In-Stream

INtegrating MainSTREAM Economic Indicators with Sustainable Development Objectives

<

SEVENTH FRAMEWORK PROGRAMME
THEME FP7-ENV-2007-1
Area 6.4.2.2. Sustainable development indicators
Grant agreement number: 211759

WP6: Costs of sustainability (general equilibrium analysis)

D 6.5: Sustainability index analysis

Francesco Bosello (*)

Elisa Portale (*)

Lorenza Campagnolo (*)

Fabio Eboli (*)

Ramiro Parrado (*)

July 2011

(*) Fondazione Eni Enrico Mattei

Table of Contents

1. Introduction ... 1

2. The ICES modelling framework ... 4

3. Indicator selection and construction for the FSI Index ... 8

4. Normalisation of indicators ... 10

4.1 Normalisation procedure: a theoretical overview ... 10

4.2 Definition of the benchmarks ... 11

5. Weighting and aggregating indicators ... 21

5.1 Weighting and aggregating: a methodological note ... 21

5.2 Weighting and aggegation procedure for the FSI ... 22

5.3 Results of the weighting procedure: the Shapley value of indicators ... 25

6 Results: FSI calculation ... 27

7 Sensitivity analysis ... 32

8 Policy application ... 35

9. Conclusions ... 39

References ... 41

Annex II: Non-Additive Measures for the Aggregation of the FEEM Sustainability Index ... 43

Robustness analysis ... 52

Determination of Shapley values for criteria ... 56

1. Introduction

Sustainability is one of the major concerns of those societies interested in understanding and governing the multi-faceted issue of development. Measuring sustainability requires the prior identification of its relevant dimensions and in particular drives a recognition of a series of indicators able to capture its diverse aspects. A comprehensive assessment of sustainability is then crucial to most of stakeholders, as it allows measuring progress, identifying policy areas to be addressed and evaluating the outcome of implemented policies.

The sustainability literature offers one particular feature that presents interesting opportunities for policy evaluations, that is the possibility to develop one single aggregated measure of sustainability.

In fact, sustainable development can be divided into three[1] or four[2] "pillars" and, for each one, a set of indicators can be defined. Theme-based indicators are probably the most used, because they are well suited to be linked to policy processes and targets. Several international institutions, among which the most known is probably the UN Commission on Sustainable Development (UN CSD), follow this approach.

Goal-oriented indicators are another category receiving increasing international acceptance as linking the choice of indicators with targets can allow an easier inter-country comparison and ranking. For instance, institutions such as the World Bank (WB), the United Nations Development Program (UNDP), the World Health Organisation (WHO) have started to use the Millennium Development Goals indicators as a standard reference. The European Union's choice of indicators itself mirrors its commitment to the Strategy for Sustainable Development (2001) and to the conclusions of the European Council (Barcelona 2002), the Plan of Implementation of the 2002 Johannesburg World Summit on Sustainable Development and the Lisbon Strategy.

Wide sets of sustainability indicators are meant to capture the complexity of the phenomenon and convey to policy makers important insights and directions, but their actual effectiveness has been limited (UN CSD 2006); current trends in sustainable development indicators show an increasing interest for core sets of headline indicators, goal-oriented indicators and aggregate indices (Pintér, Hardi and Bartelmus, 2005).

An "index" aggregates a set of indicators summarising many characteristics and aspects of a complex phenomenon. Aggregate indices have several positive aspects; they allow summarizing the relationship among the variables, facilitate communication to decision makers, and may serve as a basis for "early warning" (UN, 1995). Well-known examples of aggregate sustainability indices include the Human Development Index (HDI), the Environmental Sustainability Index (ESI), the Environmental Performance Index (EPI), the Ecological Footprint, the Index of Sustainable Economic Welfare (ISEW) and the Genuine Savings (GS).

However few issues like the construction and use of composite and aggregate sustainability indicators raise criticisms and debate in the already heated field of sustainability development

[1] Environmental, social and economic.
[2] Environmental, economic, social and institutional.

measurements. The reason is that any step in the construction of these indices are subjectivity prone, no matter the effort made as for example the choice of indicators to include, the choice of the “weights” to assign to each, the choice of the aggregation procedure. When this is the case, many criticisms can be perfectly legitimate and correct.

It is against this background, that this report describes the procedure to build, the application and results derived by a composite sustainability index proposed by FEEM within the INSTREAM project which is called in this report FEEM Sustainability Index (FSI). The analysis performed is a revisiting of INSTREAM D 6.2 (Bosello et al., 2010) where a similar investigation approach has been followed.

In particular: the FSI is constructed and computed within the framework of a recursive-dynamic general equilibrium model (ICES). The advantages of the use of this model are double:

- Firstly, the different implications for sustainability in different countries and regions can be assessed for different *future* social economic scenarios. That is: sustainability can be estimated *ex ante* and not only *ex post*.
- Secondly, the internal consistency of the model allows “by construction” a consistent integration of the different dimensions of sustainability inside the composite index. This will also encompass all feedbacks and interconnections among economic systems.

The dynamic aspect of this index along with its comprehensive nature, differentiates it from the majority of sustainability indexes available. Time is a concept that most sustainability analyses neglect either for the lack of proper instruments or for the lack of faith in the possibility to predict the future behaviour of some indicators. A (general equilibrium) model offers the possibility to evaluate future trends within a theoretically-founded framework. It works as a “consistency grid” for the indicators in time, keeping them meaningful with respect to one another

It is important here to iterate the aim of this exercise. It is not to propose “The INSTREAM composite sustainable indicator”, nor to create a new composite indicator to rival with current mainstream composite or aggregate indicators. The goals are more methodological and in line with INSTREAM objectives and they aim to test within this very specific framework the potential of composite indicators to provide synthetic measures of sustainability and deliver additional information compared to those conveyed by the “simple” GDP.

