A toolbox for measuring heterogeneity and efficiency using zonotopes

Marco Cococcioni
Department of Information Engineering
University of Pisa
Pisa, Italy
marco.cococcioni@unipi.it

Marco Grazzi
Department of Economic Policy
Università Cattolica del Sacro Cuore
Milano, Italy
marco.grazzi@unicatt.it

Le Li
School of Economics
Guangzhou College of Commerce
Guangzhou, China
lile7k@gmail.com

Federico Ponchio
Visual Computing Lab
ISTI CNR
Pisa, Italy
federico.ponchio@isti.cnr.it

Abstract. In this work, we describe the new command `zonotope`, which, by resorting to a geometry-based approach, provides a measure of productivity that fully accounts for the existing heterogeneity across firms within the same industry. The method we propose also enables assessment of the extent of multidimensional heterogeneity with applications to fields beyond that of production analysis. Finally, we detail the functioning of the software to perform the related empirical analysis, and we discuss the main computational issues encountered in its development.

Keywords: st0662, zonotope, heterogeneity measures, production analysis, multivariate Gini coefficient

1 Introduction

In this work, we present a command, `zonotope`, that allows measurement of the existing heterogeneity among multidimensional units that can also be nested in groups at different levels of aggregation. Here we document the application of the methodology in the context of production analysis in economics; however, the possible range of applications is much wider. Within the proposed field, the method also provides a measure of productivity (both at the aggregate and at the individual levels) and its change over time that allows us to relax many of the standard assumptions of the theory, which have been falsified by recent work on actual data.

Traditionally, empirical analysis in economics has suffered from the scarcity of disaggregated sources of data (that is, at the level of individual, household, enterprise, etc.) such that, in the analysis of behaviors at the micro level, much was left to theoretical analysis. This situation oftentimes required heroically simplifying assumptions on the behavior of agents, the tradeoffs they were facing, and the absence of any path dependency, among other things. Nowadays, the ever-growing availability of disaggregated data on business firms has revealed a much richer picture than that previously conjectured based on theories alone or on aggregate industry-level data.

Firms are different along most of the dimensions typically taken into consideration by economic analyses. To provide a brief account of what is at stake, consider that even firms within the same, narrowly defined industry display very different levels of productivities, in terms of both labor and total factor productivities.[1] Also, the relative input intensities are much different (Dosi and Grazzi 2006), even with relatively similar input prices. At least as relevant, such heterogeneities are persistent over time; that is, if there is any selection at work, its effects take much longer to display (Dosi 2007). The ubiquitous presence of such heterogeneity has been vividly expressed by Griliches and Mairesse (1999): "We [. . .] thought that one could reduce heterogeneity by going down from general mixtures as 'total manufacturing' to something more coherent, such as 'petroleum refining' or 'the manufacture of cement'. But something like Mandelbrot's fractal phenomenon seems to be at work here also: the observed variability-heterogeneity does not really decline as we cut our data finer and finer. There is a sense in which different bakeries are just as much different from each other as the steel industry is from the machinery industry."

The evidence recalled above presents several challenges to the standard theory of production and to the related empirical applications based on the notion of a "representative" firm or an industry production function and, of course, on the estimation of such production function itself. The observed combinations of inputs chosen by firms appear to be quite dispersed, hardly displaying any regularity resembling a conventional isoquant. Further, although output—as expected—increases in both inputs, this does happen in a nonmonotonic way (Dosi and Grazzi 2006). In addition, the degree to which firms substitute their inputs is challenged by empirical observations.

Given all that, how can one obtain measures of productivity, at both the firm and the industry levels, that do not require assumptions not met by data, and on the contrary, that take into account such pervasive heterogeneity within the industry? In this article, we present a new application, `zonotope`, that enables measurement of productivity in the presence of such firm-level heterogeneity. `zonotope` also allows for measuring the degree of heterogeneity of the firms making up the industry and assessing its variation over time. This latter feature can be applied in a wide range of fields beyond production analysis. For example, based on zonotopes and path polytopes inclusion, Andreoli and Zoli (2014) frame the dissimilarity comparisons of sets of group distributions, which can be applied into multigroup comparisons of segregation, discrimination, and mobility, as well as inequality evaluations.

To illustrate heterogeneity and its relevance, we provide some empirical evidence in figure 1 that focuses on the labor productivity distribution, where labor productivity is defined as the ratio of deflated turnover value over number of employees. The details on these two variables can be found in section 4.1.[2]

---

1. See, among others, Baily, Hulten, and Campbell (1992); Baldwin and Rafiquzzaman (1995); Bartelsman and Doms (2000); Disney, Haskel, and Heden (2003); Syverson (2011); and Dosi et al. (2016).
2. For illustrative purposes, the lines of code displayed here enable reproduction of panel (a) of figure 1 [the same holds for figure 2(a)]. Data and code (do-file) to reproduce all the evidence reported in this article are contained in the "replication" folder. More details on the data used in this article are reported in section 4.

```
. use empirical_firm_data
. sort NACE Year
. keep if NACE == "0091" | NACE == "0092"
(6,492 observations deleted)
. generate LogProductivity = log(Turnover/NumberOfEmployees)
. kdensity LogProductivity if Year==2006, nograph generate(x y009)
. kdensity LogProductivity if Year==2006 & NACE=="0091", nograph generate(y0091)
> at(x)
. kdensity LogProductivity if Year==2006 & NACE=="0092", nograph generate(y0092)
> at(x)
. label var y009 "009"
. label var y0091 "0091"
. label var y0092 "0092"
. twoway connected y009 y0091 y0092 x if x < 8 & x > 3, sort
> lc(black black black) lp(solid dash dot) mcolor(black black black)
> msymbol(plus Dh Oh) ytitle(Density) xtitle((log) Labor Productivity)
> title(Year 2006, pos(10) ring(0)) xscale(range(3(1)8)) legend(col(1)
> pos(1) ring(0))
```
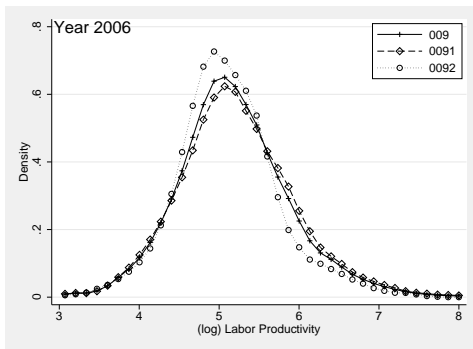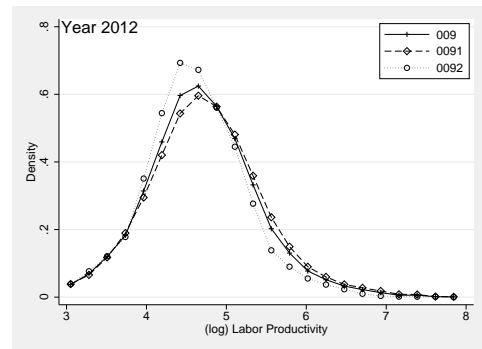

(a) Year 2006   (b) Year 2012

Figure 1. Empirical distribution of (log) labor productivity in a given 3-digit sector and two nested 4-digit sectors

   In figure 1(a), the productivity distribution of a given 3-digit sector—that is, the solid line—is sufficiently widespread to indicate the huge productivity gap between the most productive firms and the least productive ones. On the graph, productivity is measured in a log scale, which makes the heterogeneity even greater. This heterogeneity does not disappear when focusing on similar firms (the firms with a 4-digit industrial classification,[3] that is, the dashed and dotted lines); we still observe significantly different productivity levels among firms. The persistence of heterogeneity indicates not

---

3. For statistical purposes, firms are assigned to industries according to their main activity. If a firm is diversified (that is, active in more than one industry), as is often the case, then it is assigned to the sector generating more revenue. Several classifications of economic activities exist; see, for instance, Eurostat (2008) or United Nations (2008).

only that heterogeneity holds when increasing the level of disaggregation but also that it holds over time.

As recalled above, not only do firms within the same industry display very different levels of efficiency, but also they employ very diverse production techniques. This is illustrated in figure 2.

```
. use empirical_firm_data, clear
. sort NACE Year
. keep if (NACE == "0091" | NACE == "0092") & Year == 2006
(16,870 observations deleted)
. generate LogK = log(FixedAssets)
. generate LogL = log(NumberOfEmployees)
. generate LogOutput = log(Turnover)
. keep if LogK > 3 & LogK < 9 & LogL > 2 & LogL < 6
(906 observations deleted)
. bipolate LogOutput LogK LogL, collapse(mean) method(thinplatespline)
> smooth(0.1) xlevels(40) ylevels(40) saving(bipolate_result, replace)
(file bipolate_result.dta not found)
file bipolate_result.dta saved
. generate LogOutput_mean = 0
. drop LogOutput
. append using bipolate_result
(variable LogK was float, now double to accommodate using data's values)
(variable LogL was float, now double to accommodate using data's values)
(variable LogOutput_mean was float, now double to accommodate using data's
        values)
. twoway contourline LogOutput_mean LogK LogL if LogOutput_mean > 0, levels(10)
> || (scatter LogK LogL if LogOutput_mean == 0, msize(vtiny) mcolor(black)),
> title("3-digit sector 009") xtitle("Log L") ytitle("Log K")
```
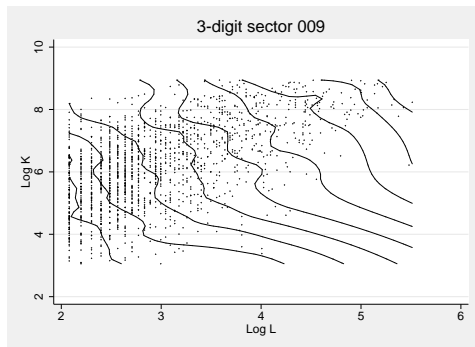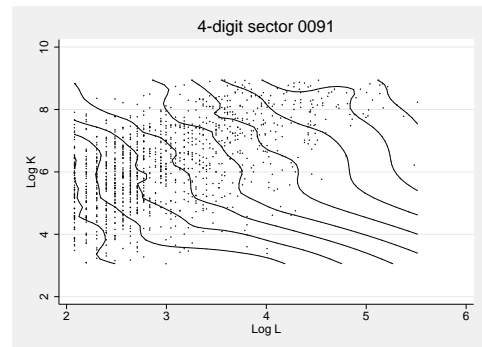

(a) 3-digit sector   (b) nested 4-digit sector

Figure 2. Contour plots of adopted techniques and output in 3- and 4-digit sectors

Figure 2(a) provides a representation of production activities of firms within a given 3-digit industry[4] assuming the standard 2-input-1-output production, where the axes represent inputs (labor and capital are proxied by number of employees and fixed assets, respectively) and the contour line displays a constant level of output as proxied by turnover value. Thus, each firm within this industry, as one observation in our empirical data sample, can in principle be represented by one point in this contour plot.

