

Enhancing non compensatory composite indicators: a directional proposal

Elisa Fusco¹

^aSOSE S.p.A.

Abstract

The construction of Composite Indicators is useful to synthesize complex social and economic phenomena, but some underlying assumptions in "classical methods", as in particular the compensability among indicators, are very strictly.

The aim of this paper is to propose an original approach that enhance non-compensatory issue by introducing "directional" penalties in a BoD model in order to consider the preference structure among simple indicators. Principal component analysis on simple indicators hyperplane allows to estimate both the direction and the intensity of the average rates of substitution.

Under an empirical point of view, our method has been applied on infrastructural endowment data in European regions.

Keywords: Composite indicators, Non-compensatory, Directional efficiency, Infrastructural endowment

JEL classification: C14, C43, H54.

1. Introduction and previous literature

Interest in Composite Indicators (CI) as a tool to support decision-makers in policy analysis context, is rapidly growing thanks to their capability to summarise multi-dimensional issues, to rank countries in benchmarking analysis and to their ease of interpretation.

On the other hand, the construction of a CI is a very complex process with multiple subsequent steps: the develop of a theoretical framework for the identification of relevant analysis dimensions, the standardisation of the simple indicators to allow comparisons, the imputation of missing data, the weighting of simple indicators and, finally, the succeeding sensitivity analysis on the robustness of the

Email address: fusco_elisa@libero.it (Elisa Fusco)

aggregation (for a complete explanation of each step, please see Freudenberg, 2003).

A critical step of the entire process, focus of the recently debate, is how to assign different importance weights in order to aggregate indicators.

In literature various statistical methods like Principal Component Analysis (PCA - Manly, 1994) and Factor Analysis (FA), Unobserved Components Models (UCM - Kaufmann et al., 1999, Kaufmann et al., 2003), Benefit of Doubt Approach (BoD - Melyn et al., 1991) or specific techniques based on expert subjective judgement like Budget Allocation Processes (BAP - Jesinghaus in Moldan et al., 1997), Analytic Hierarchy Processes (AHP - Forman, 1983, Saaty, 1987) or Conjoint Analysis (CA - Green and Srinivasan, 1978, Hair, 1995, McDaniel and Gates, 1998) have been proposed.

Most of aggregation methods above-mentioned, assume in weighting phase, the compensability among simple indicators (Bouyssou and Vansnick, 1986), namely allowing lower values in some indicators to be compensated by higher values in others. In addition, this property isn't even verified in the practical application, especially if they have to be interpreted as "*importance coefficients*" (Munda and Nardo, 2005).

In the last years, multiple solutions have been proposed to avoid this strong assumption introducing weight constraints, weighting each tensor that links the single point to the frontier (see *e.g.* Tsutsui et al., 2009) or including a penalty according to the different mix of simple indicators.

In particular, considering the third approach, Vidoli and Mazziotta (2012a) suggest to incorporate the MPVC (Method of Penalties for Coefficient of Variation - De Muro et al., 2010) idea in the basic BoD method in order to take into account the benchmark units on the frontier (as in BoD) and to penalize, in the case of non-compensatory issues, the presence of unbalanced data (as in MPVC).

In the latter method, however, given the chosen penalty criteria, the aggregate function doesn't always satisfy the *weakly positive monotonicity property* (see *e.g.* Chakravarty, 2003, Casadio Tarabusi and Guarini, 2013):

Property 1 (positive monotonicity). *Let $CI = f(I)$ an unbalanced-adjusted aggregation function of k simple indicators I , f is weakly positive monotone if for each $c > 0$, $f(I_1, \dots, I_j, \dots, I_k) \leq f(I_1, \dots, I_j + c, \dots, I_k)$.*

This means that f increases whenever any of the simple indicators increase and the others are left unchanged. Figure 1 shows how the BoD-PVC (solid

line) modifies the BoD (dashed line) level curves; in some cases the BoD-PVC doesn't satisfy the monotonicity property *e.g.* if the simple indicator I_2 increases from point B to point A, the value of CI decrease.

