DARSENA
DIPARTIMENTO DI ECONOMIA
UNIVERSITÀ DEGLI STUDI DI GENOVA

UNIVERSITÀ DEGLI STUDI
DI GENOVA

SCUOLA DI
SCIENZE SOCIALI

Policy-making and policy assessments with partially ordered alternatives

Matteo Corsi

Supervisor:

Barbara Cavalletti

To my mother and to my father

ABSTRACT

The present work collects three essays on social choice and decision-making in the presence of multiple objectives and severe informational limitations. When feasible alternatives must be ordered according to their performance under various criteria, it is typically necessary to make use of a specific functional relation and assume the implied rates of substitution between scores in different criteria. In the special case of collective choice and voting, rather than having proper rates of substitution, each individually preferred ordering of the alternatives is usually weighted according to its frequency in the population. Both decision frameworks imply the availability of extensive information about such functional relation and the proper weights of each criterion or must acknowledge a vast and arbitrary discretion to those in charge of resolving the decision process. The alternative approach herein discussed consists in applying the Pareto criterion to identify Pareto-superior alternatives in each pairwise comparison, a procedure that easily produces an incomplete ordering. Then, applying a tool of Order Theory, a complete ordering is identified from the linear extensions of the partially ordered set derived from the Pareto criterion. The claim is that this method highlights conflicts in value judgements and in incomparable criteria, allowing to search for a conflict-mitigating solution that doesn't make assumptions on the reciprocal importance of criteria or judgements. The method is actually a combination of existing but unrelated approaches in Social Choice Theory and in Order Theory and provides outcomes with interesting properties. The essays present, respectively, an axiomatic discussion of the properties of this approach and two applications to policy issues.

TABLE OF CONTENTS

FIGURES

TABLES

INTRODUCTION

1. Policy decisions in the presence of multiple criteria

In their use of executive power, governments are generally represented as having a *set* of objectives and having determined the course of action that seems more suited to achieve them. The idea of a plurality of goals underlying public decision-making is, in fact, consistent with centuries of economic theory and normative views on welfare and social justice. Classical *Utilitarianism* and *Welfarism*, the long-dominant approaches in welfare economics theory before Lionel Robbins and the 1930s revolution of ordinal utility (Sen 1995), sought to maximize a function of individual utilities, which ultimately corresponds to taking care of the broad, subjective notion of individual well-being (Kaplow 2010 p. 347) and its multidimensional structure. The *new welfare economics* tradition (Abram Bergson, Paul Samuelson, Kenneth Arrow), which reformed the problem of social welfare around the notion of ordinal utility, still conceived tastes and preferences as affecting social welfare and, therefore, justifying the multidimensional concerns of a government. Even the more recent political philosophies developed in the second half of the Twentieth Century, including the various currents of libertarian and egalitarian thinking, are perfectly consistent with multiple government goals. For instance, the self-interested agents that interact to define policies in a *public choice* framework are as likely to assemble a patchwork of objectives as a good *Rawlsian* government or a non-welfarist decision-maker concerned with Sen's *capabilities*. At the roots of the modern theory of public finance (Musgrave 1959) stand, at the very least, three main drivers of government activity: resource allocation to address market failures, income redistribution and macroeconomic stabilization.

A less uniform picture is that of how different normative approaches deal with decisions made in the presence of multiple criteria and goals. Broadly speaking, welfarist approaches adopt individual utility or well-being as the synthesis of the consequences produced by all policies aimed at all the objectives. Social welfare is then a function (*social welfare function*) of the utility level attained by all the individuals in a society[1], under a certain state of the world. Non-welfarist approaches devise a number of methods to pursue multiple objectives at the same time and to

---

[1] Specifically, a sum of individual cardinal utilities (classical welfarism), a function of individual ordinal utilities (Bergson-Samuelson social welfare function) or a function of individual preference orderings determined by individual ordinal utilities (Arrow's social welfare functions.

find which policy is more desirable in this respect. In doing so, they manipulate the arguments of the social welfare function or its functional form (or both) in order to fit the normative prescription according to which other criteria beyond individual utility affect social welfare.

Interestingly, while diverging, the paths of welfarist and non-welfarist approaches still have a common practical implication: applying either of them to policy analysis requires a precise specification of a functional form which ultimately rests on crucial prior knowledge or on purely normative assumptions. Those implications are indeed perfectly acceptable under many circumstances, but they still should not be taken for granted. Decision-making problems arise and, most often, are actually worth of analysis, when complex trade-offs could affect individual or social welfare. This is frequently the case with public policies, as long as the space for public policy is not confined to cases of unanimity. Meaningful criteria to assess policies and their expected or observed outcomes come in all sorts of measurement units and scales, sometimes defying practical attempts to aggregate, other times challenging the logical justification of an aggregation. If nothing else, it may be hard to establish beyond reasonable doubt that some apparent collective preference over a pair of states of the world should not actually be reversed. Complicating things further, the problem may arise at any scale. A local policy on public housing wait-lists based on a set of criteria, or one to assign resources to different universities or different hospitals based on performance indicators, may pose similar challenges to those implied in defining a national policy taking care of growth, social justice and emission reduction.

The motivation for this thesis comes from noticing that, in the presence of a plurality of objectives, policy-making can be affected by major problems of uncertainty and partial information that are not necessarily core interests of the multicriteria methods and of the normative approaches that are typically adopted in literature. The issue of informational paucity is somewhat conventional in Social Choice Theory (Sen 2017; J. Weymark 2013) insofar as some specific information must be excluded from a choice rule, like interpersonal comparisons of utility or irrelevant alternatives, but there seems to be no explicit attempt of defining what properties may be desirable for decision processes taking place under severe informational limitations.

The attempt made here rests on two arguments. The first is that, after a set of criteria is established, any ordering of a set of alternatives (policies, allocations, etc.) can, at a minimum, exploit (Pareto) dominance between pairs of alternatives. This is because each criterion comes with an embedded polarity (or direction) determined by the set of underlying objectives. It is a fairly standard assumption and the main open questions on it (see for instance Sen 2017 sec. 2) should be reconsidered in the light of the partial information context discussed in the remainder of this work. However, it is well known that the Pareto dominance criterion does not necessarily lead to a complete ordering of all pairs of alternatives (and of the entire set as a consequence). Then, the second argument is that the branch of discrete mathematics called *Order Theory* (Davey and Priestley 2002; Schroder 2002; Stanley 2011) has developed a substantial literature on a functional relation which, given a certain partially ordered set, identifies one and only one complete order with properties that are of notable interest for economic theory. The relation, which was first detailed in a work by Peter Winkler (Winkler 1982) but is best described as part of a wider approach discussed by a group of authors between the 1980s and the 1990s, has been alternatively indicated as *average height of an element in a partially ordered set* or, particularly in some applications, as *average rank.* Besides some passing remarks in the original body of work[2], few economic applications have been proposed, and none with a specific underlying link with economic theory. The intuition discussed in this work is that the average height of an element in a partially ordered set of policies based on the Pareto criterion is a meaningful indicator to completely order the set if no reliable information is available beyond what is required to build the partially ordered set itself.

In the following sections I will therefore discuss a possible theoretical justification for the use of average height in a partially ordered set to deal with collective choice and policy decisions, anchoring it to the existing literature on collective choice and welfare. Furthermore, I will explore two applications of this functional relation to fairly standard issues of public economics: waste taxation and the identification of the determinants of well-being. The remainder of this work is organized as follows. This introductory section will end with a brief overview of the general notation and terminology of Order Theory, leaving all the specialized notation for the following chapters. In Chapter I, the average height of an element in a partially ordered set is presented as

---

[2] See for instance the introduction of Shepp (1982).

the key part of a collective choice rule of which I discuss a simple axiomatic characterization. In this case, the ordering criteria consist of voters, individual preference profiles or individual families of social welfare functions, along the usual path of social choice theory. The axioms highlight the most relevant properties of this functional relation in a social choice framework. In the following chapters, the focus shifts from individuals as ordering criteria to ordering criteria in general. In Chapter II, the problem of evaluating the performance of municipal waste management systems in Italy is considered in the light of the multiple available indicators on one hand, and the single indicator being currently used to design an incentive mechanism on the other hand. The use of the average height functional relation for the partially ordered set of municipalities in the Province of La Spezia allows to evaluate the robustness of performance rankings moving from single to multiple performance indicators and the single indicator system is shown to have considerably regressive effects on waste taxation. In Chapter III, the OECD conceptual model of social indicators expressing national welfare is put to test in order to understand if a cross-sectional model of national subjective well-being is more effectively specified with such social indictors (and which of them perform better) rather than with GDP alone. It is found that the performance of "beyond GDP" indicators is at odds with theory and occasionally inconsistent. However, the main emphasis is on the procedure to build the "beyond GDP" model, as the average height functional relation is applied to the partial order of OECD countries according to indicators that fall within the same domain, in order to identify a synthetic indicator with desirable properties. The thesis ends with a brief concluding chapter that sums up the findings.

The content of Chapter II and Chapter III have already been published in two separate journal articles and is reported as is[3].

---

[3] I am author of sections 1, 2, 3, 4 of chapter II and sections 1, 2, 3, 4, 5 and the appendices of chapter III. A formal declaration of authorship is attached to the present thesis in fulfillment of the Italian legislation.

2. An introductory note on binary relations and Order Theory[4]

A *binary relation* on a set A is a subset $R \subseteq A \times A = \{(a, b): a, b \in A\}$. Therefore, the binary relation R consists of pairs of elements from A, and when a certain pair $(a, b) \in R$, we say that a is related to b under R. A frequently adopted convention is to replace $(a, b) \in R$ with the more compact notation aRb. The couple (A, R) consisting of a set and a binary relation is called *relational structure* and A is called the *ground set* of the relational structure. If the ground set A is finite, then the relational structure (A, R) is finite.

Binary relations of particular interest are classified with respect to the following properties.

I.	Reflexivity	$(if\ aRa\ for\ all\ a \in A)$
II.	Symmetry	$(if\ aRb \rightarrow bRa)$
III.	Anti-symmetry	$(if\ aRb\ and\ bRa \rightarrow a = b)$
IV.	Transitivity	$(if\ aRb\ and\ bRc \rightarrow aRc)$
V.	Completeness	$(if\ aRb\ or\ bRa\ for\ all\ a, b \in A)$

A binary relation R that is reflexive and transitive is called a *quasi-ordering relation*. The corresponding relational structure (A, R) is called a *quasi-ordered set*[5]. A symmetric quasi-ordering is an *equivalence relation*. An anti-symmetric quasi-ordering is an *order relation.* If an order relation is complete, it is called *linear order relation*.

From a binary relation R it is usually convenient to define three parts of particular relevance. The *asymmetric part* of R is indicated with P and is obtained by setting aPb iff aRb and $\neg(bRa)$. The *symmetric part* is indicated as aIb and is defined by setting $aRb \leftrightarrow bRa$. The *incomparability part* is indicated as $a||b$ and comes from setting $\neg(aRb)$ and $\neg(bRa)$. Intuitively, incomparability arises when a binary relation is incomplete and is absent when the binary relation satisfies completeness. Therefore, the difference between an order relation and a linear order relation is

---

[4] The content of this section is assembled from authoritative sources on binary relations and Order Theory including Stanley (2011) Schroder (2002), Davey and Priestley (2002), Foldes (1994).

[5] We conventionally assume that the properties of an order relation are not distinct from the properties of the ordered set equipped with that order relation.

that the former may imply incomparability between some pairs of elements in the ground set whereas the latter presents no incomparability.

Order relations are a particularly pervasive kind of binary relations and they are frequently denoted by the specialized symbol $\leq$ in place of the generic symbol of relation R. The effect of order relations on sets is to introduce a hierarchic structure so that $a \leq b$ is interpreted as a is less than or equal to b or, equivalently, b is greater than or equal to a. Correspondingly, the asymmetric part of the order relation is the strict order $a < b$, the symmetric part is the equality $a = b$ where a and b denote the same object. Transitivity and anti-symmetry together exclude from order relations any possibility to find cycles of the kind $a_1 \leq a_2 \leq \cdots . \leq a_n \leq a_1$ for all distinct $a_1, a_2, \dots, a_n$. The relational structure $(A, \leq)$, which consists of the ground set A and an order relation on it, is called *ordered set* or, more frequently, *partially ordered set* or *poset*, to emphasize that only in the special case of a complete order relation all pairs a, b of elements in A are also in the order relation, either as $a \leq b$ or as $b \leq a$. Partially ordered sets constitute the main, although not the only relational structures used throughout this thesis.

Before presenting some key findings in Order Theory that play a role in the following chapters, it is opportune to remind the reader that additional naming and notation conventions are extensively used in literature and are adopted when required in this thesis. In cases where there is no possible confusion, it is common to use the same letter (e.g. A) to indicate a partially ordered set as well as its ground set. When multiple binary relations satisfying *the same* properties are in use on the same ground set, they may be indexed to keep track of the differences. When multiple binary relations satisfying *different* properties are in use instead, then the specialized notation $\leq$ for order relations may be replaced with one more apt at highlighting the different types of binary relations and the term R is typically reserved for order relations in general. In extensive discussions concerning a binary relation, it is customary to indicate with $R(A)$ the binary relation over set A which is the counterpart of the ordered set (A, R). These conventions are frequently found in Social Choice Theory and I adopt them in particular in Chapter I.

I conclude this brief overview with some classical results in the study of partial orders that constitute the starting point for the functional relation I already mentioned as developed by Winkler and others and that is the main method of choice for the thesis. The results encompass what is broadly known as the field of *extension theorems* because the main point consists in defining how order relations can be extended by introducing new elements to the subset R of the Cartesian product $A \times A$.

The first result is that any order can be added an additional comparability of previously incomparable elements (Schroder 2002, Lemma 10.1), which is to say that if R is an order relation on a ground set A, a, b are two distinct elements of the ground set and $(a, b) \notin R$, then there is an order $R^{'}$ on A that contains every pair included in R and (a, b). In this sense we can say that R' contains R or, using a terminology that is more frequent in a different literature, that it is *compatible* with R. The structure of the set of all orders that contain a given R is particularly important: when ordered by inclusion, it is itself an ordered set and it is conditionally complete, with all maximal elements represented by linear orders. Following a classical theorem (Szpilrajn 1930) it is known that every element in this "ordered set of orders" stands below a maximal element and it is therefore possible to conclude that every order is contained in a linear order.

From the above it is then possible to derive the definition of *linear extension* of a partially ordered set (A, R) as a relational structure (A, L), a linearly ordered set where L is a linear order and contains R . From Szpilrajn's theorem also descends a result due to Dushnik and Miller (1941) which says that if a partially ordered set A contains non-comparable elements, then for every non comparable pair a, b there exist an *extension* R_1 in which aR_1b and an extension R_2 in which bR_2a. Finally, every partial order is the intersection of the linear orders that extend it, that is, the intersection of its linear extensions.

References

Achillas, C., N. Moussiopoulos, A. Karagiannidis, G. Banias, and G. Perkoulidis. 2013. The use of multi-criteria decision analysis to tackle waste management problems: a literature review. *Waste Management & Research*, *31*(2):115–129.

Allesch, A., and P. H. Brunner. 2014. Assessment methods for solid waste management: A literature review. *Waste Management & Research*, *32*(6):461–473.

Arrow, K. J. 1962. *Social Choice and Individual Values*. New York, New York, USA: John Wiley & Sons.

Bouyssou, D. 1986. Some remarks on the notion of compensation in MCDM. *European Journal of Operational Research*, *26*(1):150–160.

Brightwell, G. R., and W. T. Trotter. 2002. A combinatorial approach to correlation inequalities. *Discrete Mathematics*, *257*(2–3):311–327.

Brightwell, G. R., and P. Winkler. 1991. Counting linear extensions. *Order*, *8*(3):225–242.

Brüggemann, R., and P. Annoni. 2014. Average Heights in Partially Ordered Sets. *MATCH Communications in Mathematical and in Coputer Chemistry*, *71*:117–142.

Brüggemann, R., and L. Carlsen. 2011. An Improved Estimation of Averaged Ranks of Partial Orders. *MATCH Communications in Mathematical and in Computer Chemistry*, *65*:383–414.

Brüggemann, R., and G. Patil. 2011. *Ranking and prioritization for multi-indicator systems*. Springer.

Brüggemann, R., P. B. Sørensen, D. Lerche, and L. Carlsen. 2004. Estimation of averaged ranks by a local partial order model. *Journal of Chemical Information and Computer Sciences*, *44*(2):618–25.

Davey, B. A., and H. A. Priestley. 2002. *Introduction to lattices and order* (Second Edi). Cambridge, UK: Cambridge University Press.

De Loof, K., B. De Baets, and H. De Meyer. 2011. Approximation of Average Ranks in Posets. *MATCH - Communications in Mathematical and in Computer Chemistry*, *66*:219–229.

Desmond, M. 2006. Municipal solid waste management in Ireland: assessing for sustainability. *Irish Geography*, *39*(1):22–33.

Dushnik, B., and E. W. Miller. 1941. Partially Ordered Sets. *American Journal of Mathematics*, *63*(3):600–610.

Edwards, J. R., R. P. Bagozzi, R. S. Blackburn, K. a Bollen, D. Cattani, J. W. D. Jr, D. a Harrison, L. R. James, K. a Markus, E. E. Rigdon, A. H. Segars, and L. J. Williams. 2000. On the Nature and Direction of Relationships Between Constructs and Measures. *Psychological Methods*, *5*(2):155–174.

Fattore, M. 2017. Functionals and Synthetic Indicators Over Finite Posets. In M. Fattore & R. Bruggemann (Eds.), *Partial Order Concepts in Applied Sciences* (pp. 71–86). Springer

International Publishing AG.

Fishburn, P. C. 1976. Noncompensatory Preferences. *Synthese*, *33*(2):393–403.

Foldes, S. 1994. *Fundamental Structures of Algebra and Discrete Mathematics*. New York: Wiley & Sons.

Frey, B. S., and A. Stutzer. 2005. Happiness Research: State and Prospects. *Review of Social Economy*, *62*(190):207–228.

Greene, K. L., and D. J. Tonjes. 2014. Quantitative assessments of municipal waste management systems : Using different indicators to compare and rank programs in New York State. *Waste Management*, *34*(4):825–836.

Guimarães, B., P. Simões, and R. C. Marques. 2010. Does performance evaluation help public managers? A Balanced Scorecard approach in urban waste services. *Journal of Environmental Management*, *91*(12):2632–2638.

Guttman, L. 1950. The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), *Measurement and prediction.* Princeton University Press.

ISPRA. 2014. *Rapporto rifiuti urbani. Edizione 2014*. Rome.

Jorm, A. F., and S. M. Ryan. 2014. Cross-national and historical differences in subjective well-being. *International Journal of Epidemiology*, *43*(2):330–340.

Kaplow, L. 2010. *The theory of taxation and public economics* (Paperback). Princeton University Press.

Kaufman, S. M., N. Krishnan, and N. J. Themelis. 2010. A screening life cycle metric to benchmark the environmental sustainability of waste management systems. *Environmental Science and Technology*, *44*(15):5949–5955.

Kulczycka, J., L. Lelek, A. Lewandowska, and J. Zarebska. 2015. Life Cycle Assessment of Municipal Solid Waste Management – Comparison of Results Using Different LCA Models. *Polish Journal of Environmental Studies*, *24*(1):125–140.

Lerche, D., P. B. Sørensen, and R. Brüggemann. 2003. Improved estimation of the ranking probabilities in partial orders using random linear extensions by approximation of the mutual ranking probability. *Journal of Chemical Information and Computer Sciences*, *43*(5):1471–80.

Massoud, M. A., M. El-Fadel, and A. A. Malak. 2003. Assessment of public vs. private MSW management: A case study. *Journal of Environmental Management*, *69*(1):15–24.

Mendes, P., A. C. Santos, F. Perna, and M. Ribau Teixeira. 2012. The balanced scorecard as an integrated model applied to the Portuguese public service: A case study in the waste sector. *Journal of Cleaner Production*, *24*:20–29.

Musgrave, R. A. 1959. *The theory of public finance: a study in public economy*. New York: McGraw-Hill.

OECD. 2008. *Handbook on constructing composite indicators: methodology and user guide*. Brussels.

OECD. 2013. *Environmental indicators: development, measurement, and use*. Paris.

Oswald, A. J. 2008. On the curvature of the reporting function from objective reality to subjective feelings. *Economics Letters*, *100*(3):369–372.

Pattanaik, P. K. 2002. Positional Rules of Collective Decision-Making. In K. J. Arrow, A. Sen, & K. Suzumura (Eds.), *Handbook of Social Choice and Welfare volume 1* (pp. 361–394). Amsterdam: Elsevier Science B.V.

Piano regionale di gestione dei rifiuti e delle bonifiche (2015). Regione Liguria.

Sanjeevi, V., and P. Shahabudeen. 2015. Development of performance indicators for municipal solid waste management (PIMS): A review. *Waste Management & Research*, *33*(12):1052–1065.

Schroder, B. S. W. 2002. *Ordered Sets. An Introduction*. Boston: Birkhauser.

Scipioni, A., A. Mazzi, F. Zuliani, and M. Mason. 2008. The ISO 14031 standard to guide the urban sustainability measurement process : an Italian experience. *Journal of Cleaner Production*, *16*:1247–1257.

Sen, A. 1995. Rationality and social choice. *The American Economic Review*, *85*(1):1–24.

Sen, A. 2017. *Collective choice and social welfare*. Milton Keynes, UK: Penguin Random House.

Shepp, L. 1982. The XYZ conjecture and the FKG inequality. *The Annals of Probability*, *10*(3):824–827.

Shye, S. 1985. *Multiple Scaling. The Theory and Application of Partial Order Scalogram Analysis*. Jerusalem: Israel Inst of Applied Social Research.

Shye, S., and R. Amar. 1985. Partial order scalogram analysis by base coordinates and lattice mapping of items by their scalogram roles. In D. Canter (Ed.), *Facet Theory: Approaches to Social Research* (pp. 277–298). Springer Verlag.

Slemrod, J. 1996. *Tax progressivity and income inequality*. Cambridge: Cambridge University Press.

Stanley, R. P. 1981. Two Combinatorial Applications of the Alexsandrov-Fenchel inequalities. *J.Combin.Theory (A)*, *65*:56–65.

Stanley, R. P. 2011. *Enumerative Combinatorics Volume 1 second edition* (Second Edi, Vol. 1). Cambridge, UK: Cambridge University Press.

Suits, D. B. 1977. Measurement of tax progressivity. *The American Economic Review*, *67*(4):747–752.

Szpilrajn, E. 1930. Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*.

Vergara, S., and G. Tchobanoglous. 2012. Municipal solid waste and the environment: a global perspective. *Annual Review of Environment and Resources*, *37*:227–309.

Voigt, K., G. Welzl, and R. Brüggemann. 2004. Data analysis of environmental air pollutant monitoring systems in Europe. *Environmetrics*, *15*(6):577–596.

Welsch, H. 2007. Environmental welfare analysis: A life satisfaction approach. *Ecological Economics*, *62*:544–551.

Weymark, J. 2013. Social Welfare Functions. *Vanderbilt University Department of Economics Working Papers*, *VUECON-13*-(00018).

Weymark, J. A. 1984. Arrow ' s Theorem with Social Quasi-Orderings. *Public Choice*, *42*(3):235–246.

Wilson, D. C., L. Rodic, a. Scheinberg, C. a. Velis, and G. Alabaster. 2012. Comparative analysis of solid waste management in 20 cities. *Waste Management & Research*, *30*(3):237–254.

Winkler, P. 1982. Average height in a partially ordered set. *Discrete Mathematics*, *39*:337–341.

Zaccariello, L., R. Cremiato, and M. L. Mastellone. 2015. Evaluation of municipal solid waste management performance by material flow analysis: Theoretical approach and case study. *Waste Management & Research*, *33*(10):871–885.

Zopounidis, C., and P. Pardalos. 2010. *Handbook of Multicriteria Analysis*. (C. Zopounidis & P. Pardalos, Eds.). Heidelberg: Springer Verlag.

CHAPTER I

A collective choice rule based on correlation inequalities

1. Introduction

Consider the following and rather common circumstance. A collective choice supposed to emerge from the individual preferences of a population over a set of alternatives will have to be made on incomplete preference information. Obviously, the nature of the incompleteness is crucial in evaluating the opportune approaches to this kind of problem, and yet there is a number of different cases in which partial information leads to the same kind of logical set-up. For instance, suppose that, in a typical voting environment, a sufficiently large part of the relevant population is not in condition to express its preferences or that observed preferences belong to a subset of the population that represents all opinions but not proportionally. If we move outside the voting environment, just suppose that individual preferences are fully available but information on their intensity, which is supposed to matter in this particular collective choice environment, are not.

This chapter is concerned with three recurring elements of the examples above. First, given a set of alternatives, there is no rule aggregating individual preferences over those alternatives[6] that can reflect a preference ordering if it doesn't appear at least once in the population. Second, a preference profile that is observed at least once is unquestionably relevant for the identification of a collective preference because it exists in the population but, under some conditions, available information on its frequency may be irrelevant and misleading. Finally, if both the first and the second point hold as relevant for the case of interest, some important collective choice rules like majority voting or the plurality rule are certainly not appropriate. If that is the case, each pair of alternatives might still be subject to a meaningful collective ordering or might instead be incomparable following the conventional Pareto principle. In other words, the social preference could be incomplete[7].

Social choice theory has discussed a number of collective choice rules in which the social preference can be incomplete while satisfying some form of the Pareto principle. At the same time, a number of extension theorems have been studied and developed in the context of choice

---

[6] We assume the standard Arrovian setting in which individual preferences are represented by linear order relations over the set of feasible alternatives.

[7] See pg. 13.

theory[8] to represent the relation between incomplete and complete orders sharing a portion of the information embedded in each, including the well-known Pareto extension rule (Sen 2017).

The goal of this chapter is to propose a collective choice rule that extends the quasi-ordering defined by the Pareto principle to a complete order exploiting some interesting tools of order theory. Such a rule has peculiar properties that descend from working with the linear extensions of a partially ordered set related with the Pareto quasi-ordering and, furthermore, from treating them in a way that is interpretable as identifying and dealing with conflicts that generate mutually exclusive but reasonable and complete collective preferences. I define the property of a collective choice rule of this kind *balanced conflict mitigation*.

The chapter assumes that cases of partial information as those mentioned above should be taken as compelling motivation for exploiting observed individual preferences, but not the frequencies of observed preferences, to build a collective choice rule. Collective preferences are then determined as the expected position of each alternative in a linear order as if such order was randomly selected among those that are compatible with a partially ordered set that directly derives from the Pareto quasi-ordering. If the set of alternatives is finite, the probability of drawing a linear order in which a pair a, b of alternatives is ordered so that a is preferred to b is not independent of the preferences on any other pair. When that is not the case, the probability of finding a preferred to b is going to be correlated with preferences over other pairs, a result that is discussed in detail in a specific area of research on order theory

In fact, the study of the connections between correlation and order in Order Theory was a major area of research in the 1980s and the 1990s. The field was strongly indebted with the FKG theorem of statistical physics (Fortuin et al. 1971) which, in a very simple form, states that increasing events are positively correlated. Extensive reviews of these results in what was identified as the theme of *correlation inequalities* can be found in Fishburn (1992), in Trotter (1995) and, succinctly but with specific reference to the area of interest of this chapter, in Brightwell and Trotter (2002).

In 1980, Rival and Sands formulated the following conjecture (XYZ conjecture) which is strictly related to our problem: given n random variables $x_1, x_2, \dots, x_n$, with independent and uniform

---

[8] For a brief overview, see Duggan (1999)

distribution, and given partial information on the "true ordering" of such variables $\Gamma = \{x_2 < x_3, x_9 < x_7 \dots\}$, it seems plausible that, for any Γ,

$$P(x_1 < x_2 | \Gamma) \leq P(x_1 < x_2 | \Gamma, x_1 < x_3)$$

The conjecture was later proved by Shepp (1982) using the FKG theorem and, in its strict form, by Fishburn (1984) using a generalization of the same theorem. Remarkably, the motivating example by Shepp in his otherwise strictly theoretical paper is concerned with the ordering of incomes, which is suggestive of potential applications in terms of welfare and Social Choice. In the same year of Shepp's paper, a work by Winkler (1982) discussed a number of properties deriving from the XYZ conjecture, among which the formal definition of the concept of *average height of an element* x *in a partially ordered set* as the expected number of elements below x in a random linear extension of the partially ordered set. The results of Shepp and Winkler provide this chapter with the basic tool to define a collective ordering from non-unanimous individual orderings.

It is interesting to remark that order theory is almost ubiquitous in economics and partially ordered sets have a notable role in Social Choice Theory[9] but the link with correlation is only sparingly considered, if at all, in economic theory or in Social Choice Theory. However, in recent empirical literature, the interest for correlation inequalities and in particular for the concept of average height (or average rank) has apparently resurfaced in a somewhat disorderly fashion, first in literature concerned with ecological and environmental statistics as in Lerche and Sørensen (2003), in Patil and Taillie (2004) and in Bruggemann and Patil (2010). Mostly at a later stage, correlation inequalities have also occasionally been used in works that look at decision theory (Al-Sharrah 2010), at economic and social statistics (Caperna and Boccuzzo 2018; di Bella et al. 2018; Fattore 2016) and at economics (Badinger and Reuter 2015, 2017; Cavalletti and Corsi 2018). In most cases, the emphasis is on applications of average rank in a partially ordered set and the properties of this function are not discussed in the context of economic theory. A few works

---

[9] Although perhaps less so than quasi-orderings and weak orders as a consequence of considering ties as a desirable feature to model preference relations

dealing with Social Choice Theory and Voting Theory have discussed choice methods that exploit linear extensions of partially ordered sets to aggregate preferences. In Ackerman et al. (2013), individual preferences are expressed with respect to different characteristics of each alternative, with motivations that remind of Lancaster's utility model (Lancaster 1966). In that case, individual preferences are partially ordered and, among the proposed methods to aggregate them, one is to apply any voting method to the set of all linear extensions extracted from individual partially ordered preferences as if they were all ballots. This approach is related with Winkler's average height, as the authors explain, but makes full use of frequency information and assume that incompleteness descends from individual preferences that violate standard rationality assumptions rather than from non-unanimity and the Pareto principle. Furthermore, some works dealing with automated decisions and computational social choice have discussed approaches to preference aggregations in the presence of incomplete information and possibly based on applying some voting procedure on the set of all complete extensions of an incomplete collective preference (Konczak and Lang 2005; Pini et al. 2007, 2008) but with goals that are different from those underlying this chapter.

Based on this brief overview, a discussion on the properties of Winkler's average height could provide economic theory with an additional tool to deal with specific aggregation problems. Approaching it from the point of view of Social Choice Theory seems a logical strategy as the problem of incomplete preferences (and that of partial information on preferences, in particular) is a known issue that the most commonly considered collective choice rules seem ill-designed to deal with. Previous works that use linear extensions in the context of choice, as well as the structure itself of Winkler's average height in a partially ordered set, suggest the opportunity of using the notation and classification conventions of positional choice rules (Pattanaik 2002).

The remainder of the chapter is organized as follows. Section 2 presents the notation and some preliminary definitions. Section 3 translates Winkler's average height into the appropriate notation for a positional choice rule. Section 4 presents the axioms that are relevant for a collective choice rule based on Winkler's average height and, afterwards, it provides the characterization of such collective choice rule as the one that extends the strong Pareto quasi-ordering and satisfies the axiom of balanced conflict mitigation; finally, it presents an example comparing the proposed collective choice rule with other standard rules of preference aggregation.

2. Notation and definitions

2.1 Preliminaries

For notation definitions and, consequently, for much of the content of this paragraph, I closely follow Pattanaik (2002) on positional choice rules.

Therefore, N is the set of all non-negative integers, $\boldsymbol{N}$ is the class of all non-empty subsets of N and the elements $S, S', \ldots \in \boldsymbol{N}$ are called *societies*. Each element in a society is an individual i. The universal set of alternatives is indicated with X and $\boldsymbol{X}$ is the class of all non-empty subsets of X while the elements $A, A', \ldots \in \boldsymbol{X}$ are called issues. For any $A \in \boldsymbol{X}$, $\boldsymbol{V}(A), \boldsymbol{W}(A), \boldsymbol{R}(A), \boldsymbol{L}(A)$ indicate, respectively, the set of all reflexive binary relations over A, the set of all reflexive and complete binary relations over A, the set of all orderings over A and the set of all linear orderings over A. For all $A \in \boldsymbol{X}$, every $R(A) \in \boldsymbol{V}(A)$ is interpreted as a weak preference relation over A, such that $xR(A)y$ indicates that x is weakly preferred to y in the weak ordering of A. Following the typical conventions, the asymmetric part of such relations is indicated as $P(A)$ and the symmetric part as $I(A)$.

When $A = X$, the notation $\boldsymbol{V}, \boldsymbol{W}, \boldsymbol{R}, \boldsymbol{L}, R, P, I$ replaces $\boldsymbol{V}(A), \boldsymbol{W}(A), \boldsymbol{R}(A), \boldsymbol{L}(A), R(A), P(A), I(A)$.

For the sake of convenience, order relations in a general sense and incomplete order relations will both be indicated with R and strictly incomplete order relations will be signaled by specific notation accents. In that case, $\|(A)$ and $\|$ will stand for the incomparable part of the relation respectively over A and over the universal set of alternatives.

Conventionally, I assume rational individual preferences to be represented by complete orders over the universal set of alternatives. For any $S \in N$ and all non-empty subsets $\boldsymbol{W'}$ of $\boldsymbol{W}$, I indicate with $\boldsymbol{W'^S}$ the #S-fold Cartesian product of $\boldsymbol{W'}$. Each element $R_S, R'_S, \ldots$ of $\boldsymbol{W^S}$ stands for an #S-tuple of individual preference profiles $R_1, R_2, \ldots, R_i, \ldots, R_n$, one for each of the n individuals that are part of society S. I indicate each R_S with the generic term *preference profile* to clarify that I am considering the preferences of multiple individuals. With $D(\boldsymbol{W'})$ I concisely indicate the set of all (S, A, R_s) for which $S \in \boldsymbol{N}, A \in \boldsymbol{X},\ R_S \in \boldsymbol{W'^S}$.

