E C O N O M I C S  B U L L E T I N

# On the Link Between the Concepts of Kurtosis and Bipolarization

Jacques SILBER
*Bar-Ilan University*

Joseph Deutsch
*Bar-Ilan University*

Meital Hanoka
*Bar-Ilan University (Ph.D. student)*

## *Abstract*

In a paper on the measurement of the flatness of an income distribution Berrebi and Silber (1989) showed how it was possible to derive from the Gini index a measure of the degree of Kurtosis of a distribution whose definition made it quite similar to the more famous Pearson measure of Kurtosis. This note shows that it is possible to derive from the index of flatness proposed by Berrebi and Silber (1989) a measure of bipolarization that has all the important properties one would like a bipolarization index to have.

This paper was written while Jacques Silber was visiting the Laboratorio Riccardo Revelli at the Collegio Carlo Alberto, in Moncalieri, Italy. He is very thankful to the Laboratorio and in particular to its director, Bruno Contini, for their very warm hospitality.
**Citation:** SILBER, Jacques, Joseph Deutsch, and Meital Hanoka, (2007) "On the Link Between the Concepts of Kurtosis and Bipolarization." *Economics Bulletin*, Vol. 4, No. 36 pp. 1-5
**Submitted:** September 5, 2007.  **Accepted:** October 2, 2007.
**URL:** http://economicsbulletin.vanderbilt.edu/2007/volume4/EB-07D30002A.pdf

## 1. Introduction

In a critical review of the concept of Kurtosis, Balanda and MacGillivray (1988) wrote that "it is best to define kurtosis vaguely as the location- and scale-free movement of probability mass from the shoulders of a distribution into its centre and tails." As stressed by Wuensch (2007) "if one starts with a normal distribution and moves scores from the shoulders into the center and the tails, keeping variance constant, kurtosis is increased. The distribution will likely appear more peaked in the center and fatter in the tails."

In a paper on the measurement of the flatness of an income distribution Berrebi and Silber (1989) showed how it was possible to derive from the Gini index a measure of the degree of Kurtosis of a distribution whose definition made it quite similar to the more famous Pearson measure of Kurtosis. The purpose of this note is to show that it is also possible to derive from the index of flatness proposed by Berrebi and Silber (1989) a measure of bipolarization that has all the important properties one would like a bipolarization index to have. Section II recalls the main results obtained by Berrebi and Silber (1989) while Section III proves the link between their measure of the flatness of a distribution and the concept of bipolarization.

## 2. On the Measurement of the Flatness of an Income Distribution:

Pearson's (1895) famous Kurtosis index is defined as

$$K_P = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^2} \tag{1}$$

where, in our case, $x_i$ would be the income of individual $i$, $n$ the total number of individuals and $\bar{x}$ the average income in the population.

Expression (1) may be also expressed as

$$K_P = \frac{\sum_{i=1}^{n/2}(x_i-\bar{x})^4 + \sum_{i=(n/2)+1}^{n}(\bar{x}-x_i)^4}{\left[\sum_{i=1}^{n/2}(x_i-\bar{x})^2 + \sum_{i=(n/2)+1}^{n}(x_i-\bar{x})^2\right]^2} \tag{2}$$

Assuming that $x_1 \geq x_2 \geq \ldots x_i \geq \ldots x_n$ , Berrebi and Silber (1989) have proposed an alternative measure of Kurtosis defined as

$$K_G = \frac{\sum_{i=1}^{n/2}[(n-4i+2)x_i - m] + \sum_{i=(n/2)+1}^{n}[m-(4i-3n-2)x_i]}{\sum_{i=1}^{n/2}[(n-2i+1)x_i - m] + \sum_{i=(n/2)+1}^{n}[m-(2i-n-1)x_i]} \tag{3}$$

where $m$ is the median of the income distribution.

The similarity between (2) and (3) is clear. In Pearson's (1895) formulation the central value of reference is the arithmetic mean of the distribution while in the formulation suggested by Berrebi and Silber (1989) the central value is the median. But note that in both formulations the gaps with respect to the central value are given a higher weight in the numerator than in the denominator.