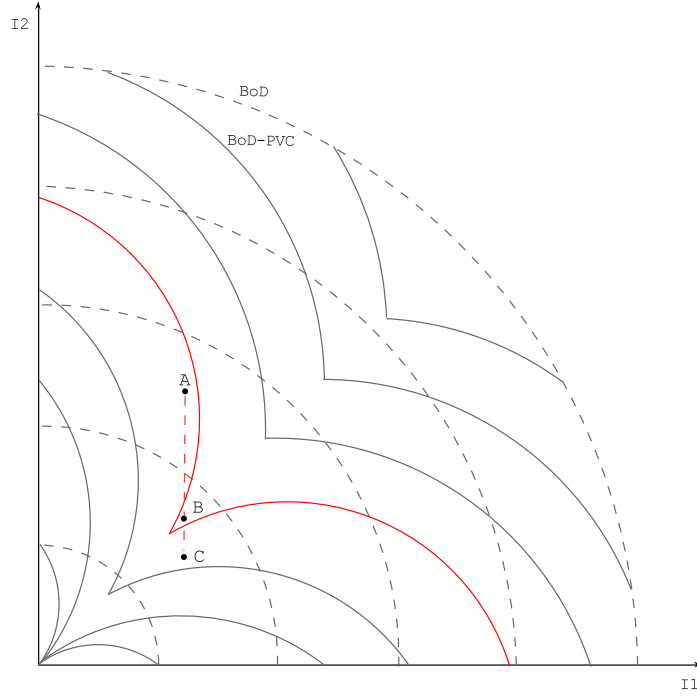


Figure 1: Comparison between BoD and BoD-PVC

In this paper, we propose a more general method starting from the BoD idea and introducing “directional” penalties with the aim to consider in the analysis the preference structure among simple indicators identified through a PCA analysis.

The paper is organized as follows. In section 2 we describe our theoretical model, the section 3 shows an empirical application on infrastructural endowment data in European regions. Conclusion and future perspective are reported in section 4.

2. A directional BoD model

In a classical production framework we consider a Decision Making Unit (DMU) i using p inputs $\mathbf{x} = (x_1, \dots, x_p) \in \mathbb{R}_+^p$ to produce q outputs $\mathbf{y} = (y_1, \dots, y_q) \in \mathbb{R}_+^q$

\mathbb{R}_+^q with a *production set* Ψ such that:

$$\Psi = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \in \mathbb{R}_+^p, \mathbf{y} \in \mathbb{R}_+^q, (\mathbf{x}, \mathbf{y}) \text{ is feasible}\} \quad (1)$$

satisfying the usual assumptions as in Shephard (1970) and in Färe et al. (1985).

The basic BoD approach is a particular case of the CCR-DEA model (Charnes et al., 1978) where x is univariate and constant equal to $\mathbf{1}$ and \mathbf{y} is a vector of k simple indicators in $[0, 1]$ (from this point denoted by \mathbf{I}).

The BoD estimator of the output efficiency score λ for a given unit o is obtained by solving the following linear program:

$$\begin{aligned} \max_{\lambda, \gamma_1, \dots, \gamma_n} \quad & \lambda \\ \text{s.t.} \quad & \lambda \mathbf{I}_o \leq \sum_{i=1}^n \gamma_i \mathbf{I}_i \\ & \gamma_i \geq 0 \end{aligned} \quad (2)$$

A strong assumption of the model (2) is the compensability among different simple indicators; given that in practical application most often exist a preference structure, we suggest to include in the BoD model (equation (2)) a "directional" penalty using the *directional distance function*¹ introduced by Chambers et al. (1998):

$$\vec{D}_T(\mathbf{x}, \mathbf{y}; \mathbf{g}) = \sup\{\beta : (x - \beta g_x, y + \beta g_y) \in \Psi\} \quad (3)$$

where $\mathbf{g} = (g_x, g_y)$ is the directional vector.

Against this background, in literature, a crucial question in a directional approach is the choice of the direction.