The major novelties respect to previous D6.2 (Bosello et al., 2010) derive from the incorporation of major suggestions provided by the INSTREAM scientific advisory board at the extent possible.

Namely: (a) strengthening the representation of the non economic pillars of sustainable development by including in the analysis additional indicators mainly referring to the environmental and social dimensions. (b) The analysis of the implication for country/region sustainability not only in a “business as usual development” up to 2020, but also of in “policy for sustainability” scenario coupling mitigation and investment in education (c) a sensitivity analysis on experts opinions.

We briefly recall the limitations of this methodology of which we are fully aware (for a more extended discussion see Bosello et al. 2010). Firstly, the sustainability assessment provided suffers from all the limitation of the CGE modelling approach (“Walrasian” full equilibrium view of the economic system, instantaneous adjustments to that equilibrium, calibration issues, simplified dynamics if any, emphasis on market values). Secondly, notwithstanding the complexity of a model, many important aspects of sustainability remain unavoidably outside its measuring ability. Albeit the efforts done, in the CGE model the economic component still dominates. Finally, subjectivity cannot be ruled out.

This said, we believe the exercise is anyway useful. Notwithstanding their limitations, economic (and CGE) models are increasingly used to support decision making in many domains. Thus it can be interesting to propose a further and innovative application of these instruments to the field of sustainability. Moreover, the modelling framework, offers a particularly appropriated environment for quantitative analyses. Thus comparing the informative content and performance of the FSI with respects to its different components and analyzing the links and trade-off between different sustainability dimensions, can be easier and more transparent.

In what follows, section 2 explains the ICES framework which simulates the baseline world economic, social and environmental scenario from 2001 up to 2020 in the future projection. Section 3 motivates the indicators' selection and explains the index construction. Section 4 presents the normalization procedure and describes the indicators' benchmark selected; section 5 highlights the methodology behind the indicators aggregation and the weighting procedure and the index tree construction and it traces sensitivity and robustness analysis of the index in the scenario. Section 6 describes the results of the FSI in the baseline scenario. Section 7 presents a sensitivity analysis conduced on the decision maker opinion. In section 8 is presented an application of the index into a climate policy scenario. Finally, section 9 presents main finding and conclusion.

2. The ICES modelling framework

The FEEM Sustainability Index (FSI) is built within the framework of a dynamic Computable General Equilibrium (CGE) model: the ICES (Inter-temporal Computable Equilibrium System) model. For a detailed description we address the interested reader to Bosello et al., (2010), here we summarize its main features. ICES is a recursive-dynamic, multi-regional CGE model of the world economy, calibrated in 2001. Data are provided by the GTAP6 database (Dimaranan et al. 2006). This includes information from 226 countries grouped in 87 regions and 57 sectors.

The model's core is based on the GTAP-E version (Burniaux and Truong, 2002) of the GTAP model (Hertel, 1997) best suited for the analysis of energy markets and environmental policies enriched with further disaggregation and data (Roson, 2003; Bigano et al., 2006).

The dynamic of the model is driven by two sources: one exogenous and the other endogenous. The first stems from exogenously imposed growth paths for some key variables - population, labour stock, labour productivity, land productivity. The second concerns the process of capital accumulation, according to which capital stock is updated over time following endogenous investment decisions.

ICES allows for a medium-term sustainability assessment until 2020. In the present version, ICES splits the world economy into 40 regions, consisting either of single countries or groups of countries (Table 1). Each economy is then detailed by 17 sectors (Table 2). This disaggregation has been chosen in order to provide an EU focus.

Table 1: ICES regional disaggregation (this study)

No.	Code	GTAP 6 Countries	No.	Code	GTAP 6 Countries
1	ARG	Argentina	21	MEX	Mexico
2	AUS	Australia	22	MEast	Bahrain, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Palestinian Territory, Occupied, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, United Arab Emirates, Yemen
3	AUT	Austria	23	NorthAfrica	Algeria, Morocco, Tunisia, Lybia, Egypt
4	Baltic	Estonia, Latvia, Lithuania	24	NZL	New Zealand
5	BNLX	Belgium, Netherlands, Luxembourg	25	POL	Poland
6	BRA	Brazil	26	PRT	Portugal
7	BUL	Bulgaria	27	RUS	Russia
8	CAN	Canada	28	ZAF	South Africa
9	CHN	China and Hong Kong	29	SEA	South-East Asia – American Samoa, Burma, Brunei Darussalam, (East) Timor Leste, Malaysia, Papua New Guinea, Philippines, (Western) Samoa, Singapore, Cambodia, Cook Islands, Fiji, French Polynesia, Guam, Kiribati, Lao People's Democratic Republic, Marshall Islands, Micronesia, Federated States of, Nauru, New Caledonia, Norfolk Island, Northern Mariana Islands, Niue, Palau, , Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu, , Vietnam , Wallis and Futuna
10	DNK	Denmark	30	ESP	Spain
11	FIN	Finland	31	SWE	Sweden
12	FRA	France	32	CHE	Switzerland
13	FSU	Former Soviet Union: Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Republic of, the former Yugoslav Republic of, Tajikistan, Turkmenistan, Ukraine, Uzbekistan	33	TUR	Turkey
14	GER	Germany	34	USA	USA
15	GBR	UK and Ireland	35	RoAfrica	Rest of Africa - Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo (Brazzaville), Cote d'Ivoire, Democratic Republic of Congo (Zaire), Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mayotte, Mozambique, Namibia, Niger, Nigeria, Reunion, Rwanda, Saint Melena, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe
16	GCM	Greece, Cyprus, Malta	36	RoAsia	Rest of Asia - Afghanistan, Bangladesh, Bhutan, Korea, Korea, Democratic People's Republic of, Macau,, Maldives, Mongolia, Nepal, Pakistan, Sri Lanka, Taiwan
17	IND	India	37	RoE	Rest of Europe – Albania, Andorra, Bosnia and Herzegovina, Faroe Islands, Gibraltar, Macedonia, Monaco, San Marino, Serbia and Montenegro(Yugoslavia)
18	IDN	Indonesia	38	RoEU	Rest of the European Union – Czech Republic, Romania, Slovakia, Slovenia
19	ITA	Italy	39	RoLA	Rest of Latin America - Anguilla, Antigua & Barbuda, Aruba, Bahamas, Barbados, Belize, Bolivia, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador , El Salvador, Falkland Islands (Malvinas), French Guiana, Grenada, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Martinique, Montserrat, Netherlands Antilles, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, Turks and Caicos, Uruguay, Venezuela, Virgin Islands, British, Virgin Islands, U.S.
20	JPN	Japan	40	NOR	Rest of the World – Bermuda, Greenland, Iceland, Liechtenstein, **Norway**, Saint Pierre and Miquelon