Before moving forward, some basic definitions are required to provide a formal introduction and some taxonomy to the problem of collective choice that will be discussed in the remainder of the chapter.

Definition 2.1 (Sen 2017, Definition 2*1) A Collective Choice Rule (CCR) is a functional relation f such that for any $R_S \in D\left(\boldsymbol{W'^S}\right)$ identifies one and only one social preference $R = f(R_1, R_2, \dots, R_i, \dots, R_n)$.

Definition 2.2 (Pattanaik 2002, Definitions 2.1.1 and 2.1.2) A Social Ranking Rule (SRR) is a function f which, for every $(S, A, R_s) \in D(\boldsymbol{W'})$ identifies one and only one $R(A) \in \boldsymbol{V}(A)$, with $\boldsymbol{W'}$ being some non-empty and balanced subset of $\boldsymbol{W}$. We indicate this with $R(A) = f(S, A, R_s)$. A Social Ranking Rule is called Social Ordering Rule (SOR) if and only if for every $(S, A, R_s) \in D(\boldsymbol{W'})$, $f(S, A, R_s)$ is an ordering over A.

3 Average height in a partially ordered set

The average height function described in Winkler (1982) and mentioned, for instance, in Shepp (1982), Brightwell and Winkler (1991) and Brightwell and Trotter (2002) has two noticeable features: it is based on a notion of height that is directly related to that of rank in positional rules and, furthermore, its domain consists of a partially ordered set. Both points deserve a brief discussion.

I first recall a standard definition of rank (or position) of an alternative $x \in A$ in a preference ordering R as

$$r(x, A, R) = \#A - \#\{j \in A: xR(A)j\} \tag{1}$$

where $\#A$ indicates the cardinality of A. This definition is canonically interpreted as referred to complete orders only. Outside of those, the definition is ambiguous and may only be applicable, for instance, to the very restricted class of partially ordered sets called *graded posets* (Stanley 2011). By the end of this chapter I will make use of the average height function to effectively generalize the notion of rank to any partially ordered set and, specifically, to partial orders that directly descend from Pareto quasi-orderings.

The second point to be discussed revolves around the fact that the domain of the average height function is not uniquely identified by a preference profile. Specifically, it is easy to see that, for all $(S, A, R_s) \in D(\boldsymbol{W'})$, any $R(A) = f(S, A, R_s)$ is an acceptable domain for the average height function as long as it is reflexive, transitive and anti-symmetric. Consequently, from definitions 2.1 and 2.2, it is clear that a straightforward application of Winkler's function cannot be a CCR or a SRR. For the moment, I will assume that $\dot{\boldsymbol{R}}(A)$ indicates the set of all reflexive, transitive and anti-symmetric binary relation over A and that $\dot{R}(A) \in \dot{\boldsymbol{R}}(A)$ is the partially ordered social preference over A that constitutes the domain of the average height function. In the following section, I will discuss $\dot{R}(A)$ in greater detail.

Let $\dot{\boldsymbol{L}}(A)$ denote the set of all linear extensions $\dot{L}_1, \dot{L}_2, \dots, \dot{L}_{\#\dot{\boldsymbol{L}}(A)}$ of $\dot{R}(A)$ over A.

Definition 3.1 (Height of an alternative). *For all* $(S, A, R_s) \in D(\boldsymbol{W'}), \dot{R}(A) \in \dot{\boldsymbol{R}}(A)$ *and* $x \in A$*, the height* H_x *of* x *is defined as the random variable of value* $f(x)$ *such that* f *is a random linear extension* $\dot{L}(A) \in \dot{\boldsymbol{L}}(A)$*, and* $f(x)$ is *given by* $\#\{j \in A: x\dot{L}(A)j\} + 1$.

The relation between H_x and the previously mentioned definition of rank is trivial. Remarkably, the distribution of H_x is unimodal (Stanley 1981).

Definition 3.2 (Average height of an alternative). *For all* $(S, A, R_s) \in D(\boldsymbol{W'}), \dot{R}(A) \in \dot{\boldsymbol{R}}(A)$ *and* $x \in A$*, the average height* h_x *of* x *is defined as*

$$h_x = \frac{\sum_{i=1}^{\#\dot{\boldsymbol{L}}(A)} \left(\#\{j \in A: x\dot{L}_i(A)j\} + 1\right)}{\#\dot{\boldsymbol{L}}(A)} \qquad (2)$$

A second formulation of (2) is possible, in which the probabilistic intuition behind this measure is made more explicit, as h_x is essentially the expected value $E(H_x)$ of the height of x in a randomly selected linear extension. Let $\dot{L}_{H_x=l} = \{\dot{L}(A) \in \dot{\boldsymbol{L}}(A): H_x = l\}$. Then

$$h_x = E(H_x) = \sum_{l=1}^{n} l \cdot Pr(H_x = l) = \sum_{l=1}^{n} l \cdot \frac{\#\dot{L}_{H_x=l}}{\#\dot{\boldsymbol{L}}(A)} \tag{3}$$

In (3), the probability of finding x at a given H_x in a randomly selected linear extension of $\dot{R}(A)$ corresponds to the ratio between the number of linear extensions in which x holds that position and the number of all linear extensions.

Finally, the following interesting property of h_x has been demonstrated by De Loof et al. (2011)

$$h_x = 1 + \sum_{j=1}^{n} Pr(H_x > H_j) = 1 + \sum_{j=1}^{n} \frac{\#\{\dot{L}(A) \in \dot{\boldsymbol{L}}(A): H_x > H_j\}}{\#\dot{\boldsymbol{L}}(A)} \tag{4}$$

In (4), the average height of x is obtained as 1 plus the ratio of linear extensions in which x is above each other element j in the set. It should be noted that the problem of counting $\#\dot{\boldsymbol{L}}(A)$ has been demonstrated to be #P-complete (Brightwell and Winkler 1991). However, a number of approximations, for $\#\dot{\boldsymbol{L}}(A)$ in general and for h_x specifically, are available in literature (Brüggemann et al. 2004; Brüggemann and Annoni 2014; Bubley and Dyer 1999; De Loof et al. 2011)

4. Characterization of a Pareto-extension rule based on h_x

4.1 The strong Pareto rule and the Pareto extension rules

From this point, I consider some properties that may be desirable in an aggregation procedure that goes from several individual preferences to one collective preference. The purpose of this work is better served by making a distinction between two groups of axioms. The first group is known to provide a social quasi-ordering of alternatives that has some opportune qualities but, among a number of limitations, it is typically incomplete. What follows is required to extend the previous social quasi-ordering into an ordering satisfying some additional properties of interest, at the price of losing some that were satisfied by the quasi-ordering.

The first group includes the following axioms, all of which are usually contemplated in the debate about the basic requirements of a CCR.

Unrestricted Domain. A CCR f satisfies Unrestricted Domain if and only if the domain of the aggregation function is the set of all possible preference profiles $\boldsymbol{W^S}$.

Independence of Irrelevant Alternatives. For all $(S, A, R_s) \in D(\boldsymbol{W'})$, a collective choice rule $f(S, A, R_s)$ satisfies Independence of Irrelevant Alternatives if and only if $[R_S/A = R'_S/A] \rightarrow [f(S, A, R_s) = f(S, A, R'_S)]$.

Strong Pareto Principle. For all $(S, A, R_s) \in D(\boldsymbol{W'})$, a collective choice rule $f(S, A, R_s)$ satisfies the strong Pareto Principle if and only if

a) $xR_iy, \forall i \in S \rightarrow xRy$ and
b) $[xR_iy, \forall i \in S \text{ } and \text{ } \exists i \in S{:}\, xP_iy] \rightarrow xPy$

Anonymity. For every permutation σ on $\{1, 2, \dots, n\}$ and every $R_S \in \boldsymbol{W^S}$, $f(S, A, R_s) = f(S, A, R_{\sigma(s)})$.

It is well known that these four axioms characterize the strong Pareto rule.

Definition 4.1 (strong Pareto rule)

The strong Pareto rule is a CCR such that, for all $(S, A, R_s) \in D(\boldsymbol{W'})$

$$\forall x, y \in A \; xRy \leftrightarrow xR_iy, \forall i \in S$$

Theorem 1 (Weymark 1984): if a CCR f satisfies Unrestricted Domain, Independence of Irrelevant Alternatives, the strong Pareto Principle and Anonymity, then f is the strong Pareto rule.

The quasi-ordering determined by the strong Pareto rule (or strong Pareto quasi-ordering) has been extensively discussed in literature and requires no further comment, except for remarking that its incompleteness represents the most obvious of its limitations as a method of collective choice. However, in the context of CCRs that seek a departure as limited as possible from the set of conditions in Arrow's famous impossibility theorem (Arrow 1962), weakening the rationality condition of completeness for admissible social rankings is an important case.
The desirable properties of orderings, if compared to the less appealing quality of quasi-orderings of presenting at least a maximal element in the set (Sen 2017) provide some context for the interest of Social Choice Theory for CCRs that are complete or more complete than the Pareto relation and that satisfy either the strong Pareto Principle or the Weak Pareto Principle at the expense of other conditions. A vast spectrum of actual choice problems are ordinarily addressed with CCRs that fall in this category. The following lemma, a generalization of the Szpilrajn extension theorem to quasi-orderings, provided the framework to look at CCRs that result in complete social rankings that are compatible (Sen 2017, Definition 1*6) with the Pareto quasi-ordering as extensions of it.

Arrow's Lemma (1951, pp. 64-68): let $\ddot{R}$ be a quasi-ordering over A, let B be a subset of A such that $\{x, y \in B \colon (x, y) \notin \ \ddot{R}, \ x \neq y\}$ and T be an ordering over B. Then, there exist an extension R of $\ddot{R}$ such that the restriction of R on B coincides with T.

Well-known CCR that extend the Pareto quasi-ordering into an ordering without further manipulation of the Arrovian axioms listed above are the proper Pareto-extension rules. They require the weakening of another condition of social rationality, that of transitive collective preferences, into quasi-transitivity. I report the definition of the Pareto-extension rule and the corresponding characterization theorem.

Definition 4.2 (Pareto-extension rule)

The Pareto-extension rule is a CCR such that, for all $(S, A, R_s) \in D(\boldsymbol{W}')$

$$\forall x, y \in A: xRy \leftrightarrow [xR_i y, \forall i \in S \;\vee \exists i \in S: xP_i y]$$

*Theorem 3*5* (Sen 1970)

A CCR with range R corresponding to the set of complete, reflexive and quasi-transitive relations of A, satisfying unlimited domain, independence of irrelevant alternatives, the strong Pareto Principle and anonymity is the necessary and sufficient condition for being the Pareto-extension rule.

Pareto-extension rules have been discussed in relation with unanimity as a crucial normative foundation for social choice, but their drawbacks were already at the center of the discussion in the same work in which the Pareto-extension rules were first characterized (Sen 1970, Theorem 5*3). In fact, the Pareto-extension rule resolves into arbitrarily imposing social indifference over alternatives on which there is no agreement and are therefore Pareto-incomparable, and this is equivalent to excluding distributional concerns from the formation of a collective preference.

4.2 Axioms and definitions for a Pareto-extension rule based on h_x

It has been argued in many instances across the literature on Social Choice Theory that the effects of combinations of conditions on the various CCR ultimately suggest a number of conflicts among attractive properties (among which the most famous is the Arrovian impossibility itself). This is obviously the case when comparing the strong Pareto rule and the Pareto-extension rule, which leads to question when it might be more acceptable for a CCR to potentially yield incomplete social preferences or, rather, complete social preferences that might be no more than quasi-transitive and are indifferent in case of disagreement among individuals. The point, however, has an even more general relevance. The same kind of conflict between the Pareto-extension rule and the strong Pareto rule is potentially attached to any rule extending the Pareto quasi-ordering and, consequently, to the set of properties emerging from its characterization. In this sense, the operation of extending the Pareto quasi-ordering to a linear order implies the larger problem of all conflicting linear extensions that are compatible with the Pareto quasi-ordering.

From the standpoint of collective choice rules, this may be an interesting issue when the available information is indeed in the form of a quasi-ordering for the lack of agreement and the lack of information required to satisfyingly resolve disagreements, but collective choice must nonetheless have the form of a complete order. In such a case, choice may take a form that could be defined as *conflict mitigation*. In a conflict-mitigating process, individual preferences are considered, an incomplete collective preference is identified, conflicts between the different possible complete orders become apparent and some prudential strategy is identified to resolve them.

The following axioms provide a definition of the conflict mitigation property and other properties that may be attractive in such circumstances. The formal structure of the axioms is directly derived from Fattore (2017) where it is cast in the framework of the axiomatic treatment of functionals on POSETS. The general point is that a desirable property of functionals on POSETS is to behave consistently on every order relation specified on a given ground set as the relation changes, starting from the linear orders over the ground set where the application of the functional is trivial. A consistent functional can then be defined as one such that the values it assumes over any subset of all the linear orders over the ground set and the value on the POSET obtained as an intersection of the linear orders in such subset are in a functional relation.

Conflict mitigation. For all $(S, A, R_s) \in D(\boldsymbol{W'})$ and all $x, y \in A$, a social ordering rule $f(S, A, R_s) = R(A)$ satisfies conflict mitigation with respect to a relevant reflexive, transitive, anti-symmetric sub-relation $\dot{R}(A)$ if and only if

$$xR(A)y \leftrightarrow r(x, \dot{R}(A)) \leq r(y, \dot{R}(A)) \tag{5}$$

and

$$r\left(x, \dot{R}(A)\right) = g\left(r\left(x,\ \dot{L}_1(A)\right), r\left(x,\ \dot{L}_2(A)\right), \dots, r\left(x,\ \dot{L}_{\#\dot{\boldsymbol{L}}(A)}(A)\right)\right) \tag{6}$$

where (5) institutes a relation between the ordering defined by f and specifically one partial order of the alternatives, by the means of the ranking of the alternatives in A according to such subrelation and (6) provides a necessary generalization of the notion of rank which is usually defined only for linear orderings (Pattanaik 2002, p. 368). The ranking functional is consistent if rankings of alternative x in all the linear extensions of $\dot{R}(A)$ are associated to the ranking of x in $\dot{R}(A)$ by means of function g.

Following Fattore (2017), it is possible to make a more stringent version of the axiom by noting that function g has a major role in shaping the conflict-mitigating social ordering rule and it seems therefore reasonable to introduce the desirable form of g within the condition.

Balanced conflict mitigation. For all $(S, A, R_s) \in D(\boldsymbol{W}')$ and all $x, y \in A$, a social ordering rule $f(S, A, R_s) = R(A)$ that satisfies conflict mitigation with respect to a relevant subrelation $\dot{R}(A)$ is balanced if and only if the functional relation g between the ranks of any element $x \in A$ in $\dot{R}(A)$ and the ranks of x in the linear extensions of $\dot{R}(A)$ satisfies the following and well known properties: *continuity* in each of its arguments, *decomposability*, *strict monotonicity*, *homogeneity* in its arguments and *traslativity*.

For the sake of clarity, I provide a formal definition of each of these (anyway rather standard) conditions on g before providing the interpretation of the axiom.

Continuity. A function g of n variables is continuous if and only if, for any point $(p_1, p_2, \dots, p_n)$ in its domain,

$$\lim_{(x_1, x_2, \dots, x_n) \to (p_1, p_2, \dots, p_n)} g(x_1, x_2, \dots, x_n) = g(p_1, p_2, \dots, p_n)$$

Decomposability. A function g is decomposable if and only if, for all $m, n = 1, 2, \ldots$ and for all vectors $\boldsymbol{x} \in [0,1]^m$ and $\boldsymbol{y} \in [0,1]^n$,

$$g_{m+n}(\boldsymbol{x}, \boldsymbol{y}) = g_{m+n}(\underbrace{g_m(\boldsymbol{x_m}), g_m(\boldsymbol{x_m}), \ldots, g_m(\boldsymbol{x_m})}_{m\ times}, \boldsymbol{y})$$

Strict Monotonicity. Let $\boldsymbol{x}, \boldsymbol{y}$ indicate two k-dimensional real vectors; function g is strictly monotone if and only if $\boldsymbol{x} < \boldsymbol{y}$ in the product order[10] over $\mathbb{R}^k$implies $g_k(\boldsymbol{x}) < g_k(\boldsymbol{y})$.

Homogeneity. A function g is homogeneous if and only if, for any real number $c \in [0,1]$ and any k-dimensional real vector $\boldsymbol{x} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_k})$ it is verified that $g(c \cdot \boldsymbol{x}) = c \cdot g(\boldsymbol{x})$.

Traslativity. A function g is traslative if and only if for any real number c and any pair of k-dimensional real vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_k), \boldsymbol{c} = (c, c \ldots, c)$ it is verified that $g(\boldsymbol{x} + \boldsymbol{c}) = g(\boldsymbol{x}) + c$.

In a balanced conflict-mitigating social ordering rule (BCMR), the ranking over the partial order is a function of the rankings over its linear extensions; sufficiently small changes[11] in the rankings of an alternative in the linear extensions produce arbitrarily small changes in the ranking of that alternative over the partial order and strictly better rankings of an alternative across all linear extensions imply a strictly better ranking of that alternative in the partial order. All these properties have either been discussed above or provide rather self-explaining advantages to a collective choice procedure. The property of decomposability states that the social ordering rule produces the same results if the function over the linear extensions is applied separately over partitions of the set of all the linear extensions and, again, on the results. As a consequence, conflict mitigation can be interpreted as a choice process in which it is possible to transparently assess the contribution of any group of potential outcomes to the overall outcome. The final two

---

[10] For *product order*, consider the following conventional definition: let A_i be a set partially ordered by order relation $\geq_i$, with $i = 1, \ldots, n$. If $x, y \in \prod_{i=1}^{n} A_i$, then $x \geq y$ should be interpreted as $x_i \geq_i y_i$ for $i = 1, \ldots, n$ and $x > y$ that $x > y$ and $x \neq y$. The order $\geq$ on $\prod_{i=1}^{n} A_i$ is called product order.

[11] That is, ignoring the inevitable granularity of rank values, which is fixed, whereas the set of all linear extensions of a partially ordered set can be arbitrarily large, much like the ground set involved.

properties provide a stricter regulation of change in the ranking of an alternative over the partial order as a consequence of changes in its rankings in the linear extensions. The property of traslativity ensures that, in a BCMR, a unit increase (or decrease) in the ranking of an alternative in all the compatible linear orders leads to a corresponding unit change in the ranking over the partial order. Changes that amount to strictly more than a unit per linear extension, therefore, produce changes of more than a unit in the partial order. Equivalently, by homogeneity, for an alternative to get a halved standing in all linear extensions exactly implies having a halved standing in the partial order. Both properties can be interpreted as technical conditions but they may nonetheless be desirable in many circumstances, particularly when the steps of conflict mitigation (identifying a partial order of preferences and then extending it) are separated by a non-trivial amount of time and/or additional information becomes available. For instance, the property of traslativity shifts a balanced conflict-mitigating ordering by c after the introduction (or the removal) of c strictly better or strictly worse alternatives. Therefore, partially ordered preferences over two disjoint sets of alternatives of this kind can be ordinally summed[12] and the ranking of an alternative in the resulting poset will be equal to its ranking in the poset of origin plus or minus the cardinality of the other poset. The property of homogeneity, instead, helps comparing a conflict-mitigating allocation of a budget among projects that are partially ordered and the various allocations implied by the compatible linear orders.

I close this section by providing the opportune definitions for a Pareto-extension rule based on correlation inequalities. Let $\ddot{R}(A)$ indicate the strong Pareto quasi-ordering over alternatives in issue A and let $\ddot{I}(A)$ indicate the reflexive, transitive, symmetric sub-relation of $\ddot{R}$ over A. Following a conventional notation, I indicate with $A/\ddot{I}$ the quotient set of $\ddot{I}(A)$ or, in other terms, the set of all equivalence classes $\{[x], [y], \dots\}$ of A with respect to $\ddot{I}(A)$, with x, y ... indicating any arbitrarily selected element within each equivalence class so that, for instance, $[x] = \{a \in A: a\ddot{I}(A)x\}$. It is immediate to see that, if all elements of $\ddot{R}(A)$ are replaced with their respective equivalence class, the quasi-ordering is turned into a partially ordered set (poset).

---

[12] See Stanley (2011, p. 283).

Definition 4.3 (strong Pareto poset)

The strong Pareto poset $\dot{R}(A/\ddot{I})$ is a reflexive, transitive, anti-symmetric binary relation over the quotient set of $\ddot{I}(A)$ such that, for all $x, y \in A$ and all $[x], [y] \in A/\ddot{I}$

$$[x]\dot{R}\big(A/\ddot{I}\big)[y] \leftrightarrow x\ddot{R}(A)y \tag{7}$$

Given definitions 3.2 and 4.3 we can easily derive $h_{[x]}$, the average height of the equivalence class of any $x \in A$ in the strong Pareto poset and, from it, the following definition

Definition 4.4 (Balanced conflict-mitigating Pareto-extension rule)

A CCR is the Shepp-Winkler Pareto-extension rule if and only if, for all $(S, A, R_s) \in D(\boldsymbol{W}')$ and for any $x, y \in A$

$$xRy \leftrightarrow h_{[x]} \geq h_{[y]} \tag{8}$$

Equivalently, and with the intent of making the structure of the rule more evident, let f, g, h indicate, respectively, the strong Pareto rule, the binary relation in (7) and the average height function in (2). A Shepp-Winkler Pareto-extension rule s is defined as the composite function

$$s = \circ \xrightarrow{f} \circ \xrightarrow{g} \circ \xrightarrow{h} \circ \tag{9}$$

From (9) it becomes clearer that s maintains a strict relationship with the strong Pareto rule. In fact, while independence of irrelevant alternatives for s is subject to the considerations in Pattanaik (2002 sec. 4.1) on that condition for positionalist rules and is only satisfied in the formulation therein proposed, the other conditions in the strong Pareto rule are clearly not affected by either g or h and therefore are satisfied by s.

4.3 Axiomatic characterization of the BCMR

I am now organized to state and prove the following theorem

Theorem 4.2: if a CCR s is an ordered extension of the strong Pareto poset and satisfies balanced conflict mitigation, then s is the balanced conflict-mitigating Pareto-extension rule.

A CCR that is at the same time an extension of the strong Pareto poset and is conflict-mitigating is one such that, in (5), $\dot{R}(A)$ is replaced with $\dot{R}(A/\ddot{I})$ and each alternative with its equivalence class. If it is balanced, then in (6) function g must be the arithmetic mean. Specifically, if g is continuous, decomposable, strictly monotone, and homogeneous, then g must be a power mean of power p in direct application of the Kolmogorov-Nagumo-De Finetti theorem (Beliakov et al. 2007). If such power mean has the traslative property, then it is immediate to see that it must be of power $p = 1$. As a consequence, we can rewrite (6) as the following relation:

$$r\left([x], \dot{R}\left(A/\ddot{I}\right)\right) = \frac{\sum_{i=1}^{\#\dot{\boldsymbol{L}}(A/\ddot{I})} r\left([x],\ \dot{L}_i(A)\right)}{\#\dot{\boldsymbol{L}}(A/\ddot{I})} \tag{10}$$

Substituting in (10) the definition of rank provided at the beginning of section 3 and making the opportune adjustments, it is easily seen that the condition above is equivalent to this:

$$\#\{[j] \in A\colon [x]\dot{R}\left(A/\ddot{I}\right)[j]\} = \frac{\sum_{i=1}^{\#\dot{\boldsymbol{L}}(A/\ddot{I})} \#\{[j] \in A\colon [x]\ \dot{L}_i\left(A/\ddot{I}\right)[j]\}}{\#\dot{\boldsymbol{L}}(A/\ddot{I})} \tag{11}$$

In turn, by adding 1 on both sides of (11), from definition 3.1 we get that the condition I derived from having a conflict-mitigating CCR based on the strong Pareto poset is equivalent to the definition of average height of an alternative in (2). Following from there, the definition of conflict-mitigating choice rule in (5) is identical to the definition of the BCMR in (8).

5. Discussion

The set of all linear extensions of the strong Pareto poset is a remarkable object that deserves attention and a proper place in social choice theory. A class of specific but not so uncommon decision-making processes appears to work precisely on that set, which can still be built with a rather tentative knowledge of individual preferences in society. Each linear extension represents a collective choice that preserves observed unanimous preferences and is obtained without acknowledging special privileges to specific individuals or without restricting the domain of acceptable individual preference orderings. As soon as a non-negative "weight" is assigned to each individual preference ordering, expressing its frequency in the population or the intensity of the preference or any value judgement that is normatively accepted, one linear extension of the strong Pareto poset is unequivocally identifiable as the collective preference of the population of interest over that set of equivalence classes of alternatives. As long as that weight is not assigned and no assumptions are made about it, all the linear extensions can only represent possible outcomes of the decision-making process reflecting conflicting value systems and interests. If conflict is not going to be resolved by assigning a weight, then it can be mitigated by acknowledging that there is no ground to suggest that one linear extension is more likely than the others. Under these conditions, the expected position of each element in a randomly selected linear extension is a cautious, balanced outcome that mitigates the conflict between different value systems in a way that satisfies a list of desirable properties. While developed with different intents, the theory on correlation inequalities and Winkler's average height functional relation provide a method to extend incomplete collective preferences in a way that is consistent with the objectives of social choice theory, at least under some specific conditions if not as a general approach.

References

Ackerman, M., S. Choi, P. Coughlin, E. Gottlieb, S. P. Choice, N. October, M. Ackerman, S. C. Peter, E. Gottlieb, & J. Wood. 2013. Elections with partially ordered preferences. Public Choice, 157(1/2):145–168.

Al-Sharrah, G. 2010. Ranking using the copeland score: A comparison with the hasse diagram. Journal of Chemical Information and Modeling, 50(5):785–791.

Arrow, K. J. 1962. Social Choice and Individual Values. New York, New York, USA: John Wiley & Sons.

Badinger, H., & W. H. Reuter. 2015. Measurement of fiscal rules: Introducing the application of partially ordered set (POSET) theory. Journal of Macroeconomics, 43(December 2014):108–123.

Badinger, H., & W. H. Reuter. 2017. The case for fiscal rules. Economic Modelling, 60(September 2016):334–343.

Beliakov, G., Pradera, A., & T. Calvo. 2007. Aggregation Functions: A Guide for Practitioners. Soft Computing (Vol. 18).

Brightwell, G. R., & W. T. Trotter. 2002. A combinatorial approach to correlation inequalities. Discrete Mathematics, 257(2–3):311–327.

Brightwell, G. R., & P. Winkler. 1991. Counting linear extensions. Order, 8(3):225–242.

Brüggemann, R., & P. Annoni. 2014. Average Heights in Partially Ordered Sets. MATCH Communications in Mathematical and in Coputer Chemistry, 71, 117–142.

Bruggemann, R., & G. P. Patil. 2010. Multicriteria prioritization and partial order in environmental and ecological statistics. Environmental and Ecological Statistics, 17:383–410.

Bubley, R., & M. Dyer. 1999. Faster random generation of linear extensions. Discrete Mathematics, 201(1–3), 81–88.

Caperna, G., & G. Boccuzzo. 2018. Use of Poset Theory with Big Datasets: A New Proposal Applied to the Analysis of Life Satisfaction in Italy. Social Indicators Research, 136(3):1071–1088.

Cavalletti, B., & M. Corsi. 2018. By diversion rate alone: The inconsistency and inequity of waste management evaluation in a single-indicator system. Papers in Regional Science, (January):1–23.

De Loof, K., B. De Baets, & H. De Meyer. 2011. Approximation of Average Ranks in Posets. MATCH - Communications in Mathematical and in Computer Chemistry, 66:219–229.

di Bella, E., L. Gandullia, L. Leporatti, M. Montefiori, & P. Orcamo. 2018. Ranking and Prioritization of Emergency Departments Based on Multi-indicator Systems. Social Indicators Research, 136(3):1089–1107.

Duggan, J. 1999. A General Extension Theorem for Binary Relations. Journal of Economic Theory, 86(1):1–16.

Fattore, M. 2016. Partially Ordered Sets and the Measurement of Multidimensional Ordinal Deprivation. Social Indicators Research, 128(2):835–858.

Fattore, M. 2017. Functionals and Synthetic Indicators Over Finite Posets. In M. Fattore & R. Bruggemann (Eds.), Partial Order Concepts in Applied Sciences (pp. 71–86). Springer International Publishing AG.

Fishburn, P. C. 1984. A correlational inequality for linear extensions of a poset. Order, 1(2):127–137.

Fishburn, P. C. 1992. Correlation in partially ordered sets. Discrete Applied Mathematics, 39:173–191.

Fortuin, C. M., P. W. Kasteleyn, & J. Ginibre. 1971. Correlation inequalities on some partially ordered sets. Communications in Mathematical Physics, 22(2):89–103.

Konczak, K., & J. Lang. 2005. Voting procedures with incomplete preferences. IJCAI-Workshop, 6.

Lancaster, K. 1966. A New Approach to Consumer Theory Author (s): Kelvin J . Lancaster Source : The Journal of Political Economy , Vol . 74 , No . 2 (Apr ., 1966), pp . 132-157 Published by : The University of Chicago Press Stable URL : http://www.jstor.org/stable/182883. Journal of Political Economy, 74(2):132–157.

Lerche, D., & P. B. Sørensen. 2003. Evaluation of the ranking probabilities for partial orders based on random linear extensions. Chemosphere, 53(8):981–992.

Patil, G., & C. Taillie. 2004. Multiple indicators, partially ordered sets, and linear extensions: Multi-criterion ranking and prioritization. Environmental and Ecological Statistics, 11:199–228.

Pattanaik, P. K. 2002. Positional Rules of Collective Decision-Making. In K. J. Arrow, A. Sen, & K. Suzumura (Eds.), Handbook of Social Choice and Welfare volume 1 (pp. 361–394). Amsterdam: Elsevier Science B.V.

Pini, M. S., F. Rossi, K. B. Venable, & T. Walsh. 2007. Incompleteness and incomparability in preference aggregation. IJCAI International Joint Conference on Artificial Intelligence, 175:1464–1469.

Pini, M. S., F. Rossi, K. B. Venable, & T. Walsh. 2008. Aggregating partially ordered preferences: impossibility and possibility results. Journal of Logic and Computation, 19(3):475–502.

Sen, A. 1970. Collective Choice and Social Welfare (1st ed.). San Francisco: Holden-Day.

Sen, A. 2017. Collective choice and social welfare. Milton Keynes, UK: Penguin Random House.

Shepp, L. 1982. The XYZ conjecture and the FKG inequality. The Annals of Probability, 10(3):824–827.

Stanley, R. P. 1981. Two Combinatorial Applications of the Alexsandrov-Fenchel inequalities. J.Combin.Theory (A), 65:56–65.

Stanley, R. P. 2011. Enumerative Combinatorics Volume 1 second edition (Second Edi, Vol. 1). Cambridge, UK: Cambridge University Press.

Trotter, W. T. 1995. Partially ordered sets. (R. Graham, M. Grotschel, & L. Lovasz, Eds.)Handbook of combinatorics. Elsevier Science B.V.

Weymark, J. A. 1984. Arrow ′ s Theorem with Social Quasi-Orderings. Public Choice, 42(3):235–246.

Winkler, P. 1982. Average height in a partially ordered set. Discrete Mathematics, 39:337–341.

CHAPTER II

By diversion rate alone. The inconsistency and inequity of waste management evaluation in a single-indicator system

1. Introduction

A substantial bulk of scientific literature deals with urban waste management systems (UWMS) and the evaluation of their performance (Allesch and Brunner 2014). The high-level parameter of evaluation is consistent across all the literature: it refers to the effectiveness in achieving environmental sustainability along the entire life cycle of the products becoming urban waste. The practical implementation of the evaluation, at least in a broad sense, is also not particularly controversial, and implies the identification of environmental performance indicators, which are generally considered as an important management tool providing decision makers with factual information on implemented strategies and credible forecasts on the consequences of future and even innovative policies (OECD 2013). Sets of indicators provide insight about if and which results are achieved in a specific territory after new strategies of waste collection are implemented and allow to compare contexts in which these processes are different. These sets of indicators are actually considered a standard and fundamental tool for policy evaluation, in particular to check if the targets of sustainability and communication towards citizenship have been reached (OECD 2013; Vergara and Tchobanoglous 2012).

A recent and extensive review of literature on performance indicators for UWMS can be found in a paper by Sanjeevi and Shahabudeen (2015), who also present a brief history of the development of performance indicators and discuss their organizational and management role. A more specialized part of the literature based on multi-criteria decision analysis is reviewed in Achillas et al. (2013)[13]. Basic indicators about waste collection and sorting generally include the amount of solid urban waste which is recycled, landfilled or reused (including composting), measured as an absolute value, as a percentage of total collection, or as a per capita amount and a classification of sorted and, sometimes, unsorted waste, according to their composition and their collection channel. Evaluations with a strong focus on basic performance indicators can be found in Wilson et al. (2012), where they are classified as indicators about public health and collection, environment and disposal, and resource recovery; in Desmond (2006); in Guimarães et al. (2010)

---

[13] See ISO14031 for a guidance on classification of performance indicators; see also Scipioni et al., (2008) for the Italian experience

and in Mendes et al. (2012), who emphasize the relation between goals and indicators in the framework of a balanced scorecard approach; in Massoud et al. (2003) in which focus is on economic indicators. Approaches with a strong environmental, biological and chemical component (which are typically required to evaluate the entire integrated cycle of waste management) have their own literature like, for instance, the papers by Kaufman et al. (2010), Kulczycka et al. (2015), and Zaccariello et al. (2015).