Some authors (see e.g. Briec and Lesourd, 1997, Färe et al., 2005) suggest to choose $\mathbf{g} = (1, \dots, 1)$ which is mathematically equivalent to seeking the Chebyshev distance l_∞ to the frontier of the technology. Bogetoft and Otto (2011)

¹The function satisfies the following properties:

- (i) $\vec{D}_T(\mathbf{x}, \mathbf{y}; \mathbf{g}) \geq 0 \Leftrightarrow x \in \Psi$ (representation);
- (ii) $\vec{D}_T(\mathbf{x} - \alpha \mathbf{g}, \mathbf{y} + \alpha \mathbf{g}; \mathbf{g}) = \vec{D}_T(\mathbf{x}, \mathbf{y}; \mathbf{g}) - \alpha$ for $\alpha \in \mathbb{R}_+$ (translation).

conversely propose four approaches: i) use the direction of the actual value of input consumption (output production) *i.e.* $g_x = x_o$ ($g_y = y_o$), ii) fix a part of the input-vector (output-vector) *i.e.* $g_x = (1, \dots, 1, 0, \dots, 0)$ ($g_y = (1, \dots, 1, 0, \dots, 0)$), iii) consider the subjective user point of view or iv) use the *potential improvements* or *multi-directional efficiency* analysis based on the bargaining theory.

All above proposals presuppose an exogenous choice of the researcher, on the contrary we propose to identify the direction vector directly, from data, estimating the endogenous preference structure among indicators getting through to PCA.

For seek of simplicity and in order to better visually illustrate our method, we consider the bivariate case of two simple indicators $\mathbf{I} = (I_1, I_2)$. Figure 2 compares the CI scores obtained with BoD (dashed line) and D-BoD (solid line) formulation in an hypothetical case in which the simple indicator I_1 is the most important in discriminating units. The two straight lines represent the directions underlying the models: $g_{BoD} = (I_1, I_2)$ and $g_{D-BoD} = (I_1, I_2 \cdot 0.5)$.

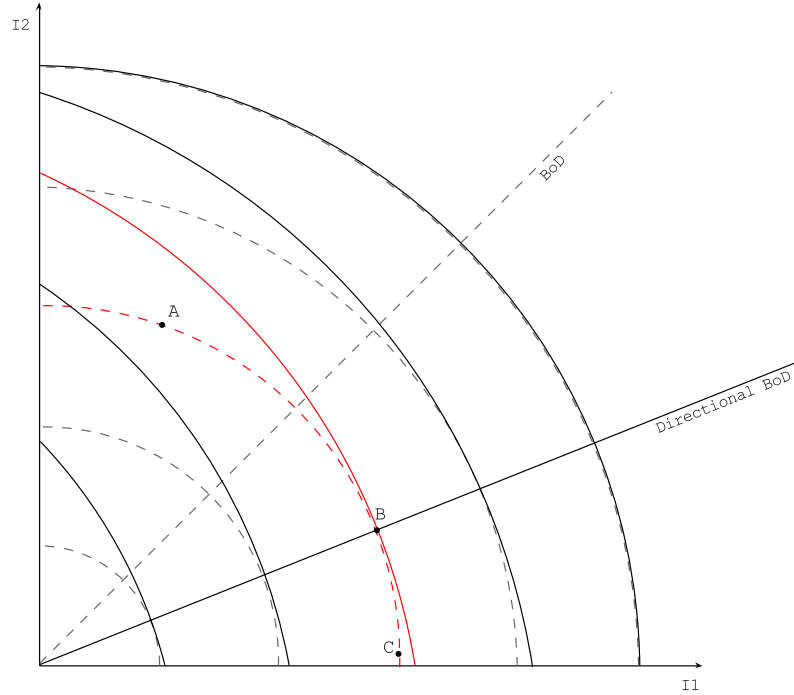


Figure 2: Comparison between BoD and Directional BoD

Given this representation, point A, B and C lie on the same level curve (the red dashed line) in a BoD model, while in a D-BoD model points A and C have a lower level of efficiency than B.