Berrebi and Silber (1989) have however shown that (3) could also be expressed as

$$K_G = \frac{1}{4}\frac{\Delta_R + \Delta_P}{\Delta} \tag{4}$$

where $\Delta$, $\Delta_R$ and $\Delta_P$ are respectively the mean difference of the whole distribution, of the distribution of the "rich" individuals, the latter being defined as those with an income higher than the median income, and of the "poor" individuals, the latter being defined as those with an income lower than the median income. More precisely $\Delta$, $\Delta_R$ and $\Delta_P$ are defined as

$$\Delta = \frac{1}{n^2}\sum_i^n \sum_j^n \left| x_i - x_j \right| \tag{5}$$

$$\Delta_R = \frac{1}{(n/2)^2}\sum_i^{n/2} \sum_j^{n/2} \left| x_i - x_j \right| \tag{6}$$

$$\Delta_P = \frac{1}{(n/2)^2}\sum_{(n/2)+1}^{n} \sum_{(n/2)+1}^{n} \left| x_i - x_j \right| \tag{7}$$

**3. The Link Between The Index of Flatness and the Measurement of Bipolarization:**

Let us now define an index $P_G$ as

$$P_G = 1 - K_G \tag{8}$$

Since the income distributions of the "rich" and of the "poor" do not overlap, it can be shown (see, Nygärd and Sandström, 1981) that in such a case

$$\Delta = (1/4)(\Delta_R + \Delta_P) + \Delta_B \tag{9}$$

where $\Delta_B$ is the between groups mean difference, the groups representing the "poor" and the "rich" as they were defined before. It is in fact easy to prove (see, Berrebi and Silber, 1989) that, since these two groups are of equal size, the between groups mean difference $\Delta_B$ , which assumes that all the "rich" earn the average income $\bar{y}_R$ of the "rich" and all the "poor" earn the average income $\bar{y}_P$ of the poor, may be expressed as

$$\Delta_B = (1/2)(\bar{y}_R - \bar{y}_P) \tag{10}$$

Combining (4), (8) and (9) we end up with

$$P_G = 1 - [(1/2)((\Delta_R + \Delta_P)/\Delta)] = [\Delta - ((1/2)(\Delta_R + \Delta_P))/\Delta]$$

$$\leftrightarrow P_G = [((1/4)(\Delta_R + \Delta_P) + \Delta_B) - ((1/2)(\Delta_R + \Delta_P))]/\Delta$$

$$\leftrightarrow P_G = [\Delta_B - ((1/4)(\Delta_R + \Delta_P))]/\Delta \tag{11}$$

$$\leftrightarrow P_G = [\Delta_B - ((1/4)(\Delta_R + \Delta_P))]/[\Delta_B + ((1/4)(\Delta_R + \Delta_P))] \tag{12}$$

Expression (12) shows clearly that $P_G$ will decrease when the within groups dispersion increases, that is, when $\Delta_R$ or $\Delta_P$ increases. In addition, since the weight of $\Delta_B$ in (12) is greater in its

numerator than in its denominator, it is also easy to see that $P_G$ will increase when the between groups dispersion increases.
These are however the two principal features of a bipolarization index which are often called, in the literature, the axioms of *Non-Decreasing Spread* and *Non-Decreasing Bipolarity* (see, Esteban and Ray, 1994, Wolfson, 1994 and 1997, Wang and Tsui, 2000, Chakravarty and Majumder, 2001 and Chakravarty et al., 2007)
We should also remember that the Gini index $G$ for the whole income distribution, the between groups Gini index $G_B$, the Gini index $G_P$ among the "poor" and the Gini index $G_R$ among the "rich" may be respectively be expressed (see, Kendall and Stuart, 1969, for a general definition of the Gini index) as

$$G = (1/2)(\Delta / \bar{y}) \tag{13}$$

$$G_B = (1/2)(\Delta_B / \bar{y}) \tag{14}$$

$$G_P = (1/2)(\Delta_P / \bar{y}_P) \tag{15}$$

$$G_R = (1/2)(\Delta_R / \bar{y}_R) \tag{16}$$

Finally, in the case of non-overlapping groups, the overall Gini index may be expressed (see, Silber, 1989b) as

$$G = G_B + G_W \tag{17}$$

where $G_W$ refers to the within groups Gini index and is written, in our case, as

$$G_W = f_P s_P G_P + f_R s_R G_R \tag{18}$$

where $f_P$ ,$f_R$ , $s_P$ and $s_R$ refer respectively to the population shares of the groups of poor and rich and to the income shares of these two groups. Since we assumed that