Table 2: ICES sectoral aggregation (this study)

No.	Code	Sector Description	Comprising GTAP Database V6 Sectors
1	**Food**	Food-related commodities	Paddy Rice, Wheat, Other Grains, Other Crops, Vegetables and Fruits, Plant Fibres, Cattle, Other Animal Products, Raw Milk, Wool, Cattle Meat, Other Meat
2	**Forestry**	Forestry	Forestry
3	**Fishing**	Fishing	Fishing
4	**Oth_Ind**	Other Industries	Beverages and Tobacco, Textiles, Wearing Appeal, Leather, Lumber, Paper and Paper Products, Fabricated Metal Products, Motor Vehicles, Other Transport Equipment
5	**Coal**	Coal	Coal
6	**Oil**	Oil	Oil
7	**Gas**	Gas, gas manufacture and distribution	Gas, Gas Distribution
8	**Oil_Pcts**	Petroleum, Coal Products	Water
9	**Electricity**	Electricity	Electricity
10	**MServ**	Market Services	Construction, Trade, Other Financial Intermediation, Insurance, Other Business Services, Dwellings, Transport
11	**Ins**	Insurance Services	Insurance and Pension Funding
12	**En_Int_Ind**	Energy Intensive Industries	Chemical Rubber Products, Non-Metallic Minerals, Iron and Steel, Non-Ferrous Metals
13	**Water**	Water	Water
14	**RD***	Research & Development	Research & Development
15	**Edu***	Education	Education
16	**Hea***	Health	Health
17	**NMServ**	Non- Market Services	Trade, Retail, Financial Intermediation, Renting

*** Additional sectors included to construct the FSI**

Note that some sectors relevant for the computation of the FSI, namely Research and Development (RD), Education (Edu) and Health (Hea), were not represented separately in the original GTAP database, though they were part of other more aggregate sectors. Thus, they had to be disentangled on purpose.

Two features of the CGE modeling approach (and in this case of the ICES model) are potentially appealing for the construction of a policy-oriented sustainability index. Firstly, the large database makes it possible to calculate the index for several regions, and to create indicators using data relative to the different sectors. Secondly, the nature of a CGE model, in which all sectors and regions are interconnected, is ideal to show how potential tradeoffs in sustainability originates and propagates through the economic system.

The possibility to project sustainability index in the future is very different from other aggregate sustainability indices, and it one of the major contributions of this work.

The sustainability performance of different regions is assessed on two different scenarios. A business as usual or reference scenario and a mitigation policy scenario.

Reference Scenario

The "reference scenario" is common across the quantitative work groups within the INSTREAM project. It assumes "medium" growth and no significant new policies up to 2020. Specifically:

- For the EU, GDP growth rates, population and labour stocks follow Carone *et al*. (2006)
- Carone et al. (2006) do not report projections for Non EU countries. However they replicate for the EU roughly half of the growth of the IPCC B2 SRES. Thus for non EU regions a GDP growth half of that of the B2 IPCC SRES scenario is assumed. Population follows UN projections.
- Land productivity follows the IMAGE model simulations (Image, 2001).

Table 3: Projected annual growth rates in the reference scenario for EU25

Regions	GDP growth rate 2004-2010	GDP growth rate 2010-2020
Austria	2.2	2.1
Belgium	2.4	2.1
Cyprus	4.3	4.1
Czech Republic	3.5	2.9
Denmark	2	1.9
Estonia	6.1	3.7
Finland	2.7	1.9
France	2.2	1.9
Germany	1.7	1.8
Greece	2.9	2
Hungary	3.7	2.8
Ireland	5.5	4.1
Italy	1.9	1.8
Latvia	7.7	4.4
Lithuania	6.5	4.3
Luxemburg	4	3
Malta	2.2	2.7
Netherlands	1.7	1.8
Poland	4.6	3.8
Portugal	1.9	2.4
Slovakia	4.6	4.2
Slovenia	3.7	2.8
Spain	3	2.5
Sweden	2.7	2.7
United Kingdom	2.8	2.5

Source: Carone *et al*. (2006)

Energy market dynamics are very important in ICES to determine the overall regional economic performances and emission profiles. The evolution of fossil fuel prices is thus a crucial variable to be considered. In the reference scenario they are close to the International Energy Outlook (EIA 2008) which highlights rather moderate increase if compared with other sources (see Table 4). This is considered more consistent with the assumption of moderate growth.