As a matter of fact, the D-BoD model rewards the combinations of I_1 and I_2 on the main direction (point B) and penalizes, in a different way, the combinations of low values of I_1 and high values of I_2 (point A) respect to combinations of high values of I_1 and low value of I_2 (point C).

We can observe that on the main direction the BoD level curve coincides with the D-BoD ones and that the two curves are overlaid on the frontier.

The proposed model is, therefore, a more general formulation of the basic BoD model where I_1 and I_2 have the same importance i.e. $g_{BoD} = (I_1, I_2)$.

Please note, finally, that this *unbalance-adjusted function* satisfies the property of *weakly positive monotonicity* emphasized in section 1 on page 3, i.e. for each $c > 0$, since $e(\mathbf{1}, I_1, \dots, I_j, \dots, I_k; \Psi, \mathbf{g}) \leq e(\mathbf{1}, I_1, \dots, I_j + c, \dots, I_k; \Psi, \mathbf{g})$.

3. European infrastructural endowment data

The D-Bod model outlined in section 2 has been applied to the terrestrial transport infrastructure endowment in European Regions following Vidoli and Mazziotta (2012b) proposal.

The data set² includes information on two simple indicators³ concerning roads (I_{Roads}) and railways (I_{Trains}) endowment for France, Germany, Italy and Spain.

²Source: Eurostat, Statistics by theme, 2012

³For variables and the method of construction of simple indicators, please see Vidoli and Mazziotta (2012b).

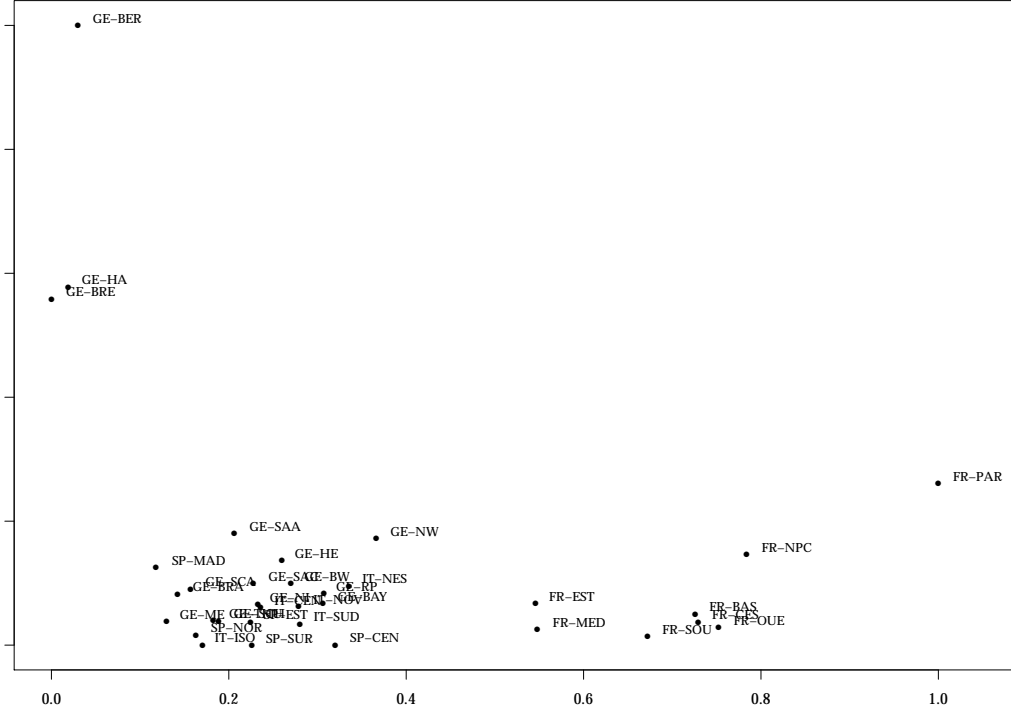


Figure 3: European infrastructure endowment - Data

Figure 3 shows three departments (GE-BER- Berlin, GE-BRE - Bremen and FR-BAS - Hamburg) with low value of I_{Roads} and high value of I_{Trains} . Table 1 and Figure 4 show results obtained with BoD and D-BoD methods. In particular, while the Spearman Index between the two approaches is very high (equal to 0.938), the average CI score of the isolated departments falls (see Table 1) from 55.86% to 14.19% in GE-BRE - Bremen and from 57.92% to 17.31% in FR-BAS - Hamburg.