$$f_P = f_R = 1/2 \tag{19}$$

and since in our case

$$s_P = (1/2)(\bar{y}_P / \bar{y}) \tag{20}$$

and

$$s_R = (1/2)(\bar{y}_R / \bar{y}) \tag{21}$$

we end up, combining expressions (18) to (21) with

$$G_W = (1/4)(\bar{y}_P / \bar{y})G_P + (1/4)(\bar{y}_R / \bar{y})G_R \tag{22}$$

If we combine now expressions (11), (14), (15), (16), (17) and (22) it is easy to show that we will end up with

$$P_G = (G_B - [(1/4)(\bar{y}_P / \bar{y})G_P + (1/4)(\bar{y}_R / \bar{y})G_R])/G$$

$$\leftrightarrow P_G = (G_B - G_W)/G \tag{23}$$

Note first the relative similarity between the definition of the bipolarization index $P_G$ given in (23) and the polarization index $P_W$ proposed by Wolfson (1992) which was defined as

$$P_W = (G_B - G_W)(\bar{y}/m) \tag{24}$$

Second note also the similarity between the polarization index suggested by Kanbur and Zhang (2001) who defined their index $P_{KZ}$ as

$$P_{KZ} = I_B /(\Sigma_g w_g I_g) \tag{25}$$

where $I_B$ refers to any between groups inequality index, $I_g$ to the corresponding inequality index within group $g$ and $w_g$ to the weight of group $g$ (generally a population weight but in the case of the Gini index it has to be the product of the population and income weight of group $g$, as indicated in Silber, 1989). In other words in the case of the Gini index, $P_{KZ}$ would be defined as

$$P_{KZ} = G_B / G_W \tag{26}$$

which is an unbounded index at the difference of the index $P_G$ proposed in this paper.
Note also the link between the indices $P_G$ and $P_{KZ}$ when the latter is defined on the basis of the Gini Index. Combining (17), (23) and (26) we derive that

$$P_G = (G_B - G_W)/(G_B + G_W) = [(G_B / G_W) - 1]/[(G_B / G_W) + 1] = (P_{KZ} - 1)/(P_{KZ} + 1) \tag{27}$$

Clearly both indices move in the same direction since $\partial P_G / \partial P_{KZ} > 0$.

## 4. Conclusion:

We have attempted in this note to show that, at least in the case of two non-overlapping groups of equal size, there was a clear link between the concept of bipolarization and that of the kurtosis of an income distribution. The analysis was limited to the case of two non-overlapping groups of equal size. It seems that the definition of the polarization index $P_G$ could be easily extended to that of several non overlapping groups but the existence of a link in such a case with the concept of kurtosis remains to be proven. The extension of the use of the index $P_G$ to the case of overlapping groups is probably more problematic and additional work is certainly required before some conclusions may be drawn.

## References

Balanda, K. P. and H. L. MacGillivray (1988) "Kurtosis: A Critical Review," *American Statistician*, **42**: 111-119.

Berrebi, Z. M. and J. Silber (1989) "Deprivation, the Gini Index of Inequality and the Flatness of an Income Distribution," *Mathematical Social Sciences* **18**: 229-237.

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

Chakravarty, S. R., A. Majumder and S. Roy (2007) "A Treatment of Absolute Indices of Polarization," *Japanese Economic Review*, **58**: 273-293.

Esteban, J.-M. and D. Ray (1994) "On the Measurement of Polarization," *Econometrica*, **62**: 819-852.

Kendall, M. G. and A. Stuart (1969) *The Advanced Theory of Statistics*, Charles Griffen: London.

Nygärd, F. and A. Sandström (1981) *Measuring Income Inequality*, Almqvist and Wiksell International: Stockholm.

Pearson, K. (1895) "Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material," *Philosophical Transactions of the Royal Society of London*, **186**: 343-414.

Silber, J. (1989) "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality," *Review of Economics and Statistics* **71**: 107-115.

Wang, Y. Q. and K. Y. Tsui (2000) "Polarization Orderings and New Classes of Polarization Indices," *Journal of Public Economic Theory*, **2**: 349-363.

Wolfson, M. C. (1994) "When Inequalities Diverge," *American Economic Review, Papers and Proceedings*, **84**: 353-358.

Wolfson, M. C. (1997) "Divergent Inequalities: Theory and Empirical Results," *Review of Income and Wealth*, **43**: 401-421.

Wuensch, K., "Skewness, Kurtosis and the Normal Curve," in *Karl Wuensch's Statistical Lessons*, available on http://core.ecu.edu/psyc/wuenschk/docs30/Skew-Kurt.doc