Table 4: Fossil fuel prices (% changes wrt 2001)

Oil Prices	2010	2020
ICES-FSI	**31.60**	**77.88**
World Energy Outlook (2008)	106.4	212.8
Intern. Energy Outlook (2008)	49.8	99.5

Gas Prices	2010	2020
ICES-FSI	**14.25**	**38.05**
World Energy Outlook (2008)	100.0	199.9
Intern. Energy Outlook (2008)	19.0	38.0

Coal Prices	2010	2020
ICES-FSI	**1.44**	**5.14**
World Energy Outlook (2008)	69.8	139.7
Intern. Energy Outlook (2008)	2.9	5.8

Policy Scenario

This consists in a unilateral EU emission reduction policy aiming , consistently with the EU climate and energy package, to curb greenhouse gas emissions the 20% below 1990 levels in 2020. The cap is imposed starting in 2010 and increases up to 2020, when the overall target is met. EU countries are allowed to trade emission within the EU to reach full efficiency in the abatement. No targets are imposed to Non-EU countries which cannot participate to the carbon market.

3. Indicator selection and construction for the FSI Index

The indicators included in the FSI (see Table 5 below) are an enlarged set compared to that considered by D6.2 (Bosello et al., 2010) on their turn a subset of those resulting from the thorough selection process described by INSTREAM D4.1 (Best el al. 2009) defining the scope of quantitative assessment within the project. Addressing the interested reader directly to Best et al. (2009), here it is anyway worth recalling briefly the outcome of the indicator selection procedure. The strategy adopted by INSTREAM was first of all to focus on a limited number of key indicators on which to perform quantitative assessment. This safeguards transparency and tractability. Then the choice has been made among those indicators which are more commonly used to measure sustainability, explicitly referring to the lists set by the Lisbon Strategy and the EU Sustainable Development Strategy. An obvious criterion for the choice has been to select those indicators within the measurement capability of modeling tools and quantitative approaches available within the INSTREAM quantitative research team. Table 5 groups the indicators in three domains or sustainability pillars. As a final result 23 indicators have been selected. 5 coincide with those of the Lisbon Strategy indicator shortlist (they are 14 in total). All of them are part of the larger set of sustainability indicators indicated by the EU Sustainable Development Strategy as key to measure progress toward sustainability.

Table 5: Indicators selection and description

SD Dimension	INDICATORS	DESCRIPTION
Economic	*GDP per capita (*) (**)*	GDP PPP / population
	Consumption per capita (*)	Consumption expenditure PPP / population
	Relative trade Balance (*)	Trade Balance / market openness *100
	Capital stock per capita (*)	Capital stock PPP/ population
	Capital stock growth rate (*)	(Capital stock(t) - Capital stock(t-1))/ Capital stock(t-1) *100
	Investment as %GDP (*)	Investment / GDP*100
	Terms of trade (*)	Value of export / Value of imports
	Total R&D expenditure as %GDP (*)	R&D expenditure / GDP *100
Environmental	*Energy Intensity (*)*	Energy Used / GDP PPP (Toe/ ml$)
	Greenhouse gases emission per capita (*)	N_2O+CH_4+CO_2 emissions / Population (Tons CO2 eq. per capita)
	Greenhouse gases Intensity (*)	N_2O+CH_4+CO_2 emissions / GDP PPP (Tons CO2 eq/ml $)
	CO_2 intensity	CO_2 emissions / Energy consumption (Tons CO2 eq/ Toe)
	Energy imported	Energy imported/ Energy Used (Toe/Toe)
	Renewable	Renewable energy consumption/ Energy Used (Toe/Toe)
	Plant biodiversity	Endangered plants/ total plants*100
	Animal biodiversity	Endangered animals/ total animals*100
	Water Use	Water use/ water available
Social	*Population growth rate (*)*	Population growth rate
	Food consumption	Household expenditure on food commodity / Household expenditure*100
	Insurance	Insurance expenditure/ GDP *100
	Education	Education expenditure/ GDP *100
	Private health expenditure	Private health expenditure/ Total health expenditure*100
	Total health expenditure	Total health expenditure/GDP*100

Note: (*) Already part of INSTREAM D6.2 (Bosello et al., 2010); (**) Part of INSTREAM selection of indicators for the quantitative analysis (Best et al., 2009)

4. Normalisation of indicators

4.1 Normalisation procedure: a theoretical overview

Normalisation is required prior to any data aggregation when, as usual, indicators in a data set have different measurement units (see Nardo, 2008).

The methods that appeared better-suited for the construction of the FSI is re-scaling applying a series of benchmarks that represent different levels of the indicators[3].

Each of these levels corresponds to a **normalised value** comprised between **0 and 1** as follows:

Table 6: Normalisation grid

0	Extremely unsustainable situation
0.25	indicator is still not sustainable but not as severely as in the previous case
0.50	a discrete level of sustainability, but still far from target
0.75	satisfactory level in the sustainability, yet not on target
1	target level, fully sustainable