In the input plane, we first plot all such points representing the firms' labor and capital combinations from empirical data and then plot the isoquants indicating the possible combination of labor and capital corresponding to the same output level. As observed, dots in the input plane are quite dispersed while the corresponding isoquants do not display any regularity resembling a conventional production function.[5] Furthermore, the output increases with the inputs in a nonmonotonic manner. To be specific, given a quantity of one input, different firms attain the same level of output with very different levels of the other input. Again notice that this type of heterogeneity does not disappear when we increase the disaggregation of the industrial classification; similar phenomena can be observed in 4-digit-level industries, as reported in figure 2(b).

To sum up, this empirical evidence suggests that, within one industry, firms not only display very different levels of productivity but also have production techniques that are more heterogeneous. As a result, accounting for such apparent differences among firms would greatly improve productivity measurement and its change over time. In the next section, we focus on this attempt.

2   The zonotope approach to production analysis

In this section, we briefly outline the geometric approach to production analysis on which we rely for the proposed software packages. For a more detailed exposition, we refer the reader to Hildenbrand (1981) and Dosi et al. (2016).

The seminal work by Hildenbrand (1981) suggests an agnostic and data-oriented approach, which—instead of estimating some aggregate production function—offers a representation of the empirical production possibility set of an industry in the short run based on actual microdata. In such a setting, it is possible to represent a firm (or, for that matter, an establishment) in the input–output space. In such a way, the production possibility set of any given industry is represented geometrically by the space formed by the finite sum of all the line segments linking the origin and the points representing each production unit, called a zonotope. Based on this zonotope framework, Dosi et al. (2016) show that by further exploiting the properties of zonotopes, it is possible to obtain rigorous measures of heterogeneity and productivity without imposing on data a model like that implied by standard production functions.

---

4. The message would not change if a different industry were picked.
5. For a more detailed discussion, refer to Hildenbrand (1981) and Dosi and Grazzi (2006).

Similarly to Koopmans (1977), Hildenbrand (1981), and now also Dosi et al. (2016), we denote the production activity, as representing the actual technique of production unit i, by a vector

$$\mathbf{a}_i = (\alpha_{i_1}, \dots, \alpha_{i_l}, \alpha_{i_{l+1}}) \in \mathbb{R}_+^{l+1} \tag{1}$$

which indicates that during the current period, this production unit, at its best, can produce $\alpha_{i_{l+1}}$ units of output by means of $(\alpha_{i_1}, \dots, \alpha_{i_l})$ units of input. Then we can define the short-run production possibilities of an industry with N units during the current period by a finite family of production activity vectors $\{\mathbf{a}_i\}_{1 \le i \le N}$. Notice that, any vector $\mathbf{a}_i$ from the collection of vectors $\{\mathbf{a}_i\}_{1 \le i \le N}$ in $\mathbb{R}_+^{l+1}$ can be associated with a line segment

$$[0, \mathbf{a}_i] = \{ s_i \mathbf{a}_i \mid s_i \in \mathbb{R}, \quad 0 \le s_i \le 1 \}$$

Further, with the assumption of $N \ge l + 1$, Hildenbrand defines the short-run total production set associated with the family $\{\mathbf{a}_i\}_{1 \le i \le N}$ as the Minkowski sum

$$\mathbf{Y} = \sum_{i=1}^{N} [0, \mathbf{a}_i]$$

of line segments generated by production activities $\{\mathbf{a}_i\}_{1 \le i \le N}$, and, more explicitly, defines the short-run feasible industry production function as the zonotope

$$\mathbf{Y} = \left\{ \mathbf{y} \in \mathbb{R}_+^{l+1} \mid \mathbf{y} = \sum_{i=1}^{N} \phi_i \mathbf{a}_i, \quad 0 \le \phi_i \le 1 \right\}$$

Hildenbrand also defines his short-run efficient industry production function within the zonotope framework. Let's project the above-defined zonotope $\mathbf{Y}$ on its first l coordinates and denote this projection as $\mathbf{D}$, which reads

$$\mathbf{D} = \{ \mathbf{u} \in \mathbb{R}_+^l \mid \exists\, x \in \mathbb{R}_+ \ s.t. \ (\mathbf{u}, x) \in \mathbf{Y} \}$$

Thus his production function $F : \mathbf{D} \to \mathbb{R}_+$ follows:

$$F(\mathbf{u}) = \max\{ x \in \mathbb{R}_+ \mid (\mathbf{u}, x) \in \mathbf{Y} \}$$

This definition implies that, given the level $u_1, \dots, u_l$ of inputs for the industry, the maximum total output could be achieved by allocating, without any restrictions, the amounts $u_1, \dots, u_l$ of inputs over the individual production units within the industry in one of the most efficient ways. However, the frontier associated with this production function does not provide any information on the actual technological setup of the whole industry. This production function could not be the focal reference either, from a positive or from a normative point of view (Hildenbrand 1981).

Within the zonotope framework, Dosi et al. (2016) define the main diagonal of a zonotope $\mathbf{Y}$ as the diagonal joining the origin $\mathbf{O} = (0, \dots, 0) \in \mathbf{Y} \subset \mathbb{R}^{l+1}$ with its opposite vertex in $\mathbf{Y}$. They call this diagonal the production activity of the industry,

because it expresses both the amount of inputs used and the output produced by the industry. To be specific, this diagonal, denoted by $\mathbf{d}_Y$, is simply the sum of individual production activities of the N production units involved in the industry, that is,

$$\mathbf{d}_Y = (\beta_1, \ldots, \beta_l, \beta_{l+1}) = \left(\sum_{i=1}^{N} \alpha_{i_1}, \ldots, \sum_{i=1}^{N} \alpha_{i_l}, \sum_{i=1}^{N} \alpha_{i_{l+1}} \right) \in \mathbb{R}_+^{l+1} \qquad (2)$$

Obviously, if all firms in one industry were to use the same technique in a given year, all the vector-firms would lie on the same line. This is the case where only one technology is adopted and all the firms within this industry are homogeneous. In this case, the associated zonotope would degenerate to one with null volume, that is, coinciding with the diagonal $\mathbf{d}_Y$. On the other hand, the maximal heterogeneity case occurs when one industry involves some firms with almost zero inputs but sufficient output and other firms with a large quantity of inputs but little output. In such a case, the generated zonotope almost becomes a parallelotope. Starting from this simple observation on these two extreme cases, it is possible to derive a rigorous measure for industry heterogeneity.

First, let $\mathbf{A}_{i_1,\ldots,i_{l+1}}$ be the matrix whose rows are vectors $(\alpha_{i_1}, \ldots, \alpha_{i_{l+1}})$ and let $\Delta_{i_1,\ldots,i_{l+1}}$ be its determinant. It is possible to compute the volume of the zonotope $\mathbf{Y}$ in $\mathbb{R}^{l+1}$ by using the formula

$$\text{Vol}(\mathbf{Y}) = \sum_{1 \leq i_1 \leq \ldots \leq i_{l+1} \leq N} |\Delta_{i_1,\ldots,i_{l+1}}| \qquad (3)$$

where $|\Delta_{i_1,\ldots,i_{l+1}}|$ is the module of the determinant $\Delta_{i_1,\ldots,i_{l+1}}$. However, the value of $\text{Vol}(\mathbf{Y})$ depends both on the unit in which inputs and output are measured and on the number of firms. To normalize the measure, the volume of zonotope $\mathbf{Y}$ generated by the production activities $\{\mathbf{a}_i\}_{1 \leq i \leq N}$ is divided by the volume of the parallelotope with diagonal $\mathbf{d}_Y = \sum_{i=1}^{N} \mathbf{a}_i$. The parallelotope is the zonotope with the largest volume if the main diagonal is fixed. Such ratio is defined as

$$G(\mathbf{Y}) = \frac{\text{Vol}(\mathbf{Y})}{\text{Vol}(\mathbf{P}_Y)} \qquad (4)$$

where $\mathbf{P}_Y$ denote the parallelotope with diagonal $\mathbf{d}_Y$ and volume $\text{Vol}(\mathbf{P}_Y)$. The normalized volume, $G(\mathbf{Y})$, is named the Gini volume.[6]

Aside from the heterogeneity measure, Dosi et al. (2016) also suggest that the angle formed by the industry production activity vector $\mathbf{d}_Y$ with the space generated by all inputs expresses the industry productivity, and the tangent of this angle can be an appropriate measure.[7] To be specific, the measure of productivity P for a given industry including N firms at the current period is

$$P = \text{tg}\{\Theta_{l+1}(\mathbf{d}_Y)\} = \frac{\sum_{i=1}^{N} \alpha_{i_{l+1}}}{||\text{pr}_{-(l+1)}(\mathbf{d}_Y)||} \qquad (5)$$

---

6. For a generalization of the measure, refer to Franciosi, Settepanella, and Terni (2020).

7. Note that such intuition can be directly expanded to the level of the single vector-firm so as to derive an appropriate measure of productivity at the level of the firm.

where for any vector $\mathbf{v} = (x_1, \ldots, x_k) \in \mathbb{R}^k$ for $k = 2, 3, 4, \ldots$, projection map $\mathrm{pr}_{-j}(\cdot)$ follows

$$\mathrm{pr}_{-j}(\mathbf{v}) : \mathbb{R}^k \to \mathbb{R}^{k-1}$$
$$(x_1, \ldots, x_k) \mapsto (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k)$$

$\Theta_j(\mathbf{v})$ represents the angle formed by the vector $\mathbf{v}$ and the space generated by all entities in vector $\mathrm{pr}_{-j}(\mathbf{v})$, and $||\mathbf{v}||$ represents the normal of the vector $\mathbf{v}$. Furthermore, similarly, given the industry input vector $\mathbf{w} = \mathrm{pr}_{-(l+1)}(\mathbf{d}_Y) \in \mathbb{R}_+^l$, we have $\Theta_i(\mathbf{w})$ and its tangent value to measure the relative intensity of input i with respect to all the other inputs for $i = 1, \ldots, l$.