Finally, we highlight that Berlin remains at the top of the ranking (see Figure 2) because on the frontier the BoD level curve coincides with the D-BoD ones.

Table 1: European infrastructure endowment: Comparison between BoD and D-BoD

NUTS2	Department	Country	Bod	D-Bod
FR-BAS	Bassin Parisien	France	0.72579	0.72579
FR-CES	Centre-Est (FR)	France	0.72935	0.72935
FR-EST	Est (FR)	France	0.54597	0.54597

Continued on Next Page.

Table 1: European infrastructure endowment: Comparison between BoD and D-BoD

NUTS2	Department	Country	Bod	D-Bod
<i>FR-MED</i>	Méditerranée	France	0.54767	0.54767
<i>FR-NPC</i>	Nord - Pas-de-Calais	France	0.78384	0.78384
<i>FR-OUE</i>	Ouest (FR)	France	0.75227	0.75227
<i>FR-PAR</i>	Île de France	France	1.00000	1.00000
<i>FR-SOU</i>	Sud-Ouest (FR)	France	0.67219	0.67219
<i>GE-BAY</i>	Bayern	Germany	0.30602	0.30602
<i>GE-BER</i>	Berlin	Germany	1.00000	1.00000
<i>GE-BRA</i>	Brandenburg	Germany	0.18692	0.14226
<i>GE-BRE</i>	Bremen	Germany	0.55858	0.14190
<i>GE-BW</i>	Baden-Württemberg	Germany	0.29817	0.26998
<i>FR-BAS</i>	Hamburg	Germany	0.57922	0.17308
<i>GE-HE</i>	Hessen	Germany	0.32695	0.25981
<i>GE-ME</i>	Mecklenburg-Vorpommern	Germany	0.13439	0.12965
<i>GE-NI</i>	Niedersachsen	Germany	0.23789	0.23278
<i>GE-NW</i>	Nordrhein-Westfalen	Germany	0.44095	0.36605
<i>GE-RP</i>	Rheinland-Pfalz	Germany	0.31083	0.30717
<i>GE-SAA</i>	Saarland	Germany	0.33017	0.20871
<i>GE-SAC</i>	Sachsen	Germany	0.26706	0.22757
<i>GE-SCA</i>	Sachsen-Anhalt	Germany	0.20525	0.15678
<i>GE-SCH</i>	Schleswig-Holstein	Germany	0.18842	0.18842
<i>GE-THU</i>	Thüringen	Germany	0.18239	0.18239
<i>IT-CEN</i>	Centro (IT)	Italy	0.23567	0.23567
<i>IT-ISO</i>	Isole	Italy	0.17049	0.17049
<i>IT-NES</i>	Nord-Est	Italy	0.34354	0.33554
<i>IT-NOV</i>	Nord-Ovest	Italy	0.27881	0.27881
<i>IT-SUD</i>	Sud	Italy	0.28023	0.28023
<i>SP-CEN</i>	Centro (ES)	Spain	0.32009	0.32009
<i>SP-EST</i>	Este (ES)	Spain	0.22460	0.22460
<i>SP-MAD</i>	Comunidad de Madrid	Spain	0.20997	0.11784
<i>SP-NOR</i>	Noreste (ES)	Spain	0.16303	0.16303
<i>SP-SUR</i>	Sur (ES)	Spain	0.22607	0.22607