This literature clarifies that, unfortunately, implementing the evaluation of a UWMS on more than a single indicator at a time can quickly become a complex issue, particularly if specific targets have to be met with respect to each indicator. What the different methods propose are, essentially, strategies to build aggregation functions that transform different sets of indicators into a synthetic measure that satisfies a certain number of goals and, as a consequence, the corresponding definition of sustainability in waste management. In doing so, each method has to make strong assumptions, the most frequent being connected either with the identification of a common unit of measurement (economic or otherwise) for the dimensions that have to be aggregated, or with the property of aggregation functions which implicitly justify at least some degree of compensation between the different dimensions being aggregated. In these frameworks, each indicator conveys its own distinctive part of information about the outcomes of waste management operations, in a set that is frequently characterized by elements of complementarity and elements of correlation that each method has to deal with using more or less sophisticated arrangements.

Nonetheless, it is only recently that the criteria for selecting any indicator over others have come under scrutiny. In their paper (2014), Greene and Tonjes measured a battery of indicators of all degrees of complexity for ten municipalities in the state of New York, showing that ordering the municipalities according to the different indicators leads to very different rankings as a result of low or even negative correlation levels among the indicator values. They also stress that, all the noted methods notwithstanding, most UWMS are ultimately evaluated according to the recycling rate even though it may be insufficient, on its own, to assess sustainability of the waste management system.

The inconsistency of rankings based on similar but not identical indicators is a well-known issue in other fields where multi-indicator analyses are common, but in the specific field of UWMS evaluation this is quite a novelty and the complexity and multi-dimensionality of the problem (with a particular reference to the problems of identification, measurement and interpretability of the sustainability requirements) is generally strongly underestimated.

In this work, the issue is studied in the context of the Italian multi-level waste taxation scheme, where the system assessment (and the parameters upon which it is conducted) has direct financial consequences on local governments and, in turn, on local taxation. The Italian national and regional legislation concerning UWMSs, in fact, is gradually designing a scheme of incentives and sanctions based on the performance of each UWMS at a municipal level, and the system is based on a single indicator, diversion percentage.

In the first part of the work, the system is analysed using 2013 waste collection figures of almost seven thousand local governments assembled by the Italian National Institute for Environmental Protection and Research (ISPRA) and the 2014 data on the 32 municipalities in the Province of La Spezia (Liguria Region). We show that assumptions about the equivalence of diversion rate and degree of environmental sustainability of waste production and collection systems, which supposedly justify the use of a single indicator, are largely untenable. Applying analytical tools developed in the context of partial order theory (Schroder 2002) to municipalities in the study area, we show that a ranking of sustainability performances based on diversion rate is not robust to the introduction of additional indicators and that a multi-indicator evaluation based on four relevant metrics of sustainability assigns substantially different rankings.

Afterwards, using the 2010-2014 data on waste collection in the Province of La Spezia, the second part of the paper makes use of a forecast of the 2019 tax burden generated by recently introduced waste taxation scheme. Presenting the forecast, it is shown that, choosing diversion rate as the single performance indicator, the scheme defines a tax function which is unrelated with environmental externalities and is inconsistent with the aims of the legislation. The current legislation is shown to be severely regressive in redistributing the burden between municipalities in the Province, thus potentially distorting fiscal competition and the residents' perception about environmental policies and the performance of local governments.

2. The study area, data and methods

2.1 The Province of La Spezia and the UWMS legislation

The main study area of this work, the Province of La Spezia, located in the easternmost part of Regione Liguria, is a Province of 32 municipalities with a population of about 220,000 and 160,889 taxpayers in year 2014. Compared to other Provinces in Italy, it has a few remarkably average or next-to average characteristics like income (which is close to the average of Northern Italy), size of the largest urban area and a slightly above-average population density. According to the Italian Institute of Statistics, its municipalities represent a wide variety of geographical features (landlocked mountainous, landlocked hilly, coastal hills) and more than half of its municipalities are classified as mountain municipalities.

Waste collection and disposal services are managed at a municipal level. In the years covered by this work, 24 out of 32 municipalities in the Province outsourced the service to a single company, Acam Ambiente S.p.A. on the basis of public service contracts, whereas the remaining 8 were managing the service in-house. Waste collection is mostly operated by waste collection vehicles on dumpsters, some of which are dedicated to sorted waste (paper, metal, plastic, glass and organic). Other types of waste (wood, fabric, waste of electric and electronic equipment and all sorts of bulky waste can be carried, at no charge, to a small number of collection centres or can be picked up by service operators at home for a fee. In just three municipalities across the province, curbside collection is active and covers the entire municipality, whereas in two more municipalities it is active in some areas. Service operators are responsible for managing the early stages of the recycling process for sorted waste. A negligible amount of unsorted waste produced in the province is treated in a single facility operated by Acam Ambiente S.p.A. to obtain refuse-derived fuel. All remaining unsorted waste is transported to landfills, which may be located in the provincial or, in case of necessity, in the regional territory, but extra-regional transportation is also allowed when no landfill is active, at the condition of specific agreements between the involved regional governments, including on the compensation fee.

The Italian legislation on waste management is a combination of national measures, which deal with general principles, and regional measures, which regulate implementation details by Regional Law (see Table 1 for the current legislation in Regione Liguria).

Table II-1. Minimal compendium of current National and Regional legislation on waste management evaluation in Regione Liguria

Law	Level	Effect
Law 549/1995	National	The "polluter pays" principle introduced in the national legislation; parliamentary mandate to government to introduce taxation on landfill disposal
Legislative Decree 22/1997	National	Introduction of the Special Tax on Landfill Disposal and definition of minimum rate
Legislative Decree 152/2006 (and its modifications)	National	Reclassification of the sources of Municipal Solid Waste; definition of minimum levels of diversion rate to be attained to avoid sanctions
Law 147/2013	National	Full cost coverage of waste management service through the new waste tax TARI
Law 221/2015	National	Sanctions for not reaching minimum levels of diversion rate now potentially referred to municipalities; gradual reductions on tax on landfill disposal if diversion targets are met
Ministerial Decree May 26th 2016 (Ministry of Environment)	National	Guidelines for calculating the diversion rate of municipal solid waste
Regional Law 20/2015 (Regione Liguria)	Regional	Introduction of a regional sanction scheme for municipalities not reaching a regional target of diversion rate
Regional Law 16/2016 (Regione Liguria)	Regional	Definition of current rate of Tax on Landfill Disposal for municipalities in Regione Liguria

Sources: Own overview of National (Italian) and Regional legislation (Regione Liguria)

The structure of incentives and sanctions changed several times and its commencement was repeatedly delayed since it was first introduced in the national legislation with Legislative Decree 22/1997, with the latest changes occurring in the first months of 2016. The overall objective has nonetheless remained the same: re-orienting UWMS in the perspective of a drastic reduction of landfill disposal and making municipalities responsible for it. Several policy tools are intended to attain this goal, among which diversion and recycling is by far the one which is given more emphasis but, according to multiple regional legislations, prevention of waste generation is expected to play an important role as well (*Piano regionale di gestione dei rifiuti e delle bonifiche* 2015). By design, the system is also clearly linked with the national regime of waste taxation introduced with Law 147/2013, which requires complete coverage of service costs at municipal level (see Art. 1, Section 654) through the local waste tax "TARI".

At present, Municipalities are responsible for controlling service provision costs and are required to fully cover those costs through TARI. Subsequently, costs are adjusted as a consequence of performance evaluation on two possible grounds.

First, a tax on landfill disposal is directly added to those costs by the national legislation, which also defines a minimum rate, a 20% surtax[14] if the municipality does not comply with a minimum diversion rate of 65% and gradually reduced rates for municipalities beyond the 65% threshold. Regional governments can decree rates above the national minimum, so the current rate for Regione Liguria is at 15 € per Ton of waste sent to landfill.

Second, regional governments are also allowed to add further targets and sanctions, which again are added to service provision costs. Municipalities in Regione Ligura that do not comply with a diversion threshold of 45% are therefore required to pay a further 25 € per Ton of unsorted urban waste exceeding the allowed percentage.

Then, as the overall cost to be covered with TARI is determined, municipalities also have some degree of freedom in redistributing the overall tax burden among their residents. Law 147/2013 indicates the general redistribution criteria, consisting of a specific tax base (real property), different determination methods of specific rates for households and other taxpayers, and some general criteria for adjustment coefficients (number of household members in the case of

---

[14] The surtax mechanism was initially targeted at the "Ambiti Territoriali Ottimali" or ATO (Optimal territorial subdivisions), a service provision-related territorial subdivision that had to be gradually activated by groups of municipalities integrating their systems of service provision. The process was cancelled and law 221/2015 redirected the surtax mechanism to single municipalities wherever the ATO had not been activated. As a consequence, the surtax can work differently in different municipalities and conclusions in this work are designed to not depend on the effect of the surtax.

households and expected waste production potential of different activities in the case of the other taxpayers). Municipalities can then set the exact repartition between households and other taxpayers and the adjustment coefficients, provided that they conform to rational criteria and aim at proportionality between waste production and tax burden. Table 2 provides a synthesis of the current taxation in Regione Liguria affected by UWMS performance evaluation.

Table II-2. Taxation affected by UWMS performance evaluation currently in force in Regione Liguria

Tax	**Tax base**	**Current rate**	**Effect on taxpayer**
TARI	Real property	Variable, up to full coverage of service provision at municipal level	Direct
Tax on Landfill Disposal	Waste sent to landfill	15 € per Ton, 18 € per Ton if below 65% of diversion rate, tax reduced by 30% to 70% if diversion above 65%	Through TARI
Regional sanction (Regione Liguria)	Waste in excess of 55% sent to landifill	25 € per Ton	Through TARI

Sources: Own compendium of National and Regional legislation

2.2 Data: 2010-2014

The main source of data is the Italian National Institute for Environmental Protection and Research (ISPRA)[15], a public body that operates under the vigilance and policy guidance of the Italian Ministry for the Environment and the Protection of Land and Sea. Among its duties, ISPRA creates and manages databases on environmental issues, including the main national database on urban solid waste, the *Catasto dei rifiuti* (urban waste cadastre) which is populated with data provided by the regional divisions of the institute and by municipalities through a specific form for environmental declaration

The reported indicators are detailed in Table 3.

---

[15] See http://www.isprambiente.gov.it

At the time of writing this paper, the database reports data for the years 2010-2014. Data for year 2013 has a good nationwide coverage, as detailed waste management indicators for 2013 cover 4158 municipalities and the main indicators (first 5 rows of Table 3) cover 8046 municipalities. According to ISPRA, missing 2013 data from municipalities that did not provide full information can be estimated to amount to 0.3% of urban waste production and 0.1% of sorted waste (ISPRA 2014). Data for 2014 was fully available for the Province of La Spezia.

Table II-3. Italian National Institute for Environmental Protection and Research, waste indicators for Italian municipalities

Indicator	Notes
Urban waste (total)	Tonnes
Sorted waste (total)	Tonnes
Percentage of sorted waste	Percent of total urban waste
Urban waste per capita	(Kg./Resid. Year)
Sorted waste per capita	(Kg./Resid. Year)
Mixed bulky waste	Tonnes
Paper and cardboard	Tonnes
Textiles	Tonnes
Wood	Tonnes
Glass	Tonnes
Metals	Tonnes
Plastics	Tonnes
Electrical and electronic waste (WEEE)	Tonnes
Hazardous and biomedical	Tonnes
Biodegradable waste	Tonnes
Other sorted waste	Tonnes

Source: ISPRA (www.isprambiente.gov.it)

In order to ensure homogeneous data, since 1997 (legislative decree 22/1997) the Italian national legislation prepared the ground for introducing a uniform methodology to calculate diversion rate and other statistics on diversion through a future governmental decree. However, the decree was only issued in 2016 (ministerial decree 78/2016) and, as a consequence, at the time of writing this paper no available data was necessarily consistent with the methodology defined in it. Between 1997 and 2016, all regional governments across Italy had to issue provisional norms providing for regionally homogeneous methods to calculate indicators on waste management, which were then materially applied by local government authorities and companies supplying waste collection and disposal services. Notably, in the study area of this paper, 24 out of 32 municipalities have the same service provider (ACAM Ambiente S.p.A), a degree of concentration that is similar to that found across Italy. On these grounds (same rules and same data collectors), we expect intra-provincial comparisons of municipal data to be sufficiently reliable. Furthermore, among its institutional duties, ISPRA is also responsible for cross-checking data from different sources, like for instance municipalities, collection service providers and disposal/landfill service providers and for conducting punctual inspections in case of discrepancies.

Data on resident population at a municipal level for years 2010-2014 was retrieved from the website of the National Institute of Statistics (ISTAT)[16]. Additional data on income and income classes in fiscal year 2013 at a municipal level was retrieved from the data archive of the Ministry of Economy and Finance[17]. Finally, data on waste tax revenue in fiscal year 2013 was retrieved from municipal budget reports of each municipality in the Province of La Spezia.

The 2013 nationwide data from ISPRA provides good initial evidence on the issue at stake. A first remark should be that, even though lagging, Italian authorities have at their disposal a multi-indicator system about environmental performances of UWMS meaning that, either directly or through some basic transformation, several of the available indicators do measure different relevant dimensions of environmental performance. While there are obvious overlaps between some indicators, particularly when they detail a specific kind of waste, many still provide relevant information as they hint at the level of sophistication of the sorted waste collection chain, at different environmental awareness levels and, most importantly, at very different kinds of environmental harm.

Given the relative wealth of information available, it seems inevitable to investigate the possible drawbacks of an evaluation system based on diversion percentage alone. In choosing it as the discriminant to define sanctions and incentives, the legislator assumed that diversion percentage

---

[16] http://demo.istat.it

[17] http://www1.finanze.gov.it/finanze2/analisi_stat/v_4_0_0/contenuti/Redditi_e_principali_variabili_IRPEF_su_base_comunale_CSV_2014.zip

would be a very good proxy for the overall level of environmental sustainability attained by a municipal UWMS. According to this implicit assumption, a high diversion percentage should accurately predict most (if not all) of the different dimensions of environmental sustainability in waste production and collection, including those concerning waste production per capita.

Contrary to expectations, though, the overall correlation between these two measures at a national level is negligible (R=0.11 in the last year available). Municipalities that in 2013 were already compliant with the 65% diversion rate goal, were no more accomplished than all others in containing the amount of waste produced per capita. Furthermore, correlation between diversion rate and unsorted waste (destined to landfill) per capita is evident but far from perfect (R=-0.73), with plenty of cases where a low diversion percentage is combined with very small amounts of waste sent to landfill. Based on the relations above, the assumed comprehensiveness of diversion percentage as an indicator cannot be confirmed. In order to make an idea of the magnitude of the issue, it can be helpful to consider that, in 2013, more than 2 million Italians lived in municipalities that, in spite of diverting less than 50% of their waste, were already producing and sending to landfill less than 159 Kg/year per capita, an amount corresponding to a very ambitious target set by Piedmont Region for 2020 and well below the target set for the same year by Liguria Region (about 182 Kg/year per capita).

In the study area, the magnitude of this effect is even larger: the two correlation coefficients in 2014, shown in Table 4, were R=0.04 between diversion percentage and urban waste per capita and R=-0.49 between diversion percentage and unsorted waste per capita.

Table II-4. Correlation coefficients of Diversion %

Performance indicator	Diversion %
Waste per capita	+0.04
Unsorted per capita	-0.49*
Sorted per capita	+0.82*
Sorted paper per capita	+0.67*
Sorted plastic per capita	+0.44*
Sorted glass per capita	+0.31
Sorted organic per capita	+0.85*
Sorted WEEE per capita	+0.53*

Notes: * p<0.05

Sources: Own calculation on data from ISPRA

Diversion percentage is a really strong correlate (at least R=±0.7) of only two of all collection statistics available (Table 4), diversion per capita and organic waste per capita. Furthermore, given the unimpressive correlations among collection indicators, it is very unlikely that environmental sustainability dimensions concerning anything else than waste collection would be well represented by diversion percentage, even though there is no available data to verify that. This preliminary data overview suggests that diversion percentage is not an adequate proxy for most of the other indicators. It is also specifically weak in representing some dimensions, like waste production per capita, that are certainly relevant to the aims of the waste disposal legislation.

In the remainder of this work, in order to compare evaluations based on diversion percentage with more multidimensional evaluations, three indicators taken from Table 4 will be used extensively: unsorted waste per capita, sorted organic waste per capita and sorted waste of electric and electronic equipment (WEEE) per capita. Unsorted waste per capita is a particularly relevant indicator as one of the goals of waste legislation is to reduce the amount of waste sent to landfill and, in the absence of functioning materials recovery facilities, unsorted waste and waste sent to landfill are very similar amounts. Organic waste per capita is relevant for three different reasons: in terms of weight, it is the single most important category of sorted waste were organic waste collection is working; it is also a potentially critical kind of waste in terms of environmental hazard; furthermore, the management strategy is typically distinct from that of all other sorted materials. Finally, WEEE sorting is relevant both because of the environmental hazard and because sorting is only sparingly active across the Province of La Spezia and can possibly be considered as an indicator of above-average effort.

2.3 Data projections to 2019

Unfortunately, the 2010-2014 data, by itself, is not sufficient to answer our research questions. The latest changes in the scheme of incentives and sanctions will come into force in 2016 and will affect taxpayers in 2017 based on the 2016 diversion rates. With respect to that, earlier data have two significant limitations. First, performance evaluations from 2016 onwards will arguably reflect not only the 2014 indicator level but also its recent trend, meaning that municipalities with low diversion rates might be catching up quickly, even more so because of the looming enforcement of the new regional sanction from 2016. Second, any evaluation on fairness depends on which municipalities in the study area qualify for incentives or sanctions; any consideration based on 2014 could be misleading as very few municipalities, by then, reached the 65% (or even just the 45%) threshold while, on the other hand, some were sufficiently near to close the gap in a few years.

In order to make considerations based on a context in which the new legislation regime is fully operational, year 2019 seems an ideal choice. At that point, the latest changes in legislation will have been active for four years, allowing municipalities that are substantially increasing their diversion rate, to reach the target. It will also be the last year of the 45% regional threshold, as it will increase to 65% in 2020. As a consequence, making estimates for the years following 2020, would imply assumptions on how municipalities are going to adjust to the impending 65% target, which is currently out of reach for almost all municipalities in the study area at the current rate of improvement.

Based on such considerations, in this paper 2010-2014 data are used to produce a straightforward 2019 forecast by OLS estimation for all relevant waste collection indicators, together with the estimates for 2019 population[18] (in order to calculate per capita indicators) provided by the Italian institute of official statistics (ISTAT) and the taxation burden produced by the new scheme.

A detailed discussion on what can be inferred from 2019 forecasts using various analytical tools is in the following sections. Even without introducing specific tools, though, forecasts provide further compelling evidence about the questionable logic of the scheme. In Figure 1 the black line represents the cost of an additional ton of waste sent to Landfill in the study area (the main vertical axis on the left) as a function of diversion rate. The dashed line represents the same cost in municipalities where the 20% surtax is applied (see note 1, pg. 3 and Table 2). The joint effect of the regional sanction and the tax on landfill disposal makes the cost a decreasing function of the diversion rate until 45% diversion is reached. Then the cost is constant until the 65% of diversion is achieved, after which it starts decreasing again, in steps. The data points around the

---

[18] See http://demo.istat.it

line are the municipalities in the study area, positioned according to their diversion rate and their amount of landfill disposal per capita (the secondary vertical axis on the right). Municipalities on the left have to pay for every additional ton of waste sent to landfill with a relatively high increase in unit cost[19] under the rationale that their waste management system, having a low diversion percentage, is not working well. Nonetheless, the sanction decreases steeply and, at 45% of diversion rate, it becomes about 60% of the theoretical maximum.

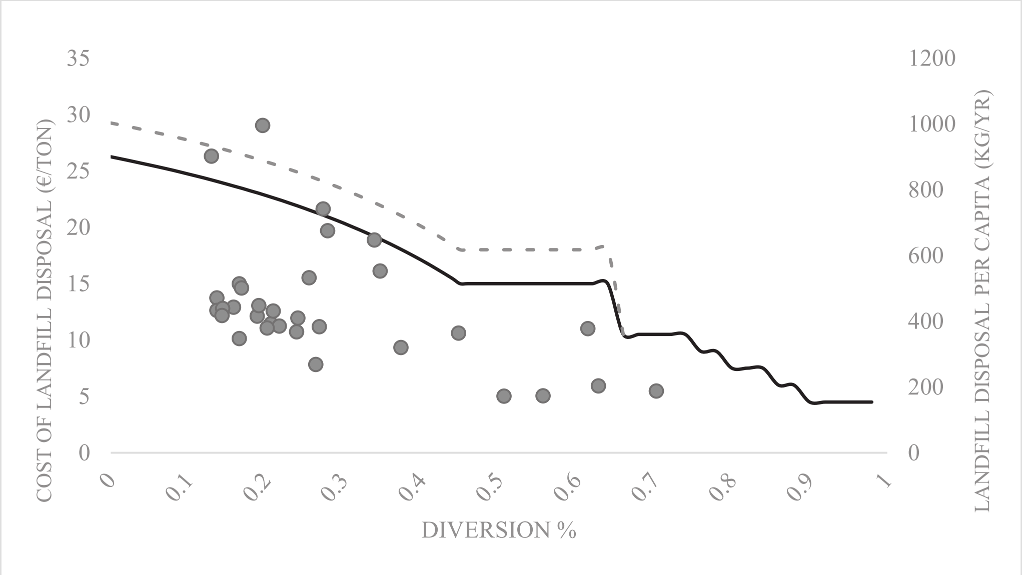

FIGURE 1 Province of La Spezia, cost of landfill disposal, diversion rate and landfill per capita
Notes: The dark line represents the cost faced by municipalities in the study area to send an additional ton of waste to landfill (left vertical axis), as a function of their diversion rate. The dotted line integrates the cost with the national surtax for municipalities not belonging to ATOs and below 65% of diversion rate (see footnote 2). Each circle represents a municipality in the study area, plotted according to the amount of waste sent to landfill per capita (right vertical axis) and its diversion rate. Notice the cluster of municipalities (bottom left) with average-to-low landfill use per capita and yet facing costs that are very close to the maximum.

Two inconsistencies appear evident from the figure. Municipalities with an average or high diversion rate can increase landfill use per capita indefinitely at a very bland rate of 4.5-15 €/ton, provided that their diversion rate is not affected. Municipalities on the flat section of the cost

---

[19] Notably, the highest potential cost corresponding to a municipality with no diversion, is 29 €/ton. It is a diminutive sanction compared to that imposed by the corresponding tax in the United Kingdom and the rate of which is 82.60 £/ton. This apparently confirms the scarce interest of the Italian legislation for this specific indicator of environmental performance.

curve can increase landfill use even at the expense of their diversion rate, provided that they remain above the 45% threshold. The second and most relevant point is that even in this relatively small dataset we have a practical example of a municipality sending less than 350 Kg/year per capita to landfill, but paying a cost that is almost 60% higher than another municipality with a landfill usage of almost 380 Kg/year per capita. More generally, the cost of landfill disposal as function of diversion rate is broadly proportional to the effect that diversion percentage has on the use of landfill per capita in about half the municipalities in the study area. The cluster of points in the lower left part of the figure indicates municipalities for which the cost is, instead, disproportionate with respect of such a relevant environmental indicator.

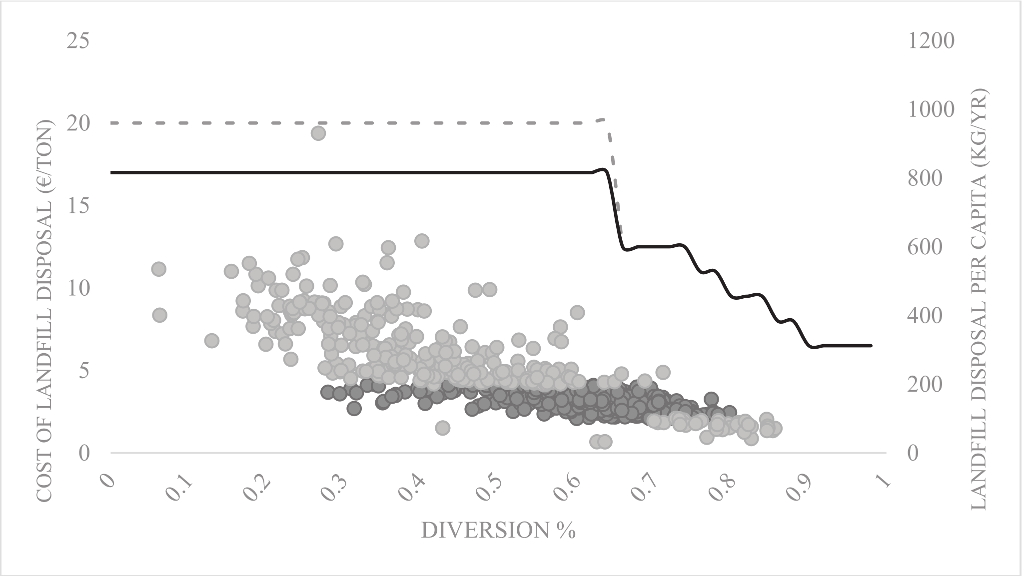

FIGURE 2 Disposal, diversion rate and landfill per capita

Notes: The dark line represents the cost faced by municipalities in Lombardy to send an additional ton of waste to landfill (left vertical axis), as a function of their diversion rate. The dotted line integrates the cost with the national surtax for municipalities not belonging to ATOs and below 65% of diversion rate (see footnote 2). Each circle represents a municipality in Lombardy, plotted according to the amount of waste sent to landfill per capita (right vertical axis) and its diversion rate. Darker dots are meant to highlight municipalities with a very moderate use of landfill. As in the study area, they are subject to all sorts of unit cost, from very low to the highest possible.

In Figure 2, the same situation is verified with respect to the municipalities of Lombardy and observed 2013 data. Without additional regional sanctions, the cost curve as a function of the diversion rate is flat until the 65% threshold. Even so, its shape seems to have little relation with the effect that diversion has on landfill use and, as a consequence, different costs hit municipalities sending about the same amount of waste to landfill per capita in an apparently random fashion. The darker dots in Figure 2 identify municipalities sending an amount between 100 Kg/year and 200 Kg/year per capita to landfill. For some of those municipalities, the cost of an additional ton of waste sent to landfill is 17 €, while others pay as little as 9.5 €. In sum we found no relation, either in the study area or in a much larger and representative dataset, that the tax function is proportional to landfill use and, consequently, related to the social cost associated with landfill disposal, with the local governments paying wildly different fees to cover the same cost. This seems hard to reconcile with the overall aims of the legislation and has the look of an unintended consequence.

2.4 Methods

In the remainder of the paper, we discuss two distinct issues that require application of as many different methods. First, we aim at comparing the sustainability ranking of UWMS in the study area and determined by diversion rate, with the ranking produced by a multi-indicator system. The latter can be obtained following different approaches, among which one can be found in partial order theory, after noting that a set of UWMS, each identified by a vector of observed indicators, is in fact a partially ordered set. We interpret partial order theory as the ideal framework to build a multi-indicator evaluation both because of its solid mathematical background (Schroder 2002) and its existing applications in multi-indicator evaluations (Brüggemann and Patil 2011) but also because it is less information demanding and less reliant on strong assumptions than other available methods (see section 3). Specifically, we borrow three different tools from partial order theory. Hasse diagrams (Schroder 2002 p. 6) provide visual representations of partially ordered sets and are particularly useful for data exploration and for visual comparison. The ξ matrix is a representation in matricial form of dominance in a partially ordered set (Brüggemann and Patil 2011 p. 22) and it is also a very practical tool to count the number of ordered pairs of element in a partially ordered set. This makes it useful to track the changes in such number between a diversion rate-based ranking and a multi-indicator ranking.

Finally, average height in partially ordered sets (Winkler 1982) is a tool that reduces a partially ordered set into a linear order and thus allows a complete ranking.

In the subsequent part of the paper, we discuss the effects of a single-indicator system on fairness in waste taxation. We make use of the well-known Suits index (Suits 1977) which is designed to represent tax progressivity using the same mathematical tools of Gini's index.

3. Comparing single and multi-indicator evaluations of sustainability performance

3.1 From a multi-indicator system to a metric for performance evaluation

In order to study the possible drawbacks of an evaluation based on a single indicator, it can be helpful to compare its results with those that stem from a proper multi-indicator system. If the evaluation based on diversion percentage is an objective assessment of environmental performance, considering other critical indicators should not alter the results dramatically. Otherwise, good and bad performances would only reflect our single-indicator arbitrary choice rather than some general concept of environmental sustainability.

An ideal multi-indicator system in the context of waste management would provide information on more than just waste collection. It would arguably cover the activities of transport, treatment, disposal, the efficiency of the whole industrial process and more. As we mentioned before, though, available data only covers some indicators about waste collection. Even so, when the multi-indicator system consisting of currently available data will be made comparable to the indicator about diversion percentage, the differences will prove to be remarkable.

Multi-indicator systems can be used to support performance evaluations in a multiplicity of ways. In a dashboard environment, the levels of all indicators concerned are separately represented to provide an overview of the situation. However, benchmarks and thresholds are considered indicator by indicator and the dashboard by itself cannot provide a metric of overall performance. More compact metrics of performance are usually obtained through a dimensional reduction of the multi-indicator system, which is generally performed as a linear combination (and occasionally a geometric combination) of the original indicators. In spite of being so common, though, this approach is not always feasible or appropriate for the context. Linear and geometric combinations of indicators allow, respectively, full and non-constant compensation between indicators, meaning that a poor performance in one dimension can be offset by a positive performance in another one (OECD 2008). This implies the existence of a system of weights that make the different indicators comparable, and the acceptability of a rate of substitution between them, but both assumptions may easily be questionable whenever a monetization of the indicators is not immediately possible.

Alternatively, dimensional reduction can be performed without linear or geometric combinations of indicators and without compensability among indicators, thanks to methods developed in the field of discrete mathematics and in order theory. Such methods have easily found applications in chemistry and environmental research (Brüggemann et al. 2004; Brüggemann and Patil 2011; De Loof et al. 2011; Voigt et al. 2004), where e.g. a ranking of areas according to their pollution level makes little sense if considering more than one pollutant implies accepting the notion of a rate of substitution between different chemicals.

The main contributions of this field of research consist in the interpretation of multi-indicator systems as partially ordered sets, the development of improved methods to visually study partially ordered sets through Hasse diagrams and, finally, applying the measure of average height in a partially ordered set (Winkler 1982) and developing some approximations of it.

3.2 Partially ordered sets, Hasse diagrams, Local Partial Order Models and their application to UWMS evaluation

In Order Theory, two elements h, k in a set S are comparable if it is possible to establish a hierarchy between them (Schroder 2002), i.e. if it is possible to state that:

$$h \geq k \qquad V \qquad h < k \tag{1}$$

The set S is linearly (or totally) ordered if all of its n elements are comparable. In a more general case, S is a partially ordered set (POSET) if at least two of its elements are comparable.

A typical example of a POSET is a data matrix $P_{n \times m}$ containing n elements described by m attributes $a_1 \dots a_m$. In the corresponding POSET P, each element k is represented by its profile $a^{(k)} = a_1^{(k)}, a_2^{(k)} \dots, a_m^{(k)}$ and, given the definitions of comparability and POSET, P must contain at least a couple ($a^{(h)}$, $a^{(k)}$) for which $a^{(h)} \geq a^{(k)}$, which is the case if $a_t^{(h)} \geq a_t^{(k)}$ $\forall$ $1 \leq t \leq m$.

A group of UWMSs can be described by a POSET in which the profile of each system consists of the row vector of its performance indicators. The results will possibly indicate that some systems are performing better than some others but not every possible couple of systems will necessarily end up being comparable. In some cases, it might not be possible to decide which in the couple is performing better.

This analysis can be very practically represented through a Hasse diagram (Schroder 2002 p. 7). A Hasse diagram represents each profile of the POSET as a node of a directed graph where all edges are directed upwards, and the direction can therefore be omitted. The structure of the graph depends on the notion of covering relation. A profile $a^{(h)} > a^{(k)}$ is said to cover $a^{(k)}$ if there exists no third profile $a^{(z)}$ in the POSET for which $a^{(k)} < a^{(z)} < a^{(h)}$. The graph of the Hasse diagram is then built according to the following rules:

1. If $a^{(k)} < a^{(h)}$, then $a^{(k)}$ is below $a^{(h)}$ in the drawing;
2. If $a^{(h)}$ covers $a^{(k)}$, an edge is included between the two.

Generally, profiles are arranged vertically in levels that depend on the structure of the graph and, when a profile could be located in several levels, the highest possible is selected. Figure 3 presents an example of Hasse diagram with just six profiles: $a^{(g)}$ covers $a^{(k)}$, $a^{(h)}$, $a^{(l)}$ and $a^{(r)}$, all of which cover $a^{(s)}$.