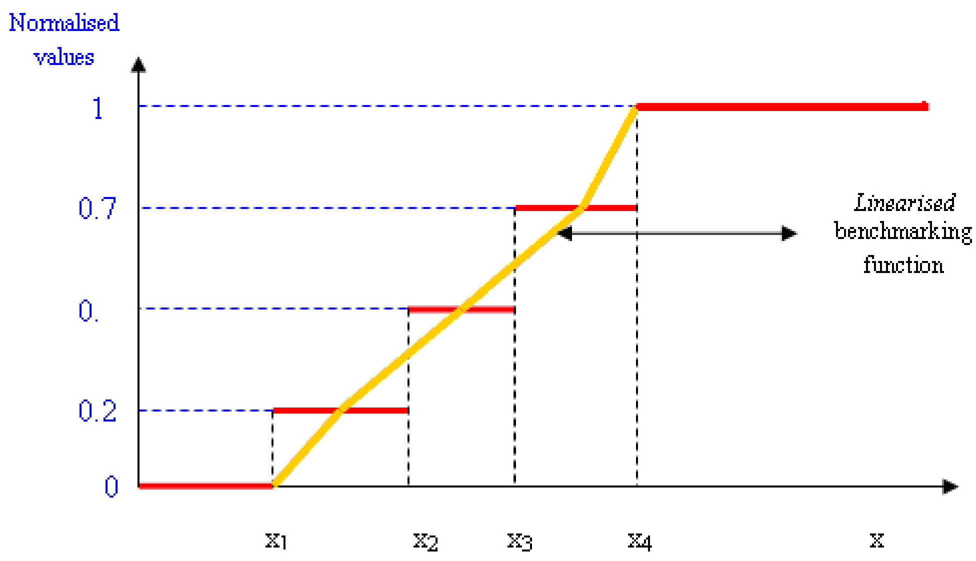
The normalised value corresponds to a specific level of the indicator, defined according to reliable and authoritative literature and international legislation sources, in order to increase the acceptability of the methodology. A different benchmark has been designed for each indicator according to different European and international sources, both at a technical and at a policy level. Whenever possible, the objectives outlined in the EU Sustainable Development Strategy or in the Lisbon Strategy have been used to define one or more level of the benchmarking function. In all other cases broader EU policy objectives and international standards from established institutions such as the OECD or the World Bank have been taken as primary source of information.

The five benchmark values for each indicator define four closed intervals and two open intervals that form a step function, whose form is different according to the values chosen (linear increasing, quadratic, etc). In order to avoid the discontinuity of a step function, each step has been “linearised”, taking the mean values of two subsequent intervals and interpolating, thereby creating a continuous step function as presented in Figure 1.

---

[3] Ranking or using categorical scales would not allow to keep the quantitative aspect of the indicators, and percentages of variations do not clarify whether the variation is sustainable or not. Standardisation using mean and standard deviation has also been excluded as it produces normalised indicators that are not comparable across different years, but only across countries in one specific year, finally rescaling in o -1 mostly shows the relative evolution of the value of the indicator in time across countries, rather than providing relevant boundaries. Furthermore, the reference minimum and maximum levels change over the years, making it difficult intertemporal comparison of sustainability.

Figure 1: the construction of the benchmarking function



The following section gives an overview of the five benchmark values proposed for each indicator and explains in detail how these values have been derived.

4.2 Definition of the benchmarks

At this stage each indicator is evaluated singularly, meaning that any interaction with other indicators is disregarded at the moment. Indicators are benchmarked to values that indicate different levels of performance in absolute terms, *per se*.

It is important to note that the benchmark functions defined are used to normalise all the projections up to 2020, that is the time horizon for this work, and that only one benchmarking function has been defined for all countries, including developed and developing ones. Even if we acknowledge that these may appear as oversimplifications, both decisions have a very important operative upside: using a common measure, which is stable in time, allows to emphasise the progress towards/from more sustainable levels of the indicators. Nonetheless, selecting one single benchmarking function is not a very simple task, even with the support provided by the numerous literature and policy sources. Therefore, the intervals defined by the five benchmark values derived by accredited sources have sometimes been modified in order to conform to the complexity of the situations to assess. These modification were also needed in order to capture the potential variability of the indicators across countries and create intervals able to show the modifications of the country ranking in time. Incorporating a very ambitious policy objective, well above most countries possibilities could be very meaningful in terms of sustainability, but if the intervals defined for the normalisation are not within a reasonable reach of the current situations the index will not be able to assess progress in time. This feature will be particularly important when evaluating policy implications in a further section.

The description of the benchmarking procedure is structured as follows. Firstly, a general definition of the benchmarking criteria will be given for every indicator, providing further explanations in the case in which the original source had been modified. Then, the benchmark table is shown.

1. GDP PER CAPITA

The benchmarking procedure for this indicator is limited to the appraisal of the economic situation; therefore, the higher scores are given to the higher level of GDP per capita, avoiding any evaluation of scarcity and optimal allocation of resources. GDP p.c. is measured in PPP (nominal) – US$ 2001. The definition of the five intervals corresponding to the five benchmarking values have been selected using as upper bound the average GDP per capita of the European Union in 2005, namely to 23500 Euro (34,446.30 US$).

DEFINITION	BENCHMARK VALUE
x<1500	0
1500≤x<15000	0.25
15000≤x<25000	0.50
25000≤x<35000	0.75
x≥35000	1

2. CONSUMPTION EXPENDITURE PER CAPITA

Consumption expenditure per capita is a clearer measure of average material well-being than GDP per capita, which is more concentratd on the production side. As for GDP the definition of the five intervals corresponding to the five benchmarking values have been selected using as upper bound the average per capita consumption of the European Union in 2005.

DEFINITION	BENCHMARK VALUE
x<1000	0
1000≤x<8000	0.25
8000≤x<16000	0.5
16000≤x<25000	0.75
x>25000	1

3. RELATIVE TRADE BALANCE

Relative trade balance (RTB) is an indicator of competitiveness. This indicator compares the trade balance (exports minus imports) to the market openess (exports plus imports). RTB highlights the ability to sell home goods in foreign markets (exports exceeds imports, i.e. a trade surplus) or the excessive dependence from imported goods (expenditure on imported goods is higher than export revenues, i.e. a trade deficit). There are no specific targets for trade balances but it is preferible to have a positive value of the rate, which means a trade surplus condition. The average value for the benchmark corresponds to a neutral trade position (export equal to imports); upper and lower bound values guarantee a symmetric distribution of country specific performances across time.

