

Mosler, Karl

Article — Published Version

Commentary on “From unidimensional to multidimensional inequality: a review”

METRON

Provided in Cooperation with:

Springer Nature

Suggested Citation: Mosler, Karl (2020) : Commentary on “From unidimensional to multidimensional inequality: a review”, METRON, ISSN 2281-695X, Springer, Milan, Vol. 78, Iss. 1, pp. 51-54,
<https://doi.org/10.1007/s40300-020-00165-7>

This Version is available at:

<https://hdl.handle.net/10419/288770>

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.



<https://creativecommons.org/licenses/by/4.0/>



Commentary on “From unidimensional to multidimensional inequality: a review”

Karl Mosler¹

Received: 17 February 2020 / Accepted: 28 February 2020
© The Author(s) 2020, corrected publication 2021

Abstract

These remarks supplement the paper of Andreoli and Zoli from a practical view-point, that is, of a data analyst who wishes to compare distributions of socio-economic endowments regarding their inequality.

Keywords Multidimensional inequality · Crossing Lorenz curves · Dominance tests

The authors of this paper are to be praised for their well-written survey. They provide a comprehensive account of the measurement of economic inequality in one and many dimensions, focussing on Lorenz-like orderings. In addition, several fresh views on multidimensional inequality are given regarding the consequences of aggregation, the meaning of correlation increasing transfers, and the connection of inequality with the dissimilarity of multivariate distributions.

My remarks supplement their work from a practical viewpoint, that is, of a data analyst who wishes to compare distributions of socio-economic endowments regarding their inequality. The analyst draws conclusions from data about income, wealth, education and other attributes of the populations under inquiry. But these data are more or less complete, reliable and accurate. Their empirical distribution has to be seen as an unknown “true” distribution values plus an eventual noise or contamination.

With one attribute, empirical Lorenz (or generalized Lorenz) curves are determined and point-wise compared, in order to decide whether one population is more unequal than the other. However, empirical Lorenz curves, say of income, rather often cross and exhibit no clear order. Such data tells us that the curves either are “really” non-ordered (i.e. there exist Schur-convex functions that rank them in different directions) or that they are ordered but the noise overlays the ordering.

To cope with crossing Lorenz curves, a weakening of the usual Lorenz dominance is often in place. Several approaches offer themselves:

- Shifting to a weaker relation that has a specific meaning, like higher degree dominance.
- Restricting the order to an “essential part” of the data.
- Restricting the comparison to one (or several) inequality indices.

✉ Karl Mosler
kmosler@uni-koeln.de

¹ Institute of Econometrics and Statistics, Universität zu Köln, Koeln, Germany

- Building a stochastic model of data generation and performing a statistical test for Lorenz order.

When two empirical Lorenz curves, L_X and L_Y , cross, this often happens in their tails. If they are ordered in their “middle part”, say

$$L_X(t) \leq L_Y(t) \quad \text{for all } t \in [t_0, t_1],$$

we consider them as *quantile-restricted Lorenz ordered* or (t_0, t_1) -Lorenz ordered [9]. Of course, $[t_0, t_1]$ must include a large enough interval, covering e.g. the *coefficient of minimal majority* $MM(X) = L_X^{-1}(0.5)$. Arguments backing this approach are:

- Data in the tails are often less reliable and precise, compared to those in the middle.
- Tail data may be missing due to reluctant answering behavior.
- Sampling schemes of Official Statistics usually exclude the very poor and the very rich.
- Recipients of middle incomes appear to be most decisive in elections.
- The actual size of very rich incomes, beyond some level, is of little public interest; similarly, that of the very poor.

Observe that, if all people in the lower part of the population, up to quantile t_0 , have the same income, and all people in the upper part beyond quantile t_2 have the same income, then the (t_0, t_1) -restricted generalized Lorenz order is equivalent to the unrestricted generalized Lorenz order.

Many authors have constructed tests for *generalized Lorenz order (or distribution equality) against non-order*, $H_0 : GL_X(t) \leq GL_Y(t)$ for all t (or $H_0 : GL_X(t) = GL_Y(t)$ for all t) against not H_0 , among them [7,11], and [3]. However, to statistically establish the Lorenz order, reverse tests are needed, testing *non-order against generalized Lorenz order*, an approach, which has been pursued by [5] and [4].