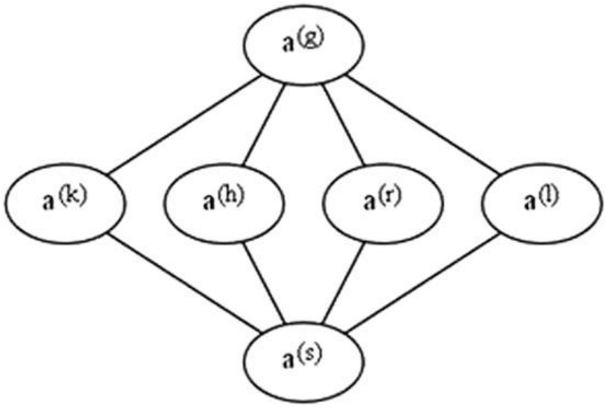

FIGURE 3 Hasse diagram of a POSET of six profiles

The pairwise ordering of profiles that produces the Hasse diagram can also be represented in matricial form, a strategy that can be very convenient to compare different partial orders. Given a POSET of n profiles $a^{(1)}, a^{(2)}, \dots, a^{(n)}$, we can represent it as a square matrix $\xi = (\xi_{ij})$ with $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, n$ where $\xi_{ij} = 1$ if $a^{(i)} > a^{(j)}$ and $\xi_{ij} = 0$ otherwise. The ξ matrix of the POSET in Figure 3 can therefore be represented as follows.

$$\xi = \begin{array}{c} \\ g \\ k \\ h \\ r \\ l \\ s \end{array} \begin{array}{c} \begin{array}{cccccc} g & k & h & r & l & s \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \tag{2}$$

In this work, the ξ matrix is used as a diagnostic tool to measure the loss of comparability in a POSET when the original set of attributes determining the order is expanded, so that if POSET P_1 is ordered according to m attributes, POSET P_2 is ordered according to $m + 1$ attributes and so forth. The number of comparable pairs in P_1 is:

$$\frac{\sum_i \sum_j \zeta_{1,ij}}{2} \tag{3}$$

Therefore, the loss of comparability between P_1 and P_2 can be defined as:

$$\frac{\sum_i \sum_j (\zeta_{1,ij} - \zeta_{2,ij})}{2} \tag{4}$$

In the context of a performance evaluation process, there are two advantages in representing the group of UWMSs to be evaluated, as a POSET. The first is to identify performance rankings that do not depend on just one in a set of multiple relevant and non-perfectly correlated indicators. Such rankings can be represented with Hasse diagrams and provide valuable and multidimensional information even though they probably contain some incomparability and the ranking can therefore be incomplete.

The second advantage is that a partial order can be reduced to a linear order, without resorting to any linear or geometric combination of the indicators, by the way of Local Partial Order Models (LPOMs) (Brüggemann and Patil 2011). Given a POSET P, (see Figure 3), the linear extensions of P (Schroder 2002 pp. 161–163) are all the linear orders compatible with P, meaning that they preserve all order relations found in the POSET. In the case of P, the POSET contains nine order relations between profiles: $a^{(g)} > a^{(s)}, a^{(k)}, a^{(h)}, a^{(l)}, a^{(r)}$ and $a^{(s)} < a^{(k)}, a^{(h)}, a^{(l)}, a^{(r)}$. A linear extension L of P is, therefore, identified by any ordering of all incomparable profiles $a^{(k)}$, $a^{(h)}, a^{(l)}, a^{(r)}$ provided that the order relations with $a^{(g)}$ and $a^{(s)}$ are preserved (see Figure 3). The linear extensions of a POSET can be used to extract valuable information about its structure. Formally, the position of a profile $a^{(k)}$ in the v-th linear extension, counting from the bottom, is called height $H(a^{(k)}, L_v)$ of $a^{(k)}$. The average height of $a^{(k)}$ across all linear extensions is then:

$$hav\left(a^{(k)}\right) = \frac{\sum_{v=1}^{LT} H\left(a^{(k)},\, L_v\right)}{LT} \tag{5}$$

where LT is the number of all linear extensions in the POSET (Brüggemann and Patil 2011). If $hav\left(a^{(k)}\right)$ is known, it amounts to a linear order that is compatible with the partial order of the POSET. However, the number of all linear extensions in a POSET is $1 \leq LT \leq n!$ (Brüggemann and Carlsen 2011) and determining all of them quickly becomes computationally impossible as n grows (Brüggemann et al. 2004).

Several methods to approximate $hav(a^{(k)})$ have been proposed[20] among which LPOMs, that are being progressively developed to allow approximations with the least possible distortion. We will make use of the first LPOM to be developed, called LPOM0 (Brüggemann et al. 2004), which was empirically shown (Brüggemann and Annoni 2014) to be not necessarily a worse approximation than the others. Looking at the POSET in Figure 3 as an example, the average height of, say, profile $a^{(k)}$, depends mostly on profiles that are certainly ordered above it $\left(a^{(g)}\right)$ and on profiles certainly ordered below it $\left(a^{(s)}\right)$. The exact order of profiles above and below is irrelevant, as any linear extension of the subsets above and below $a^{(k)}$ will produce the same outcome. Consequently, the average height of $a^{(k)}$depends on the method selected to merge incomparable profiles $a^{(h)}, a^{(l)}, a^{(r)}$ with the others. The approximation performed by LPOM0 consists in considering all the incomparable profiles as a single object that can be assigned to any position above or below $a^{(k)}$.

Let $\left(a^{(k)}\right)$ and $F\left(a^{(k)}\right)$, or O and F for brevity, be the principal downset and upset of $a^{(k)}$, i.e., respectively, the subset of the POSET P containing $a^{(k)}$ and all $a^{(h)} \leq a^{(k)}$ or all $a^{(h)} \geq a^{(k)}$. Furthermore, let $U\left(a^{(k)}\right)$, or U, be the subset of P containing all elements incomparable to $a^{(k)}$. Finally, let $\|P\|$ denote the number of profiles contained in the POSET P. LPOM0 approximates the height of each profile to a weighted average of its height when the block of incomparable profiles are merged above it, and a weighted average of its height when the block of incomparable profiles is merged below it.

---

[20] For an extensive list with detailed comparison see Brüggemann and Patil (2011) and Brüggemann et al. (2005).

$$hav(a^{(k)}) \cong \frac{\|O\|}{\|O\| + \|F\|}(\|O\| + \|U\|) + \frac{\|F\|}{\|O\| + \|F\|}\|O\| \tag{6}$$

In (6), the first product represents the height of $a^{(k)}$ if U is merged below it (the sum of profiles in O and U) and is weighted by the probability of such arrangement, that is, the share of all spots available for U that are actually below $a^{(k)}$. Similarly, the second product represents the height of $a^{(k)}$ if U is merged above it (just the number of profiles in O) weighted by the share of spots available above $a^{(k)}$. For a discussion on the possible distortions of this approximation see Brüggemann and Carlsen (2011).

The quantity in (1) is, therefore, our partial order-derived synthetic indicator of UWMS performance. LPOM0 of user defined datasets can be performed online on www.PyHasse.org.

3.3 Results

The results for 2019 data projections are shown in Figure 4[21]. On the left, from top to bottom, are represented Hasse diagrams built on four performance indicators (Figure 4a.: diversion percentage, unsorted waste per capita, organic waste per capita and WEEE per capita), on three indicators (Figure 4b.: without organic waste) and on two indicators (Figure 4c.: with just diversion percentage and unsorted waste per capita). On the right is the Hasse diagram based on diversion percentage alone (Figure 4d.), which corresponds to a total order. As by default in Hasse diagrams, all indicators are consistently directed so that high values (in this case, good performances) are at the top of the diagram and bad performances are at the bottom, while performances on the same level are incomparable. The number of levels is a broad indicator of the amount of comparability and, by representing the same group of UWMS according to an increasing set of indicators, we also make sure that, moving clockwise from Figure 4d. to Figure 4a., the order relations that depended on diversion percentage but are inconsistent with the other indicators are progressively dropped.

---

[21] The graphs are obtained by using the output of PyHasse as input for the graph visualization software Graphviz (http://www.graphviz.org)

Moving from a single indicator to four, some rankings are relatively stable. Three of the top five municipalities remain at the top level and three of the bottom five municipalities remain somewhere close to the bottom. However, the overall effect is a marked decrease of comparability. In Figure 5, the progressive loss of comparability is tracked in correspondence with the introduction of each additional indicator. In the case of total order, the number of comparable pairs that can be identified in our set is 496, while the number falls to 187 when four indicators are used, corresponding to less than 38% of the original level of comparability.

What looked like a straightforward ranking in a one-dimensional world is in fact a problematic partial order because the four performance indicators are not consistent with each other. As a consequence, whatever "compliance line" is drawn to discriminate UWMS that conform to multidimensional parameters of environmental sustainability from those that are not performing well enough, the results would be substantially different from those due to the 50% of diversion rate threshold.

Based on the amount of comparability left after increasing the number of indicators to four, we get only about six levels of performance. Each level except the lowest has a robust degree of internal incomparability meaning that, for example, municipalities performing well in a four-dimensional frame of reference are doing well for different reasons. Even though Hasse diagrams and other partial order tools can only illustrate incomparabilities and not provide causal explanations, incomparability of municipal services (and particularly between the best performing municipalities) may hint to the fact that each indicator is a challenge of different severity depending on the characteristics of a municipality. In turn, this means that choosing one indicator over the others as the key parameter implies the arbitrary allocation of an advantage for some UWMS and that of a disadvantage for others.

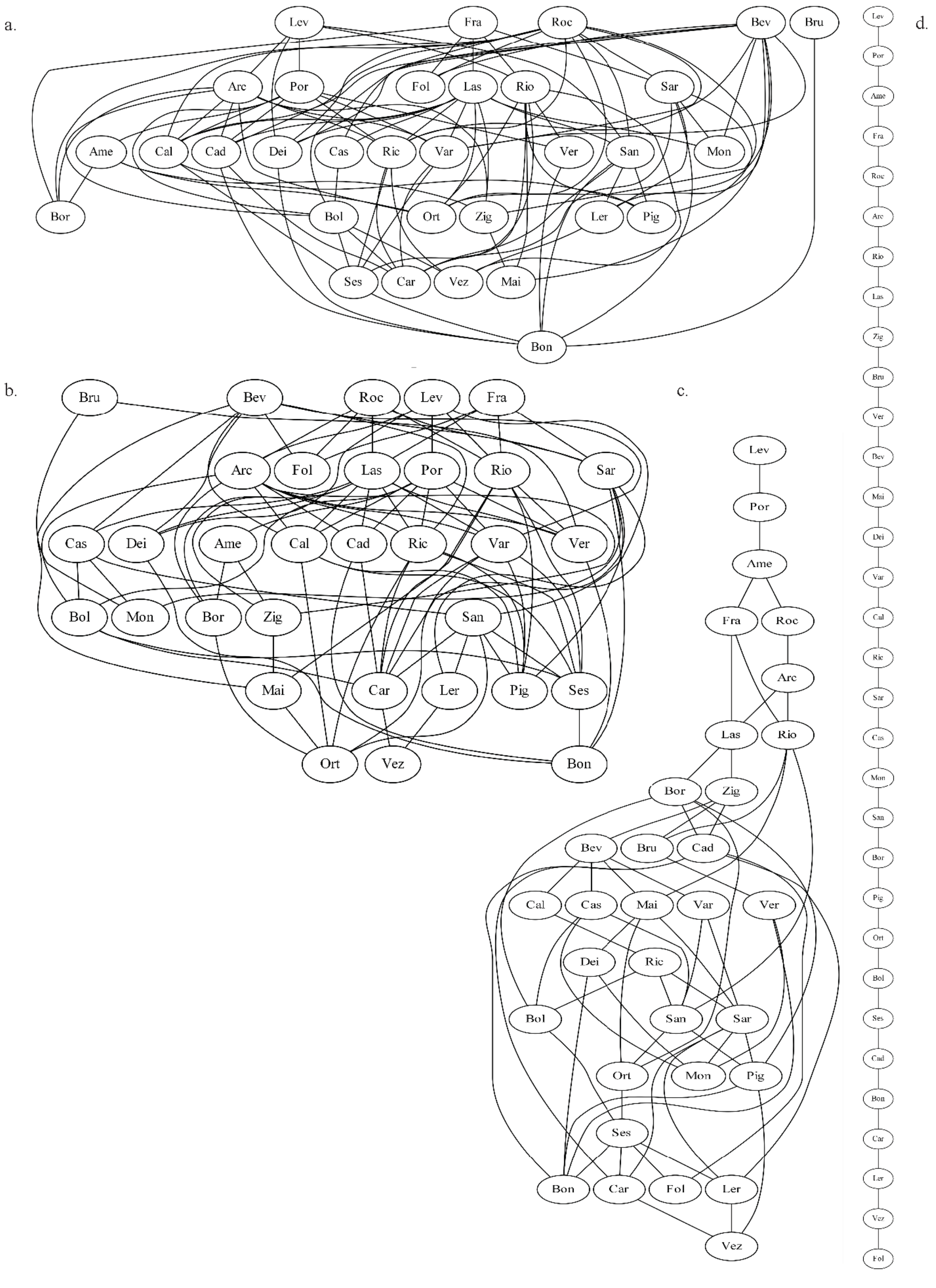

FIGURE 4 Province of La Spezia. Partial orders of UWMS performances: (a) ranked according to diversion rate, landfill per capita, organic per capita and WEEE per capita; (b) ranked according to diversion rate, landfill per capita and organic per capita; (c) ranked according to diversion rate and landfill per capita; (d) ranked according to diversion rate (linear order)

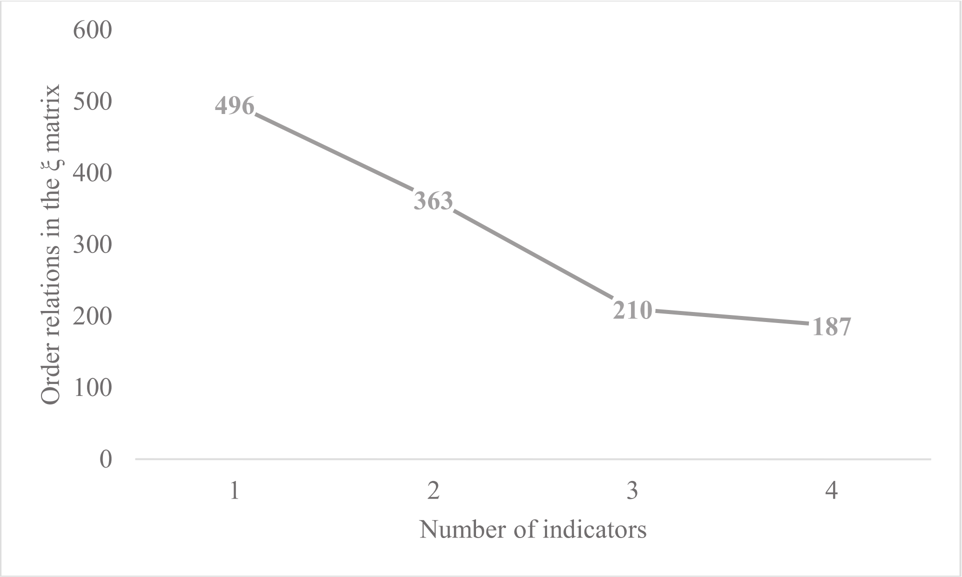

FIGURE 5 Loss of order relations following an increase in dimensionality. When linearly ordered according to diversion rate, the UWMS in the study area form 496 comparable pairs. Progressively introducing landfill per capita, organic per capita and WEEE per capita in the ranking procedure, the resulting partial orders have a number of comparable pairs that is reduced to 187

In Table 5, UWMS performances are ranked according to diversion percentage and compared with a ranking based on a local partial order model with four attributes (LPOM 4 ranking), which is a linear extension of the partial order represented in Figure 4a. As the Hasse diagrams were suggesting, a linear order based on four indicators is very different from that produced by diversion percentage and, as anticipated above, the two rankings make it clear that there is no meaningful threshold that can be placed to discriminate good and bad performances in both rankings without conspicuous differences in the composition of the good group and the bad group.

All this considered, the assumption that diversion percentage is an adequate proxy for all available information on environmental sustainability achieved by UWMS seems ungrounded and its consequences on evaluation are sizeable.

Table II-5. UWMS ranking by diversion percentage and by local partial order model with 4 attributes

Municipality	Rank (Div %)	Rank (LPOM 4)	Municipality	Rank (Div %)	Rank (LPOM 4)
Ame	3	11	Lev	1	2
Arc	6	7	Mai	13	28
Bev	12	4	Mon	20	24
Bol	25	17	Ort	24	28
Bon	28	32	Pig	23	30
Bor	22	26	Por	2	6
Bru	10	10	Ric	17	14
Cal	16	18	Rio	7	8
Car	29	26	Roc	5	3
Cad	27	16	San	21	12
Cas	19	13	Sar	18	9
Dei	14	20	Ses	26	25
Fol	32	21	Var	15	14
Fra	4	1	Ver	11	19
Las	8	5	Vez	31	31
Ler	30	21	Zig	9	21

Source: Own calculation based on data from ISPRA

4 Fairness of the single-indicator evaluation

4.1 Fairness of the 2019 estimated tax burden

Diversion percentage as a stand-alone performance indicator for UWMS is an arbitrary choice with vast consequences on the results of the evaluation. But are these consequences randomly spread across the population? And more specifically, can we expect no consequences on the fairness of the waste taxation system when environmental performance is arbitrarily represented with diversion percentage and when fiscal sanctions and incentives are determined according to it?

In our study area and more generally in Italy, incentives and sanctions based on diversion percentage directly affect the distribution of the tax burden among the municipalities rather than among the individuals. In fact, the current legislation imposes full cost coverage of waste management services at municipal level through TARI, and such costs are reduced by incentives and increased by sanctions. As a consequence, taxpayers in a municipality affected by sanctions will collectively suffer a higher tax burden and taxpayers in a municipality enjoying the incentives will get a decreased burden. Then, at the level of each municipality, further national and local norms regulate the redistribution of the tax burden between residents according to loosely redistributive principles. However, as income groups are not proportionally distributed across all municipalities, it is rather evident that intra-municipal redistribution will not be enough to compensate inter-municipal distortions, if there are any.

This seems remarkable, as waste taxation is one of the (few) backbones of fiscal autonomy in Italy. Sizeable differences in the local taxation burden can affect fiscal competition and citizens' perception about the performance of local governments. As long as the diversion rate of specific waste categories depends on socio-economic factors, territorial fairness is also affected and potential effects could be more severe precisely where territorial inequality is greater already. All these consequences could arguably be accepted if they were justified by relevant and consistent policy objectives, much less so if the policy goal (protecting the environment) is not consistent with the parameter that is ultimately responsible for creating such consequences.

Based on these considerations, we argue that the size and nature of inter-municipal distortions should be explicitly determined.

4.2 SUITS index and its application to evaluate the fairness of waste taxation in the study area

Different methods are available to evaluate the progressiveness of a tax. Distributional tables offer a general understanding of how the tax impacts different income classes, while synthetic indices provide a sharper summary statistic to allow comparison of different situations or alternatives (Slemrod 1996). Among the latter, the Suits index (Suits 1977) is a standard of tax policy analysis. The index is a close relative of the Lorenz Curve and the Gini concentration ratio: it compares cumulative income with cumulative tax burden. It is defined as:

$$S = \frac{(K - L)}{K} = 1 - \left(\frac{L}{K}\right) \tag{7}$$

where K, L can be represented as areas, like in Figure 6, delimited by the diagonal and by the tax burden function. Formally, $K = 0.5$ while L is defined as:

$$L = \int_0^1 T(y)\ dy \tag{8}$$

where T is the tax burden and y is income. In the case of a progressive tax, as in Figure 6, L is smaller than K, and the index value is then $0 < S \leq 1$, corresponding to the white area represented in the figure. In the case of a perfectly proportional tax, the tax burden function coincides with the diagonal starting in (0, 0), so that $L = K = 0.5$ and the index $S = 0$. In the case of a regressive tax, $K < L$ and the index value is $-1 \leq S < 0$.

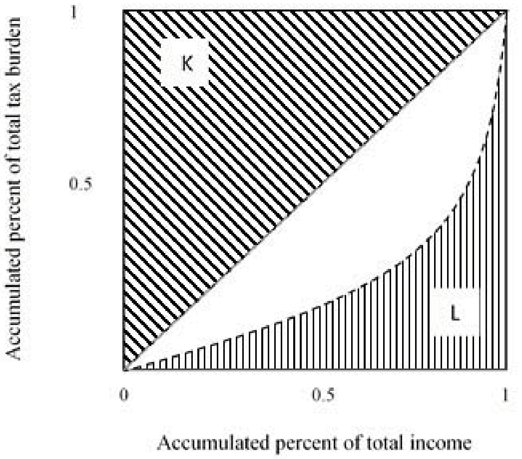

FIGURE 6 Graphical representation of the Suits index of a progressive tax
Notes: The white area is identified as the difference between the areas K and L. The Suits index can then be obtained dividing it by the area K. As long as K - L is a positive quantity (i.e. as long as the curve delimiting L is convex) the Suits index is positive as well and the tax under consideration is progressive.

The index is particularly effective in comparing how different tax schemes distribute the burden in relation with income. In the case presented here, one side of the comparisons can include the special tax on landfill disposal and the regional sanction, but also the combination of the two, to represent the performance-related taxation scheme. The other side can be represented with the observed tax burden at the latest date available (which is year 2014). We will also make a case for including, as a term of comparison, an entirely hypothetical poll tax by which every taxpayer in the study area would pay the same share of the 2019 tax burden determined by the tax on landfill disposal and the regional sanction regardless of the place of residence, so that the tax would amount to:

$$C = \frac{\sum_{i=1}^{32} T_i}{n} \tag{9}$$

where T_i is the tax burden generated by municipality i and n is the number of taxpayers across the entire province. Available data would not allow a straightforward application of the index, a relatively common situation with the Suits index. The index is intended to cumulate individual incomes and individual tax burdens, whereas available information is frequently more aggregated than that. In order to get values for accumulated percent of the tax burden, we forecast the 2019 burden generated by the upcoming scheme for each municipality. Furthermore, it was possible to obtain figures for the 2014 burden generated in each municipality by the current waste taxation from the municipalities' budgets. We have no data on the intra-municipal burden distribution, which is subject to national legislation, local regulations and, most importantly, to the distribution, at individual level, of land and real estate ownership. However, this is a relatively unimportant limitation, as the scheme we analyse only discriminates between municipalities and not between individuals. Given that intra-municipal burden distribution is generally informed by some degree of progressivity, we just expect the actual individual tax burden to be somewhat less regressive than the inter-municipal distribution, proportionally, for all taxes. A more significant issue is that the income level of each taxpayer is also not available. Nonetheless, data from the Ministry of Economy and Finance provide the average income of residents at municipal level. Along with it, more detailed information quantify, again at municipal level, the number of taxpayers falling into eight different income classes and the overall income of that class. From this, it is obviously possible to build a municipal average income for each income class and assign each taxpayer with the average income of his/her income class.

As a consequence of these adjustments, each of our data points consists of one taxpayer with a tax burden equal to the average municipal tax burden and an income that is either equal to the average municipal income or to the average municipal income of the income class they belong to.

4.3 Results

In order to assess the effects of the scheme of incentives and sanctions, six separate Suits indexes have been calculated using the overall average municipal income (see Table 6) and the respective curves are represented in Figure 7.

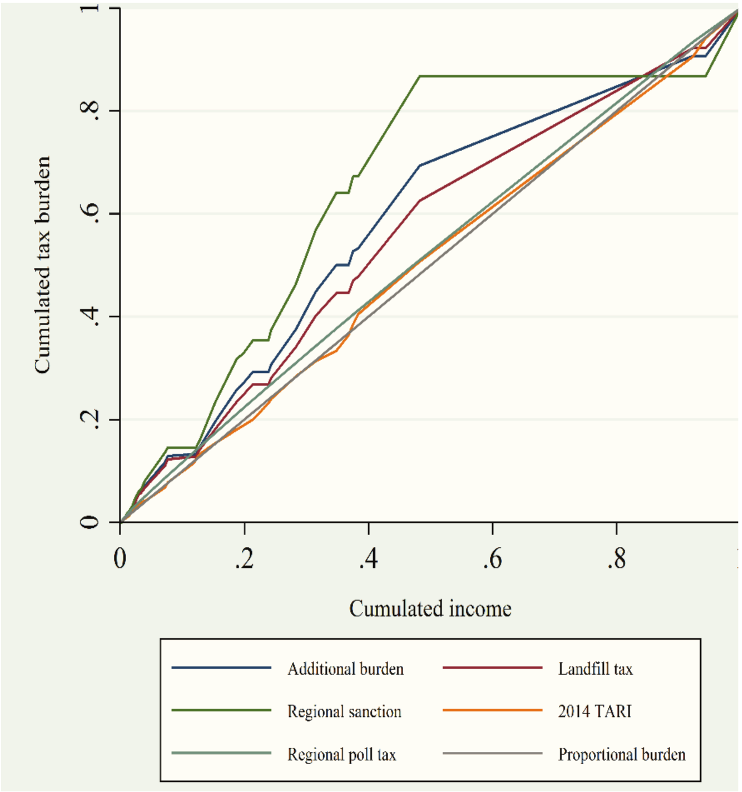

FIGURE 7 Suits indexes of waste taxation for the municipalities in the study area
Notes: All individuals are assigned with the average income of their municipality and, therefore, cumulated income is sorted by municipality on the horizontal axis. The grey line represents a proportional distribution of the tax burden between the municipalities. The yellow line represents the distribution of the tax burden in 2014, prior to the waste taxation reform discussed in this paper. The sage green line depicts a hypothetical poll tax with a total revenue that is equivalent to the projected 2019 tax burden and that we propose as a notable example of regressive taxation. The red line and the green line represent the projected 2019 national landfill tax and the regional sanction. The blue line is the combination of both.

As a consequence of using one average income for each municipality, the horizontal axis is sorted by income and by municipality conjointly, meaning that all taxpayers from the municipality with the lowest average income populate the first section of the axis, followed by all taxpayers of the second lowest and so forth.

Table II-6. Suits indexes of waste taxation

Tax	Suits index
Tax on landfill disposal	-0.11992
Regional sanction	-0.30379
Additional tax burden	-0.17158
2014 TARI	-0.0054
Poll tax	-0.03975

Source: Own calculation with data from the Ministry of Economy and Finance (MEF)

The grey diagonal represents a proportional distribution of the tax burden between all municipalities in the study area and a SUITS index of 0. The two curves closer to the grey diagonal represent, respectively, the 2014 TARI tax (i.e. waste taxation without the new scheme) and the hypothetical 2019 provincial poll tax introduced in the previous section and are the benchmark for the performance related scheme. The 2014 TARI is the observed distribution of the tax burden across municipalities: its fairness depends on how average income, population and waste taxation were distributed across municipalities in the study area at the latest available date. Furthermore, given the prescription of full cost coverage, the 2014 TARI Suits index shows the fairness implications of full cost coverage. The 2019 poll tax represents the fairness of the additional burden introduced by the scheme if all discrimination between municipalities was removed. It is nonetheless a typically regressive form of taxation and an unlikely policy solution, as it clearly violates the legislation principle according to which the polluter pays. However, it provides, by comparison, a very convenient measure of how much unfairness in projected 2019 environmental sanctions is due to discriminating municipalities according to their diversion percentage. At this stage, the Suits indexes of both are very close to full proportionality, indicating that the variability of average income across municipalities is not high enough to generate substantial regressiveness. The performance related taxation scheme, on the other hand, shows regressive effects from mild, in the case of the tax on landfill disposal, to moderate in the case of the regional sanction. It is noteworthy that much of the emerging inequity is concentrated on mid-income municipalities, so that about the poorest half of the municipalities suffer over 80% of the tax burden.

This first step of the analysis has the advantage of being directly focused on a comparison between municipalities, but it also comes with some limitations. The detail level of regressiveness effects, using just the average municipal income, is not very accurate. Furthermore, the disproportionate effect of the largest urban area, the city of La Spezia, makes conclusions less robust.

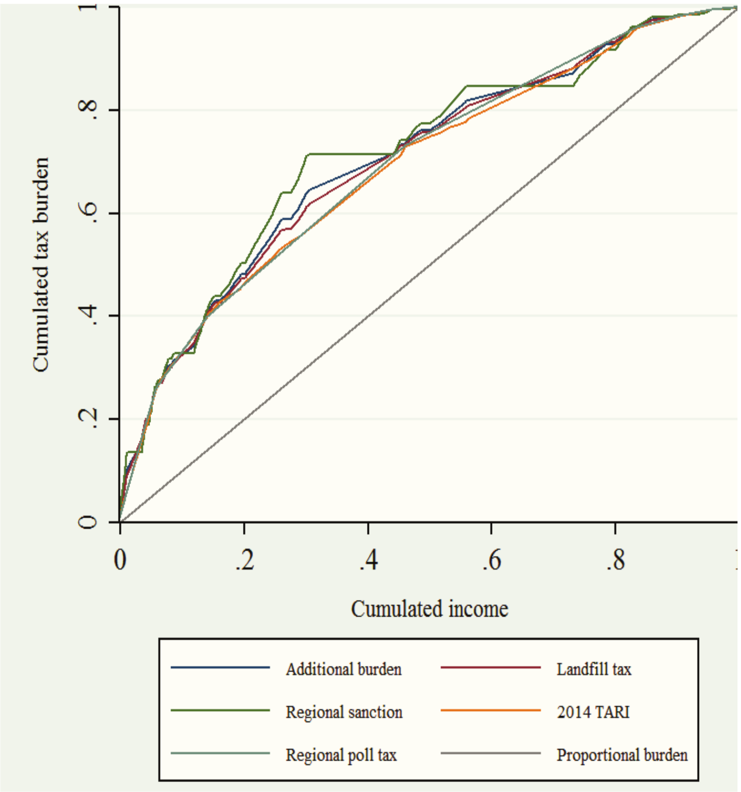

FIGURE 8 Suits indexes of waste taxation for the municipalities in the study area
Notes: All individuals are assigned with the average income of their income group in their municipality and, therefore, cumulated income is no longer sorted by municipality on the horizontal axis. The grey line represents a proportional distribution of the tax burden between the municipalities. The yellow line represents the distribution of the tax burden in 2014, prior to the waste taxation reform discussed in this paper. The sage green line depicts a hypothetical poll tax with a total revenue that is equivalent to the projected 2019 tax burden and that we propose as a notable example of regressive taxation. The red line and the green line represent the projected 2019 national landfill tax and the regional sanction. The blue line is the combination of both.

Therefore, in Figure 8, the analysis is replicated with the full amount of information available about personal income, moving the focus from inter-municipal inequity to inter-personal inequity before intra-municipal redistribution of the tax burden.

Table II-7. Suits indexes of waste taxation

Tax	Suits index
Tax on landfill disposal	-0.41684
Regional sanction	-0.43897
Additional tax burden	-0.40784
2014 TARI	-0.38166
Poll tax	-0.39261

Source: Own calculation with data from the Ministry of Economy and Finance (MEF)

The resulting analysis is much finer grained and reflects the additional source of inequality as a result of the fact that the distribution of average income of income classes is much more complex than what the overall average municipal income could tell.

As a consequence of this more detailed approach, the horizontal axis is now sorted so that individuals of the lowest income class in the municipality where such class has the lowest average income populate the first section of the axis, followed by individuals of the lowest income class in the municipality where such class has the second lowest average income and so on.

When a more detailed distribution of income is taken into account, the Suits index of the 2014 TARI tax (see Table 7) is -0.38 and that of the capitation tax is -0.39, meaning that a distribution of the tax burden between municipalities according to their costs of waste management provision is a bit more progressive than a capitation tax. In any case, at this point, the two taxes show a rather severe degree of unfairness, as the lowest quintile of income distribution pays for about 50% of the tax burden. Finding that TARI is less regressive than a capitation tax even without intra-municipal redistribution is not overly surprising: it seems reasonable that less affluent municipalities tend to control their costs more tightly. It is remarkable, instead, that the difference is not great even if a capitation tax is so severely regressive and it is unclear if the further, intra-municipal redistribution of the burden that we are not taking into account can ultimately compensate this.

The red curve represents the forecast of the 2019 special tax on landfill disposal. The tax is more regressive than both of our benchmarks, with a Suits index of -0.41. By visual inspection of Figure 8, it is possible to identify that it is the second quintile of the income distribution that pays much of the price of increased regressiveness, followed by the third quintile.

The green curve marks the cumulated burden of the forecast 2019 regional sanction. It closely follows the burden distribution of the special tax for landfill disposal, but in fact there are subtle differences: municipalities with more than 45% of diversion don’t pay at all, whereas municipalities below the threshold pay a sanction that depends both on total waste production and on sorted waste. As a matter of fact, the regional sanction is more regressive than the tax on landfill disposal and the most regressive of the group, with a Suits index of -0.44.

The blue curve describes the combination of the tax on landfill disposal and the regional sanction, that is, the forecasted increase in tax burden due to the scheme of incentives and sanctions and it Suits index is -0.42.

Before the intra-municipal redistribution of the tax burden takes place, all tax schemes are very distant from proportionality. Specifically, the poll tax is markedly regressive, taxation based on full cost coverage is slightly less regressive and incentives and sanctions based on diversion percentage are more regressive. However, while full cost coverage responds to a policy goal, diversion percentage is a very poor proxy of environmental sustainability in waste management and the resulting increase in regressiveness seems unjustified.

5 Comments

As service providers, waste management systems have a clear institutional goal (sustainable waste disposal) but no direct performance indicator. The concept of sustainability in waste management has to be split into disparate elements and only then, some assembly of the available measurable elements can be put to use as a performance metric.

Even so, with the assemblage of different indicators being necessary to get to a synthetic evaluation parameter, this remains in all evidence a multi-dimensional problem. The lack of a strong, positive correlation between waste management indicators suggests, on one hand, that simplified evaluations with just one or few dimensions can be very inaccurate in measuring actual sustainability. When put to the test, more than 60% of performance ranks in our set of waste management systems proved to be a consequence of the choice of one specific performance indicator over others and, consequently, unrelatable to overall sustainability. On the other hand, it raises a red flag because unidimensional evaluations are likely to be not just bad but also biased. A taxation scheme built on a single indicator favors some and penalizes others, potentially affecting the distribution of the tax burden between municipalities and between income groups. It also affects policies, at the advantage of those local governments following a certain sustainability approach (like increasing diversion) over those emphasizing a different one (e.g. reducing waste production). Even if any of these distortions may be acceptable in the light of a policy objective like sustainable waste disposal, when there is no coincidence between sustainability and its selected proxy, the rationale for tolerating the distortions becomes unclear. If the burden of waste taxation is adjusted by national and regional governments according to environmental performance evaluations, biased evaluations may subtract resources to communities that are not necessarily managing urban waste badly but have a lower income, ultimately making improvements harder to achieve. When the adjustment affects the distribution of the burden between local governments rather than between individuals, as in the case presented here, this could also have consequences on fiscal competition, inducing location patterns on households and on economic activities that could reinforce the bias.