DEFINITION	BENCHMARK VALUE
X<-10	0
-10≤x<-5	0.25
-5≤x<0	0.50
0≤x<5	0.75
x≥5	1

4. Capital Stock per capita

Conditions for achieving economic sustainability impose that per capita capital stock is high enough to allow economic growth; this depends on a consistent investment effort. The higher the amount of capital stock per capita, the higher the score. The distribution of values across normalisation scale has been defined according to results of our simulations, given that capital accumulation is endogenous in the model and in absence of targets from exogenous sources.

DEFINITION	BENCHMARK VALUE
x<12000	0
12000≤x<40000	0.25
40000≤x<80000	0.50
80000≤x<120000	0.75
x≥120000	1

5. Capital Stock growth rate

Capital stock growth rate depends on accumulation of investments through time, then high level of investments contributes to increase capital stock and then to speed up growth rate. Capital stock and investment are important component in GDP growth, so good scores are attributed to highest capital stock growth rate.

DEFINITION	BENCHMARK VALUE
x<0%	0
0%≤x<2%	0.25
2%≤x<4%	0.50
4%≤x<6%	0.75
x≥6%	1

6. Investments as % GDP

An increase of business investment over GDP is considered good as capital is a key production factor and is produced through investment. The higher the share of investments over GDP, the higher the score in terms of sustainability. The definition of the five intervals corresponding to the five benchmarking values have been selected using as upper bound the average share of investment in European Union in 2005, namely 17% of GDP.

DEFINITION	BENCHMARK VALUE
x<8%	0
8%≤x<11%	0.25
11%≤x<14%	0.5
14%≤x<17%	0.75
x>17%	1

7. Terms of Trade

Terms of trade indicator can be summarized as the total value of exports over the total value of imports. A value higher than 1 of this indicator mirrors a trade surplus condition. This is a beneficial position and stimulates the competitiveness of economic systems. The average benchmark value corresponds to a neutral trade position (export equal to imports). Once again, there is no official or recognized attribution of targets for sustainable terms of trade thus also in this case the five benchmarking values has been chosen according to simulation results.

DEFINITION	BENCHMARK VALUE
x<0.8	0
0.8≤x<0.9	0.25
0.9≤x<1	0.50
1≤x<1.1	0.75
x≥1.1	1

8. Research and Development

The benchmark values for this indicator, already included in the EU SDS and in the Lisbon Strategy list, have been devised starting from the target set at the 2000 European Council in Barcelona, which suggests to increase spending of R&D innovation with the aim of approaching 3% of GDP by 2010. Starting from this target, 3% of GDP has been set as a benchmark for the maximum normalised score, with the lower bound set for expenditures lower than 0.2%.

DEFINITION	BENCHMARK VALUE
x<0.2%	0
0.2%≤x<1%	0.25
1%≤x<2%	0.50
2%≤x<3%	0.75
x≥3%	1

9. Energy intensity

Energy is essential for economic and social development. However, energy use affects resource availability and the environment. Improving energy efficiency and decoupling economic development from energy use are important sustainable development objectives. This indicator is present in the EU SDS and in the Lisbon Strategy. Energy intensity is calculated as units of energy used per unit of GDP PPP in Tons of oil equivalent/$m.

The world energy intensity in 2006 is 218 Toe/$mil GDP-PPP (WRI 2006).The benchmark for energy intensity has been set giving the highest score to countries which can obtain 150 Tons oil equivalent/$m unit of energy and the lowest to those obtaining 300 Tons oil equivalent/$m or more units of energy.

DEFINITION	BENCHMARK VALUE
x<150	1
150≤x<200	0.75
200≤x<250	0.50
250≤x<300	0.25
x≥300	0

10. GHG EMISSION PER CAPITA

According to IEA (2009) the annual average per capita emission level is equal to 16 Tons of CO_2 equivalent per capita in 2008. Considering future trend in population, in order to take at least stable GHG concentrations, the optimal level of GHG emission per person/year should be no more than 1.1 Tons CO_2 equivalent per capita (IPCC, 2007) which is also broadly consistent with the EU target of a reduction of the 20% by 2020 with respect to 1990. The benchmark for GHG emission per capita has been set giving the highest score to countries which reports less than 4 unit of GHG per capita and the lowest to those producing 20 or more units of GHG per capita.

DEFINITION	BENCHMARK VALUE
x<4	1
4≤x<9	0.75
9≤x<13	0.50
13≤x<20	0.25
x≥20	0

11. GHG EMISSIONS INTENSITY PER GDP

It is probably the main established indicator measuring the “climate-change performance” of an economic system. It has an intrinsic paramount policy relevance as the target for climate change mitigation policies have been establishedin the Kyoto Protocol and debated in the post Kyoto process. It can offer also important distributional insights on climate change costs and benefits. The world GHG intensity in 2006 is 529 Toe/$mil GDP-PPP (WRI 2006). The benchmark for GHG intensity has been set giving the highest score to countries which produce less than 400 unit of GHG emissions and the lowest to those have a rate of 1300 or more units of GHG emissions per GDP.

DEFINITION	BENCHMARK VALUE
x<400	1
400≤x<700	0.75
700≤x<1000	0.50
1000≤x<1300	0.25
x≥1300	0

12. CO_2 INTENSITY

CO_2 intensity measures the quantity of carbon dioxide (CO_2) released into the atmosphere on the total energy consumption in the country. The average rate of percentage variation of the carbon intensity of energy consumption in Europe is 2.1% (EEA). The upper bound for the normalisation exercise has been set at 1.2%. The other ranges were defined according to the projection distributions and the fact that the value for this indicator has to be decreasing in the future because of increasing in renewable share of energy and a decreasing of CO_2 emissions.