A few remarks are needed on the nature of the measure, its relation to similar techniques, and the scope of applicability. Notice that the volume of zonotope, as well as the angles, are measures and not estimates. Further, although the measure of productivity bears some similarity, mostly because of the nonparametric nature, with data envelopment analysis in production (see, among others, Farrell [1957], Charnes, Cooper, and Rhodes [1978], and Simar and Zelenyuk [2011]), the two methods are clearly distinct; the zonotope approach considers all firms, not just those on the frontier. Although both approaches are data driven and nonparametric, the emphasis of data envelopment analysis is, for the industry, to construct the efficient frontier by enveloping the data and, for the individual firm, to proxy its efficiency (for an application to Stata, see Ji and Lee [2010] and Badunenko and Mozharovskyi [2016]). Empirically, the traditional deterministic approach faces some issues, which have been addressed by recent work. To be specific, Daraio and Simar (2007) deal with the sensitivity to measurement errors and outliers, and Cazals, Florens, and Simar (2002) and Daraio and Simar (2005) propose robust frontiers. In addition, there exists a command that deals with such issues; see Belotti et al. (2013).

As recalled above, the zonotope approach provides a general framework not necessarily limited to industry heterogeneity and productivity analysis. For example, Aruka (2017) argues that the zonotope framework provides a different view of the production set and discloses new possibilities for modeling international trade. In the latter field, he suggests a further generalization of the model so that it is possible to jointly assess more than three countries and three commodities, thus abandoning the limiting special case of two countries and two commodities. In a rather different domain, the software application we are presenting here can be used to compute measures of multivariate disparity similar to those outlined in Koshevoy and Mosler (1996, 1997).

In sum, the proposed framework allows the investigation of the production and productivity dynamics of firms while taking into consideration the existing heterogeneity across firms, which is high and persistent even within the same industry. The geometry-based approach, on one side, enables us to relax many of the current existing assumptions, which have often been shown to lack empirical support. On the other side, while it allows for providing a representation in the production space with several inputs and outputs, on two or three dimensions, it is a way to provide an illustration of the trend of the industry and of individual firms.

To reinforce the last point, we take advantage of the simple one-input, one-output setting, as depicted in figure 3.
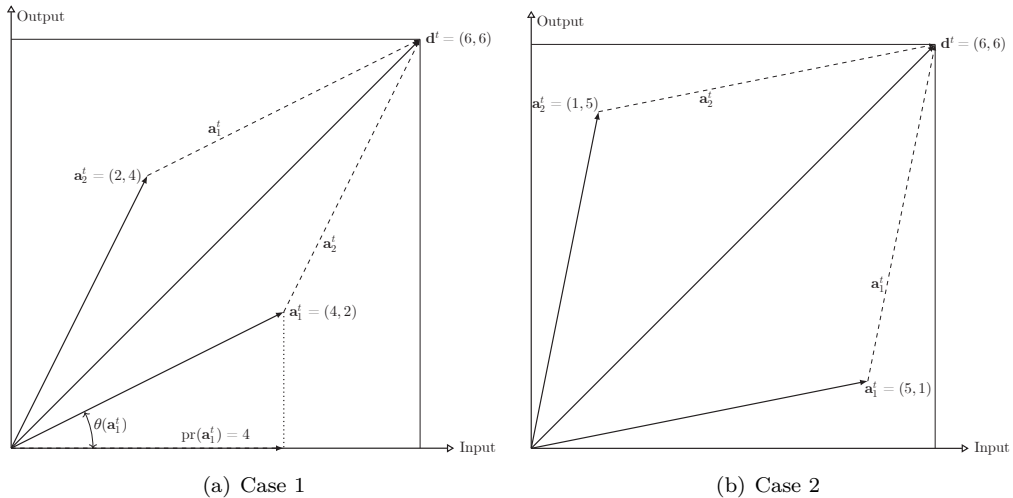

(a) Case 1   (b) Case 2

Figure 3. One-input, one-output examples of zonotopes $\mathbf{Y} = \sum_{i=1}^{2} [0, \mathbf{a}_i^t] \subset \mathbb{R}^2$

There are two firms, represented by $\mathbf{a}_1^t$ and $\mathbf{a}_2^t$ as defined in (1). In figure 3(a), $\mathbf{a}_1^t = (4, 2)$ indicates that firm 1 produces 2 units of outputs using 4 units of inputs. The industry production activity, as the sum of vector-firms defined in (2), is represented by $\mathbf{d}^t = (6, 6)$. In figure 3(b), we have identical industry production activity but with more firm heterogeneity. Comparing these two plots, the higher heterogeneity is consistent with the bigger parallelogram (the areas are equal to 12 and 24, respectively), which indicates the possibility of using the volume of zonotope defined in (3) to measure the heterogeneity. However, as discussed, one still needs a more refined measure, by applying some normalization [as in (4)] to ensure that differences are not due to, for instance, different units of measure, which might well happen in the real world. More in detail, the area is normalized by dividing it by the area of the corresponding square. This graphical illustration can also provide the intuition behind the productivity measures for the industry [defined in (5)] and for the firm [defined in (6)] in $\mathbb{R}^2$. In this respect, figure 3(a) shows that the tangent value of $\Theta(\mathbf{a}_1)$ is indeed the ratio of firm 1's output over its input, that is, labor productivity, which is one of the most popular productivity measures when accounting for only one input.

We conclude the section with table 1, which summarizes the key concepts of the proposed zonotope method.

Table 1. Key concepts of the proposed method

Key concept	Details
Individual production activity $\mathbf{a}_i$	It is a vector defined in $\mathbb{R}_+^{l+1}$. Intuitively, it represents the multiple-input, one-output production activity of one firm whose first lth elements are inputs while the $(l+1)$th is an output
Zonotope $\mathbf{Y}$	It provides a description of the whole industry obtained by combining all firms. Formally, it is the Minkowski sum of line segments generated by firm production activities $\{\mathbf{a}_i\}_{1 \le i \le N}$.
Gini $G(\mathbf{Y})$	It is the normalized volume of the zonotope, and it provides a measure for industry heterogeneity; the bigger the index, the larger the differences across firms.
Productivity $P = \text{tg}\{\Theta_{l+1}(\mathbf{d}_Y)\}$	In a two-inputs, one-output case, the "steeper" the vector-firm (or vector-industry) the more productive the firm (or industry). More formally, it is the tangent value of the angle formed by the vector and the input-hyperplane.

3   The zonotope command

In this section, we introduce the command `zonotope`.[8]

3.1 Syntax

The syntax of the command to compute the zonotope is

`zonotope` *varlist* $\big[$ *if* $\big]$ $\big[$ *in* $\big]$ $\big[$, <u>verbose</u> $\big]$

where the option `verbose` requests that the program print on screen all the computed quantities (see below). Instructions for installing the package under different operating systems are provided in section 8.

---

8. To ease future developments, the software is also available at https://github.com/zonotopes.

3.2 Description

The `zonotope` command requires a list of vector variables, where the last variable is the output and the other variables are the inputs of the relation we want to analyze. All the vector variables must have the same length. All the variables can be seen as columns of a single matrix, `gen` (that is, the matrix of generators); therefore its ith row coincides with the ith generator $\mathbf{a}_i$, as defined in (1). The number of columns in this matrix represents the dimension where the zonotope lies, $(l+1)$. In the `zonotope` command, the generators are assumed to be nonnegative; thus, all entries of this matrix should be nonnegative.

3.2.1 Output and return value

The `zonotope` command returns two vectors, diagonal and tangents, and one matrix, `gen` (it contains the generators actually used, which is important when the `if` or `in` qualifier is used). In addition, it returns several scalar values, such as

- `nrow`: the number of generators actually used (that is, N)

- `ncol`: the number of variables (it coincides with $l+1$ if the program has been called with l input variables and one output variable)

- `etMIN`: the elapsed time (expressed in minutes)

- `S1`, `S2`, ..., `S8`: statistics that are detailed below

When the `zonotope` command is called with the `verbose` option, it shows all the returned variables (both vectors and scalars) on the screen.

All the scalars returned by `zonotope` can be accessed using the `r()` command. For example, the volume of the zonotope can be displayed with the command

```
. display r(S1)
```

and the elapsed time can be displayed with

```
. display r(etMIN)
```

3.2.2 Output vector: diagonal

The output vector `diagonal` contains $\mathbf{d}_Y$, the geometric diagonal of the zonotope, which, according to (2), is defined as the sum of generator $\mathbf{a}_i$ for $i = 1, \ldots, N$. Specifically,

$$\mathbf{d}_Y = (\beta_1, \ldots, \beta_l, \beta_{l+1}) = \left(\sum_{i=1}^{N} \alpha_{i_1}, \ldots, \sum_{i=1}^{N} \alpha_{i_l}, \sum_{i=1}^{N} \alpha_{i_{l+1}} \right)$$

Clearly, it is an $(l+1)$-dimensional (row) vector. It can be easily displayed on screen (or reused) using `matrix list diagonal`.

3.2.3   Output vector: `tangents`

The output vector `r(tangents)` contains the tangent of the angle formed by each generator and the input space. Thus, it is an N-dimensional (column) vector. It can be easily displayed on screen (or reused) using `matrix list tangents`.