Fairness considerations aside, the most evident consequence of selecting diversion rate as the single performance indicator is encouraging measure fixation on an indicator that is mostly blind to the environmental consequences of landfill disposal: given a certain diversion rate, any amount of waste sent to landfill is possible with limited consequences. This is likely to produce effects regardless of taxation, as local governments concentrate their efforts to pursue improvements in the key indicator selected by the legislation and, all the while, the relative visibility of the indicator also affects their reputation with the voters. Under these premises, in many supposedly virtuous and sustainable municipalities, dismayed residents will suddenly have to face, at some point in time, landfill expansion and landfill exhaustion, with all the related issues.

In the specific case of the Italian scheme of incentives and sanctions, some efficiency concerns, although not extensively discussed in this work, seem justified as well. Raising the cost that municipalities face for waste management by the way of regional and national landfill taxes, can be justified in the light of negative externalities generated by landfill disposal. Setting the tax rate according to diversion percentage, though, practically removes the Pigouvian relation between the social cost generated by each municipality, which is a function of its use of the landfill, and the rate payed for it.

The legislators have several alternative approaches to look at, should they decide to embrace some form of multi-dimensional evaluation, and they should arguably refer to the vast literature on multi-criteria decision making (Zopounidis and Pardalos 2010) to ponder such alternatives. The most straightforward, but also the most demanding in terms of information, is the selection of an eco-economical model providing the weighting parameters for a weighted composite indicator of sustainability in waste management. This implies a more precise definition of sustainability, strong assumptions about the aggregation function underlying the composite indicator and even stronger about compensation of trade-offs between different performance indicators. However, it also provides a synthetic measure that can be used to define a tax rate. Other methods require less extensive assumptions about the relation between performance indicators and sustainability but still need weights, as is the case with non-compensatory aggregation procedures (Bouyssou 1986; Fishburn 1976). It should also be noted that they only produce rankings which may be unsuitable to define tax rates. The previously introduced notion of average height in partially ordered sets may also be useful as it requires neither weights, nor assumptions on aggregation functions and there are applications in literature where it is used for the similar task of comparing pollutants (Brüggemann and Annoni 2014). Further research is required to introduce information on distance in partially ordered sets, without which only ranks can be obtained.

Multi-indicator evaluations may have different degrees of complexity and, because of that, they may in some cases be less appealing for decision makers. However, complexity in evaluation reflects complexity in the objectives of the organization and therefore it seems reasonable to ultimately base evaluations on a small group of proper performance indicators. The availability and quality of data is also a requirement and, again, this doesn’t seem to lead to higher costs as at least some performance indicators are already being collected and, in the case of diversion rate, they are extensively used. Furthermore, diversion rate is in itself a rather complex measure which requires the same information that would be needed for a very reasonable two-indicator evaluation (diversion rate and landfill disposal per capita).

The shortcomings of performance evaluation based on diversion rate may affect a relatively small portion of waste taxation, but the potential distortions on policies are remarkable and waste legislations would certainly benefit from a multi-indicator approach to sanctions and incentives for waste management performance.

References

Achillas, C., N. Moussiopoulos, A. Karagiannidis, G. Banias, and G. Perkoulidis. 2013. The use of multi-criteria decision analysis to tackle waste management problems: a literature review. *Waste Management & Research*, *31*(2):115–129.

Allesch, A., and P. H. Brunner. 2014. Assessment methods for solid waste management: A literature review. *Waste Management & Research*, *32*(6):461–473.

Arrow, K. J. 1962. *Social Choice and Individual Values*. New York, New York, USA: John Wiley & Sons.

Bouyssou, D. 1986. Some remarks on the notion of compensation in MCDM. *European Journal of Operational Research*, *26*(1):150–160.

Brightwell, G. R., and W. T. Trotter. 2002. A combinatorial approach to correlation inequalities. *Discrete Mathematics*, *257*(2–3):311–327.

Brightwell, G. R., and P. Winkler. 1991. Counting linear extensions. *Order*, *8*(3):225–242.

Brüggemann, R., and P. Annoni. 2014. Average Heights in Partially Ordered Sets. *MATCH Communications in Mathematical and in Coputer Chemistry*, *71*:117–142.

Brüggemann, R., and L. Carlsen. 2011. An Improved Estimation of Averaged Ranks of Partial Orders. *MATCH Communications in Mathematical and in Computer Chemistry*, *65*:383–414.

Brüggemann, R., and G. Patil. 2011. *Ranking and prioritization for multi-indicator systems*. Springer.

Brüggemann, R., P. B. Sørensen, D. Lerche, and L. Carlsen. 2004. Estimation of averaged ranks by a local partial order model. *Journal of Chemical Information and Computer Sciences*, *44*(2):618–25.

Davey, B. A., and H. A. Priestley. 2002. *Introduction to lattices and order* (Second Edi). Cambridge, UK: Cambridge University Press.

De Loof, K., B. De Baets, and H. De Meyer. 2011. Approximation of Average Ranks in Posets. *MATCH - Communications in Mathematical and in Computer Chemistry*, *66*:219–229.

Desmond, M. 2006. Municipal solid waste management in Ireland: assessing for sustainability. *Irish Geography*, *39*(1):22–33.

Dushnik, B., and E. W. Miller. 1941. Partially Ordered Sets. *American Journal of Mathematics*, *63*(3):600–610.

Edwards, J. R., R. P. Bagozzi, R. S. Blackburn, K. a Bollen, D. Cattani, J. W. D. Jr, D. a Harrison, L. R. James, K. a Markus, E. E. Rigdon, A. H. Segars, and L. J. Williams. 2000. On the Nature and Direction of Relationships Between Constructs and Measures. *Psychological Methods*, *5*(2):155–174.

Fattore, M. 2017. Functionals and Synthetic Indicators Over Finite Posets. In M. Fattore & R. Bruggemann (Eds.), *Partial Order Concepts in Applied Sciences* (pp. 71–86). Springer International Publishing AG.

Fishburn, P. C. 1976. Noncompensatory Preferences. *Synthese*, *33*(2):393–403.

Foldes, S. 1994. *Fundamental Structures of Algebra and Discrete Mathematics*. New York: Wiley & Sons.

Frey, B. S., and A. Stutzer. 2005. Happiness Research: State and Prospects. *Review of Social Economy*, *62*(190):207–228.

Greene, K. L., and D. J. Tonjes. 2014. Quantitative assessments of municipal waste management systems : Using different indicators to compare and rank programs in New York State. *Waste Management*, *34*(4):825–836.

Guimarães, B., P. Simões, and R. C. Marques. 2010. Does performance evaluation help public managers? A Balanced Scorecard approach in urban waste services. *Journal of Environmental Management*, *91*(12):2632–2638.

Guttman, L. 1950. The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), *Measurement and prediction.* Princeton University Press.

ISPRA. 2014. *Rapporto rifiuti urbani. Edizione 2014*. Rome.

Jorm, A. F., and S. M. Ryan. 2014. Cross-national and historical differences in subjective well-being. *International Journal of Epidemiology*, *43*(2):330–340.

Kaplow, L. 2010. *The theory of taxation and public economics* (Paperback). Princeton University Press.

Kaufman, S. M., N. Krishnan, and N. J. Themelis. 2010. A screening life cycle metric to benchmark the environmental sustainability of waste management systems. *Environmental Science and Technology*, *44*(15):5949–5955.

Kulczycka, J., L. Lelek, A. Lewandowska, and J. Zarebska. 2015. Life Cycle Assessment of Municipal Solid Waste Management – Comparison of Results Using Different LCA Models. *Polish Journal of Environmental Studies*, *24*(1):125–140.

Lerche, D., P. B. Sørensen, and R. Brüggemann. 2003. Improved estimation of the ranking probabilities in partial orders using random linear extensions by approximation of the mutual ranking probability. *Journal of Chemical Information and Computer Sciences*, *43*(5):1471–80.

Massoud, M. A., M. El-Fadel, and A. A. Malak. 2003. Assessment of public vs. private MSW management: A case study. *Journal of Environmental Management*, *69*(1):15–24.

Mendes, P., A. C. Santos, F. Perna, and M. Ribau Teixeira. 2012. The balanced scorecard as an integrated model applied to the Portuguese public service: A case study in the waste sector. *Journal of Cleaner Production*, *24*:20–29.

Musgrave, R. A. 1959. *The theory of public finance: a study in public economy*. New York: McGraw-Hill.

OECD. 2008. *Handbook on constructing composite indicators: methodology and user guide*. Brussels.

OECD. 2013. *Environmental indicators: development, measurement, and use*. Paris.

Oswald, A. J. 2008. On the curvature of the reporting function from objective reality to subjective feelings. *Economics Letters*, *100*(3):369–372.

Pattanaik, P. K. 2002. Positional Rules of Collective Decision-Making. In K. J. Arrow, A. Sen, & K. Suzumura (Eds.), *Handbook of Social Choice and Welfare volume 1* (pp. 361–394). Amsterdam: Elsevier Science B.V.

Piano regionale di gestione dei rifiuti e delle bonifiche (2015). Regione Liguria.

Sanjeevi, V., and P. Shahabudeen. 2015. Development of performance indicators for municipal solid waste management (PIMS): A review. *Waste Management & Research*, *33*(12):1052–1065.

Schroder, B. S. W. 2002. *Ordered Sets. An Introduction*. Boston: Birkhauser.

Scipioni, A., A. Mazzi, F. Zuliani, and M. Mason. 2008. The ISO 14031 standard to guide the urban sustainability measurement process : an Italian experience. *Journal of Cleaner Production*, *16*:1247–1257.

Sen, A. 1995. Rationality and social choice. *The American Economic Review*, *85*(1):1–24.

Sen, A. 2017. *Collective choice and social welfare*. Milton Keynes, UK: Penguin Random House.

Shepp, L. 1982. The XYZ conjecture and the FKG inequality. *The Annals of Probability*, *10*(3):824–827.

Shye, S. 1985. *Multiple Scaling. The Theory and Application of Partial Order Scalogram Analysis*. Jerusalem: Israel Inst of Applied Social Research.

Shye, S., and R. Amar. 1985. Partial order scalogram analysis by base coordinates and lattice mapping of items by their scalogram roles. In D. Canter (Ed.), *Facet Theory: Approaches to Social Research* (pp. 277–298). Springer Verlag.

Slemrod, J. 1996. *Tax progressivity and income inequality*. Cambridge: Cambridge University Press.

Stanley, R. P. 1981. Two Combinatorial Applications of the Alexsandrov-Fenchel inequalities. *J.Combin.Theory (A)*, *65*:56–65.

Stanley, R. P. 2011. *Enumerative Combinatorics Volume 1 second edition* (Second Edi, Vol. 1). Cambridge, UK: Cambridge University Press.

Suits, D. B. 1977. Measurement of tax progressivity. *The American Economic Review*, *67*(4):747–752.

Szpilrajn, E. 1930. Sur l’extension de l’ordre partiel. *Fundamenta Mathematicae*.

Vergara, S., and G. Tchobanoglous. 2012. Municipal solid waste and the environment: a global perspective. *Annual Review of Environment and Resources*, *37*:227–309.

Voigt, K., G. Welzl, and R. Brüggemann. 2004. Data analysis of environmental air pollutant monitoring systems in Europe. *Environmetrics*, *15*(6):577–596.

Welsch, H. 2007. Environmental welfare analysis: A life satisfaction approach. *Ecological Economics*, *62*:544–551.

Weymark, J. 2013. Social Welfare Functions. *Vanderbilt University Department of Economics Working Papers*, *VUECON-13*-(00018).

Weymark, J. A. 1984. Arrow ' s Theorem with Social Quasi-Orderings. *Public Choice*, *42*(3):235–246.

Wilson, D. C., L. Rodic, a. Scheinberg, C. a. Velis, and G. Alabaster. 2012. Comparative analysis of solid waste management in 20 cities. *Waste Management & Research*, *30*(3):237–254.

Winkler, P. 1982. Average height in a partially ordered set. *Discrete Mathematics*, *39*:337–341.

Zaccariello, L., R. Cremiato, and M. L. Mastellone. 2015. Evaluation of municipal solid waste management performance by material flow analysis: Theoretical approach and case study. *Waste Management & Research*, *33*(10):871–885.

Zopounidis, C., and P. Pardalos. 2010. *Handbook of Multicriteria Analysis*. (C. Zopounidis & P. Pardalos, Eds.). Heidelberg: Springer Verlag.

CHAPTER III

‘‘Beyond GDP’’ Effects on National Subjective Well-Being of OECD Countries

1. Introduction

Since the end of World War II, maximizing GDP has been the primary policy goal of almost every country around the world. Nevertheless, over the same period, some politicians, economic experts and scientists have warned about the limitations of GDP as an indicator and sometimes have spoken against its use, particularly as a *proxy* of well-being. Notoriously, one of the initial proponents of GDP, Nobel Prize economist Simon Kuznets, denied that national welfare could be directly inferred from GDP, a measure of national income (Kuznets 1934). Starting in the 1970s, works by Richard Easterlin, Fred Hirsch and Tibor Scitovsky cast doubts on the relationship between income, consumption and well-being (Diener 2000; Welsch 2009) and, consequently, on the effects of GDP-increasing policies on well-being. Some of their conclusions have been challenged by influential economists[22]. Even though the issue can be considered controversial (Easterlin 2015; Stevenson and Wolfers 2013), a branch of economics now assumes that income is neutral to well-being at least for income levels that satisfy basic needs. Several personalities have pledged for a more diverse set of targets for public policies (among which, famously, Robert F. Kennedy in a 1968 speech on Gross National Product) or for a more multi-dimensional approach to welfare measurement, an endeavor spanning from a well-known work by Nordhaus and Tobin (1972) to the recent Stiglitz-Sen-Fitoussi report (Stiglitz et al. 2010). Many works of this group typically start from a theory of why income does not explain well-being under certain conditions and propose a new dimension of well-being that is then tested empirically. This is the case, for example, with all the recent literature on relational goods (Becchetti et al. 2008, 2011; Bruni and Stanca 2008; Gui and Stanca 2010; Stanca 2009): according to the theory, income stops "buying" well-being at a certain level because every additional unit is obtained with longer working times and is used to buy a greater proportion of positional goods (which do not increase well-being) at the expense of free time which would otherwise be spent acquiring relational goods, that positively influence well-being. Other authors, more drastically, have claimed that GDP should

---

[22] See Stevenson and Wolfers (2008), the subsequent comments by Becker, Rayo and Krueger (2008) and, with different arguments, Oswald (2008).

be replaced, rather than integrated, as a policy target (Costanza et al. 2009; Kahneman et al. 2004; Kubiszewski et al. 2013).

In recent times, national and international institutions have cautiously moved in the direction suggested by GDP critics: a vast choice of indicators has been developed to complement GDP in measuring well-being, as a reference for policy decisions and to inform the policy debate and engage citizens (like the UN's Human Development Index, the Cobb-Daly Genuine Progress Indicator, or OECD's Better Life Index and, in Italy, ISTAT's BES). The recent formulation of sustainable development goals seems to indicate that a policy shift of some sorts has been embraced. The process has been called "beyond GDP" approach (Costanza et al. 2014, 2009; Fleurbaey 2009).

Among the proponents of the "beyond GDP" framework, there is substantial agreement on one point: well-being must be a multidimensional construct and a correct identification of its dimensions and their interactions can support better policy decisions. Beyond this, however, many aspects remain unclear.

1) The endeavor of measuring well-being has yet to identify or build a single, agreed indicator to measure overall progress in well-being (for an extensive review of the different metrics see OECD (2013) but some measures of perceived well-being derived from large-scale surveys, like "happiness", "life-satisfaction" and "subjective well-being" have raised considerable interest;
2) So far, national and international initiatives building on these measures have treated perceived well-being as one dimension alongside others in the multidimensional concept of well-being (OECD 2013, 2014); however, proponents of such measures haven't been as stringent in characterizing the role of perceived well-being (Diener 2000; Kahneman et al. 2004; Oishi and Diener 2014); as a matter of fact, SWB literature frequently presents perceived well-being as a much encompassing proxy of well-being[23] and as a response to ill-advised GDP-maximizing policies;

---

[23] "Unlike economic indicators, which locate a person's well-being primarily in the material realm of marketplace production and consumption, well- being indicators assess the full range of inputs to the quality of life, from social

3) In fact, there is no consensus on the validity of these measures (Bond and Lang 2014; Diener et al. 2013; Heine et al. 2002; Jorm and Ryan 2014; Krueger and Schkade 2008) and there is some disagreement about their degree of equivalence (Jurado and Perez-Mayo 2012; Welsch 2009); furthermore, while the distinctness of "beyond GDP" approaches from orthodox utilitarianism is a very relevant qualification for some proponents of the "beyond GDP" approach, in the case of perceived well-being the distinction is rather blurred because the point of view is typically hedonic[24];
4) There is no consensus on the components or determinants of well-being even after a great variety of plausible dimensions have been proposed, looking at the spheres of socioeconomic welfare, psychology and lifestyle, human/relational environment, natural environment and more;
5) The lack of consensus depends to some degree on the longstanding and unresolved discussion on whether different forms of capital (economic and social capital, manmade and natural capital, etc.) are substitutes or complements (or any combination of the two)[25];
6) Correspondingly, it also depends on disagreements about potential trade-offs between different determinants of well-being and their substitutability so that, in general, it is unclear how the components should be compared and if they should be compared at all;
7) Finally, existing estimates about the determinants of well-being are frequently affected by scarce multidimensionality or, otherwise, by endogeneity issues.

At present, a "beyond GDP" approach requires, at least, advancements in two directions.

The first is a better specification of which indicator of well-being can be the object of rigorous measurements and, consequently, of maximization efforts; this goes beyond the scope of our work.

---

relationships to spirituality and meaning, from material consumption to feelings of relaxation and security", (Diener and Tov 2012, p. 3)

[24] Attempts of putting subjective well-being at the center of policy decision making have well documented foundations in Bentham's quantitative hedonism (Kahneman and Riis 2005). For a distinction between hedonic and eudaimonic research on well-being see, for instance, Deci and Ryan (2008)

[25] See for example Beckerman (1994) and Daly (1995) on manmade and natural capital.

In order to test the feasibility of one specific measure alternative to GDP, we will adopt subjective well-being (SWB) as a reference indicator[26] because of the influential opinions supporting its use as a major policy target (Diener and Tov 2012; Kahneman et al. 2004). The measure of SWB we adopt has the distinctive quality of being part of the OECD database of social indicators (OECD 2014), which is also the source of much of the data used throughout this work, and thus has the advantage of being homogeneous in terms of methodology and that of having documented institutional relevance[27]. In the short term, these qualities seem to make this metric the more likely well-being indicator to have an impact on policy decision. The choice of some other metric of perceived well-being, as those discussed in point 1, above, or those mentioned in OECD (2013), would obviously be possible, with results not necessarily equivalent to those presented in this paper, even though some additional issues in terms of reliability and validity should arguably be addressed in that case (OECD 2013, p. 28-59). Additionally, it is common ground that the use of any subjective measure of well-being as the criterion to assess specific policies can be criticized by philosophers and economists on many grounds, including objections to hedonic well-being. This, we argue, is a matter for future research: drastic departures from the utilitarian paradigm implied by SWB would require consensus on a strong and testable alternative conceptual model, but such consensus has yet to be reached.

The second advancement required for going "beyond GDP", once the decision is made on some metric of well-being, is the identification of convincing evidence on the dependence of such indicator from any policy-controlled measurable attribute (of individuals or communities) other than GDP. This last point is the focus of our work.

Our intent is to model the average national value of SWB across OECD member countries using three large sets of indicators on society, environment and time-use, which broadly cover many of the proposed multidimensional integrations to GDP. The innovative content of our proposal concerns, first, the specification of the well-being problem and consists in applying a method, derived from partial order theory and recently applied to socio-economic issues (Fattore 2016),

---

[26] For the source of data and the metric of SWB used here, see section 2.1.

[27] In other words, it has already been used as a major component of the "quantitative evidence on the social situation" (OECD 2014, p. 78) used by very relevant institutions.

which we argue to be more effective than a battery of indicators or, at the opposite, of a synthetic indicator. The specification we provide is not meant to be conclusive, meaning that better explanatory variables can ultimately be found, but the method is a good benchmark for future research. Furthermore, the results we obtain modeling SWB with the three sets of explanatory indicators provide at least some conclusive evidence on the prospects of improving perceived well-being by acting on the typical institutional indicators concerning "beyond GDP" dimensions.

The paper is organized as follows.

In section 2, we present the indicators used in this work and the thematic framework according to which they are divided, followings the example of an OECD database on social indicators and extending it to environmental and time-use indicators. We also discuss at some length the properties of data concerning the specific metric of SWB that we are using and we make the case for modeling it as a discrete variable.

In section 3, we provide a brief overview of the method, which is then detailed in subsection 3.2, where we present Hasse diagrams, partial order scalograms with base coordinates and local partial order models; all three are partial order methods that we combine in order to deal with multidimensionality. In subsection 3.3 we briefly present ordered logistic regression models and post-estimation tests for the validity of these models assumptions and for the comparison of different models.

In section 4, we present results for all steps of our method. First, we show the output of Hasse diagrams and POSAC, which are graphical representations of dimensional reduction procedures that preserve partial order. In our work, these methods are used to analyze the sets of indicators and to identify variables that have to be dropped in order to ensure that each thematic group of indicators is partially ordered to an acceptable degree. Then, a further dimensional reduction is performed using local partial order models, this time with the intent of using the output dimensions for the subsequent parts of the analysis. Finally, the section presents the outcome of several ordered logit models that use SWB as the dependent variable and the dimensions produced in the previous steps as explanatory variables.

In section 5 we discuss the results, focusing on post-estimation tests, on the interpretation of the ordered logit models, on the statistical significance and on the sign of the coefficients.

In section 6 we look at the broader meaning of our results, drawing conclusions in terms of validity of different indicators and thematic groups of indicators as drivers of SWB and commenting on the effectiveness of the corresponding theories about well-being; we also discuss the implications of our results when it comes to deciding if a multidimensional definition of SWB is a better policy target than GDP and which policy recommendations are likely to come from its use.

Our results are consistent with the view that SWB is multidimensional and some factors other than GDP clearly emerge, so that a statistical model predicting SWB in OECD countries does not strictly require an income indicator; however we found no evidence in support of several theoretical assumptions concerning "beyond GDP" approaches and about half of those "beyond GDP" factors that we found significant are unlikely to lead to major policy shifts.

2. Data

Four criteria guided the selection of data for this research. First, we identified a single source, in order to obtain the best possible consistency in collection methods. As a source, OECD also ensures a strong theoretical background to all indicators published on its website and has less missing data than other comparable sources. Second, data referred to OECD countries are generally clustered around the top section of the curve of SWB as a function of national GDP, i.e. around the countries that, based on the Easterlin paradox, should have little benefit for additional income and much benefit from other things. Third, we selected datasets and specific indicators that we see as credible representatives of the universe of all measures conceived or proposed to adjust or replace GDP as a measure of national progress. There is no ambition to make a full taxonomy of such measures and to include all of them in the attempt to model the national levels of SWB. A selected, multidimensional group of a few dozen seems enough to evaluate if existing knowledge about "beyond GDP" factors affecting SWB is substantive and established. Finally, data selection is organized in the perspective of dimensional reduction. Groups of the selected indicators are supposed to be related to common underlying constructs or latent variables and such relation should have a well-defined theoretical background. That is the case when the indicators are already organized in groups in their database of origin and the group is labelled with a very precise theoretically defined concept. This assumption is even more important because of the methodology applied in this work, and this is discussed at length in the methods section. Needless to say, as our intent is assessing the performance of "beyond GDP" indicators, data on GDP *per capita* will be occasionally used for comparison purposes.

2.1 Social indicators

The OECD database of Social Indicators (OECD-DSI) consists of five groups of five national indicators each. The indicators cover 34 member countries and are included in the OECD annual publication "Society at a Glance"[28] and the groups, reported in Table 1, represent the domains of general socio-demographic context, self-sufficiency, equity, health and social cohesion; out of the five groups, at least four clearly imply an underlying construct. At a later stage of this work, the first group will lose a couple of variables, and those left will be clearly related to the sphere of demography, so from this point on we will use the more accurate term "demography" (dem) to label the group.

Notably, the social cohesion group includes an indicator which is a national average of survey responses concerning life satisfaction on an 11-step ladder (from 0 to 10). It comes from the Gallup World Poll and is one of the standard reference indicators for SWB, even though there are terminological discrepancies in literature concerning the equivalence of terms like subjective well-being, life satisfaction and happiness (Frey and Stutzer 2005; Jorm and Ryan 2014; Welsch 2007). This specific indicator is a very appropriate choice as the response variable in our models. The reference year for the SWB indicator is 2012 and this dictates the reference year for the entire work. Actually, each indicator has its own national collection schedule and, consequently, in some cases, the latest data are older or more recent than 2012. We expect this to have negligible effects on our results as the indicators will be converted in ranks. Occasional missing data is

---

[28] http://www.oecd.org/social/statistics.htm . See the publication for definitions and methodology.

handled with pairwise deletion in all correlation analyses; as for model inputs, missing ranks (due to missing data) are estimates based on the ranking of the same country in the most correlated variable of the same group.

Table III-1. The OECD database of Social Indicators (OECD-DSI)[a]

Context (Dem)	**Self sufficiency (Selfsuf)**	**Equity**	**Health**	**Social cohesion (Soccohes)**
Household disposable income [b]	Employment [e]	Income inequality [b]	Life expectancy [c]	Tolerance [d]
Fertility [c]	Unemployment [e]	Poverty [b]	Perceived health [c]	Confidence in institutions [d]
Net migration [c]	NEET Youth [d]	Living on benefits [c]	Suicide [c]	Safety and crime [d]
Marriages and civil partnerships [b]	Expected years in retirement [d]	Social spending [e]	Public health expenditure [c]	Helping others [d]
Old age support rate [d]	Education spending [b]	Recipients of out-of-work benefits [b]	Public coverage for healthcare [c]	

[a] All indicators are at the latest update as in the mentioned 2014 OECD publication.
[b] Reported year is 2010
[c] Reported year is 2011
[d] Reported year is 2012
[e] Reported year is 2013

The SWB indicator has some remarkable characteristics that should be mentioned. Observed values in reference year vary between 4.7 and 7.8 and, if ordered, the average difference between consecutive countries is below 0.1. In the entire time series between 2001 and 2014, values are confined between a minimum of 3.9 in 2012 Greece and a maximum of 8.4 in 2014 Denmark and very few countries ever fall outside of the interval between 5 and 7; yearly changes of more than 0.5 are very rare (less than 2%); only 7% of the countries changed by more than 1 point in 13 years and 66% of countries changed by less than 0.5.

In other words, national life satisfaction is extremely persistent; in fact, in most countries, its partial autocorrelation is statistically significant at one or more time lags.

There is ongoing debate on whether SWB measures are comparable across different cultures and languages (Jorm and Ryan 2014; OECD 2013). Actually, the evolution of SWB in OECD countries looks very much as if it depended on some factor that is approximately constant in the timespan of a decade, precisely like culture and language. However, several authors (Ed Diener and Oishi 2000; Ferrer-i-Carbonell and Frijters 2004) claim that there is enough evidence to assume that a 1-step distance on a 11-step Likert scale is meaningful in cross-language and cross-culture comparisons. Finding a solution to this dilemma is beyond the scope of our work; in any case, existing literature provides some evidence that the bias is not extensive (OECD 2013) in micro data, at a national level between groups, and up to a 1-step difference for a discrete variable, but there is no ground for assuming that validity holds in the case of differences of about 0.1 on a continuous variable. We argue that, even if language and culture were not cause of a significant bias, fractional differences across nations should be handled with great care by social scientists. Therefore, we will convert the SWB variable into:

a) an approximately equal-width discrete variable with three levels: one for scores of 7 and above, one for scores of 6 and above but below 7 and one for scores of 5 or above but below 6, plus Hungary (4.7);
b) an equal-width discrete variable with seven levels, with each level of width 0.5, starting from 4.5.

The results we present in the paper show no sensitivity to the number of levels in terms of coefficient signs of statistically significant variables and minor sensitivity in terms of which coefficients are statistically significant. As further evidence that the discretization is not a source of bias, linear regressions using the continuous variable, and stepwise procedures to identify the strongest regressors in such a linear regression provide results that are consistent with our models.

Furthermore, we want to control for the possibility that current differences in SWB between countries are the result of events long past. Consequently, we will use a lag variable called swb_0509 which consists of observed values of SWB between 2005 and 2009 as reported in the 2010 Gallup Global Wellbeing report.

2.2 Environmental indicators

Environmental indicators (OECD-DEI) are built from the OECD online data library[29]. In this case using the entire database was not possible: a number of indicators and entire categories are difficult to compare across countries, not easily relatable to SWB or too thematically narrow. Furthermore, the quality of data is unquestionably lower of that in OECD-DSI because of missing data. As a consequence of the limitations of this dataset we had to drop a country (Israel) and some countries have one or more indicators that were last updated several years ago. In spite of this, it was possible to build two additional groups of indicators that deserve to be part of a multidimensional study of SWB (Table 2).

Table III-2. The OECD database of Environmental Indicators (OECD-DSI)

Pollution and consumption of non-renewables	**Green policy and green production**
Population exposure to air pollution by fine particles (PM2.5) [b]	Environmental taxation (Per Capita) [f]
Pesticides produced (per unit of GDP) [c]	Renewable energy production (%)
Municipal waste (Per Capita) [d]	Wastewater treatment (% connected) [g]
Freshwater abstractions (Per Capita) [e]	Organic farming (% agricultural land under certified management) [h]

[a] All indicators are relative to 2013 unless otherwise indicated

[b] Three years average of reference years 2010-2012

[c] Reported year is 2010 except for Luxembourg (1999), for Australia, Canada, Greece and Spain (2006), for the United States (2007), for Chile and Iceland (2008), for Japan and New Zealand (2009)

[d] Reported year is 2012 except for Australia and Chile (2009), for Japan (2010) and for Canada (2009, household municipal waste only)

[e] Reported year is 2012 where not otherwise specified; Austria, (1995), Iceland, South Korea (2005), Finland (2006), Norway, Portugal (2007), Italy (2008), Belgium, Greece, Ireland, Japan (2009), Germany, New Zealand, Sweden, Turkey, United States (2010), Australia (2011), Czech Republic, Luxembourg, Mexico, Poland, Slovak Republic, Slovenia (2013)

[f] On revenue of fiscal year 2012, reported year for Greece is 2011

[g] Reported year is 2004 (Australia, estimate), 2008 (United States), 2009 (Belgium, Canada, Italy, Portugal), 2010 (Germany, Iceland, United Kingdom), 2011 (Chile, Ireland, Japan), 2012 (Austria, Estonia, Greece, Netherlands, New Zealand, Slovak Republic, Spain, Turkey)

[h] Reported year is 2007 (Greece), 2008 (Italy), 2012 (Austria, Canada, Chile, Denmark, Iceland, Mexico, South Korea, Spain)

---

[29] http://www.oecd-ilibrary.org/environment/data/oecd-environment-statistics_env-data-en

The first group consists of four indicators that measure different forms of pollution and consumption of non-renewable resources. The second group measures the intensity of environmental policies and some kinds of "green" production. The need for two different groups is rather evident: based on the existing literature on SWB and environment, we expect the first group to be inversely related to SWB whereas the second should be directly related; we also anticipate the first group to strongly favor poor and developing countries while the second tilts the analysis towards the developed countries.

2.3 Time-use indicators

The database of time use indicators (OECD-DTUI) was extracted from the OECD gender data portal[30]; methodological documentation on the national time use surveys from which this OECD database is built can be found in Miranda (2011).

Table III-3. The OECD database of Time Use Indicators (OECD-DTUI)[a]

Uses compatible with production of relational goods	Uses non compatible with production of relational goods
Care for household members	Travel to and from work/study
Participating /attending events	Routine housework
Visiting or entertaining friends	Sleeping
Eating and drinking	Tv or radio at home

[a] Reported year for Portugal is 1999, for Estonia and Hungary is 1999-2000, for Slovenia is 2000-2001, for Denmark is 2001, for Germany is 2001-2002, for Poland is 2003-2004, for Belgium, Ireland and United Kingdom is 2005, for Netherlands is 2005-2006, for Australia and Turkey is 2006, for Austria and Italy is 2008-2009, for France, Korea and Mexico is 2009, for Finland, New Zealand and Spain is 2009-2010, for Canada, Norway, Sweden and United States is 2010, for Japan is 2011.

---

[30] http://www.oecd.org/gender/data/

In this case we have a further drop in data quality because of missing data and because the results are only approximately comparable, as the primary sources are national surveys. The database is nonetheless the standard reference for scientific literature on time use. As in the case of the OECD-DEI, the OECD-DTUI is organized in groups that are not particularly useful for our research, mixing indicators of relational use of time with others that suggest non-relational use or have no implications on relationality. Table 3 presents our reorganization of some indicators in two groups that seem the most effective to represent relational and non-relational time use. These indicators are available for a total of 26 OECD countries.

