.

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

Alderson, A. S., J. Beckfield and F. Nielsen, "Exactly How Has Income Inequality Changed? Patterns of Distributional Change in Core Societies." *International Journal of Comparative Sociology*, 46, 405-423, 2005.

Berry, A., F. Bourguignon, and C. Morrisson, "Changes in the World Distributions of Income Between 1950 and 1977," *Economic Journal*, 93, 331–50, 1983.

Brandolini, A. and A. B. Atkinson, "Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study," *Journal of Economic Literature*, 39, 771–99, 2001.

Chakravarty, S. R. and C. D'Ambrosio, "Polarization Orderings Of Income Distributions," *Review of Income and Wealth*, 56, 47–64, 2010.

Dallinger, U., "The Endangered Middle Class? A Comparative Analysis of the Role Public Redistribution Plays", *Journal of European Social Policy*, 23(1), 83–101, 2013.

Everitt, B. S. and G. Dunn, *Applied Multivariate Data Analysis*, Edward Arnold, London, 2001.

Foster, J., J. Greer, and E. Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica,* 52, 761–66, 1984.

Gini, C., Sulla misura della concentrazione e della variabilità dei caratteri. Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti, LXXIII, parte II, 73, 1203–48, 1914.

Jenkins, S. P., "Distributionally-Sensitive Inequality Indices and the GB2 Income Distribution," Review of Income and Wealth, 55, 392–98, 2009.

———, "INEQDECO: Stata Module to Calculate Inequality Indices with Decomposition by Subgroup," Statistical Software Components S366007, Boston College Department of Economics, revised 19 Apr 2001, 1999/2010.

Pressman, S., "The Decline of the Middle Class: An International Perspective," *Journal of Economic Issues*, 41, 181–200, 2007.

Yitzhaki, S., "Stochastic Dominance, Mean Variance, and Gini's Mean Difference," *American Economic Review*, 72, 178–85, 1982.

———, "On an Extension of the Gini Inequality Index," *International Economic Review*, 24, 617–28, 1983.

Wolfson, M. C., "Stasis amid Change: Income Inequality in Canada, 1965-1983," *Review of Income and Wealth*, 32, 337–69, 1986.

1. The Wolfson index is chosen here because it is standard in the literature, even though more reliable propositions exist (Alderson et al., 2005, Chakravarty and D’Ambrosio, 2010). [↑](#footnote-ref-1)
2. A log-symmetric definition such as .75 to 1.33 might be preferred, but the .75 to 1.25 of the median definition is far more common in the literature (Pressman, 2007). Working on quintile dynamics, Dallinger (2013) found similar variations in the different sub-strata of the middle classes. [↑](#footnote-ref-2)
3. In the social sciences, PCA is a very common tool for the multidimensional descriptive synthesis of continuous variables (Everitt and Dunn, 2001: chap. 3, 48sqq.). PCA extracts (via the diagonalization of the correlation matrix of “active variables”, here the selected inequality indicators) a hierarchy of complementary axes 1, 2, 3, etc. from higher to lower levels of variance. Figure A2a1 presents the scores (correlations) between axes 2 and 3 and the indicators. [↑](#footnote-ref-3)