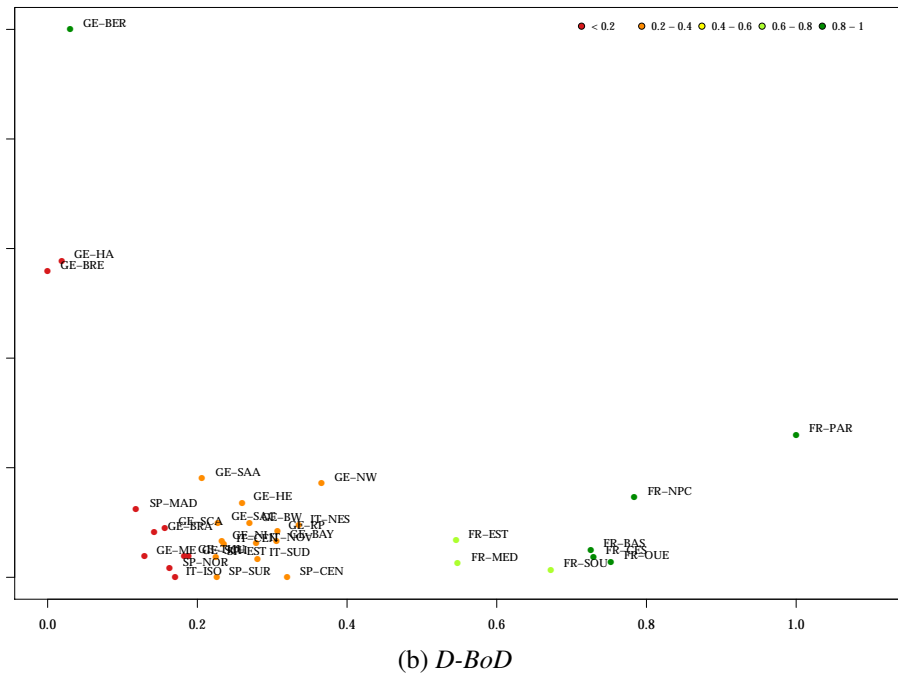
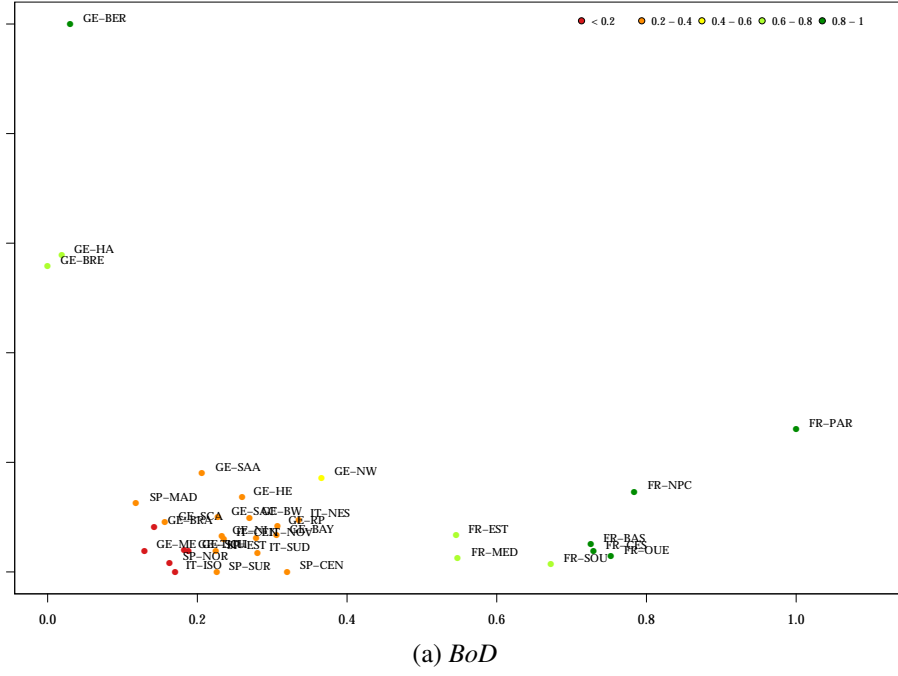


Figure 4: European infrastructure endowment: Comparison between BoD and D-BoD

4. Conclusions

In this paper we have presented a new approach for the construction of CIs that enhance non-compensatory issue by correcting the BoD index with “directional” penalties.

Our future remarks are the introduction of the *order-m* approach in order to obtain a more robust estimator.