3. Methods

3.1 Overview of the methodology

Our intent is to transform the available datasets of indicators into a compact group of meaningful and clearly interpretable regressors and to explore models of SWB based on them. In order to avoid endogeneity issues, which quickly become uncontrollable in highly multidimensional models, we perform a dimensional reduction, shrinking the set to a manageable number of variables. However, in doing so, we neither use weighted linear aggregations, which imply arbitrary weightings, nor the common data-driven methods of dimensional reduction, which frequently incur in the issue of having, as output, nameless and artificial dimensions that cannot justify one or few specific policy strategies over others. We instead apply a method, derived from the theory of partially ordered sets, which requires no assumptions about the importance of the indicators aggregated in a single dimension and, at the same time, produces a synthetic dimension which has no ambiguity of interpretation because it is, as much as possible, order-preserving with respect to ranks built on all the indicators synthesized in it. As explained in the following subsection, the main requirement for this process is that there is a relatively consistent latent dimension underlying each group. From this, we obtain a set of variables that form a plausible model of policy-controllable drivers of SWB; we compare the predictions of national average SWB made with such model with those that can be made using GDP alone. This process will be organized according to the steps presented in Table 4.

Table III-4. A step-by-step scheme of the methodology

Step 1	**Analysis and identification of adequate partial orders**	Hasse diagrams and partial order scalogram analyses with base coordinates are used jointly, as graphical tools, to study if each group of indicators (see Table 1-3) can be dimensionally reduced to a single dimension representing the concept in the group label (e.g. self-sufficiency). If the graphical representation of the new dimension shows a reasonable amount of partial order, the process moves to the following step. If not, the output is used to identify which indicator is most inconsistent with the existence of a single latent variable (based on weak monotonicity coefficients) and we drop that indicator
Step 2	**Definition of the final set of independent variables**	The groups resulting from the previous phase are finally transformed into a final set of independent variables using local partial order models
Step 3	**Estimation of ordered logit models**	Discretized SWB is estimated across OECD countries using the set of variables defined in the previous step as regressors in ordered logit models

The first step uses partial order tools designed to allow dimensional reduction, but its actual goal is to study if each group of indicators (corresponding to each column in Tables 1-3) allows to create some partial ordering of countries. That ordering is, in fact, a latent variable behind the group of indicators. The interpretation of such latent variable is relatively easy: it can be identified with the concept expressed in the name of each group. When the ordering is insufficient, a variable is identified as the least consistent with the existence of a single latent variable and is dropped; then the procedure is repeated. It should be noted that, at this stage, partial order tools are not used to obtain the quantitative value of the latent variable for the following analysis: we are just interested in a graphical representation of the partial order.

The second step consists of performing dimensionality reduction through a different partial order method. The OECD-DSI will be reduced to a group of five indicators, following the OECD classification, whereas the OECD-DEI will be reduced to two indicators following our own interpretation of the available data and the OECD-DTUI will be reduced to two indicators that best represent the concepts of production/consumption of relational and non-relational goods. The third step will consist of modelling and model comparison.

3.2 Partial Order Methods

In Partial Order theory, a set is defined as "linearly (or totally) ordered" when all the pairs of its elements are linked by an order relation. More generally, we have a "partially ordered set" (or POSET) when some pairs of its elements are linked by an order relation, but not necessarily all of them. That is the case with many multi-indicator systems, when their purpose is to provide a single ranking for a set of elements defined by their scores in several different attributes. In such a set, two elements can be ordered if all the attribute scores of one element are higher of all the attribute scores of the other. Otherwise, the two elements are called "incomparable" and, because of their presence, a complete ranking of the set to which they belong would be impossible without further elaboration.

A great variety of methods is available to force some complete ranking on a POSET, from composite indicators to Principal Component Analysis and Multidimensional Scaling. However, there are remarkable benefits in deliberately applying, instead, Partial Order Methods both for the analysis and for the definition of a complete ranking; a comprehensive overview of the available tools can be found in Brüggemann and Patil (2011), whereas this paper will focus on those that are more relevant for our research issue.

1. Partial Order Methods like Local Partial Order Model (LPOM) allow dimensionality reduction in the multi-indicator system down to a single dimension, exactly like weighted linear aggregations and other methods to produce composite indicators, but they require no arbitrary weighting; at the same time, tools like Hasse diagrams and partial order scalograms provide substantial awareness and control over compensation;
2. Compared to common data-driven methods to reduce dimensionality like Principal Components Analysis and Multidimensional Scaling, LPOM is just as effective in reducing multicollinearity, sensitivity to model specification and overfitting issues in

models built on multi-indicator systems, but it produces more interpretable, less ambiguous outputs;

3. Partial order can be applied without substantial *a priori* assumptions on the direction of causality between observed dimensions and output dimensions, which means that it is appropriate in reflective models, where multiple indicators are measured to make inferences on an underlying latent variable influencing them, as well as in formative models, where a convenient combination of multiple indicators is used to generate a synthetic measure[31].

Using Partial Order Methods on multi-indicator systems implies two assumptions. The indicators included in the system should ultimately refer to one underlying construct, either reflectively or formatively. In other words, they should "sample, or cover, the different aspects of a well-defined content universe" (Raveh and Landau 1993). A ranking based on indicators that have no common content universe (e.g. a ranking of countries based on average income and number of colors in their flag) may still be possible but would have no meaning. Furthermore, the relation between indicators and the underlying construct must be uniform in its direction from high to low (Shye 1985). The first assumption works very well with the OECD-DSI: the 24 social indicators are organized in five groups, at least four of which can be easily referred to an underlying construct. The other two datasets were selected so that four more groups with as many underlying constructs could be identified in the domains of environmental and time-use data. The second assumption holds only in some cases and occasionally we will have to adjust the indicator to make directions uniform with respect to the underlying construct.

Dimensionality reduction, regardless of the method of choice, can significantly benefit from some preliminary analytical steps allowed by Partial Order Methods, the first being the Hasse diagram (Brüggemann and Patil 2011).

Suppose that matrix $P_{n \times m}$ describes a POSET of n elements with m attributes $a_1 \dots a_m$. Each element k can be represented by its profile $a^{(k)} = a_1^{(k)}, a_2^{(k)} \dots, a_m^{(k)}$and, given the definition of

---

[31] See Edwards et al. (2000) for a definition of formative and generative models in the development of constructs (indicators).

partial order, $a^{(k)} < a^{(h)}$ if $a_t^{(k)} < a_t^{(h)}$ ∀ 1 < t < m. Furthermore, profile $a^{(h)}$ is said to cover $a^{(k)}$ if there exists no third profile $a^{(z)}$ in the POSET for which $a^{(k)} < a^{(z)} < a^{(h)}$. A Hasse diagram represents each profile of the POSET as a node of a directed graph where all edges are directed upwards (and the direction can therefore be omitted) and is built according to the following rules:

1) If $a^{(k)} < a^{(h)}$, then $a^{(k)}$is below $a^{(h)}$in the drawing;
2) If $a^{(h)}$ covers $a^{(k)}$, an edge is included between the two.

Generally, profiles are arranged vertically in levels that depend on the structure of the graph and, when a profile could be located in several levels, the highest possible is selected. Figure 1 presents an example of Hasse diagram with just six profiles: $a^{(g)}$ covers $a^{(k)}$, $a^{(h)}$, $a^{(l)}$ and $a^{(r)}$, all of which cover $a^{(s)}$.

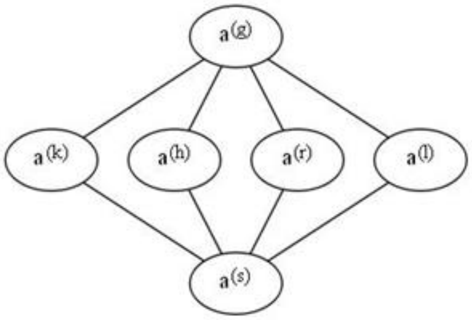

FIGURE 1 Hasse diagram

Notably, in this example four profiles are not connected by an edge and lay on the same vertical level. This is the representation of incomparability in Hasse diagrams: we know that the three profiles are below $a^{(g)}$ and they are above $a^{(s)}$but, based on available data, we cannot decide an order among them. As exploratory tools, Hasse diagrams are particularly helpful because of the number of levels they identified in a POSET. In fact, a POSET with n elements and n levels is a totally ordered set and the interdependence between the attributes is very high; if the POSET of a multi-indicator system has almost as many levels as elements, the multi-indicator system is redundant as most of the information would be still available in a system with just one indicator. At the opposite, if a multi-indicator system has a POSET with n elements and 1 level (with $n > 1$), all of its elements are incomparable and all of its indicators are not co-graduated; more generally, a POSET with very few levels provides little information on orderability and it is not particularly useful for the purpose of dimensional reduction with any dimensional reduction method. Based on these premises, before proceeding with dimensional reduction on the three datasets, we will make sure that each group of indicators to be aggregated has a POSET with a number of levels below n and reasonably greater than 1.

Hasse diagrams provide useful visual information on the amounts of comparability in a POSET, but little in terms of interpretation. Ideally, before moving on with dimensionality reduction, we want to know more, particularly about incomparability: what makes specific pairs of profiles incomparable and how much incomparable they are. This is possible through a second Partial Order tool called POSAC or Partial Order Scalogram with base Coordinates[32].

POSAC is a tool for exploratory data analysis that extends the unidimensional Guttman scale (Guttman 1950). It provides a graphical representation of the same matrix $P_{n \times m}$ discussed above and consisting of the n elements of a POSET and their m attributes $a_1 \ldots a_m$ (with m>2) in a two-dimensional space. While POSAC is, in itself, a method of dimensionality reduction, it is designed to provide an interpretable graphical representation rather than a formal description of the two resulting dimensions, and therefore it is mostly useful in the exploratory phase of the analysis. Compared to Principal Components Analysis, instead of trying to preserve distances, POSAC tries to preserve, as much as possible, order relations and incomparabilities between the elements (Brüggemann and Patil 2011).

First, in each of the m columns of $P_{n \times m}$, observed data are converted in ranks[33]. From this point on, only the relations between the observations are considered and not their magnitude, making the method robust to outliers and stable in front of slight changes in input data, and making different scales between attributes irrelevant. Conventionally, POSAC ranks observations so that small scores are assigned a high ranking, but the opposite result can be easily obtained and, as clarified in the results section, associating small scores with low rankings provides more intuitive

---

[32] A detailed formalization of the mathematical background and the algorithm of POSAC can be found in Shye and Amar (1985) and Shye (1985) and the algorithm it is currently supported in the statistical package Systat.

[33] At this stage, indicators can be ranked so that the direction of their relation with the underlying construct is uniform.

results and is therefore the method of choice in this work. Essentially, the POSAC algorithm calculates the weak monotonicity coefficients of all attributes and identifies the two attributes a_i and a_j that have the least positive correlation. Afterwards, each element k is assigned a profile $a^{(k)} = a_1^{(k)} \dots a_i^{(k)} \dots a_j^{(k)} \dots, a_m^{(k)}$ consisting of its ranking in each attribute. The maximal and minimal profiles, determined as:

$$a^M \equiv \max_{p \in P} a_1^{(p)} \max_{p \in P} a_2^{(p)} \dots \max_{p \in P} a_n^{(p)}$$

and

$$a^m \equiv \min_{p \in P} a_1^{(p)} \min_{p \in P} a_2^{(p)} \dots \min_{p \in P} a_n^{(p)}$$

where P is the observed population, are added to the set of observed profiles and placed on a Cartesian space in (1, 1) and (0, 0) respectively. The initial placement (x_k, y_k) of profile $a^{(k)}$ is subsequently determined by solving the following set of simultaneous linear equations:

$$\begin{cases} x_k + y_k = \sum_{t=1}^{m} a_t^{(k)} & (1) \\ x_k - y_k = a_i^{(k)} - a_j^{(k)} & (2) \end{cases}$$

The coordinates are then appropriately translated to fit into [0, 1]. Finally, the placement is improved through iterative two-step cycles that approximate the maximization of the number of profile pairs which are correctly represented in their order or incomparability relation (Shye 1985).

From the definition of orderable pair, we know that

$$a^{(k)} > a^{(h)} \leftrightarrow x_k > x_h \text{ and } y_k > y_h \qquad (3)$$

$$a^{(k)} \parallel a^{(h)} \leftrightarrow \begin{cases} x_k \geq x_h \text{ } and \text{ } y_k \leq y_h \\ or \\ x_k \leq x_h \text{ } and \text{ } y_k \geq y_h \end{cases} \tag{4}$$

So, by construction, we expect from (1) and (3) that profiles with larger scores in all attributes will be placed closer to (1,1), towards the top right corner of the POSAC space, whereas profiles with smaller scores in all attributes will be closer to (0, 0). Form (2) and (4), instead, we can see how POSAC deals with profiles that have relatively large scores in some attributes and small scores in others and are, therefore, incomparable. Profiles with large scores of a_i and small scores of a_j will be placed close to the bottom right corner and, at the opposite, small scores of a_i and large scores of a_j will have them placed closer to the top left corner. Graphically, as in Figure 2, order and incomparability with respect to one profile $a^{(h)}$ can be identified for each other profile by looking at its position in a new Cartesian space with origin $a^{(h)}$.

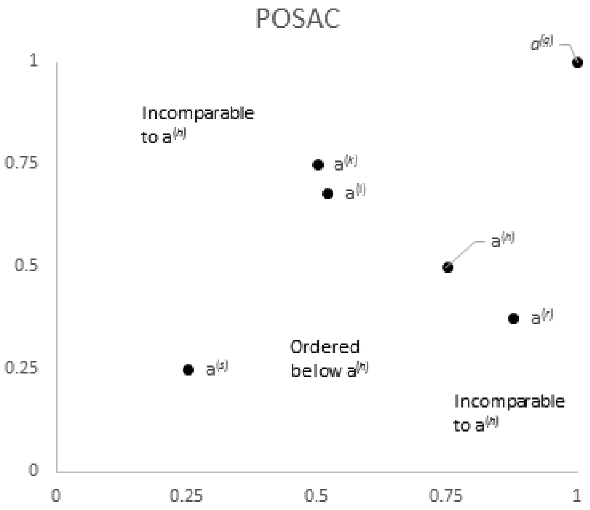

FIGURE 2 POSAC of partially ordered set represented in figure 1

A further support to POSAC output analysis, particularly when an entire set is considered, is given by the following:

$$J = X+Y \tag{5}$$

$$L = X-Y \tag{6}$$

Partial Order literature refers to the J (or Joint) axis as the diagonal from (0, 0) to (1, 1) and the L (or Lateral) axis as the diagonal from (1, 0) to (0, 1). The two equations (5) and (6) describe the actual output dimensions of POSAC[34], and the score of each profile on J and L is hence determined. The J axis has some substantial advantages over the output of a PCA: there are no factor loadings to consider in size or direction: if two comparable profiles have a different score in J, the one with the highest score has better scores in *all* the original indicators. Visually, groups of comparable profiles will be broadly aligned with the J axis and POSAC outputs that show general alignment with J are closer to an ordered set. At the opposite, ample alignment with L identifies a set that is largely not orderable.

As we did with Hasse diagrams before, we can now point to the features of POSAC that are most relevant to this research. First, the position of profiles along the J axis is a good visual approximation of a total ranking. Second, the shape taken by the cloud of profiles along the J and the L axes can be used to cross-check indication on excess or lack of comparability in the POSET. Third and most importantly, some remarkable information on the nature of incomparability that could not be seen in the Hasse diagram in Figure 1 can be inferred from the graph: now we can say that $a^{(k)}$ and $a^{(r)}$ are not just incomparable, but also very different because they have opposite rankings in the two least positively correlated attributes. Obviously, by looking at the coefficients of weak monotonicity, we can identify which are the attributes that create such a difference. At the same time, $a^{(k)}$ and $a^{(l)}$, while incomparable, are much less different and form a cluster of profiles with relatively similar characteristics compared to the others.[35]

A convenient application of Partial Order Methods to reduce a multi-indicator system to a single dimension is represented by the family of Local Partial Order Models (LPOMs) (Brüggemann and Patil 2011). Such models make use of the concepts of linear extension of a POSET and average height of a profile.

---

[34] Clearly, by determining the coordinates (x, y) of each profile, POSAC attributes that profile a score on the J axis and a score on the L axis.

[35] In a Hasse diagram, the four incomparable profiles can be put side to side in any configuration and the method itself does not provide the information required to distinguish between the incomparability of $a^{(k)}$ and $a^{(r)}$ and that of $a^{(k)}$ and $a^{(l)}$

Given a POSET P, like that represented as a Hasse diagram in Figure 1, a linear extension L of P is any linear order, of the same attributes of P, which preserves all the order relations found in P. In the case of P (Figure 1) the POSET contains nine order relations between profiles: $a^{(g)} > a^{(s)}, a^{(k)}, a^{(h)}, a^{(l)}, a^{(r)}$ and $a^{(s)} < a^{(k)}, a^{(h)}, a^{(l)}, a^{(r)}$. Essentially, a linear extension will be obtained by attributing any order to all incomparable profiles $a^{(k)}, a^{(h)}, a^{(l)}, a^{(r)}$ provided that the order relations with $a^{(g)}$ and $a^{(s)}$ are preserved (see Figure 3). Any partial order that is not a linear order will therefore have more than one linear extension and the average position of a profile in all linear extensions will depend on (and will carry substantial information about) the structure of the POSET. Formally, if $H(a^{(k)}, L_v)$ is the height of a profile $a^{(k)}$ in the v-th linear extension, the average height of $a^{(k)}$ is:

$$hav\left(a^{(k)}\right) = \frac{\sum_{v=1}^{LT} H\left(a^{(k)}, L_v\right)}{LT} \tag{7}$$

where LT is the number of all linear extensions in the POSET (Brüggemann and Patil 2011). If $hav\left(a^{(k)}\right)$ is a known quantity, the POSET can be linearly ordered according to it as it would happen with a composite indicator but without any need for arbitrary weights (Brüggemann and Carlsen 2011).

However, the number of all linear extensions in a POSET is $1 \leq LT \leq n!$ (Brüggemann and Carlsen 2011) and determining all of them is a combinatorial exercise (Lerche et al. 2003) that quickly becomes computationally impossible as n grows (Brüggemann et al. 2004).

In order to overcome this issue, several methods to approximate $hav(a^{(k)})$ have been proposed[36] among which LPOMs, that are being progressively developed to allow approximations with the least possible distortion effects (Brüggemann and Carlsen 2011). For the sake of simplicity and given the findings in Brüggemann and Annoni (2014) we will make use of the first LPOM to be developed, called LPOM0 (Brüggemann et al. 2004), which was empirically shown to be not necessarily a worse approximation than the others. Looking at the POSET in figures 1, 3 as an example, the idea behind the approximation is the following: the average height of, say, profile

---

[36] For an extensive list with detailed comparison see Brüggemann and Patil (2011) and Brüggemann et al. (2005).

$a^{(k)}$, depends first of all on profiles that are certainly ordered above it $\left(a^{(g)}\right)$ and on profiles certainly ordered below it $\left(a^{(s)}\right)$. The precise order of profiles above and profiles below is irrelevant, i.e. any linear extension of the subsets above and below $a^{(k)}$ will produce the same outcome. Then, the average height of $a^{(k)}$depends on the method selected to merge incomparable profiles $a^{(h)}, a^{(l)}, a^{(r)}$ with the others. The approximation consists in treating all the incomparable profiles as a single object that can be assigned to any position above or below $a^{(k)}$.

Let $\left(a^{(k)}\right)$ and $F\left(a^{(k)}\right)$, or O and F for brevity, be the principal downset and upset of $a^{(k)}$, i.e., respectively, the subset of the POSET P containing $a^{(k)}$ and all $a^{(h)} \leq a^{(k)}$ or all $a^{(h)} \geq a^{(k)}$. Furthermore, let $U\left(a^{(k)}\right)$, or U, be the subset of P containing all elements incomparable to $a^{(k)}$. Finally, following a common convention in LPOMs literature, let $\|P\|$ denote the number of profiles contained in the POSET P. LPOM0 approximates the height of each profile to a weighted average of its height when the block of incomparable profiles are merged above it, and a weighted average of its height when the block of incomparable profiles is merged below it.

$$hav\left(a^{(k)}\right) \cong \frac{\|O\|}{\|O\|+\|F\|}\left(\|O\|+\|U\|\right)+\frac{\|F\|}{\|O\|+\|F\|}\|O\| \tag{8}$$

In (8), the first product represents the height of $a^{(k)}$ if U is merged below it (the sum of profiles in O and U) and is weighted by the probability of such arrangement, that is, the share of all spots available for U that are actually below $a^{(k)}$. Similarly, the second product represents the height of $a^{(k)}$ if U is merged above it (just the number of profiles in O) weighted by the share of spots available above $a^{(k)}$. For a discussion on the possible distortions of this approximation see Brüggemann and Carlsen (2011).

The quantity in (8) is, therefore, our partial order-derived synthetic indicator for each group of indicators identified in the section about data. LPOM0 of user defined datasets can be performed online on www.PyHasse.org.

3.3 Modeling and model comparison

The transformation of data on national average SWB in OECD countries (see section on data) made our response variable an ordinal discrete variable with three or seven response categories.

The covariates built with LPOM0 in the previous section are continuous variables. An appropriate regression model for ordinal dependent variables is the ordered logistic regression or proportional odds model (McCullagh 1988), provided that the relationship between all pairs of outcome groups is the same (hence, the proportionality of odds); in that case, a single set of regression coefficients can be used to describe the whole set. Some tests to check if this assumption holds are available, including the score test (Greene 2011) the Wolfe-Gould test (Wolfe and Gould 1998) and the Brant test (Brant 1990), but they are known to be "anti-conservative" and to frequently lead to reject the assumption (O'Connell 2006, p. 26), particularly when the circumstances are not ideal, like in the case of small sample size or when the set of covariates is large. If the assumption does not hold, the multinomial logistic model (Greene 2011) estimates a different set of coefficients for each outcome group. Given the inherent complexity of interpretation of the multinomial logistic model, the models presented in the results section are, in general, ordered logistic models. For each model, we provide the results of the score test, the Wolfe-Gould test and the Brant test, in order to discuss their compliance with the proportionality of odds assumption. All the main models in tables 5-8 have also been estimated as multinomial logistic models to cross-check the sign and the statistical significance of the coefficients; for the sake of brevity, the results are not provided, but any relevant discrepancy with the ordered logistic models will be mentioned.

4. Results

4.1 Preparing and performing dimensional reduction

In this subsection, we present the results of the first two steps of our work. In order to keep our exposition as compact as possible, we will only include the Hasse diagrams and the POSAC outputs for those groups of indicators that could not be reduced to a single dimension without dropping some variables. Hasse diagrams and POSAC outputs for the remaining groups of indicators can be found in the supplementary material. Figure 3 and 5 represent, respectively, the POSAC output and the Hasse diagram of 34 OECD member countries based on their profiles in five indicators that the OECD calls "context indicators". As anticipated earlier, this is the least homogeneous group in the OECD-DSI and, as expected, incomparability is very high: we have just three levels for 34 profiles, several countries are incomparable to all others and the POSAC output is mostly aligned along the L axis. One could argue that there is no single underlying construct linking all five of these indicators. We drop household disposable income and net migration from the group on the grounds that they are significantly correlated and they describe a dynamic that is probably captured just as well by GDP. The resulting set of three indicators is more homogeneous: fertility, marriages and civil partnerships and old age support rate together provide hints on some important demographic characteristics of a country (e.g. about age distribution) and those demographic characteristics in turn indicate some sort of potential sustainability of formal and informal welfare systems and, therefore, sustainability of income in the log term. Figure 4 and figure 6 show the POSAC output and the Hasse diagram of this reduced set of three indicators. We now have seven levels and a more complex POSAC output with some alignment to the J axis as well. Remarkably, some strong and apparently solid economies like Germany, Denmark, Austria and Switzerland get very poor scores. This result will be commented later, together with the results of other variables and the estimated models.

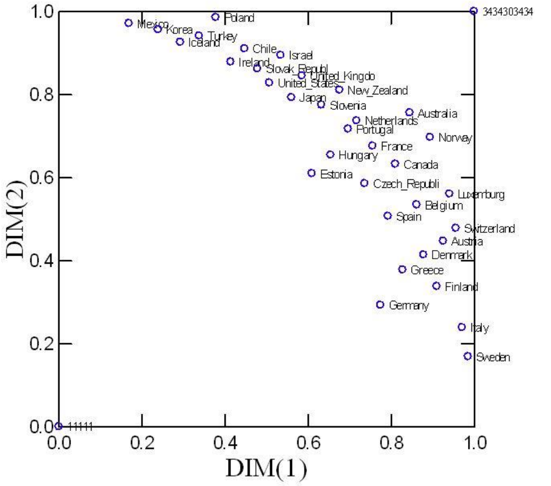

FIGURE 3 POSAC of OECD member countries based on five "Context indicators"

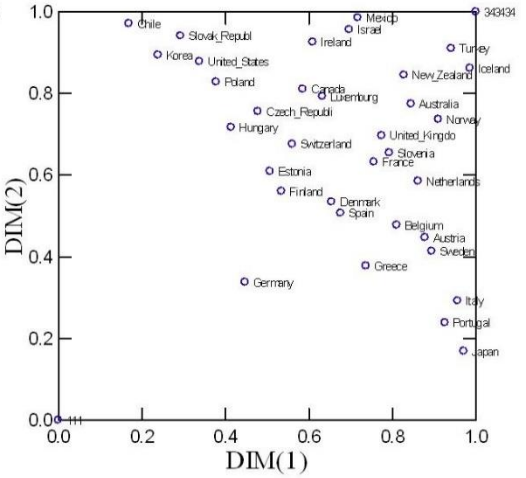

FIGURE 4 POSAC of OECD member countries based on three "Context indicators"

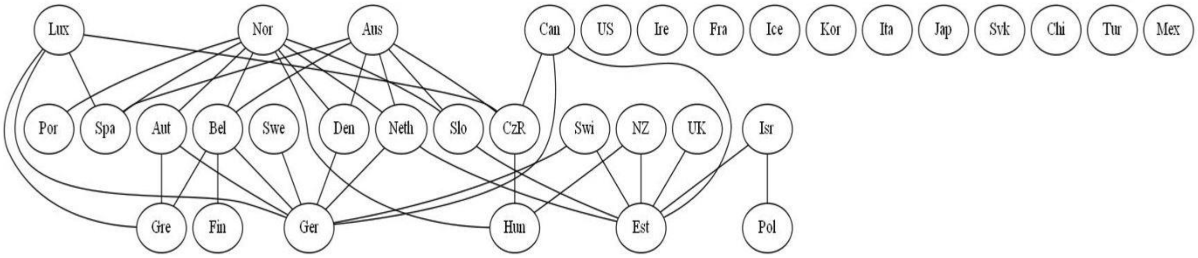

FIGURE 5 Hasse diagram of OECD member countries based on five "Context indicators"

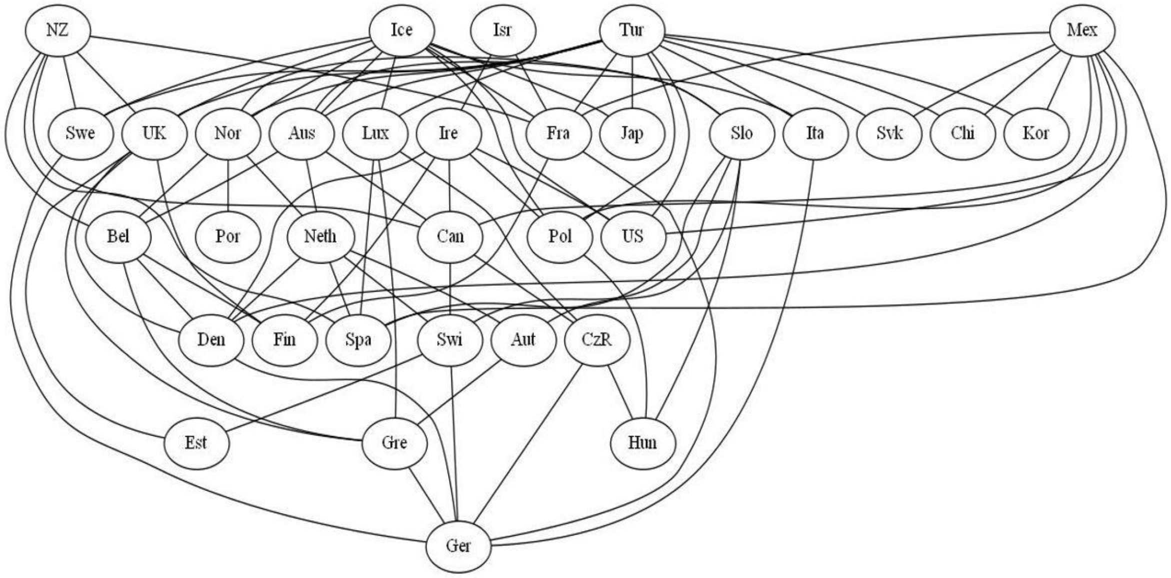

FIGURE 6 Hasse diagram of OECD member countries based on three "Context indicators"

Profiles ordered according to self-sufficiency indicators[37] show reasonable comparability, are organized around four levels[38] and, consequently, a complete ranking based on self-sufficiency as defined in the OECD-DSI appears reasonable. However, the degree of incomparability should not be overlooked: in our ranking, among the countries with average-to-low levels of self-sufficiency, some allow long retirement periods but have high unemployment rates (e.g. Greece, Spain, Italy, France and Belgium) while others have reasonable unemployment rates but allow a very short retirement period (Mexico, Chile, Korea or Japan). A complete ranking will therefore assume that, based on the OECD indicators, these two factors can compensate each other in terms of resulting self-sufficiency.

The five equity indicators in OECD-DSI produce a good degree of comparability: we have six levels for the 34 countries in the Hasse diagram[39], and the POSAC output shows long chains of comparable profiles organized in three or four groups[40]. Broadly speaking, some countries (like Slovak Republic, Slovenia, Poland and Korea) have low levels of income inequality and poverty, but only 57% of their unemployed population, on average, qualifies for out-of-work benefits. At the opposite, Australia, Spain, Portugal and the United States have greater income inequality and poverty but, on average, 84% of their unemployed population get the benefits. All in all, the five equity indicators can reasonably be reduced to a single dimension.

---

[37] See Figure 1 of supplementary material.

[38] See Figure 3 of supplementary material.

[39] See Figure 2 of supplementary material.

[40] See Figure 4 of supplementary material.