DEFINITION	BENCHMARK VALUE
x<1.2%	1
1.2%≤x<1.6%	0.75
1.6%≤x<1.8%	0.50
1.8%≤x<2.2%	0.25
x≥2.2%	0

13. ENERGY IMPORTED

This indicator measures the ratio between energy imported and energy consumed in Toe. In particular the "Green Paper - Towards a European strategy for the security of energy supply" by the European Commission reports that if no measures are taken, in the next 20 to 30 years 70% of the Union's energy requirements (the current energy imported share is 50%) will be covered by imported products in all sectors of the economy. This indicator is related to the energy security problem which is one of the current global challenges. These statements have been at the basis of the definition of the sustainable level of imports in a range between 22 and 45% of overall energy consumption.

DEFINITION	BENCHMARK VALUE
x<9%	1
9%≤x<22%	0.75
22%≤x<45%	0.50
45%≤x<60%	0.25
x≥60%	0

14. RENEWABLE

Clean energy is generate from renewable sources as hydropower, geothermal, wind and solar power[4]. A suitable benchmark for this indicator can be found in the EU policy of recent years proposed in the package "Climate action and renewable energy package". The EU is committed to reducing its overall emissions to at least 20% below 1990 levels by 2020, and is ready to scale up this reduction to as much as 30% under a new global climate change agreement when other developed countries make comparable efforts. It has also set itself the target of increasing the share of renewable in energy use to 20% by 2020.
Therefore, the ambitious target of 30% for a new global climate change agreement has been chosen as upper bound for this normalisation exercise; even if this target may seem to ambitious for some countries, it must be kept in mind that the level of this indicator as projected by the model tends to overestimate future trends in clean energy consumption. This is due to the construction of this indicator within the model.

DEFINITION	BENCHMARK VALUE
x<0.5%	0
0.5%≤x<4%	0.25
4%≤x<10%	0.50
10%≤x<30%	0.75
x≥30%	1

[4] This indicator is present in the WDI list and values by World Bank for this indicator in the period 2005-2008 are quite different among countries. Low income and lower-middle income countries have a share of 4% of total energy use; middle income countries have a share of 5%. Upper-middle and high income countries have an average of respectively 8% and 14% of Clean Energy of total energy use.

15 and 16. BIODIVERSITY INDEX OF PLANTS (BIP) AND BIODIVERSITY INDEX OF ANIMALS (BIA)

The literature on biodiversity focuses on best practices and goals that are not directly measurable through such an index, therefore more general objectives have been at the core of the definition of the benchmarking values. One of the Lisbon objectives devised in 2001 committed EU countries to try and halt biodiversity loss by 2010. A wider biodiversity strategy was already introduced in 1998, characterised by general non-quantitative objectives, which nevertheless pushed in favour of nature-conservation strategies. The number of endangered species is linked to the model so that over time it is possible to see how, due to human activities, there are more and more plants and animals species in danger of extinction.

BIODIVERSITY INDEX PLANTS

DEFINITION	BENCHMARK VALUE
$x<6$	1
$6 \leq x<25$	0.75
$25 \leq x<50$	0.50
$50 \leq x<85$	0.25
$x \geq 85$	0

BIODIVERSITY INDEX ANIMALS

DEFINITION	BENCHMARK VALUE
$x<6$	1
$6 \leq x<10$	0.75
$10 \leq x<15$	0.50
$15 \leq x<20$	0.25
$x \geq 20$	0

According to a study published on Nature (Thomas et al, 2004) relative to 1,103 animal and plant species in sample regions covering some 20% of the Earth surface, 15 to 37% of the species risk extinction because of climate change scenarios at 2050. In order to proxy for climate change, the number of endangered species has been inversely linked to emissions of CO_2. The defining criterion for the benchmarking values has been simply to apply the concept to "halt biodiversity loss"(Lisbon 2001). The analysis of the numerical data of the GTAP-ICES model for 2001 has provided the basis for building the benchmarking values: the average of better performing countries' scores has been used as upper bound and the remaining intervals have been decided respectively.

17. USE OF RENEWABLE WATER RESOURCES

The indicator shows to what extent freshwater resources are already used, and the need for adjusted supply and demand management policy. The indicator varies between countries and in time as a function of climate, population, and economic and institutional capacity to manage water resources and demand. According to the World Business Council for Sustainable Development, water stress/scarcity occurs where there is not enough water for all uses, whether agricultural, industrial or domestic; when the amount withdrawn is so great that water supplies are no longer adequate to satisfy all human or ecosystem requirements. The extensive literature on the sustainable management of water resources unfortunately does not provide a quantification of the ideal pressure on resources. From a strictly environmental perspective, this pressure should not exceed full capacity. To get close to this, 90% has been chosen as lower bound for this indicator and 1% is the upper bound.

DEFINITION	BENCHMARK VALUE
x<1%	1
1%≤x<15%	0.75
15%≤x<40%	0.50
40%≤x<90%	0.25
x≥90%	0

18. GROWTH RATE OF POPULATION

Population growth can be related to many different contexts. Here, it has been evaluated only with respect to social pressures it may determine. Countries that exhibit a too high population growth rate are expected to face social tensions in terms of resource uses and distribution. Therefore, the highest scores have been assigned to countries with a stable (slightly declining or increasing) level of population growth.