3.2.4   Output matrix: `gen`

The output matrix `gen` contains the generators actually used, as filtered by `if` and `in` qualifiers. It can be easily displayed on screen (or reused) using `matrix list gen`.

We report below a brief description of the eight statistics mentioned above. Some of them have a clear economic interpretation because they directly correspond to the key concepts of the proposed method, while others are necessary intermediate steps toward the measure of interest.

- **Statistic S1: Volume**

    Given N generators $\mathbf{a}_i \in \mathbb{R}_+^{l+1}$, we can generate one zonotope, again denoted as $\mathbf{Y}$. Thus, we can compute its volume as

    $$\texttt{S1} \equiv \mathrm{Vol}(\mathbf{Y})$$

    where $\mathrm{Vol}(\cdot)$ follows (3). We report this volume as `S1`, which is printed on screen and also returned as an output scalar. `S1` (Volume) provides only a preliminary idea of industry heterogeneity, but it represents an intermediate step to compute the more refined measure of the Gini index.

- **Statistic S2: Diagonal's norm**

    The norm of the diagonal $||\mathbf{d}_Y||$ is computed as the square root of the sum of the squares of all the components of the diagonal, that is,

    $$\texttt{S2} \equiv ||\mathbf{d}_Y|| = \sqrt{\sum_{i=1}^{l+1} \beta_i^2}$$

    The length of the diagonal of the zonotope represents the "size" of the industry; the longer this vector, the bigger the industry is.

- **Statistic S3: Sum of squared norms of all the generators**

    As indicated by the name, we first compute for each generator the sum of the square of each of its components, and then we sum over all generators.

    $$\texttt{S3} = \sum_{i=1}^{N} \left(\sum_{j=1}^{l+1} \alpha_{i_j}^2 \right)$$

- **Statistic S4: Gini index**

  According to (4), this Gini index is computed as the ratio between the volume of the zonotope and the product of the components of the diagonal. We rewrite this industry heterogeneity measure as follows:

  $$\texttt{S4} \equiv G(\mathbf{Y}) = \frac{\text{Vol}(\mathbf{Y})}{\text{Vol}(\mathbf{P}_Y)} = \frac{\texttt{S1}}{\prod_{j=1}^{l+1} \beta_j}$$

  where $\mathbf{P}_Y$ denotes the parallelotope with diagonal $\mathbf{d}_Y$. The normalized volume of the zonotope, as already discussed, is a more refined measure of industry heterogeneity. The bigger this index, the more heterogeneous the industry.

- **Statistic S5: Tangent of angle formed by diagonal and input space**

  Given the diagonal vector reported as $\mathbf{d}_Y$, according to (5), we can further compute the tangent of the angle formed by the diagonal and its input space. We report this industry productivity measure as $\texttt{S5}$, as follows:

  $$\texttt{S5} \equiv \text{tg}\left\{\Theta_{l+1}(\mathbf{d}_Y)\right\} = \frac{\beta_{l+1}}{\sqrt{\sum_{j=1}^{l} \beta_j^2}}$$

  Intuitively, it provides our measure for industry productivity. In a two-inputs, one-output setting, the "steeper" the vector, the more productive is the industry.

- **Statistic S6: Cosine against output**

  $\texttt{S6}$ reports the cosine of the complementary angle of $\Theta_{l+1}(\mathbf{d}_Y)$ as

  $$\texttt{S6} = \frac{\beta_{l+1}}{\sqrt{\sum_{j=1}^{l+1} \beta_j^2}}$$

  Because of the complementary angle relationship, $\texttt{S5}$ and $\texttt{S6}$ are connected as follows:

  $$\texttt{S5} = \frac{\texttt{S6}}{\sqrt{1 - (\texttt{S6})^2}}$$

- **Statistic S7: Cosine of diagonal projected on input plane with x axis**

  The angle formed by the x axis and the projection of the diagonal in the input plane measures the relative intensity of the first input related to the other inputs. We report its cosine value as

  $$\texttt{S7} = \frac{\beta_1}{\sqrt{\sum_{j=1}^{l} \beta_j^2}}$$

- **Statistic S8: Volume against cube of the norm of the diagonal**

  The last statistic is the ratio between the volume of the zonotope and the cube of the norm of the diagonal.

  $$\texttt{S8} = \frac{\mathrm{Vol}(\mathbf{Y})}{||\mathbf{d}_Y||^3} = \frac{\texttt{S1}}{(\texttt{S2})^3}$$

  This definition of volume is restricted to the boundary edges of the cone of all possible vectors, and as such, it considers only the most diverse vectors. It attempts to measure the maximal diversity of the field (hence, it takes into account only the boundary vertices) by measuring how "wide" the cone is.

Other than these statistics, the command also provides the elapsed time, expressed in minutes.

3.3   A working example of the zonotope command

In this section, we provide a step-by-step working example that shows the use of the `zonotope` command, with comments on the related output. The general framework, as before, is production analysis.

We start by loading the dataset into Stata and listing the first five rows:

```
. use empirical_firm_data, clear
. list in 1/5
```

	Year	NACE	Number_s	FixedA_s	Materi_t	Turnover	ID
1.	2006	0001	19	2969.52	25513.89	29127.35	145725
2.	2006	0001	9	920.53	1622.95	2796.362	328840
3.	2006	0001	22	631.447	2721.53	4316.787	485189
4.	2006	0001	35	2620.777	11364.47	13958.58	227378
5.	2006	0001	11	982.525	677.13	1362.247	491850

The first column indicates the year, and the second reports the firm's sector of main activity. Columns 3 to 6 report, respectively, number of employees, fixed assets, material cost, and turnover. The last column prints a fake ID of the firm. In the interest of replicability, we are providing artificially generated data that display the same distributional properties as the "real" data used in other works, such as Dosi et al. (2016) and Dosi et al. (2021). Each row in the dataset represents a different firm in a given year and industry.

As usual in the literature, the number of employees and fixed assets are used as proxies for labor and capital inputs, respectively, and turnover is used for output. In this two-inputs, one-output setting, the `zonotope` command allows analysis of the industry in $\mathbb{R}^3$. For example, focusing on a specific year and sector, one can type

```
. zonotope NumberOfEmployees FixedAssets Turnover if NACE == "0001"
> & Year == 2006, verbose
```

We will go through the results step by step to provide specific comments about the results relevant to the investigation.

```
----------------------------------------------------------------------
ZONOTOPE LIBRARY VER 1.3
----------------------------------------------------------------------
INPUT: SET OF GENERATORS
N. of dimensions (including the output): 3
N. of generators: 160
----------------------------------------------------------------------
```

The above, the first part of the output, after informing us about the version of the package, displays the number of dimensions of the current analysis (3 in the example) and the number of generators that in this case are firms (160).

```
Computation started (it can take a while) ...
----------------------------------------------------------------------
OUTPUT VECTORS
Diagonal (as a row vector):
>
5531 459015 2.84636e+006
```

The second part of the output, shown above, reports the production activity of the whole industry, the diagonal of the zonotope, as defined by (2), which is the result of the aggregation of the 160 generator-firms. The total number of employees in this industry is 5,531, while the total fixed assets and turnover of this industry are equal to 459,015 and 2,846,360 thousand Euros, respectively.

```
Tangent between each input generator and the input space:
9.80857
3.03763
6.8322
5.32565
1.38639
   (output omitted)
```

For the sake of exposition, we reported only the tangent for the first five firms of this example. As will be discussed in section 4.2, we extend the definition of industry productivity, (5), to individual firm productivity, as proxied by the angle that the vector-firm forms with the input plane. For example, firm 145725 produces 29,127.35 thousand Euros of output from 19 employees and 2,969.52 thousand Euros of fixed assets. According to (6), productivity of firm 145725, as proxied by the tangent of the above-mentioned angle, is

$$\frac{29127.35}{\sqrt{19^2 + 2969.52^2}} = 9.80857$$

Recall that a higher value of the tangent of the angle suggests a higher productivity because the vector-firm is steeper; that is, the angle that the vector forms with the input plane is larger. Figure 3 provides a simple illustration of this, in a one-input, one-output setting.

As for many other indicators of efficiency that use multiple inputs, the value of the measure alone is not of great use. It is much more relevant instead to perform a comparison across firms within the same industry or an intertemporal comparison of the same unit over time. This is the sort of comparison that is proposed in table 2 in section 4.1.

```
------------------------------------------------------------------------
OUTPUT STATISTICS
S1: Total volume:  1.46438e+015
S2: Diagonal norm: 2.88313e+006
S3: Sum of squared norms: 2.9266e+011
S4: Gini index: 0.202645
S5: Tangent of angle btw. diagonal and the input plane: 6.20056
S6: Cosine against output: 0.987243
S7: Cosine of proj. of diagonal on input plane with x axis: 0.0120488
S8: Volume against the cube of the norm of the diagonal: 6.11027e-005
------------------------------------------------------------------------
```

As shown above, the `zonotope` command then produces the eight statistics described in section 3.2. For example, `S2` reports the length of the diagonal, which measures the size of the industry, accounting for two inputs and one output.

`S4` reports the Gini index, which measures the heterogeneity of the sector, while `S5` reports industry productivity as defined in (5). We strongly suggest using the measure of heterogeneity of a sector (or, in a different context, of the relevant aggregate) to make cross-sectional or intertemporal comparison. This is how the measures of heterogeneity are used in Dosi et al. (2016) and Dosi et al. (2021). The same argument applies to `S5` and tangents, that is, the tangent value of the angle between the diagonal and the input plane. For example, it is difficult to conclude whether one industry in a specific country is highly productive based only on the magnitude of `S5`. But if we have the `S5` productivity levels for the same industry in two different countries, then it is possible to determine which country is more productive in a given industry. Or if we have the `S5` productivity level for one industry in one specific country over time, it is then possible to determine whether that country has increased productivity in that industry. In section 4.1, we provide a more detailed example regarding the dynamics of heterogeneity and productivity for a few selected industries over time.