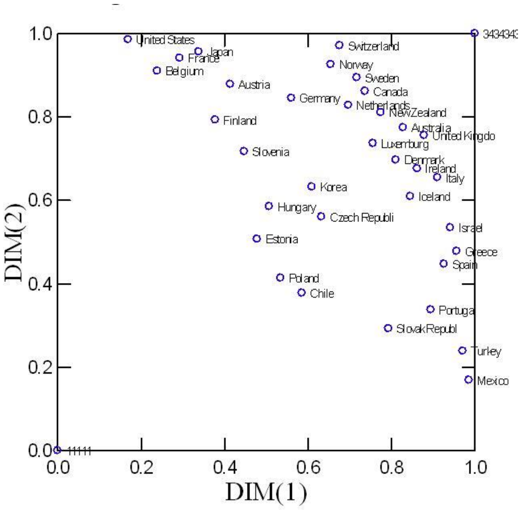

FIGURE 7 POSAC of OECD member countries based on five "Health indicators"

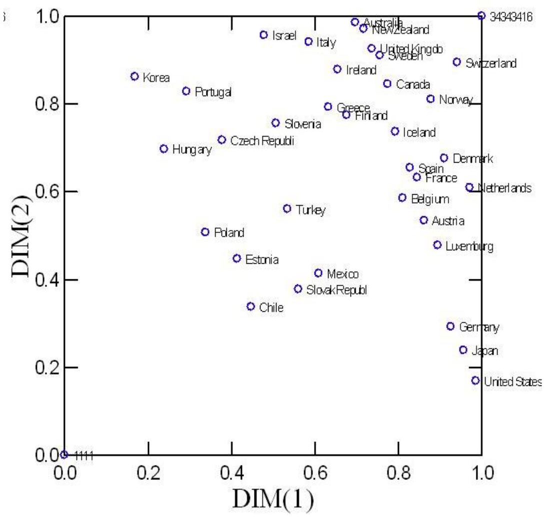

FIGURE 8 POSAC of OECD member countries based on four "Health indicators"

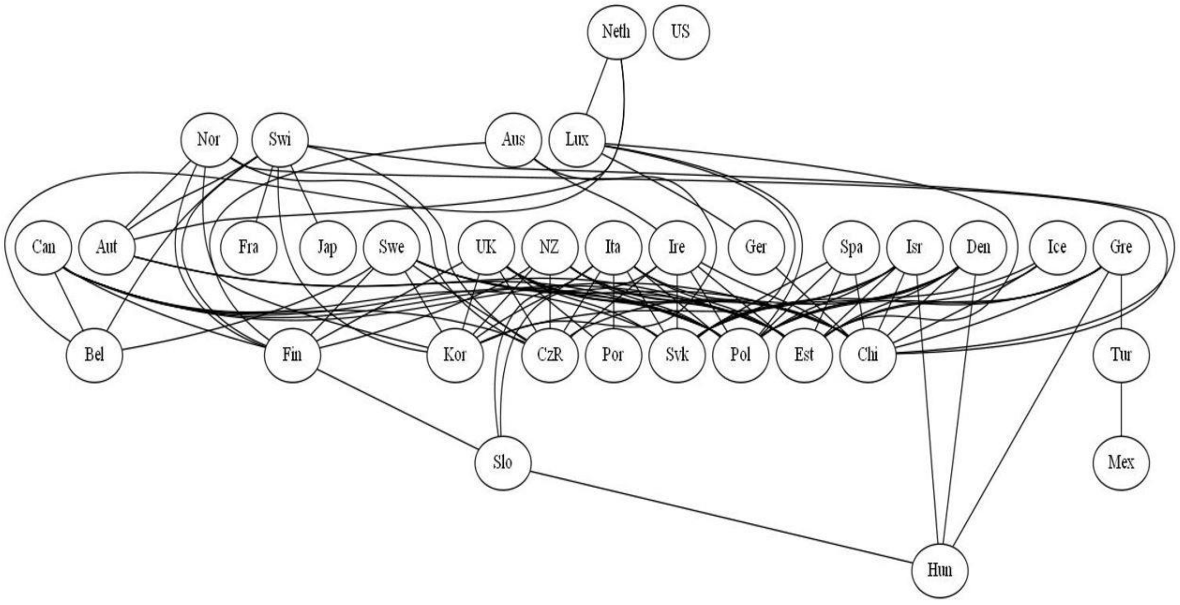

FIGURE 8 Hasse diagram of OECD member countries based on five "Health indicators"

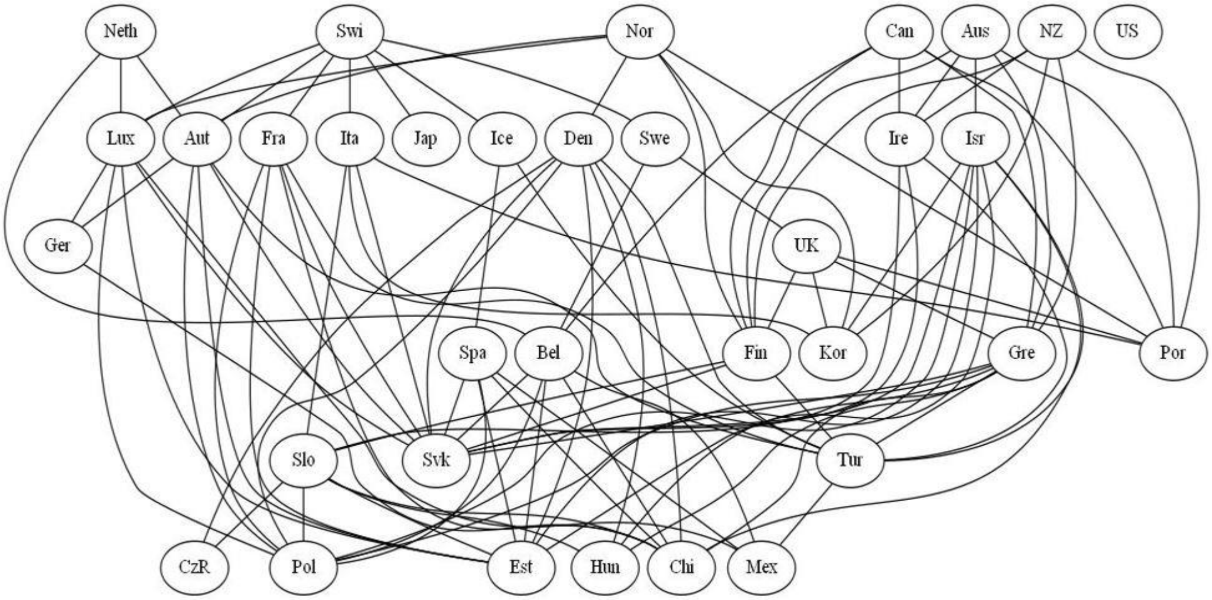

FIGURE 9 Hasse diagram of OECD member countries based on four "Health indicators"

The original set of five health indicators produces six levels in the Hasse diagram in figure 9 but half of the levels have just one or two members and the overall amount of incomparability is high. The POSAC output in figure 7 is correspondingly aligned, in two different groups, parallel to the L axis. We conclude that it is not reasonable to look for a single underlying "health" factor behind all five indicators. The best results in terms of comparability, if we sacrifice one variable, are obtained by dropping suicide rate (figures 8 and 10): the number of levels in the Hasse diagram does not increase, but comparability grows and we obtain long chains of comparable profiles as the new POSAC output illustrates. The main source of the remaining incomparability is due to public health expenditure and public health coverage: countries clustered around the bottom right corner have high public expenditure and, sometimes, low levels of coverage, while countries near the top left corner tend towards universal coverage.The set of four social cohesion indicators[41] can be represented in a Hasse diagram with six levels[42], comparability is rather strong

---

[41] The reader should keep in mind that the social cohesion indicators provided by OECD originate from surveys and report a subjective feeling of the respondent, so it could be appropriate to define this as "perception of social cohesion".

[42] See Figure 6 of supplementary material.

and that is reflected in the POSAC output, provided with the additional material[43]. Two groups of countries with average-to-low social cohesion are less aligned with the J axis: Turkey, Israel, Slovak Republic and Mexico have poor rankings in perception of personal safety from crime and good rankings in confidence in institutions. At the opposite, Slovenia, Iceland, Japan and, to a lesser degree, Spain, have good rankings in personal safety but poor rankings in confidence in institutions. Our interpretation is that these are two different paths for communities suffering from some dysfunction to stay away from the bottom of the social cohesion ranking: soft power (less conflict and less discipline) and hard power (more conflict and more discipline). Their incomparability means that the set of OECD indicators cannot intrinsically lead to choose one model over the other.

The four indicators on pollution and consumption of non-renewable resources (cnrr) can be represented in a Hasse diagram with four levels[44] and the degree of comparability is low. As illustrated by the POSAC output in the supplementary materials[45], a rather large group of countries (Slovak Republik, Czech Republic, Poland, Hungary, Korea and Turkey) have a very limited production of municipal waste but high concentrations of particulate. At the opposite end of the spectrum, Ireland, New Zealand and Denmark have low concentrations of particulate but produce high amounts of municipal waste. The link between the indicators and the underlying latent variable in this case is inconsistent. We will provide an example of a more consistent formulation of this environmental regressor in the supplementary material, appendix A.1. In any case, the two variants of the regressor provide almost identical results in models[46].

The four indicators on green policies and production (gpp) create a Hasse diagram with just four levels[47]. The amount of incomparability is not so high to exclude that this category of indicators can be related to a single underlying construct but, as the POSAC output shows[48], we have to deal with groups of countries with very different situations. A cluster of OECD member states (Korea,

---

[43] See Figure 5 of supplementary material.

[44] See Figure 7 of supplementary material.

[45] See Figure 8 of supplementary material.

[46] See Table 1 of supplementary material

[47] See Figure 9 of supplementary material.

[48] See Figure 10 of supplementary material.

Luxembourg, Netherlands, Australia) have good scores in wastewater treatment and, sometimes, in environmental taxation, but the share of energy produced from renewable resources is low, *ceteris paribus*. Other countries, at the opposite (like Iceland, Chile, Slovenia and Portugal) rely on larger shares of renewable energy but they either treat much less than 70% of their wastewater or they have little environmental taxation or both.

The final remarks concern two sets of indicators of relational and non-relational uses of time. The first set features very high levels of incomparability[49]: based on available indicators, the idea that we can rely on a latent construct about relationality in time use is doubtful. As a matter of fact, there is much more difference in how time use is allocated between different relational activities than in the average amount of time dedicated to relationality in different OECD countries. The set of non-relational uses of time is less affected by incomparability, with a Hasse diagram of four levels[50], but as the POSAC output makes clear[51], yet again the latent construct is barely visible from the available indicators. In the case of these last two sets of indicators, however, we are not going to reduce the number of variables in order to fine-tune the latent construct. As anticipated earlier, the two sets of indicators derived from the OECD-DTUI are presented here with the specific intent of testing the theory on relational goods. Indeed, time-use indicators and their relation with SWB make for an interesting field of enquiry, the development of which is at a much earlier stage compared to those concerning social or even environmental indicators. There is no theory on relationality and SWB to refer to if we find that some relational uses of time affect SWB and others don't or do but in an unpredictable way. Precisely for this reason, we will estimate our models with indicators of relational and non-relational uses of time based on the full set of indicators. Furthermore, while commenting the results, we will make clear that more substantive evidence of a relation between time use and SWB could probably be found with better indicators, better data and a stronger theory.

---

[49] See Figures 9 and 11 of supplementary material.

[50] See Figure 14 of supplementary material.

[51] See Figure 12 of supplementary material.

Table III-5. Pearson correlation coefficients

	TriSWB	**HeptaSWB**	**gdp**
TriSWB	1		
HeptaSWB	0.9499*	1	
gdp	0.6053*	0.6213*	1
dem†	0.4014*	0.3730*	0.1310
selfsuf†	0.7014*	0.7640*	0.7197*
equity†	0.3320	0.3697*	0.5340*
health†	0.6505*	0.6948*	0.5983*
soccohes†	0.7455*	0.7756*	0.7405*
swb_0509	0.8436*	0.8846*	0.5194*
cnrr†	-0.2015	-0.2388	-0.1331
gpp†	0.3060	0.3763*	0.3711*
relational†	0.4805*	0.4929*	0.2723
nonrelational†	-0.4796*	-0.6053*	-0.5595*

* $p<0.05$

Table 5 presents the correlation coefficients between the nine regressors built through dimensional reduction thanks to the method of local partial order model (identified with a dagger) plus one lag term (swb_0509) and three other variables: GDP *per capita* (gdp), a tripartite discrete indicator of SWB (TriSWB) and a heptapartite discrete indicator of SWB (HeptaSWB).

These coefficients are the logical conclusion of our process of dimensionality reduction. The 24 social indicators collected by OECD can be reduced to five, while strictly controlling the levels of comparability; the resulting indicators are generally correlated with the indicators of SWB but the correlation is particularly strong in the case of self-sufficiency and social cohesion and weak or absent in the case of equity. In our preliminary analysis, we had found that average-to-low levels of self-sufficiency and social cohesion, as defined by the OECD indicators, are compatible with different unbalanced policy approaches. In other words, mediocre results in terms of self-sufficiency may come from less unemployment and shorter retirement or more unemployment and longer retirement, while mediocre social cohesion may still be attained by dysfunctional communities exercising soft power or hard power. The OECD indicators do not inherently lead to choose one model over the other in these couples of unbalanced approaches. However, the correlation coefficients with SWB after the dimensional reduction are rather strong, meaning that we found rather substantive evidence that:

1) balanced approaches seem definitely better;
2) both unbalanced approaches to self-sufficiency seem to allow at least some wellbeing;
3) both unbalanced approaches to social cohesion seem to allow at least some wellbeing

A number of environmental indicators have been reduced to two: an indicator of pollution and unsustainable consumption and an indicator of green economy and policy. The correlation with SWB is null or marginal, suggesting that environmental policies might have no perceptible effect on SWB in the OECD countries or, if they have one, it is not captured by a standard set of environmental indicators. Several time use indicators have been reduced to two dimensions that are somewhat correlated with SWB: in the case of relational uses of time, the correlation is positive but weak, whereas in the case of non-relational uses of time it is negative and stronger. As expected, the lag term is the best correlate to SWB indicators. Table 5 presents two more elements worth mentioning. First, none of the indicators is a particularly accurate predictor of SWB except, possibly, the lag term. This is consistent with the hypothesis that SWB is multidimensional. Second, GDP *per capita* is a marginally worse predictor of only a few other indicators that, incidentally, correlate very well with it. This finding seems to suggest that our perception of well-being may consist of more than just GDP, but “beyond GDP” effects will probably point to adjusting, not dramatically overturning, GDP-based policy decisions.

4.2 Models

Tables 6 and 7 present 12 ordered logit models of SWB based on the social indicators built from the OECD-DSI, at times completed with GDP *per capita* and the lag term of SWB. Odd numbers indicate models in which the dependent variable (TriSWB) has three possible outcomes. Models indicated with even numbers have the same set of independent variables as the model to their left, but the dependent variable (HeptaSWB) has seven possible outcomes. Models (1-4) are straightforward estimates of SWB using the full OECD-DSI dataset, with or without the lag term. Models (5-12) are more parsimonious models obtained with forward and backward stepwise selection of the regressors. The Bayesian Information Criterion (BIC)[52] statistic shows that including all social variables as in models (1-4) increases complexity without great improvements in model fit. The tests on the proportionality of odds assumption are frequently discordant but coefficient signs of the statistically significant regressors are consistent throughout all models and consistent with estimates from multinomial logit models that we do not include for brevity.

---

[52] See Greene (2011, pg. 160); BIC is a criterion for model selection that balances the likelihood of the model with the number of parameters in it. A model with a lower BIC should be preferred over a similar model with a higher BIC as the introduction of additional parameters in the second model did not provide sufficient improvements in likelihood.

Table III-6. Ordered logit models with social

VARIABLES	(1) TriSWB	(2) HeptaSWB	(3) TriSWB	(4) HeptaSWB	(5) TriSWB	(6) HeptaSWB
dem	0.126**	0.0829*	0.176**	0.147***	0.215**	0.104**
	(0.0616)	(0.0431)	(0.0841)	(0.0537)	(0.0989)	(0.0454)
selfsuf	0.160***	0.156***	0.140*	0.159***		
	(0.0619)	(0.0502)	(0.0742)	(0.0607)		
equity	-0.0359	-0.0304	0.0151	0.0138		
	(0.0534)	(0.0420)	(0.0844)	(0.0486)		
health	0.00200	0.0259	-0.133	-0.0836	-0.193*	-0.0441
	(0.0549)	(0.0457)	(0.0991)	(0.0572)	(0.117)	(0.0504)
soccohes	0.120*	0.115**	0.115	0.0915		
	(0.0646)	(0.0551)	(0.100)	(0.0685)		
swb_0509			0.238**	0.211***	0.275**	0.243***
			(0.112)	(0.0576)	(0.110)	(0.0601)
gdp					0.315**	0.106***
					(0.157)	(0.0395)
Constant cut1	3.388**	-0.865	9.985**	3.806*	17.17***	5.795***
	(1.560)	(1.327)	(3.990)	(1.994)	(6.593)	(2.200)
Constant cut2	6.990***	1.745*	18.41**	7.400***	26.27***	9.117***
	(2.149)	(1.016)	(7.493)	(2.060)	(10.09)	(2.422)
Constant cut3		3.107***		9.438***		10.61***
		(1.156)		(2.391)		(2.606)
Constant cut4		5.193***		12.77***		13.15***
		(1.423)		(3.006)		(2.997)
Constant cut5		6.906***		17.22***		17.60***
		(1.627)		(4.168)		(4.403)
Constant cut6		9.946***		21.46***		21.58***
		(1.996)		(4.728)		(4.885)
N	34	34	34	34	34	34
STATISTICS						
Log-likelihood	-17.20	-38.43	-10.87	-27.15	-10.56	-30.07
LR Chi^2 test	37.35	45.50	50.01	68.06	50.63	62.23
Pseudo R^2	0.521	0.372	0.697	0.556	0.706	0.509
BIC	59.08	115.65	49.95	96.62	42.28	95.41
TESTS OF PROPORTIONALITY OF ODDS ACROSS RESPONSE CATEGORIES (P>Chi^2)						
Wolfe-Gould	0.002	0.000	0.020	0.017	0.192	0.207
Brant	1.000	1.000	1.000	1.000	0.650	1.000
Score	0.009	0.041	0.045	0.310	0.410	0.587

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table III-7. Ordered logit models with social regressors

VARIABLES	(7) TriSWB	(8) HeptaSWB	(9) TriSWB	(10) HeptaSWB	(11) TriSWB	(12) HeptaSWB
dem	0.137**	0.116***	0.142**	0.121***		
	(0.0666)	(0.0446)	(0.0644)	(0.0435)		
selfsuf	0.120*	0.139**	0.135**	0.162***		
	(0.0679)	(0.0562)	(0.0648)	(0.0544)		
soccohes	0.0465	0.0612			0.0988	0.115**
	(0.0630)	(0.0524)			(0.0700)	(0.0554)
equity					-0.0203	-0.0218
					(0.0573)	(0.0409)
swb_0509	0.158**	0.176***	0.167***	0.182***	0.166***	0.176***
	(0.0632)	(0.0479)	(0.0611)	(0.0463)	(0.0544)	(0.0432)
Constant cut1	7.737***	2.899*	7.808***	2.833*	5.131***	1.428
	(2.547)	(1.565)	(2.464)	(1.533)	(1.890)	(1.468)
Constant cut2	13.73***	6.361***	13.48***	6.313***	9.661***	4.272***
	(4.404)	(1.673)	(4.160)	(1.627)	(2.925)	(1.347)
Constant cut3		8.286***		8.237***		5.560***
		(1.956)		(1.894)		(1.473)
Constant cut4		11.22***		11.22***		7.778***
		(2.380)		(2.364)		(1.804)
Constant cut5		15.05***		14.59***		10.65***
		(3.374)		(3.150)		(2.421)
Constant cut6		19.32***		18.65***		14.56***
		(4.001)		(3.753)		(3.013)
N	34	34	34	34	34	34
STATISTICS						
Log-likelihood	-12.07	-28.27	-12.34	-28.99	-15.11	-33.44
LR Chi^2 test	47.61	65.83	47.06	64.39	41.53	55.48
Pseudo R^2	0.664	0.538	0.656	0.526	0.579	0.453
BIC	45.30	91.80	42.32	89.72	47.85	98.62
TESTS OF PROPORTIONALITY OF ODDS ACROSS RESPONSE CATEGORIES (P>Chi^2)						
Wolfe-Gould	0.094	0.423	0.099	0.382	0.088	0.075
Brant	0.518	1.000	0.241	0.000	0.266	0.000
Score	0.159	0.733	0.145	0.713	0.131	0.306

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table III-8. Ordered logit models with social and environmental regressors

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
VARIABLES	TriSWB	HeptaSWB	TriSWB	HeptaSWB	TriSWB	HeptaSWB	TriSWB	HeptaSWB
dem	0.203*	0.0914*	0.123	0.121**	0.123	0.128**		
	(0.121)	(0.0510)	(0.0869)	(0.0542)	(0.0799)	(0.0519)		
selfsuf			0.154*	0.220***	0.174**	0.238***		
			(0.0807)	(0.0722)	(0.0754)	(0.0694)		
equity							-0.0180	-0.0249
							(0.0665)	(0.0442)
health	-0.184	-0.00786						
	(0.144)	(0.0538)						
soccohes			0.0717	0.0487			0.150*	0.150**
			(0.0870)	(0.0708)			(0.0884)	(0.0653)
swb_0509	0.294**	0.289***	0.221**	0.246***	0.228**	0.251***	0.201***	0.211***
	(0.131)	(0.0764)	(0.102)	(0.0719)	(0.0970)	(0.0693)	(0.0745)	(0.0569)
gdp	0.313*	0.112***						
	(0.169)	(0.0421)						
cnrr	-0.0609	-0.0810	-0.0672	-0.146**	-0.0931	-0.152***	-0.0464	-0.0766*
	(0.0780)	(0.0502)	(0.0765)	(0.0581)	(0.0762)	(0.0570)	(0.0653)	(0.0461)
gpp	-0.0289	-0.0741	-0.112	-0.0900*	-0.0970	-0.0866*	-0.103*	-0.0670
	(0.0778)	(0.0513)	(0.0717)	(0.0502)	(0.0637)	(0.0491)	(0.0627)	(0.0419)
Constant cut1	16.09**	4.270*	7.178**	0.315	6.610**	0.207	4.340*	-0.180
	(7.213)	(2.422)	(3.658)	(1.875)	(3.112)	(1.858)	(2.402)	(1.779)
Constant cut2	25.76**	8.181***	15.28**	4.562**	14.01**	4.450**	10.20***	3.070*
	(11.28)	(2.735)	(6.905)	(2.036)	(5.658)	(1.971)	(3.816)	(1.647)
Constant cut3		9.918***		7.234***		7.094***		4.566***
		(2.942)		(2.337)		(2.261)		(1.745)
Constant cut4		12.56***		11.21***		11.19***		6.955***
		(3.349)		(2.938)		(2.902)		(2.044)
Constant cut5		18.26***		16.35***		15.79***		10.93***
		(5.178)		(4.521)		(4.181)		(3.044)
Constant cut6		22.37***		21.20***		20.59***		14.95***
		(5.691)		(5.234)		(4.939)		(3.576)
N	33	33	33	33	33	33	33	33
STATISTICS								
Log-likelihood	-10.08	-27.76	-10.04	-23.41	-10.41	-23.65	-12.58	-30.30
LR Chi^2 test	50.04	64.32	50.12	73.04	49.39	72.54	45.05	59.24
Pseudo R^2	0.713	0.537	0.714	0.609	0.704	0.605	0.642	0.494
BIC	48.14	97.48	48.06	88.77	45.29	85.77	49.63	99.07
TESTS OF PROPORTIONALITY OF ODDS ACROSS RESPONSE CATEGORIES (P>Chi^2)								
Wolfe-Gould	0.003	0.051	0.003	0.026	0.075	0.054	0.036	0.012
Brant	1.000	1.000	1.000	-	1.000	1.000	0.377	0.000
Score	0.349	0.556	0.083	0.420	0.300	0.008	0.127	0.377

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table III-9. Ordered logit models with social and time-use regressors

VARIABLES	(21) TriSWB	(22) HeptaSWB	(23) TriSWB	(24) HeptaSWB	(25) TriSWB	(26) HeptaSWB	(27) TriSWB	(28) HeptaSWB
dem	0.415	0.213**	0.0961	0.125**	0.0959	0.125**		
	(0.260)	(0.0963)	(0.0727)	(0.0620)	(0.0720)	(0.0617)		
selfsuf			0.0942	0.138*	0.103	0.144**		
			(0.0810)	(0.0745)	(0.0786)	(0.0713)		
equity							-0.0520	-0.0603
							(0.0739)	(0.0585)
health	-0.589	-0.199*						
	(0.406)	(0.120)						
soccohes			0.0353	0.0230			0.0973	0.117
			(0.0777)	(0.0776)			(0.100)	(0.0895)
swb_0509	0.540*	0.403***	0.140*	0.253***	0.152**	0.260***	0.152**	0.231***
	(0.309)	(0.139)	(0.0725)	(0.0859)	(0.0685)	(0.0834)	(0.0731)	(0.0749)
gdp	0.783	0.324**						
	(0.527)	(0.152)						
relational	0.299	0.0658	0.0263	0.0124	0.0372	0.0178	-0.00793	0.0116
	(0.216)	(0.0797)	(0.0902)	(0.0730)	(0.0853)	(0.0705)	(0.0888)	(0.0662)
nonrelational	-0.155	-0.219*	-0.0821	-0.225**	-0.0856	-0.227**	-0.107	-0.238**
	(0.161)	(0.132)	(0.102)	(0.104)	(0.100)	(0.103)	(0.109)	(0.107)
Constant cut3		17.24**		6.135**		6.149**		2.024
		(7.243)		(3.059)		(3.016)		(2.714)
Constant cut4		21.00***		10.29***		10.40***		5.066*
		(7.790)		(3.734)		(3.703)		(2.999)
Constant cut5		31.62***		15.63***		15.66***		9.769**
		(11.88)		(5.242)		(5.208)		(4.081)
Constant cut6		38.00***		21.59***		21.56***		15.60***
		(13.03)		(6.456)		(6.424)		(4.990)
Constant cut1	35.30	9.208	5.038*	-0.912	5.086*	-0.892	1.991	-3.532
	(21.91)	(5.731)	(3.061)	(2.710)	(2.987)	(2.679)	(3.086)	(3.016)
Constant cut2	57.01*	15.04**	10.77**	3.542	10.83**	3.568	7.061*	0.321
	(34.37)	(6.974)	(4.587)	(2.880)	(4.531)	(2.843)	(4.031)	(2.738)
N	26	26	26	26	26	26	26	26
STATISTICS								
Log-likelihood	-8.027	-15.61	-10.46	-16.40	-10.56	-16.45	-11.32	-19.07
LR Chi^2 test	38.96	62.29	34.09	60.71	33.88	60.62	32.36	55.37
Pseudo R^2	0.708	0.666	0.620	0.649	0.616	0.648	0.588	0.592
BIC	42.12	70.32	46.98	71.90	43.93	68.73	45.45	73.98
TESTS OF PROPORTIONALITY OF ODDS ACROSS RESPONSE CATEGORIES (P>Chi^2)								
Wolfe-Gould	0.013	0.387	0.002	0.324	0.001	0.129	0.004	0.123
Brant	1.000	1.000	1.000	-	1.000	-	1.000	-
Score	0.072	0.000	0.032	0.003	0.017	0.000	0.056	0.006

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Self-sufficiency (selfsuf), social cohesion (soccohes) and GDP per capita (gdp) have consistently positive coefficients and tend to be significant in most models, but their statistical significance is reduced when they appear together, indicating that the information they provide might be partially redundant. The demographic variable (dem) is consistently significant and positive in all models. The coefficients of equity have erratic signs and they are never statistically significant. Health is usually not significant as well and tends to have an unexpected negative sign. Models (13-20) in table 8 are extensions of models (5-12) with the addition of two environmental regressors. The coefficients of the social regressors are generally stable (with a slight reduction of the significance of dem) so we can broadly confirm the results found above. The variable concerning the green economy and legislation (gpp) is never significant with a tendency of having an unexpected negative sign. The variable related to polluting and unsustainable consumption is negative but significant only in models with seven outcomes. The inconsistent significance of this regressor, along with its being uncorrelated with SWB indicators, suggests great caution; we will provide our interpretation in the discussion section. Models (21-28) in table 9 are a different extension of models (5-12), this time with the addition of relational and non-relational uses of time. The social regressors lose some significance, particularly in models with a response variable with three outcomes (TriSWB). This is not entirely surprising: both time use variables have decent or good correlation with other social variables and with GDP per capita (hence, part of the information can be thought as overlapping). The time use regressors themselves provide very different results. Relational use of time is non-significant in all models. At the opposite, non-relational use of time is significant in models with the heptapartite response variable, and consistently negative in all models.

5. Discussion

Our findings are consistent with the view, shared by many, that SWB must indeed be multidimensional: no single factor, be it GDP or beyond GDP, is a particularly accurate predictor of national average SWB, if taken alone. Actually, as table 5 showed, GDP still works better than many proposed "beyond GDP" factors and not dramatically worse than the best of them. We didn't test all possible "beyond GDP" factors, though, and some not covered by our work might prove more clearly superior to GDP.

After moving from an unidimensional to a multidimensional approach, we found that national SWB can be predicted with reasonable accuracy by making use of some parsimonious selection of three regressors representing self-sufficiency, social cohesion and GDP *per capita*[53]. As the three regressors seem to reduce each other's statistical significance, they arguably convey much shared information; in fact, including all three in a model is not justified by the improvement of the model. Our interpretation of this block of variables is that income stability in all phases of the life cycle (i.e. self-sufficiency) and contained social conflict produce an environment in which a higher average SWB is more likely. Furthermore, self-sufficiency and social cohesion combined explain national SWB better than GDP *per capita*.

This group of regressors has an interesting counterpoint in demography. Stable, fertile families in a dynamic demographic environment are frequently associated with relatively higher national SWB. In some circumstances, countries have both large networks of family and solidarity ties and high levels of self-sufficiency and social cohesion, a context that our models indicate as the best possible in terms of well-being. More frequently, countries seem to trade demographic dynamism with self-sufficiency and social cohesion and vice versa. The relation between demography and development has been studied for a long time; in this case we find a relation between demography and subjective well-being that could also be an important parameter for policy decisions. Policymakers should be aware that social cohesion and income stability, combined as they frequently are, are likely to be more relevant than demography in lifting the perception of well-being but, in the absence of demographic dynamism and particularly in the long run, the level of SWB is not sustainable and it is likely going to decrease.

Equity is not correlated with SWB and provides no useful information to predict its national level. Based on the OECD indicators combined in our equity regressor, we must conclude that, in OECD well-being data, there are no signs of aversion for inequality or preference for inequality-reducing social transfers. Since equity had at least some weak correlation with SWB in table 4, we suppose

---

[53] As defined in the data section of this work.

that any effect of equity is better explained by other variables, like self-sufficiency or social cohesion.

When other social regressors are considered, health is not significant or weakly significant but negative. There are several possible explanations for this and all should be empirically tested. In any case, given the results, other socio-economic factors are much more likely to directly affect the perception of well-being at a national level.

Our attempt of expanding the best social models with environmental indicators provided mixed results. We found no evidence that green policies and the green economy (as they could be inferred from the available indicators) have any relation with SWB. The quality and the amount of data are relatively poor and future research could point to green policies (and the appropriate indicators) that are more immediately perceived by the population. This notwithstanding, the link between environmental-friendly development and SWB is not as generalized and evident as one might expect. Our indicator on non-sustainable production and consumption, instead, works rather well in some models with seven potential outcomes, in spite of having no correlation with the indicators of SWB. This could depend on some unidentified source of bias in the model or on the fact that, all other things being equal, pollution and non-sustainable consumption matter to SWB at least to some degree. In either case, the effect is much less direct than with other indicators and future research should probably assume that what exactly channels environmental issues into SWB effects is rather unclear when looking at standard indicators.

The attempt of adding time-use indicators to social models provided more conclusive results. The amount of time dedicated to relational activities has no apparent connection with SWB in the OECD countries after social factors are taken into account. That is, more time devoted to social interaction does not correspond to more SWB unless the social context is better as well in terms of low levels of social conflict, a good disposition towards the others and widespread economic self-sufficiency. Meanwhile, a better social context corresponds to more SWB regardless of the amount of relational time so that, apparently, a socio-economic model of SWB is just as good without specific information about relational time.

Our indicator of non-relational uses of time, instead, appears to be linked with national SWB, at least in models with seven potential outcomes: more time dedicated to these activities corresponds to lower average national SWB. Two possible interpretations arise, particularly while trying to be consistent with the findings concerning the other time-use regressor. The first is that SWB may be affected by the quantity of non-relational time but, at the same time, by the quality rather than the quantity of relational time. The second is that some uses of time may be connected with national average SWB for reasons other than the production of relational goods. Either way, based on our results, there is no reason to expect that policies increasing the amount of relational time would have any effect on SWB.

6. Conclusion

Some authors claim that maximizing SWB is a more meaningful social objective than maximizing GDP and that other factors beyond income play a major role in defining well-being. In this work, we studied two issues connected with this claim, looking at the context of OECD member countries. We looked at the crowded category of proposed, "beyond GDP" policy-controlled factors, searching for evidence that some might be major drivers of national average SWB. We also compared any such effect with that of GDP, in order to evaluate if these factors have a better chance of leading to a maximization of SWB than GDP itself.

In our analyses, we made use of Partial Order Methods that have been rarely (if ever), applied to this field of study. They seem particularly appropriate to the case, as SWB is generally theorized as strongly multidimensional while standard modeling strategies require a great deal of compromise when working with many potential regressors and non-trivial levels of multicollinearity. This approach can be applied to more well-being metrics and can be extended to much more complex conceptual models, providing an alternative to methods that sacrifice interpretability or lose control of compensation between different dimensions.

Our results support the view that the national perception of well-being depends on a multidimensional set of determinants: GDP, just like any other "beyond GDP" indicator at our disposal, taken alone, explains little of SWB variability across countries. Even in multivariate models, we don't have to explicitly include GDP in the set of regressors to obtain reasonable predictions. Nonetheless, after testing a rather vast field of alternative indicators, we found only two groups of variables that consistently anticipate higher national SWB: self-sufficiency and social cohesion, demography and non-relational use of time. The two groups are not mutually exclusive, but they lead to conclusions that are better kept distinct. The first group, consisting of self-sufficiency and social cohesion, while being more accurate than GDP alone, is going to favor small or moderate adjustments to GDP-led policies rather than the kind of dramatic overturn that some expect from "beyond GDP" approaches. At the opposite, the second group might point to greater policy shifts but, at present, it is not particularly clear in which direction. Demography does respond, over long periods, to some policy decisions, but the time scale is clearly well beyond the reach of one government mandate. Time-use is probably manageable in shorter time spans through policy, but having to discard the role of relational time-use, we are left without a definite theory of why time-use matters and in which way.

We also found SWB to be extremely persistent over long periods of time, and that more than a few conclusions in our models are based on the assumption that small differences in the average value of SWB are independent from cultural and linguistic differences that surveys cannot avoid. Even if this second reason for concern is left aside, the modest variability of well-being indicators in periods of ten or even fifteen years, makes the improvement of this indicator a very impervious target for policy-makers in the timespan of one mandate.

Improving the current policy goals of governments with "beyond GDP" approaches seems a sensible intent. SWB, we argue, provides information that can be used to fine-tune the decisions of policymakers and put focus on factors like self-sufficiency that have an unquestionable effect on SWB but can be underestimated by GDP-driven approaches. However, we found no evidence that individual preferences look totally different if they are inferred from perceived well-being rather than from revealed preferences in a market system. On the contrary, based on our results, claims that SWB would lead to a paradigm shift and to an overturn of GDP-driven policies seem to linger on many ineffective indicators and on theories based on inconsistent factual evidence.

As we reminded throughout the paper, other metrics of SWB might possibly provide different results and, maybe, other indicators in more complex model frameworks could ultimately identify as significant some specific elements of theories that we found ineffective. With this work, we have presented an approach that can hopefully support further research in that direction and help overcoming some multidimensionality issues. In the absence of stronger evidence, though, we argue that SWB is unlikely to respond to many of the policies that "beyond GDP" literature is promoting and, consequently, well-being research should envision a more substantial paradigm change, shifting its attention to measures lying outside the hedonic sphere and unrelated to individual preferences, or focus on the less radical agenda of policy adjustment that emerges from individual preferences.

References

Becchetti, L., Pelloni, A., & Rossetti, F. (2008). Relational goods, sociability, and happiness. *Kyklos*, 61(3), 343–363.

Becchetti, L., Trovato, G., & Londono Bedoya, D. A. (2011). Income, relational goods and happiness. *Applied Economics*, 43(3), 273–290.

Becker, G., Rayo, L., & Krueger, A. B. (2008). Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. Comments and Discussion. *Brookings Papers on Economic Activity*, 88–102.

Beckerman, W. (1994). "Sustainable Development": Is it a Useful Concept? *Environmental Values*, 3(3), 191–209.

Bond, T. N., & Lang, K. (2014). The Sad Truth About Happiness Scales. NBER Working Papers 19950

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, 46(4), 1171–1178.