References

- Bogetoft, P., Otto, L., 2011. Benchmarking with Data Envelopment Analysis, Stochastic Frontier Analysis and R. International series in operations research & management science. Springer New York.
- Bouyssou, D., Vansnick, J.-C., 1986. Noncompensatory and generalized noncompensatory preference structures. *Theory and Decision* 21, 251–266.
- Briec, W., Lesourd, J.-B., 1997. Metric distance, and profit functions: Some duality results. Tech. rep., Universite Aix-Marseille III.
- Casadio Tarabusi, E., Guarini, G., 2013. An unbalance adjustment method for development indicators. *Social Indicators Research* 112 (1), 19–45.
- Chakravarty, S. R., 2003. A generalized human development index. *Review of Development Economics* 7 (1), 99–114.
- Chambers, R. G., Chung, Y., Färe, R., 1998. Profit, Directional Distance Functions, and Nerlovian Efficiency. *Journal of Optimization Theory and Applications* 98 (2), 351–364.
- Charnes, A., Cooper, W., Rhodes, W., 1978. Measuring the efficiency of decision making units. *European journal of operational research* 2(4), 429 – 444.
- De Muro, P., Mazziotta, M., Pareto, A., 2010. Composite indices of development and poverty: An application to mdgs. *Social Indicators Research* 104(1), 1–18.
- Färe, R., Grosskopf, S., Lovell, C., 1985. *The Measurements of Efficiency of Production*. Studies in Productivity Analysis. Springer.
- Färe, R., Grosskopf, S., Noh, D., Weber, W., 2005. Characteristics of a polluting technology: theory and practice. *Journal of Econometrics* 126 (2), 469–492.
- Forman, E., 1983. *The analytic hierarchy process as a decision support system*.
- Freudenberg, M., 2003. Composite indicators of country performance: A critical assessment. Tech. rep., OECD Science, Technology and Industry Working Papers 2003/16, OECD, Directorate for Science, Technology and Industry.
- Green, P. E., Srinivasan, V., 1978. Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research* 5 (2), 103–23.
- Hair, J., 1995. *Multivariate data analysis with readings*. Prentice Hall International Editions. Prentice Hall.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2003. Governance matters III: Governance indicators for 1996-2002. Development and Comp Systems 0308001, EconWPA.
- Kaufmann, D., Kraay, A., Zoido-lobatón, P., 1999. Aggregating governance indicators. In: *WB Policy Research Paper*. p. 2195.
- Manly, B., 1994. *Multivariate Statistical Methods: A Primer*. Chapman and Hall.
- McDaniel, C., Gates, R., 1998. *Contemporary marketing research*. West Pub. Co.

- Melyn, W., Moesen, W., te Leuven (1970-). Centrum voor Economische Studiën, K. U., 1991. Towards a Synthetic Indicator of Macroeconomic Performance: Unequal Weighting when Limited Information is Available. Public economics research papers. Katholieke Universiteit Leuven.
- Moldan, B., Billharz, S., Matravers, R., 1997. Sustainability indicators: a report on the project on indicators of sustainable development. SCOPE Series. Wiley.
- Munda, G., Nardo, M., 2005. Constructing consistent composite indicators: the issue of weights. Tech. rep., EUR 21834 EN, European Commission.
- Saaty, R., 1987. The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling* 9 (3–5), 161 – 176.
- Shephard, R., 1970. Theory of cost and production functions. Princeton studies in mathematical economics. Princeton University Press.
- Tsutsui, M., Tone, K., Yoshida, Y., 2009. Technical efficiency based on cost gradient measure. Tech. rep., Discussion Paper 09-14, GRIPS Policy Information Center, Tokyo.
- Vidoli, F., Mazziotta, C., 2012a. Constructing a variance weighted bod (benefit of doubt) composite indicator and its confidence interval through bootstrap: an application to italian infrastructural endowment.
- Vidoli, F., Mazziotta, C., 2012b. Robust weighted composite indicators by means of frontier methods: An application to european infrastructure endowment. XXXIII Conferenza Italiana di Scienze Regionali, Roma, AISRe.