The `S7` statistic reports the cosine value of the angle[9] formed by the projection of the diagonal on the input plane, that is, the angle formed by the industry input vector $(5531, 459015)$ with the x axis (which corresponds to the number of employees). The value 0.0120488 indicates that this angle is almost $\pi/2$, telling us that the industry adopts many more fixed assets compared with employees.

The last bit of output is displayed as follows:

```
Elapsed time (MIN):
0.001383
```

---

9. The value of this angle is highly influenced by the choice of the unit of measure. For a discussion of this issue, see Dosi et al. (2016) and Dosi et al. (2021).

This is the computation time. As we will explain more in section 5, building the zonotope of a set of generators has an exponential complexity (that is, it is very time consuming, especially in a high dimension and when the number of generators is high).

It is possible to display (or reuse) the computed results as follows:

```
----------------------------------------------------------------------
 Now you can display (or reuse) the computed results in this way:

   display r(S1)  (for the volume)
   display r(S2)  (for the norm of the diagonal)
   ...
   display r(S8)  (volume against the cube of the norm of the diagonal)

   matrix list diagonal  (for the vector containing the diagonal)
   matrix list tangents  (for the vector with the tangents)
   matrix list gen       (for the actually used generators, after if/in)
----------------------------------------------------------------------
```

4 Empirical analysis

In this section, we provide three additional, more-advanced examples of using `zonotope` to perform economic investigations. In section 4.1, based on firm-level data, we compute the heterogeneity and productivity levels of industries as suggested by Dosi et al. (2016). In section 4.2, we compute firm-level measures of productivity. And in section 4.3, based on the income and expenditure data of Namibia, we compute the multivariate Gini coefficient as a measure for inequality as suggested by Mosler (1994) and Koshevoy and Mosler (1997).

4.1 Assessing industry heterogeneity

In this section, we conduct some empirical investigations on industry heterogeneity and productivity. We do this by using larger sets of data to replicate the standard situation faced during empirical analysis on firm-level data, where firms are grouped according to their sector of main activity, at different levels of aggregation.

For replication of all the analyses, we use the same data introduced in the previous section. For each firm in the selected industries, number of employees and fixed assets are chosen to proxy inputs and turnover is chosen for output. All values except for the number of employees are assumed to be in thousands of Euros and deflated at the 4-digit `NACE` level with the year 2010 as benchmark to perform intertemporal comparison. We refer to the 4-digit `NACE` sector simply as 0001, 0002, etc. When it is necessary to aggregate the 4-digit sectors into 3-digit sectors, we use 0091, 0092, and 009, as we did in figures 1 and 2.

4.1.1 The two-inputs, one-output case

Traditionally, the most standard setting in which to investigate production is that which assumes production activity with two inputs and one output; number of employees and fixed assets are chosen to be the proxies for inputs and turnover value is chosen for the output. For each industry in one specific year, the normalized zonotope volume and the tangent value of the angle formed by the industry production vector and its input plane are easily computed as S4 and S5, which can be used to measure the industry heterogeneity and productivity level for a specific industry at a given point in time. Columns (2) and (3) of table 2 report these two results for selected industries in 2006, while column (1) reports the number of firms within that industry that year. To get these three variables for a specific industry in a specific year—for example, sector 0001 in 2006—we run the following commands:

```
. use empirical_firm_data, clear
. zonotope NumberOfEmployees FixedAssets Turnover if Year == 2006
> & NACE == "0001", verbose
```

From the available results, we select

```
--------------------------------------------------------------------
ZONOTOPE LIBRARY VER 1.3
--------------------------------------------------------------------
INPUT: SET OF GENERATORS
N. of dimensions (including the output): 3
N. of generators: 160
--------------------------------------------------------------------
   (output omitted)
S4: Gini index: 0.202645
S5: Tangent of angle btw. diagonal and the input plane: 6.20056
   (output omitted)
```

Table 2. Computation results for Gini coefficient and productivity among selected sectors and years—$\mathbb{R}^3$ case

Sector	Year 2006			Year 2009			Year 2012		
	Obs (1)	Gini (2)	tg (3)	Obs (4)	Gini (5)	tg (6)	Obs (7)	Gini (8)	tg (9)
0001	160	0.203	6.201	222	0.233	4.239	346	0.264	4.483
0002	115	0.126	6.617	134	0.119	6.780	162	0.138	9.292
0003	14	0.016	1.561	31	0.020	1.803	70	0.062	1.385
0004	386	0.153	3.014	431	0.190	2.077	603	0.200	2.937
0005	57	0.090	5.744	69	0.087	3.116	142	0.114	4.590
0006	142	0.155	4.078	168	0.170	2.660	212	0.254	5.859
0007	43	0.098	0.864	52	0.112	0.892	68	0.123	1.265
0008	47	0.091	3.402	53	0.101	2.292	69	0.122	3.107
0009	38	0.073	1.495	41	0.088	1.060	47	0.080	2.204
0010	18	0.084	1.276	26	0.092	1.048	40	0.201	1.995
0011	65	0.103	3.834	85	0.172	3.249	126	0.168	3.487
0012	177	0.204	3.102	189	0.155	2.153	259	0.164	3.153
0013	54	0.042	2.372	58	0.063	1.539	82	0.076	1.796
0014	225	0.137	3.376	245	0.125	2.218	368	0.126	2.755
0015	28	0.001	0.899	33	0.001	0.725	45	0.001	0.751
0016	45	0.026	1.288	44	0.022	0.912	62	0.042	0.875
0017	86	0.072	3.713	89	0.089	1.486	121	0.094	2.292

The chosen normalization strategy for the volume of the zonotope seems to be effective, because there is no apparent relation between the number of generator-firms and the Gini coefficients. For example, in 2006, there are 386 firms in sector 0004 and 160 firms in sector 0001. However, given the larger number of firms in sector 0004, we do not necessarily expect its Gini coefficient (the normalized volume of the zonotope) to be bigger than that of sector 0001. Indeed, based on our data sample, the Gini coefficient of sector 0001 is larger: 0.203 compared with 0.153 for sector 0004.

After taking into account the effect due to the number of firms, we notice that the heterogeneity levels are different among different industries. For example, in 2006, the Gini coefficients vary from 0.001 for sector 0015 to 0.204 for sector 0012. Similarly, as indicated in column (3), the productivity levels among different sectors are different.

From columns (4) to (6) and from columns (7) to (9), we report similar results in the years 2009 and 2012, respectively. This allows us to explore the dynamic of industry heterogeneity and productivity over time. Most of the selected sectors share an upward trend in their heterogeneity levels. For example, the heterogeneity level of sector 0001 increases from 0.203 in 2006 to 0.233 in 2009, and again to 0.264 in 2012. As for the productivity, for most of the industries, we report a decrease from 2006 to 2009 and an increase from 2009 to 2012.

The same analysis can be performed in the case of four dimensions, for instance in the three-inputs, one-output case.[10] In the interest of space, results are not reported here, but they can be obtained with the following commands:

```
. use empirical_firm_data, clear
. zonotope NumberOfEmployees FixedAssets MaterialCost Turnover if
> Year == 2006 & NACE ==  "0001", verbose
```

4.2   A geometric approach to firm-level productivity

As indicated in (5), Dosi et al. (2016) propose as a measure for industry productivity the tangent of the angle formed by the industry production activity vector $\mathbf{d}_Y$ with the space generated by all inputs. The approach can also be easily applied to individual firms. Assume firm i's production activity follows (1). Then its productivity level can be measured by the tangent of $\Theta_{l+1}(\mathbf{a}_i)$, that is, the angle formed by the firm production activity vector $\mathbf{a}_i$ with the space generated by all inputs. One possible measure for the productivity of firm i, denoted by p_i, follows:

$$\mathrm{p}_i = \mathrm{tg}\left\{\Theta_{l+1}(\mathbf{a}_i)\right\} = \frac{\alpha_{i_{l+1}}}{||\mathrm{pr}_{-(l+1)}(\mathbf{a}_i)||}, \quad i = 1, \ldots, N \tag{6}$$

The `zonotope` command computes this productivity for each firm and returns all of them as the vector `tangents`. Based on the same firm-level data used above, we compute the productivities of 12 selected firms (which all report data in all years) from industry sector 0010 in 2006 and report them in the column "Year 2006" of table 3. This newer measure is also consistent with the previous literature in that we still observe relevant heterogeneous productivity at the firm level. To compute the productivity (6) for the 12 firms,[11] we run the following commands:

```
. use empirical_firm_data, clear
. destring ID, replace
ID: all characters numeric; replaced as long
. egen selected = anymatch(ID), values(40029 84229 92290 96040 117252 140293 162
> 695 214062 214655 227040 233379 257599)
. keep if selected
(19,263 observations deleted)
. sort ID NACE Year
. zonotope NumberOfEmployees FixedAssets Turnover if
> Year == 2006 & NACE == "0010", verbose
  (output omitted )
```

---

10. In this case, material cost is chosen to be the proxy for the third input.
11. For illustrative purposes, the code only displays results for the year 2006. To generate columns for the other years, simply change the `if` qualifier. The do-file provided with the package generates the complete table 3 as reported below.

Among the available results, we focus on that referring to firm productivity:

```
Tangent between each input generator and the input space:
1.34676
6.91531
0.399336
5.51919
0.851644
5.66018
11.7067
23.4575
13.3015
4.76769
0.46963
0.962496
```

We then compute productivity of these firms in 2009 and 2012 and report the results in the respective columns of table 3. Note that, in accordance with previous findings, firm-level productivity displays persistence, and hence the large within-industry dispersion does not vanish over time.

Table 3. Productivity of selected firms from sector 0010—$\mathbb{R}^3$ case

Firm ID	Year 2006	Year 2009	Year 2012
1	1.347	2.470	4.353
2	6.915	6.704	13.996
3	0.399	0.178	0.292
4	5.519	4.470	8.272
5	0.852	0.603	0.997
6	5.660	3.336	5.394
7	11.707	14.092	14.949
8	23.457	10.902	4.380
9	13.302	19.282	3.519
10	4.768	5.417	11.143
11	0.470	0.622	0.762
12	0.962	1.090	1.393

4.3 An application to inequality

The `zonotope` command is not limited to firm-level data to compute industry heterogeneity and productivity at the industry and firm levels. In this section, we show that the command is also helpful for computing the multivariate Gini coefficient, which measures inequality among a group of N units with $l > 1$ attributes.