Brüggemann, R., & Annoni, P. (2014). Average Heights in Partially Ordered Sets. *MATCH Communications in Mathematical and in Computer Chemistry*, 71, 117–142.

Brüggemann, R., & Carlsen, L. (2011). An Improved Estimation of Averaged Ranks of Partial Orders. *MATCH Communications in Mathematical and in Computer Chemistry*, 65, 383–414.

Brüggemann, R., & Patil, G. (2011). Ranking and prioritization for multi-indicator systems. New York. Springer.

Brüggemann, R., Simon, U., & Mey, S. (2005). Estimation of averaged ranks by a local partial order model. *MATCH Communications in Mathematical and in Computer Chemistry*, 54, 498–518.

Brüggemann, R., Sørensen, P. B., Lerche, D., & Carlsen, L. (2004). Estimation of averaged ranks by a local partial order model. *Journal of chemical information and computer sciences*, 44(2), 618–25.

Bruni, L., & Stanca, L. (2008). Watching alone: Relational goods, television and happiness. *Journal of Economic Behavior and Organization*, 65(3-4), 506–528.

Costanza, R., Hart, M., Talberth, J., & Posner, S. (2009). Beyond GDP : The Need for New Measures of Progress Beyond GDP : The Need for New Measures of Progress. *The Pardee papers*, (4), 1–47.

Costanza, R., Kubiszewski, I., Giovannini, E., Lovins, H., McGlade, J., Pickett, K., et al. (2014). Time to leave GDP behind. *Nature*, 505, 283–285.

Daly, H. E. (1995). On Wilfred Beckerman's Critique of Sustainable Development. *Environmental Values*, 4(1), 49–55.

Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: an introduction. *Journal of Happiness Studies*, 9, 1–11.

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34–43.

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research*, 112(3), 497–527.

Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being*. (pp. 185–218).

Diener, E., & Tov, W. (2012). National Accounts of Well-being. In K. C. Land, M. Sirgy, & A. C. Michalos (Eds.), *Handbook of Social Indicators and Quality of Life Research* (pp. 137–157). Dordrecht: Springer.

Easterlin, R. (2015). *Happiness and Economic Growth: The Evidence.* Springer. Netherlands.

Edwards, J. R., Bagozzi, R. P., Blackburn, R. S., Bollen, K. a, Cattani, D., Jr, J. W. D., et al. (2000). On the Nature and Direction of Relationships Between Constructs and Measures. *Psychological methods*, 5(2), 155–174.

Fattore, M. (2016). Partially Ordered Sets and the Measurement of Multidimensional Ordinal Deprivation. *Social Indicators Research*, 128(2), 835-858.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How Important is Methodology for the Estimate of the Determinants of Hapiness? *The Economic Journal*, 114(497), 641–659.

Fleurbaey, M. (2009). Beyond GDP: The Quest for a Measure of Social Welfare. *Journal of Economic Literature*.

Frey, B. S., & Stutzer, A. (2005). Happiness Research: State and Prospects. *Review of Social Economy*, 62(190), 207–228.

Gallup. (2010). *Global wellbeing. The behavioral economics of GDP Growth*. Gallup Inc., Washington DC

Greene, W. H. (2011). *Econometric Analysis* (Fifth Edit.). Prentice Hall.

Gui, B., & Stanca, L. (2010). Happiness and relational goods: Well-being and interpersonal relations in the economic sphere. *International Review of Economics*, 57(2), 105–118.

Guttman, L. (1950). The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), *Measurement and prediction*. Princeton University Press.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of personality and social psychology*, 82(6), 903–918.

Jorm, A. F., & Ryan, S. M. (2014). Cross-national and historical differences in subjective well-being. *International Journal of Epidemiology*, 43(2), 330–340.

Jurado, A., & Perez-Mayo, J. (2012). Construction and Evolution of a Multidimensional Well-Being Index for the Spanish Regions. *Social Indicators Research*, 107(2), 259–279.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004). Toward national well-being accounts. In *American Economic Review* (Vol. 94, pp. 429–434).

Kahneman, D., & Riis, J. (2005). Living , and thinking about it : two perspectives on life. In F. Huppert, N. Baylis, & B. Kaverne (Eds.), *The science of well-being: Integrating Neurobiology, Psychology and Social Science* (pp. 285–304). Oxford University Press.

Krueger, A. B., & Schkade, D. A. (2008). The reliability of subjective well-being measures. *Journal of Public Economics,* 92, 1833–1845.

Kubiszewski, I., Costanza, R., Franco, C., Lawn, P., Talberth, J., Jackson, T., & Aylmer, C. (2013). Beyond GDP: Measuring and achieving global genuine progress. *Ecological Economics*, 93, 57–68.

Kuznets, S. (1934). *National income*, 1929-32.

Lerche, D., Sørensen, P. B., & Brüggemann, R. (2003). Improved estimation of the ranking probabilities in partial orders using random linear extensions by approximation of the mutual ranking probability. *Journal of chemical information and computer sciences*, 43(5), 1471–80.

McCullagh, P. (1988). Regression models for ordinal data. *Journal of the royal statistical society. Series B (Methodological)*, 42, 109-142.

Miranda, V. (2011). Cooking, Caring and Volunteering: Unpaid Work Around the World. *OECD Social, Employment, and Migration Working Papers*, (116), 0-1, 3–4, 6–39.

Nordhaus, W. D., & Tobin, J. (1972). Is Growth Obsolete? In William Nordhaus and James Tobin. *Economic Research: Retrospect and Prospect Vol 5: Economic Growth* (Vol. 5).

O'Connell, A. (2006). *Logistic Regression Models for Ordinal Response Variables*. Sage Publications.

OECD (2013). *Guidelines on Measuring Subjective Well-being*. OECD Publishing.

OECD (2014). *Society at a Glance 2014: OECD Social Indicators*. OECD Publishing.

OECD. http://www.oecd-ilibrary.org/environment/data/oecd-environment-statistics_env-data-en. Accessed November 2015.

OECD. http://www.oecd.org/gender/data/. Accessed November 2015.

Oishi, S., & Diener, E. (2014). Can and Should Happiness Be a Policy Goal? *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 195–203.

Oswald, A. J. (2008). On the curvature of the reporting function from objective reality to subjective feelings. *Economics Letters*, 100(3), 369–372.

Raveh, A., & Landau, S. (1993). Partial order scalogram analysis with base coordinates (POSAC): Its application to crime patterns in all the states in the United States. *Journal of quantitative criminology*, 9(1), 83–99.

Shye, S. (1985). *Multiple Scaling. The Theory and Application of Partial Order Scalogram Analysis*. Jerusalem: Israel Inst of Applied Social Research.

Shye, S., & Amar, R. (1985). Partial order scalogram analysis by base coordinates and lattice mapping of items by their scalogram roles. In D. Canter (Ed.), *Facet Theory: Approaches to Social Research* (pp. 277–298). Springer Verlag.

Stanca, L. (2009). With or without you? Measuring the quality of relational life throughout the world. *Journal of Socio-Economics*, 38(5), 834–842.

Stevenson, B., & Wolfers, J. (2008). Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. *Brookings Papers on Economic Activity*, 2008, 1–87.

Stevenson, B., & Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation? *American Economic Review*, 103(May 2013), 598–604.

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris: Commission on the Measurement of Economic Performance and Social Progress.

Welsch, H. (2007). Environmental welfare analysis: A life satisfaction approach. *Ecological Economics*, 62, 544–551.

Welsch, H. (2009). Implications of happiness research for environmental economics. *Ecological Economics*, 68(11), 2735–2742.

Wolfe, R., & Gould, W. (1998). An approximate likelihood-ratio test for ordinal response models. *Stata Technical Bulletin* (Vol. 42).

CONCLUDING REMARKS

Policy decisions under severe informational limitations

In the previous chapters, the complexity of decision-making has been represented in the form of the various legitimate preferences of the individuals in a population (Chapter I) or, rather, in different possible combinations of performance indicators that, once assembled, should represent an overall level of performance (Chapter II and III). In Chapter I, the multiplicity of goals was presented as a matter of fact of theoretical relevance (collective choice is based on individual preferences by definition); in Chapter II, multiple criteria are considered as a diagnostic tool to ponder the consequences of using a single criterion; in Chapter III, the complexity of a multi-level dataset is represented by the multidimensionality of each of its domains. In all cases, what prevented a straightforward reduction of complexity was the assumption that the scores of each alternative under different criteria defy aggregation unless substantial exogenous information allows it.

The circumstances in which this is relevant seem to go beyond those illustrated in this thesis. In the field of collective choice, one can think of many scenarios in which a simple functional relation between the ranking of an alternative in all individual preferences and its ranking in the collective preference doesn't seem such a good idea. Think of decision-making in a country in which preferences on fundamental issues are strictly correlated with ethnic self-identification and minorities know in advance that they are set to systematically lose majority decisions and have little incentive in participating to collective decision-making. All solutions that apply to these cases implicitly assume that the balance between different individual orderings cannot be established exclusively as a consequence of the frequency of its appearance in the population. About the same argument can be made for all processes in which local interests have to be balanced with general interests and bargaining solutions are not effective. In the more general case in which criteria do not necessarily consist of individual preferences, direct trade-offs between criteria can easily appear vacuous, either when the actual trade-off is not yet known and thus informational limitations are temporary, or when it is intrinsically impossible to weigh down an exact trade-off because it depends on an unpredictable future state of matters, on uncertain value judgements, on non-linear and unknown processes. For instance, even though we may admit some trade-off between growth and environmental quality when both are within a given range, outside of that range any trade-off may be extremely unpredictable as a consequence of feedbacks and non-linearities.

In all these cases, the complexity of the decision-making process, which corresponds to multiple objectives and concerns, results in conflicting orderings in which some alternatives dominate others and some are just incomparable. The partially ordered set of alternatives derived from the Pareto criterion, though, represents more than a failed attempt at ordering: it provides information on those conflicts between alternative orderings, alternative interests and possible solution. Specifically, the height of an alternative in this partially ordered set varies between a maximum and a minimum in a range that represents all possible ranks assumed by that alternative depending on any meaningful package of exogenous information that, for instance, endows criteria with weights.

This approach is unlikely to provide a major alternative to the standard approaches at decision-making: the information that it assumes as unavailable is sufficiently important that it may appear as unacceptable to even start a decision-making process without having it secured. At the same time, a wide variety of situations may require a decision no matter what and, in that case, the approach discussed in this work provides a unified and rigorous method as an alternative to *ad hoc* solutions that are frequently found in practice.

APPENDICES

Supplementary material for Chapter III

Appendix A.1

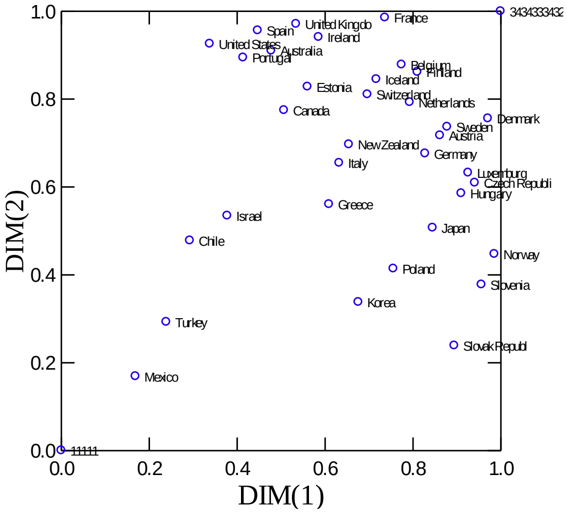

FIGURE 1 POSAC of OECD member countries based on five "Selfsuf indicators"

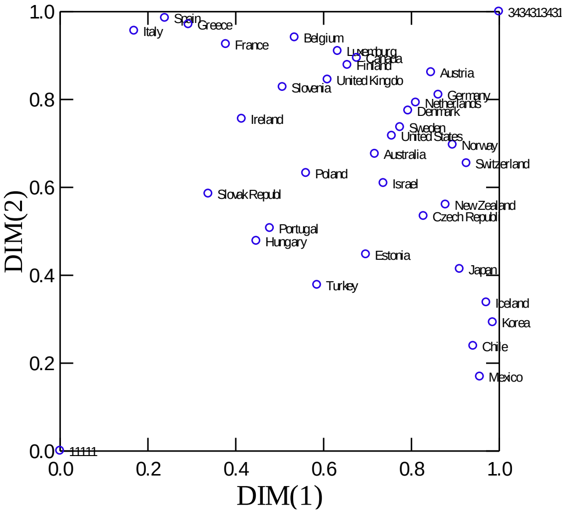

FIGURE 2 POSAC of OECD member countries based on five "Equity indicators"

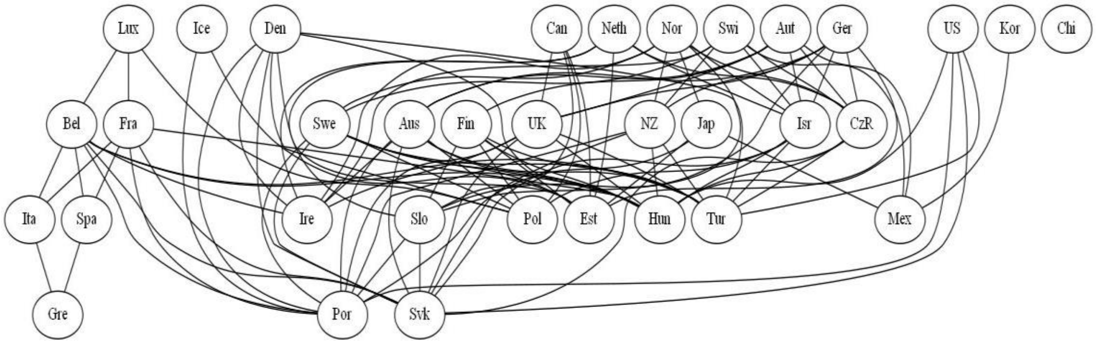

FIGURE 3 Hasse diagram of OECD member countries based on five "Selfsuf indicators"

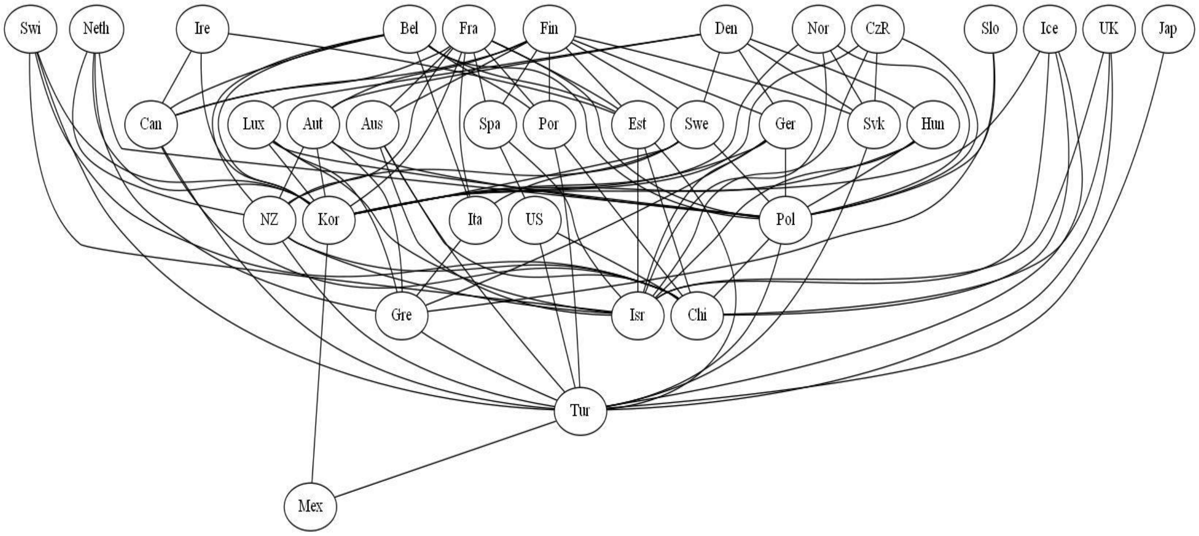

FIGURE 4 Hasse diagram of OECD member countries based on five "Equity indicators"

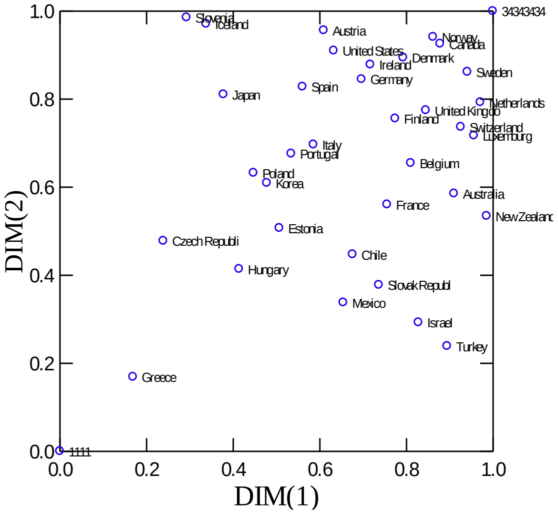

FIGURE 5 POSAC of OECD member countries based on four "Soccohes indicators"

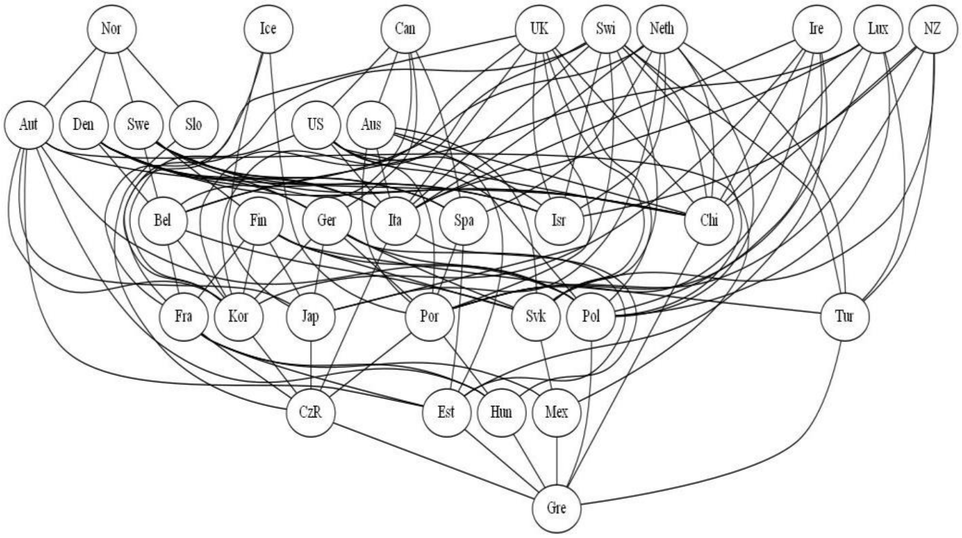

FIGURE 6 Hasse diagram of OECD member countries based on four “Soccohes indicators”

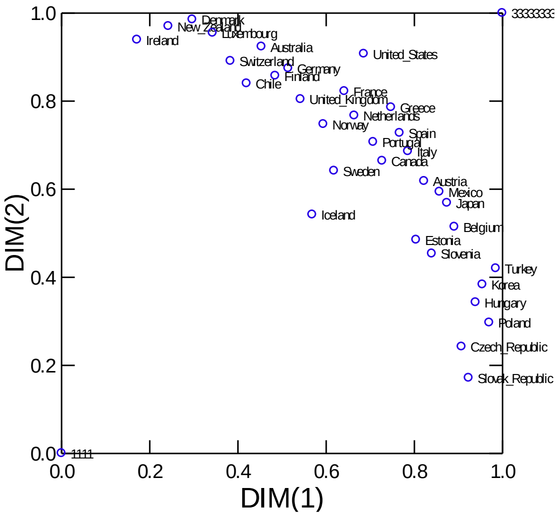

FIGURE 7 POSAC of OECD member countries based on five “Cnrr indicators”

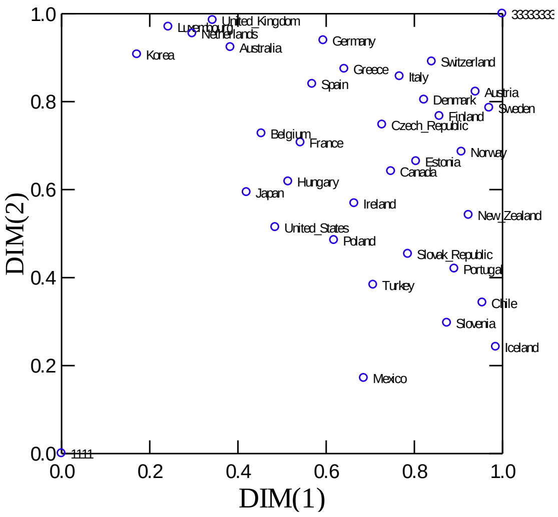

FIGURE 8 POSAC of OECD member countries based on five “Gpp indicators”

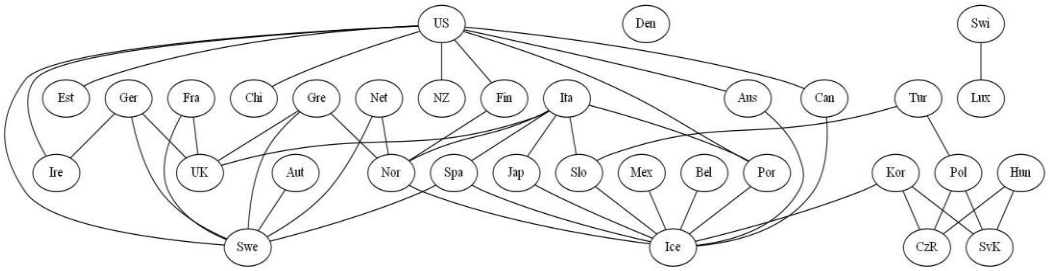

FIGURE 9 Hasse diagram of OECD member countries based on “Cnrr indicators”

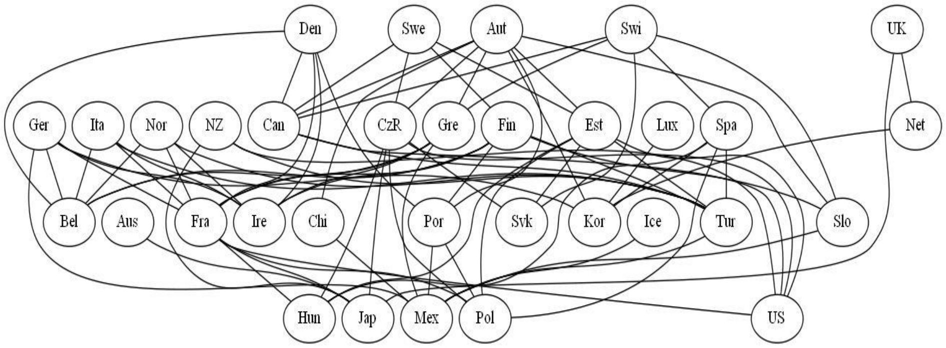

FIGURE 10 Hasse diagram of OECD member countries based on “Gpp indicators”

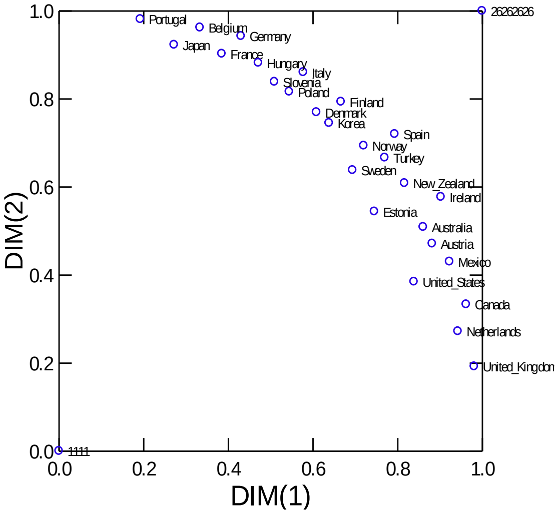

FIGURE 11 POSAC of OECD member countries based on “relational indicators”

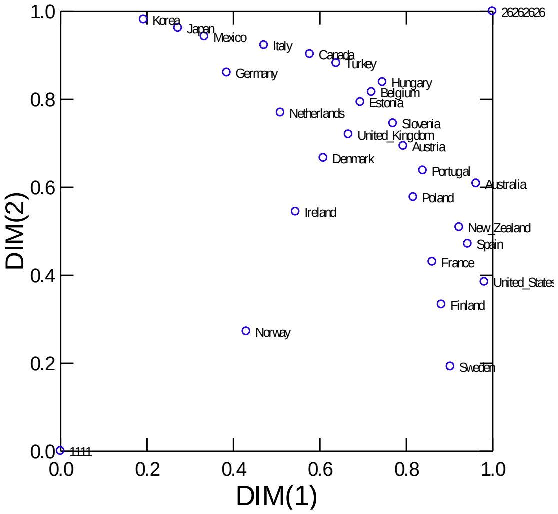

FIGURE 12 POSAC of OECD member countries based on “Nonrelational indicators”

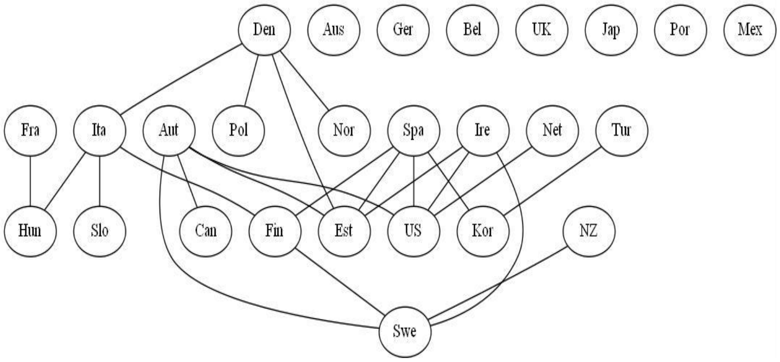

FIGURE 13 Hasse diagram of OECD member countries based on “Relational indicators”

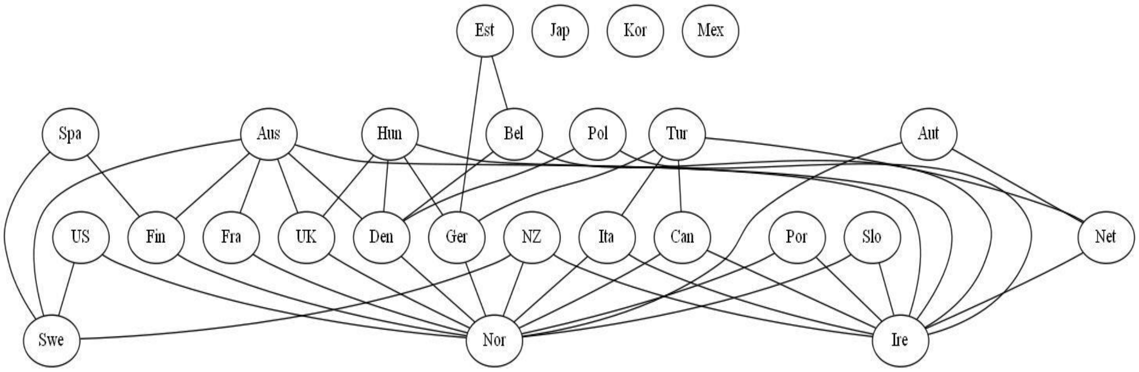

FIGURE 14 Hasse diagram of OECD member countries based on “Nonrelational indicators”

Appendix A.2

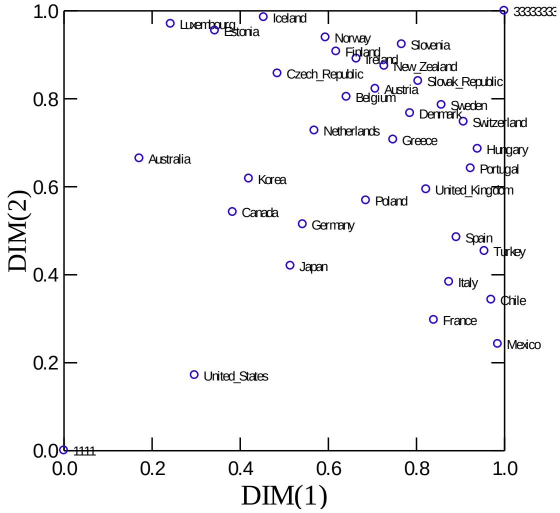

FIGURE 15 POSAC of cnrr indicators (with CO2 replacing Particulate)

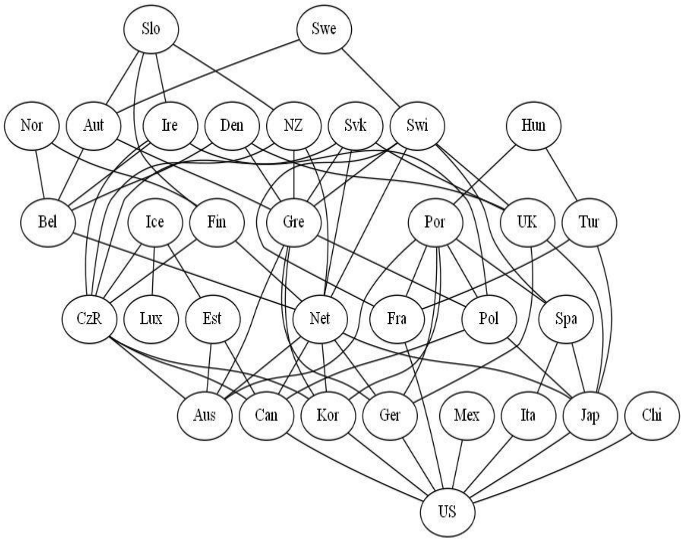

FIGURE 16 Hasse diagram based on cnrr indicators (with CO2 replacing Particulate)

The four indicators on pollution and consumption of non-renewable resources (cnrr) including CO2 can be represented in a Hasse diagram with six levels and the degree of comparability is good. As illustrated by the POSAC output above, some countries with average to poor environmental performances (Mexico, Chile, France, Italy, Turkey and Spain) have a mild climate and relatively large agricultural sectors that consume water and produce pesticides; for arguably different reasons they also have a larger production of municipal waste. Other countries with comparably mediocre performances (Iceland, Estonia, Luxembourg, Norway, Finland and Czech Republic) have better results in these indicators but produce a larger amount of CO2 per capita, possibly as a consequence of very different climatic conditions.

Table A-1. Ordered logit models with social and environmental regressors (including CO2)

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
VARIABLES	TriSWB	HeptaSWB	TriSWB	HeptaSWB	TriSWB	HeptaSWB	TriSWB	HeptaSWB
dem	0.234*	0.0973*	0.124	0.124**	0.132*	0.131**		
	(0.123)	(0.0523)	(0.0849)	(0.0535)	(0.0739)	(0.0512)		
selfsuf			0.137*	0.225***	0.161***	0.248***		
			(0.0815)	(0.0754)	(0.0745)	(0.0724)		
equity							0.00611	-0.0328
							(0.0636)	(0.0472)
health	-0.215	-0.00350						
	(0.151)	(0.0540)						
soccohes			0.100	0.0566			0.155*	0.159**
			(0.0815)	(0.0655)			(0.0865)	(0.0661)
swb_0509	0.282**	0.306***	0.219**	0.259***	0.212***	0.264***	0.200***	0.216***
	(0.115)	(0.0815)	(0.0985)	(0.0738)	(0.0803)	(0.0702)	(0.0744)	(0.0571)
gdp	0.342*	0.122***						
	(0.178)	(0.0434)						
cnrr	0.0759	-0.100*	0.0304	-0.153**	0.00212	-0.160**	0.0595	-0.0833
	(0.104)	(0.0570)	(0.105)	(0.0658)	(0.0903)	(0.0652)	(0.0903)	(0.0529)
gpp	-0.0165	-0.0633	-0.118	-0.0735	-0.0859	-0.0676	-0.126*	-0.0512
	(0.0763)	(0.0493)	(0.0793)	(0.0488)	(0.0656)	(0.0469)	(0.0725)	(0.0418)
Constant cut1	19.02**	5.641**	8.880**	2.056	8.172**	2.031	6.191**	0.698
	(8.282)	(2.399)	(4.058)	(1.740)	(3.206)	(1.697)	(2.730)	(1.560)
Constant cut2	28.30**	9.300***	16.87**	6.021***	14.79***	5.962***	11.85***	3.672**
	(12.04)	(2.821)	(7.129)	(2.051)	(5.237)	(1.959)	(4.095)	(1.493)
Constant cut3		11.09***		8.446***		8.377***		5.130***
		(3.087)		(2.349)		(2.248)		(1.638)
Constant cut4		13.82***		12.63***		12.67***		7.547***
		(3.541)		(3.158)		(3.081)		(1.953)
Constant cut5		19.89***		17.71***		17.12***		11.53***
		(5.562)		(4.684)		(4.314)		(3.003)
Constant cut6		23.84***		22.26***		21.60***		15.41***
		(6.011)		(5.306)		(4.966)		(3.510)
N	33	33	33	33	33	33	33	33
STATISTICS								
Log-likelihood	-10.13	-27.43	-10.42	-23.75	-11.26	-24.14	-12.61	-30.38
LR Chi^2 test	49.95	64.99	49.36	72.34	47.68	71.57	44.99	59.09
Pseudo R^2	0.711	0.542	0.703	0.604	0.679	0.597	0.641	0.493
BIC	48.23	96.81	48.81	89.46	47.00	86.74	49.70	99.22
TESTS OF PROPORTIONALITY OF ODDS ACROSS RESPONSE CATEGORIES (P>Chi^2)								
Wolfe-Gould	0.002	0.004	0.028	0.073	0.010	0.016	0.004	0.002
Brant	1.000	1.000	0.000	-	0.000	1.000	1.000	0.133
Score	0.192	0.250	0.031	0.013	0.019	0.008	0.068	0.170

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1