Next let us consider multidimensioned Lorenz orderings of inequality. Our remarks concern *positive price majorization* (in the terminology of [1]), while other multivariate extensions of usual Lorenz order can be similarly treated.

Let X and Y have *multivariate* empirical distributions in \mathbb{R}^d and define: X is less unequal than Y in *positive price majorization*, if $p'X$ is less unequal than $p'Y$ in univariate Lorenz order, for every $p \in \mathbb{R}_+^d$. With other words, for every non-negative *price vector* $p \in \mathbb{R}_+^d$, the *budget* $p'X$ is less unequal than the budget $p'Y$. Positive price majorization has many names in the literature, among them *price Lorenz order* and *positive directional majorization*.

With univariate data, two empirical Lorenz curves are straightforwardly checked whether one dominates the other. With multivariate data, this task is comparatively simple if we are able to restrict to a few prices p . Otherwise it comes out to be more involved. To check the data for positive price majorization, we may use a characterization of the order by upper regions.

The distribution of X is fully characterized by its *zonoid central regions* [6],

$$D_\alpha(X) = \left\{ \sum_{i=1}^n \theta_i x_i : \sum_{i=1}^n \theta_i = 1, 0 \leq \alpha \theta_i \leq \frac{1}{n} \right\}, \quad \alpha \in [0, 1].$$

The regions form nested convex sets, which can be regarded as *inter-quantile regions*, ranging from

$$D_1(X) = \left\{ \frac{1}{n} \sum_{i=1}^n x_i \right\} \quad \text{to} \quad D_0(X) = \text{conv}(x_1, \dots, x_n).$$

Positive price majorization is characterized as follows [8]:

Proposition 1 *X is less unequal than Y in positive price majorization if and only if*

$$D_\alpha(\tilde{X}) \oplus \mathbb{R}_+^d \subset D_\alpha(\tilde{Y}) \oplus \mathbb{R}_+^d \text{ for all } 0 \leq \alpha \leq 1. \tag{1}$$

Here \oplus denotes the *Minkowski sum*, $K \oplus R = \{x + y : x \in K, y \in R\}$. The set

$$D_\alpha(\tilde{X}) \oplus \mathbb{R}_+^d = \left\{ z : z \geq \sum_{i=1}^n \theta_i \tilde{x}_i : \sum_{i=1}^n \theta_i = 1, \quad 0 \leq \alpha \theta_i \leq \frac{1}{n} \right\}$$

will be mentioned as a *zonoid upper region*. Observe that

$$D_\alpha(\tilde{X}) \subset D_\alpha(\tilde{Y}) \tag{2}$$

if and only if every vertex of $D_\alpha(\tilde{X})$ is a convex combination of the vertices of $D_\alpha(\tilde{Y})$.

The set of vertices of a zonoid trimmed region can be exactly calculated by a *breadth-first search algorithm* [10]. The normal of each *facet* is identified by exactly d data points, which yield its *ridges*. As we have to verify (1) and not (2), instead of considering arbitrary normals in \mathbb{R}^d we consider only non-negative ones. From the ridges, information about the *adjacency* of the facets is extracted. The facets are arranged through a *tree-based order*, by which the whole surface can be traversed efficiently with a minimal number of computations. The algorithm has been programmed in C++ and is available in the R-package *WMTRegions* [2]. The task is simplified by dropping less important attributes (if there are any), which means setting one or more prices, p_j , to 0.

Like the univariate Lorenz order, its multivariate extension can be quantile-restricted. Zonoid regions are multivariate analogues of inter-quantile intervals. "Outer data" are excluded from further analysis by comparing zonoid upper regions only above some minimum level α_0 . *X is less unequal than Y in the quantile-restricted price Lorenz order if*

$$D_\alpha(\tilde{X}) \oplus \mathbb{R}_+^d \subset D_\alpha(\tilde{Y}) \oplus \mathbb{R}_+^d \text{ for all } \alpha_0 \leq \alpha \leq 1.$$

Finally, zonoid central (as well as upper) regions can be regarded as set-valued statistics. For this, the data are considered in a probabilistic setting. Positive price majorization and its characterization by upper regions can be extended to any *random vectors X and Y* in \mathbb{R}^d that have finite expectations; for details see [8]. We regard the data (x_1, x_2, \dots, x_n) as the realization of a random sample (X_1, X_2, \dots, X_n) from X . Then the empirical region

$$D_\alpha(X_1, X_2, \dots, X_n) \tag{3}$$

is a set-valued statistic that estimates the central region $D_\alpha(X)$.