4.3.1 Introduction to multivariate Gini coefficient

Following Mosler (1994) and Koshevoy and Mosler (1996), we focus on two multivariate Gini coefficients. To do this, we denote by ω_i^j the quantity of attribute j owned by unit i for $i = 1, \ldots, N$ and $j = 1, \ldots, l$. Thus, we have vectors

$$\mathbf{w}_i = (\omega_i^1, \ldots, \omega_i^j, \ldots, \omega_i^l) \in \mathbb{R}_+^l$$

for $i = 1, \ldots, N$. Correspondingly,

$$\widetilde{\omega}_i^j = \frac{\omega_i^j}{\sum_{i=1}^N \omega_i^j}$$

indicates the share of attribute j out of the total amount owned by unit i. Vectors $\widetilde{\mathbf{w}}_i$ follow as

$$\widetilde{\mathbf{w}}_i = \left(\widetilde{\omega}_i^1, \ldots, \widetilde{\omega}_i^j, \ldots, \widetilde{\omega}_i^l\right) = \left(\frac{\omega_i^1}{\sum_{i=1}^N \omega_i^1}, \ldots, \frac{\omega_i^j}{\sum_{i=1}^N \omega_i^j}, \ldots, \frac{\omega_i^l}{\sum_{i=1}^N \omega_i^l}\right) \in \mathbb{R}_+^l$$

for $i = 1, \ldots, N$.

Mosler (1994) introduces a multivariate Gini index by using the volume of the Lorenz zonotope (LZ) as the Minkowski sum

$$\mathbf{LZ} = \sum_{i=1}^N \left\{ 0, \left(\frac{1}{N}, \widetilde{\mathbf{w}}_i\right) \right\} \tag{7}$$

of line segments generated by $\{(1/N, \widetilde{\mathbf{w}}_i)\}_{1 \leq i \leq N}$. The `zonotope` command provides this volume, that is, Vol(**LZ**). Note, however, that Koshevoy and Mosler (1996) point out that Vol(**LZ**) would be 0 when two attributes are similarly distributed among N units or one attribute is equally distributed among all units.[12] Hence, Koshevoy and Mosler (1997) suggest that instead of using the volume of the LZ, one should use the volume of the lift zonoid "expanded" by an l-dimensional cube in $\mathbb{R}^{l+1}$. They propose the multivariate volume–Gini index as

$$R_v = \frac{1}{2^l - 1} \sum_{s=1}^l \left\{ \sum_{1 \leq j_1 \leq \cdots \leq j_s \leq l} \mathrm{Vol}(\mathbf{Z}^{j_1, \ldots, j_s}) \right\} \tag{8}$$

where zonotopes are defined as the Minkowski sum

$$\mathbf{Z}^{j_1, \ldots, j_s} = \sum_{i=1}^N \left\{ 0, \left(\frac{1}{N}, \widetilde{\mathbf{w}}_i^{j_1, \ldots, j_s}\right) \right\}$$

---

12. Let's construct one $N \times (l+1)$ matrix where row i is $(1/N, \widetilde{\mathbf{w}}_i)$. Indeed, if the rank of this matrix is smaller than $(l+1)$, we would have the volume of the zonotope equal to 0. Aside from the two possibilities we highlight in the text, note that if we have $N < l + 1$, we also have a rank smaller than $(l+1)$. However, empirically, we generally have N big enough to allow the row rank to be bigger than $(l+1)$.

of line segments generated by vectors $\{(1/N, \widetilde{\mathbf{w}}_i^{j_1,\ldots,j_s})\}_{1 \leq i \leq N}$ and $\widetilde{\mathbf{w}}_i^{j_1,\ldots,j_s} \in \mathbb{R}_+^s$ is obtained from the $j_1,\ldots,j_s$ components of vector $\widetilde{\mathbf{w}}_i$.

To show the necessity of resorting to this further measure, we use a toy example detailed in appendix A. We focus on x and z, representing two attributes among the 10 units. Using the `zonotope` command, we can easily compute the volume of the corresponding **LZ** defined in (7) as 0.053. Now assume we have a third attribute that still coincides with z. We would then have a volume of **LZ** equal to 0. On the other hand, the multivariate volume–Gini index defined in (8) is able to provide a result of 0.144, even when there are identically distributed attributes.

To get this volume-Gini index, we need to use the `zonotope` command to compute volumes of different $(2^l - 1)$ zonotopes. If we construct a sample of normalized data—that is, an $N \times (l+1)$ matrix with each row as $(1/N, \widetilde{\mathbf{w}}_i)$, where every column sums up to 1—to get each $\mathrm{Vol}(\mathbf{Z}^{j_1,\ldots,j_s})$, the `zonotope` command computes corresponding columns of data samples and reports the results in `S1`. However, recall that the `zonotope` command also reports the normalized volume of the zonotope in `S4`, which facilitates the process.

4.3.2 Empirical application of multivariate Gini coefficient

We present an empirical exercise based on actual data from the Namibia Household Income and Expenditure Survey 2009/2010 (The Namibia Statistics Agency 2013).[13] This dataset includes detailed information about income and expenditures of 9,643 households in Namibia from 2009 to 2010. For each household, we compute the income per capita (IPC) and expenditure per capita (EPC) based on the income, expenditure, and number of people in the household.[14] Notice, in particular, that the survey reports only the main source of income for each household. As a result, we have two attributes of each household, IPC and EPC, either of which can be used for computing the traditional univariate Gini coefficient.

In many studies, the univariate Gini coefficient based on IPC is adopted as the measure of inequality. We report it at the bottom of column (1) of table 4. Correspondingly, the univariate Gini coefficient based on EPC is reported at the bottom of column (2). The univariate Gini coefficient indicates a higher inequality level based on expenditure (0.780) than income (0.611). As we discuss above, these two attributes can be taken into account simultaneously by computing the volume of the corresponding LZ and multivariate volume–Gini index. We report them at the bottom of columns (3) and (4), respectively. It is not possible to compare the values of the univariate and multivariate

---

13. We downloaded the data file `NAM_2009_HIES_v01_M_v01_A_PUF_STATA8.zip` on June 20, 2019, with MD5 value 90461521c1059fa1df3bf97e6b2dc797 from https://nsa.org.na/microdata1/index.php/catalog/6/study-description.

14. According to the suggestion from the original dataset, a person 16 years old and beyond contributes 1 to the number, a person between the ages of 6 and 15 contributes 0.75, and a person 5 years old or younger contributes 0.5.

Gini indices, because they are constructed in different ways.[15] However, it is meaningful to compare one of the indexes cross-section or over time.

Table 4. Various Gini coefficients among households with different income sources in Namibia from 2009 to 2010

| Income source | Univariate Gini | | Multivariate Gini | | |
| | IPC | EPC | Vol($\mathbf{LZ}$) | R_v | Obs |
	(1)	(2)	(3)	(4)	(5)
Salaries / wages	0.558	0.700	0.155	0.471	4,953
Farming	0.656	0.866	0.206	0.576	2,014
Business activities	0.651	0.799	0.213	0.554	785
Employment pension	0.581	0.777	0.253	0.537	121
Cash remittances	0.465	0.645	0.127	0.413	282
State old pension	0.422	0.619	0.135	0.392	986
Other, specify	0.666	0.864	0.288	0.606	502
Total	0.611	0.780	0.190	0.527	9,643

Let us conduct a cross-section analysis now. Households are classified into seven groups according to their main source of income. We report the univariate and multivariate Gini coefficients between columns (1)–(4) in table 4 and the number of households per group in column (5). Notice that groups with different income sources display different inequality levels. According to the univariate Gini coefficient based on IPC, the least unequal group is the one relying on state old pension. This seems reasonable because state old pensions, compared with other sources, are expected to be more equally distributed among receivers.

The next thing to notice is related to the potential bias associated with Vol($\mathbf{LZ}$) as a measure of inequality, because it is sometimes not consistent with the evidence stemming from the univariate Gini coefficients. For example, the group "Business activities" seems more unequal when compared with "Employment pension" according to the univariate Gini coefficient and volume-Gini index R_v, while Vol($\mathbf{LZ}$) indicates the opposite. This exemplifies one possible drawback of Vol($\mathbf{LZ}$) as a measure of inequality. When IPC and EPC are more correlated, the value of Vol($\mathbf{LZ}$) is smaller.[16] Thus, it is not impossible that two highly correlated attributes, both pointing to high-level inequality according to the univariate Gini coefficient, would point to low inequality when based on Vol($\mathbf{LZ}$). Intuitively, Vol($\mathbf{LZ}$) becomes smaller because of the correlation between two such attributes. In this case, although the two univariate Gini coefficients of group "Business activities" are bigger than those of "Employment pension", because

15. Both volumes of the corresponding LZ and multivariate volume–Gini index degenerate to the univariate Gini coefficient when $l = 1$. However, when $l > 1$, they are different from the univariate Gini coefficient.
16. The extreme case would be when two attributes are identically distributed among all units, and Vol($\mathbf{LZ}$) goes to 0.

the correlation between IPC and EPC is bigger for the former (0.814 versus 0.652), the Vol(**LZ**) of "Business activities" becomes smaller than that of "Employment pension". The volume-Gini index is considered to be more robust than Vol(**LZ**) not only when there are identically distributed attributes but also when there are highly correlated attributes.

We provide the code to generate the fourth row of table 4. First, we generate the relevant variables, that is, IPC and EPC, from the raw data.

```
. use 7a0_nhies_2009_2010_household_dataset.dta, clear
. keep if cMainInc == 5
(9,535 observations deleted)
. generate incomepercapita = hIncComp/ConFac2
. generate exppercapita = hF1ExpSm/ConFac2
. egen total = total(incomepercapita), by(cMainInc)
. replace incomepercapita = incomepercapita/total
(121 real changes made)
. drop total
. egen total = total(exppercapita), by(cMainInc)
. replace exppercapita = exppercapita/total
(121 real changes made)
. drop total
. egen NumOfObs = count(cMainInc), by(cMainInc)
. generate Frequency = 1/NumOfObs
```

Then, we compute the univariate Gini coefficient based on IPC and EPC by using the community-contributed command `inequal7` (Kerm 2001).

```
. matrix Result = J(8,6,0)
. * compute univariate Gini based on income per capita
. inequal7 incomepercapita, returnscalars
  (output omitted)
. matrix Result[4, 2] = r(gini)
. * compute univariate Gini based on expenditure per capita
. inequal7 exppercapita, returnscalars
  (output omitted)
. matrix Result[4, 3] = r(gini)
```

Finally, we compute Vol(**LZ**) and volume-Gini index R_v based on IPC and EPC, using the package to compute the volume of different zonotopes. Additionally, we use the community-contributed commands `tuples` (Luchman, Klein, and Cox 2006), `matsum()` (Weesie 1997), and `frmttable` (Gallup 2012).

```
. tuples incomepercapita exppercapita, di
tuple1: exppercapita
tuple2: incomepercapita
tuple3: incomepercapita exppercapita
. matrix tem = J(`ntuples',1,0)
. forvalues i = 1/`ntuples' {
  2. di `i'
  3. zonotope Frequency `var'`tuple'i
  4. if `i' == `ntuples' {
  5. matrix Result[4, 4] = r(S1)
  6. matrix Result[4, 6] = r(nrow)
  7. }
  8. matrix tem[`i', 1] = r(S1)/(2^2 - 1)
  9. }
(output omitted)
. matsum(tem), col(ColsumTem)
. matrix Result[4, 5] = ColsumTem
. frmttable, statmat(Result) sdec(0,3,3,3,3,0) title("Gini Namibia Case")
(output omitted)
```

5   Analysis of zonotope computing time

Computing the volume of the zonotope given the list of its generators can potentially be time consuming, because its computational complexity is $\mathcal{O}(N^l)$, where N is the number of generators and $(l + 1)$ is the dimension of each generator (that is, its length). Thus, this algorithm falls within the exponential category; that is, its time does not scale in a polynomial way but in an exponential way. In particular, when N is kept constant, it is clearly exponential in the number of dimensions $(l + 1)$.

Table 5 provides the elapsed times for a varying number of variables (from 2 to 6) and a fixed number of generators ($N = 200$). On the contrary, table 6 provides the elapsed times for a varying number of generators (from 50 to 250) and a fixed number of variables $[(l + 1) = 6]$. We ran the experiments on an Intel CPU i7 4-cores on a Windows 7 operating system, with 32 GB of RAM using Stata/IC. The data from tables 5 and 6 are charted on figures 4 and 5, respectively. From these figures, we can clearly see how the time complexity of computing the volume of the zonotope is exponential. To empirically prove that the time complexity is N^l, we have plotted in figure 6 the fifth root of the elapsed times, when the number of variables $(l + 1)$ is equal to 6. The fact that the graph is a perfect straight line proves that the dependency is a power of 5, as expected.

Table 5. Elapsed times for varying numbers of variables and 200 generators

Number of variables	Number of generators	Elapsed time
2	200	0.00017 min.
3	200	0.00157 min.
4	200	0.11477 min.
5	200	6.75457 min.
6	200	4h and 40.612 min.

Table 6. Elapsed times for varying numbers of generators and 6 variables

Number of variables	Number of generators	Elapsed time
6	50	0.126 min.
6	100	5.648 min.
6	150	54.736 min.
6	200	4h and 40.612 min.
6	250	16h and 16.389 min.
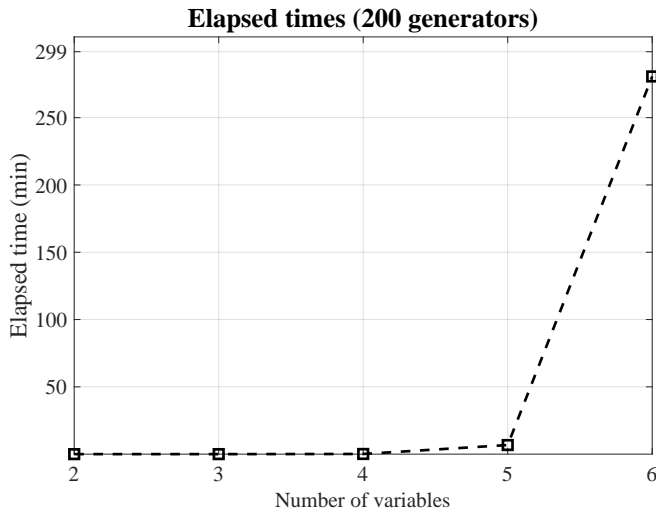
Figure 4. Elapsed time as a function of the number of variables, for 200 generators
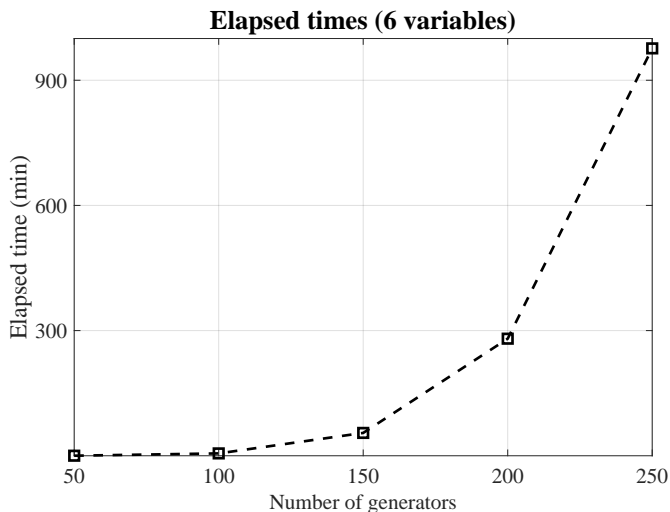
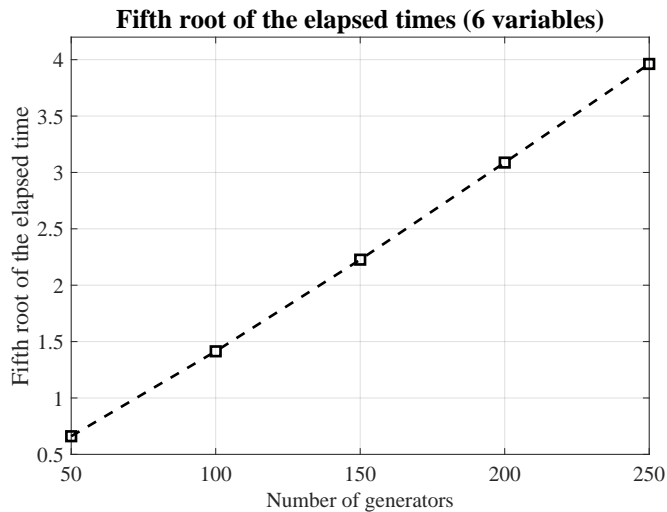Figure 5. Elapsed time as a function of the number of generators, for 6 variables

Figure 6. The fifth root of the elapsed time as a function of the number of variables, for 6 variables. The fact that the graph is a perfect straight line empirically proves that the computational complexity is N^l, because $(l+1) = 6$ in our case ($N = 200$, of course).

6 Conclusions and future work

The recent and increasing availability of disaggregated economic data challenges many of the standard theoretical assumptions, yet there is still urgent need for adequate tools to deal with these emerging rich and complex empirical observations.

The `zonotope` package provides a rigorous way to perform empirical analysis while taking advantage of some of these emerging properties. In this work, we proposed an application to production analysis in which, thanks to the proposed methodology, it is possible to relax most of the standard assumptions that do not find support in the data. Firms, and economic agents in general, are quite different from each other in many respects: size, productivity, propensity to innovate and export, etc. Such differences do not vanish over time; selection, if at work, takes longer to exert its effects. Even focusing on firms within the same narrowly defined industrial sector does not help in reducing heterogeneity. In this context, the `zonotope` package enables assessment of the level of intra-industry heterogeneity and measurement of the level of productivity, and its variation over time, without imposing strong assumptions on the actual observations.

The software package we introduced is ready for applications in other domains; the zonotope framework itself is already used in other fields, for example, still within economic analysis, to assess inequality.