Zonoid central regions (3) satisfy a *Law of Large Numbers*. They are strongly consistent estimators of $D_\alpha(X)$, viz.

$$D_\alpha(X) = H\text{-}\lim_{n \rightarrow \infty} D_\alpha(X_1, X_2, \dots, X_n).$$

H-lim means limit in the *Hausdorff distance* $\delta_H(K, R) = \max_{p \in S^{d-1}} |h_K(p) - h_R(p)|$, where $h_K(p) = \max_{x \in K} p'x$, $p \in S^{d-1}$, is the *support function* of a convex body $K \subset \mathbb{R}^d$. More explicitly,

$$P\left[\lim_{n \rightarrow \infty} \max_{p \in S^{d-1}} |h_{D_\alpha(X_1, \dots, X_n)}(p) - h_{D_\alpha(X)}(p)| = 0\right] = 1.$$

Moreover, a *Central Limit Theorem* holds for the empirical regions (3); see [8].

This allows for statistical inference about positive price majorization, specifically about the inclusion of zonoid regions at a given level α . Let $\alpha \in]0, 1[$. Given two samples (X_1, \dots, X_n) from X and (Y_1, \dots, Y_m) from Y we aim at securing the hypothesis

$$H_1 : D_\alpha(\tilde{X}) \oplus \mathbb{R}_+^d \subset D_\alpha(\tilde{Y}) \oplus \mathbb{R}_+^d \quad \text{against the null } H_0 : \text{not } H_1.$$

H_1 is equivalent to $h_{D_\alpha(\tilde{X})}(p) \leq h_{D_\alpha(\tilde{Y})}(p)$ for all $p \geq 0$, that is, by $h_{D_\alpha(\tilde{X})}(p) = \frac{1}{\alpha} E[p' \tilde{X}]$, to

$$H_1 : E[p' \tilde{X}] \leq E[p' \tilde{Y}] \quad \text{for all } p \geq 0,$$

which may be tested through a proper bootstrap procedure.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Andreoli, F., Zoli, C.: From unidimensional to multidimensional inequality: a review. *Metron* (2020). <https://doi.org/10.1007/s40300-020-00168-4>
2. Bazovkin, P., Mosler, K.: An exact algorithm for weighted-mean trimmed regions in any dimension. *J. Stat. Softw.* **47**(13), 1–29 (2012)
3. Davidson, R., Duclos, J.-Y.: Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica* **68**(6), 1435–1464 (2000)
4. Davidson, R., Duclos, J.-Y.: Testing for restricted stochastic dominance. *Econ. Rev.* **32**(1), 84–125 (2013)
5. Kaur, A., Rao, B.P., Singh, H.: Testing for second-order stochastic dominance of two distributions. *Econ. Theory* **10**(5), 849–866 (1994)
6. Koshevoy, G., Mosler, K.: Zonoid trimming for multivariate distributions. *Ann. Stat.* **25**(5), 1998–2017 (1997)
7. McFadden, D.: Testing for stochastic dominance. In: *Studies in the Economics of Uncertainty*. Springer, New York, pp. 113–134 (1989)
8. Mosler, K.: *Multivariate Dispersion, Central Regions and Depth: the Lift Zonoid Approach*. Springer, New York (2002)
9. Mosler, K.: Restricted Lorenz dominance of economic inequality in one and many dimensions. *J. Econ. Inequal.* **2**(2), 89–103 (2004)
10. Mosler, K., Lange, T., Bazovkin, P.: Computing zonoid trimmed regions in dimension $d > 2$. *Comput. Stat. Data Anal.* **53**, 2500–2510 (2009)
11. Schmid, F., Trede, M.: A Kolmogorov-type test for second-order stochastic dominance. *Stat. Probab. Lett.* **37**(2), 183–193 (1998)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.