7 Acknowledgments

8 Program and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 22-1
. net install st0662      (to install program files, if available)
. net get st0662          (to install ancillary files, if available)
```

To update the `zonotope` command to the latest version, type

```
. net cd http://www.iet.unipi.it/m.cococcioni/zono_stata_1p3
. net install zonotope, replace force
```

To run the demo, execute the following do-file:

```
. run "`c(sysdir_plus)'/z/zonotope_demo"
```

If the demo runs correctly, you are done.

However, it may fail, especially if you are using an operating system for which the plugin has not been generated yet.

In this case, the plugin must be generated from scratch, as explained next.

8.1   How to generate the zonotope plugin

You first need to download the C++ source code from the following git repository:

https://github.com/zonotopes/zono_cpp

It contains everything needed to generate the Stata plugin for most platforms (see below).

To generate the plugin, the following things are required:

- a C++ compiler

- the CMake utility

- (optional) the GIT utility (it is useful for downloading the tree of the whole project or specific subprojects)

The CMake utility is required to compile the plugin as a shared library. This shared library, named `zonotope2.plugin` or `zonotope3.plugin` (depending on the Stata version), is used by the ado-file associated with the `zonotope` command.

The ado-file is named `zonotope.ado`. Together with demo files, additional datasets and the help file can be downloaded from

https://github.com/zonotopes/zono_stata

The plugin generation has been tested on Windows 10 64-bit using Visual Studio 15, on Linux Ubuntu 18.04 64-bit using GCC, and on Mac OS 10.12.3 (16D32) 64-bit using the Mac OS Sierra operating system, with LLVM C++ compiler.

If you have troubles with the plugin, please open an issue on the associated GitHub repository. We will do our best to help you.

9   References

Andreoli, F., and C. Zoli. 2014. Measuring dissimilarity. Working Paper 23/2014, University of Verona, Department of Economics.

Aruka, Y. 2017. Some new perspectives on the inter-country analysis of the world production system. *Evolutionary and Institutional Economics Review* 14: 467–498. https://doi.org/10.1007/s40844-017-0085-2.

Badunenko, O., and P. Mozharovskyi. 2016. Nonparametric frontier analysis using Stata. *Stata Journal* 16: 550–589. https://doi.org/10.1177/1536867X1601600302.

Baily, M. N., C. Hulten, and D. Campbell. 1992. Productivity dynamics in manufacturing establishments. *Brookings Papers on Economic Activity: Microeconomics* 1992: 187–267. https://doi.org/10.2307/2534764.

Baldwin, J. R., and M. Rafiquzzaman. 1995. Selection versus evolutionary adaptation: Learning and post-entry performance. *International Journal of Industrial Organization* 13: 501–522. https://doi.org/10.1016/0167-7187(95)00502-1.

Bartelsman, E. J., and M. Doms. 2000. Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature* 38: 569–594. https://doi.org/10.1257/jel.38.3.569.

Belotti, F., S. Daidone, G. Ilardi, and V. Atella. 2013. Stochastic frontier analysis using Stata. *Stata Journal* 13: 719–758. https://doi.org/10.1177/1536867X1301300404.

Cazals, C., J.-P. Florens, and L. Simar. 2002. Nonparametric frontier estimation: A robust approach. *Journal of Econometrics* 106: 1–25. https://doi.org/10.1016/S0304-4076(01)00080-X.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429–444. https://doi.org/10.1016/0377-2217(78)90138-8.

Daraio, C., and L. Simar. 2005. Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis* 24: 93–121. https://doi.org/10.1007/s11123-005-3042-8.

———. 2007. Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of Productivity Analysis* 28: 13–32. https://doi.org/10.1007/s11123-007-0049-3.

Disney, R., J. Haskel, and Y. Heden. 2003. Entry, exit and establishment survival in UK manufacturing. *Journal of Industrial Economics* 51: 91–112. https://doi.org/10.1111/1467-6451.00193.

Dosi, G. 2007. Statistical regularities in the evolution of industries: A guide through some evidence and challenges for the theory. In *Perspectives on Innovation*, ed. F. Malerba and S. Brusoni, 153–186. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511618390.009.

Dosi, G., and M. Grazzi. 2006. Technologies as problem-solving procedures and technologies as input–output relations: Some perspectives on the theory of production. *Industrial and Corporate Change* 15: 173–202. https://doi.org/10.1093/icc/dtj010.

Dosi, G., M. Grazzi, L. Li, L. Marengo, and S. Settepanella. 2021. Productivity decomposition in heterogeneous industries. *Journal of Industrial Economics* 69: 615–652. https://doi.org/10.1111/joie.12252.

Dosi, G., M. Grazzi, L. Marengo, and S. Settepanella. 2016. Production theory: Accounting for firm heterogeneity and technical change. *Journal of Industrial Economics* 64: 875–907. https://doi.org/10.1111/joie.12128.

Eurostat. 2008. NACE Rev. 2 Statistical classification of economic activities in the European Community. Methodologies and working papers, Eurostat.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A* 120: 253–290. https://doi.org/10.2307/2343100.

Franciosi, M., S. Settepanella, and A. Terni. 2020. The robustness of the generalized Gini index. ArXiv Working Paper No. arXiv:2007.12924. https://arxiv.org/abs/2007.12924.

Gallup, J. L. 2012. A programmer's command to build formatted statistical tables. *Stata Journal* 12: 655–673. https://doi.org/10.1177/1536867X1201200406.

Griliches, Z., and J. Mairesse. 1999. Production functions: The search for identification. In *Econometrics and Economic Theory in the Twentieth Century: The Ragner Frisch Centennial Symposium*, ed. S. Steiner, 169–203. Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL521633230.006.

Hildenbrand, W. 1981. Short-run production functions based on microdata. *Econometrica* 49: 1095–1125. https://doi.org/10.2307/1912746.

Ji, Y.-B., and C. Lee. 2010. Data envelopment analysis. *Stata Journal* 10: 267–280. https://doi.org/10.1177/1536867X1001000207.

Kerm, P. V. 2001. inequal7: Stata module to compute measures of inequality. Statistical Software Components S416401, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s416401.html.

Koopmans, T. C. 1977. Examples of production relations based on microdata. In *The Microeconomic Foundations of Macroeconomics*, ed. H. G. C., 144–178. London: Palgrave Macmillan. https://doi.org/10.1007/978-1-349-03236-5_6.

Koshevoy, G., and K. Mosler. 1996. The Lorenz zonoid of a multivariate distribution. *Journal of the American Statistical Association* 91: 873–882. https://doi.org/10.1080/01621459.1996.10476955.

Koshevoy, G. A., and K. Mosler. 1997. Multivariate Gini indices. *Journal of Multivariate Analysis* 60: 252–276. https://doi.org/10.1006/jmva.1996.1655.

Luchman, J. N., D. Klein, and N. J. Cox. 2006. tuples: Stata module for selecting all possible tuples from a list. Statistical Software Components S456797, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s456797.html.

Mosler, K. 1994. Majorization in economic disparity measures. *Linear Algebra and Its Applications* 199: 91–114. https://doi.org/10.1016/0024-3795(94)90343-3.

Simar, L., and V. Zelenyuk. 2011. Stochastic FDH/DEA estimators for frontier analysis. *Journal of Productivity Analysis* 36: 1–20. https://doi.org/10.1007/s11123-010-0170-6.

Syverson, C. 2011. What determines productivity? *Journal of Economic Literature* 49: 326–365. https://doi.org/10.1257/jel.49.2.326.

The Namibia Statistics Agency. 2013. Namibia Household Income and Expenditure Survey 2009/2010 (NAM_2009_HIES_v01_M_v01_A_PUF). https://nsa.org.na/microdata1/index.php/catalog/6/study-description.

United Nations. 2008. International Standard Industrial Classification of All Economic Activities Revision 4. Statistical papers series m no. 4/rev.4, UN—Department of Economic and Social Affairs.

Weesie, J. 1997. dm49: Some new matrix commands. *Stata Technical Bulletin* 39: 17–20. Reprinted in *Stata Technical Bulletin Reprints*. Vol. 7, pp. 43–48. College Station, TX: Stata Press.

About the authors

Marco Cococcioni is an associate professor at the Department of Information Engineering at the University of Pisa in Italy. He obtained his PhD from the same university in 2004.

Marco Grazzi is a professor at the Department of Economic Policy at the Università Cattolica del Sacro Cuore in Milano, Italy. He earned his PhD at Scuola Superiore Sant'Anna in 2006.

Le Li is an associate professor at the School of Economics at Guangzhou College of Commerce in China. He got his PhD in economics from Scuola Superiore Sant'Anna in 2016.

Federico Ponchio is a researcher at ISTI within the Italian CNR, where he is affiliated with the Visual Computing Laboratory. After graduating in math at the Scuola Normale Superiore in Pisa, he got his PhD at Clausthal TU.

A   A simple toy example

As a simple example, we enter the following commands:

```
. webuse test, clear
. list
```

We now have the following dataset loaded into Stata memory:

	x	y	z
1.	1	10	2
2.	2	9	4
3.	3	8	3
4.	4	7	5
5.	5	6	7
6.	6	5	6
7.	7	4	8
8.	8	3	10
9.	9	2	1
10.	10	1	9

The dataset contains only 10 observations and 3 variables: x, y, and z.

We now enter the following command:

```
. zonotope x y z if x > 3, verbose
```

This command builds the zonotope on input variables x and y, using variable z as the output variable. With the if statement, only the observations satisfying the condition x > 3 (the first input variable must be greater than 3) will be considered.

The following result will be displayed:

```
------------------------------------------------------------------------
ZONOTOPE LIBRARY VER 1.3
------------------------------------------------------------------------
INPUT: SET OF GENERATORS
N. of dimensions (including the output): 3
N. of generators: 7
------------------------------------------------------------------------
Computation started (it can take a while) ...
------------------------------------------------------------------------
OUTPUT VECTORS
Diagonal (as a row vector):
>
49 28 46
Tangent between each input generator and the input space:
0.620174
0.896258
0.768221
0.992278
1.17041
0.108465
0.895533
------------------------------------------------------------------------
OUTPUT STATISTICS
S1: Total volume:  4818
S2: Diagonal norm: 72.808
S3: Sum of squared norms: 867
S4: Gini index: 0.0763405
S5: Tangent of angle btw. diagonal and the input plane: 0.815085
S6: Cosine against output: 0.631799
S7: Cosine of proj. of diagonal on input plane with x axis: 0.868243
S8: Volume against the cube of the norm of the diagonal: 0.0124833
------------------------------------------------------------------------
Elapsed time (MIN):
0.000000
------------------------------------------------------------------------
DONE!
------------------------------------------------------------------------
 Now you can display (or reuse) the computed results in this way:

   display r(S1)  (for the volume)
   display r(S2)  (for the norm of the diagonal)
   ...
   display r(S8)  (volume against the cube of the norm of the diagonal)

   matrix list diagonal  (for the vector containing the diagonal)
   matrix list tangents  (for the vector with the tangents)
   matrix list gen       (for the actually used generators, after if/in)
------------------------------------------------------------------------
```

As is explained in greater detail in section 5, building the zonotope of a set of generators has an exponential complexity (that is, it is very time consuming, especially in a high dimension and when the number of generators is high). Therefore, we decided to implement it in C++, to reduce the run time as much as possible. In section 8, we provide step-by-step instructions for compiling the C++ source code to create the plugin (a binary file with extension .plugin). More precisely, our new command zonotope loads and then calls the C++ Stata plugin.