DEFINITION	BENCHMARK VALUE
x<-0.8%	1
-0.8%≤x<0.2%	0.75
0.2%≤x<0.8%	0.5
0.8%≤x<1.2%	0.25
x≥1.2%	0

19. FOOD RELEVANCE IN PRIMARY CONSUMPTION

This indicator of poverty is based on the considerations arising from the empirical evidence supporting Engel's law, according to which the proportion of a nation's income spent on food is a good index of the nation's welfare. The lower the proportion, the more prosperous the nation. This indicator is expressed as the ratio of food consumption to the consumption of all primary goods (food, non energy-intensive goods). The benchmarking function was derived observing the initial distribution of this indicator across countries in the GTAP database, following the general indication that the lower the share of food consumption is, the higher the relative income the country can count upon.

DEFINITION	BENCHMARK VALUE
x<2%	1
2%≤x<7%	0.75
7%≤x<20%	0.5
20%≤x<40%	0.25
x≥40%	0

20. EXPENDITURE IN INSURANCE AND PENSIONS

This indicator includes only private-funded insurances and pensions. The individuals are mainly focused on avoiding the welfare losses by using the precautionary savings in order to secure individuals finances in the future. Households are subject to several sources of risk (in earnings, health, and mortality). Markets to insure those risks are often limited or do not exist. The main way to self-insure against them is to accumulate a buffer stock of wealth (Cagetti M., 2003). This indicator reflects the reasoning behind this publication, concluding that a reliance on the private sector for what concerns insurance and pensions should be encouraged especially because of the increase in uncertainty about the future path of income. As in previous cases the ranges defined for the normalisation have been set in order to be applicable to the projection produced by the model.

DEFINITION	BENCHMARK VALUE
x<0.1%	0
0.1%≤x<0.5%	0.25
0.5%≤x<1.3%	0.5
1.3%≤x<2.5%	0.75
x≥2.5%	1

21. EDUCATION

There is no standard international target for GDP spent on education. A general target referenced, but not sanctioned by international conventions or agreements, suggests that countries should devote at least 5% of GDP to education. The education indicator included in this work cannot deal with the problem of poverty alleviation, but only with the amount of wealth spent in education regardless on how well it performs in terms of social wealth. Therefore, as middle bound a benchmark between 4.5 and 5.5% of GDP has been chosen and the other benchmarking values have been defined starting from this value and observing the distribution of this indicator in the GTAP database.

DEFINITION	BENCHMARK VALUE
x<2%	0
2%≤x<4.5%	0.25
4.5%≤x<5.5%	0.5
5.5%≤x<9%	0.75
x≥9%	1

22. HEALTH EXPENDITURE BY PRIVATES ENTITIES

This indicator describes the percentage of total private health expenditure on the overall health expenditure. The European policy literature has been at the centre of the definition of the benchmark values for this indicator.

The European Health Strategy defines as one of its aspirations that of “enabling health for all”. The ideal level was set at 15%, because assuring a high level of health care system has to be one of the most important Government Issue. The other levels are consequential according to the distribution observed in the results.

DEFINITION	BENCHMARK VALUE
x<15%	1
15%≤x<30%	0.75
30%≤x<45%	0.5
45%≤x<60%	0.25
x≥60%	0

23. OVERALL HEALTH EXPENDITURE

This indicator considers the overall health expenditure on GDP (nominal). The WHO (1981) in the "Global Strategy for All by the Year 2000" stated that at least 5% of the Gross National Product should have been spent on health.
Considering that the sustainable benchmark needs to be applied to all countries, the lower limit chosen is 4%, close to the target introduced by WHO in 1981, while for the upper limit a triple value (12%) has been chosen. Empirical research confirms that the OECD average in health expenditure as a percentage of 2005 was 9%. However, this share varied considerably across OECD countries, ranging from around 6% in Korea, Poland and Mexico up to 15.3% of GDP in the case of the United States.

DEFINITION	BENCHMARK VALUE
x<4%	0
4%≤x<7%	0.25
7%≤x<9.5%	0.5
9.5%≤x<12%	0.75
x≥12%	1

5. Weighting and aggregating indicators

The normalised scores allow direct comparison between the different aspects of sustainability, but can say little on the overall appraisal of a country. The contribution of this work is also the construction of an index, which aggregates indicators related to different dimensions of sustainable development into one single measure, easily comparable across countries and in time. Yet, aggregating indicators that are very different in nature into one single measure is a very delicate process.

5.1 Weighting and aggregating: a methodological note

To aggregate indicators, it is necessary to assign a weight, to determine their relative importance to the final composite index. Once weights are defined, different techniques can be used to combine the weighted indicators into one single measure[5]. A broadly used aggregation technique is the **Equally Weighted Average (EWA)**, in which all indicators are given the same weight. It assumes a perfect substitutability among indicators and does not capture the relative importance of different sustainability indicators (Saisana et al., 2005), nonetheless it is considered the most transparent methodology (Yale Center for Environmental Law and Policy, 2005).

In fact, perfect substitutability very often fails to be satisfied. In these circumstances the **Non-Additive Measures** approach **(NAM)** is sufficiently general to cover a lot of preference structures of the Decision Makers (also called DM) and allows for the modeling of many types of interactions going from the compensative to the substitutive attitude of the DM. The first indicates that her/his satisfaction is high only if all the criteria are satisfied (this can be represented by the logical conjunction operator "AND"). The second that her/his satisfaction is high if at least one of them is high (this can be represented by the logical disjunction operator OR).

How to obtain these weights? They can be obtained as usual by means of questionnaires. In practice an exponentially increasing numerical complexity is determined by the number of weights involved. Comparing NAM and EWA approaches: if *n* is the number of the criteria or indicators, the **EWA** approach, only needs to determine *n* weights, a **NAM** approach requires the specification of 2^n weights, i.e. one for every possible combination of the *n* criteria (see Annex II for further detail).

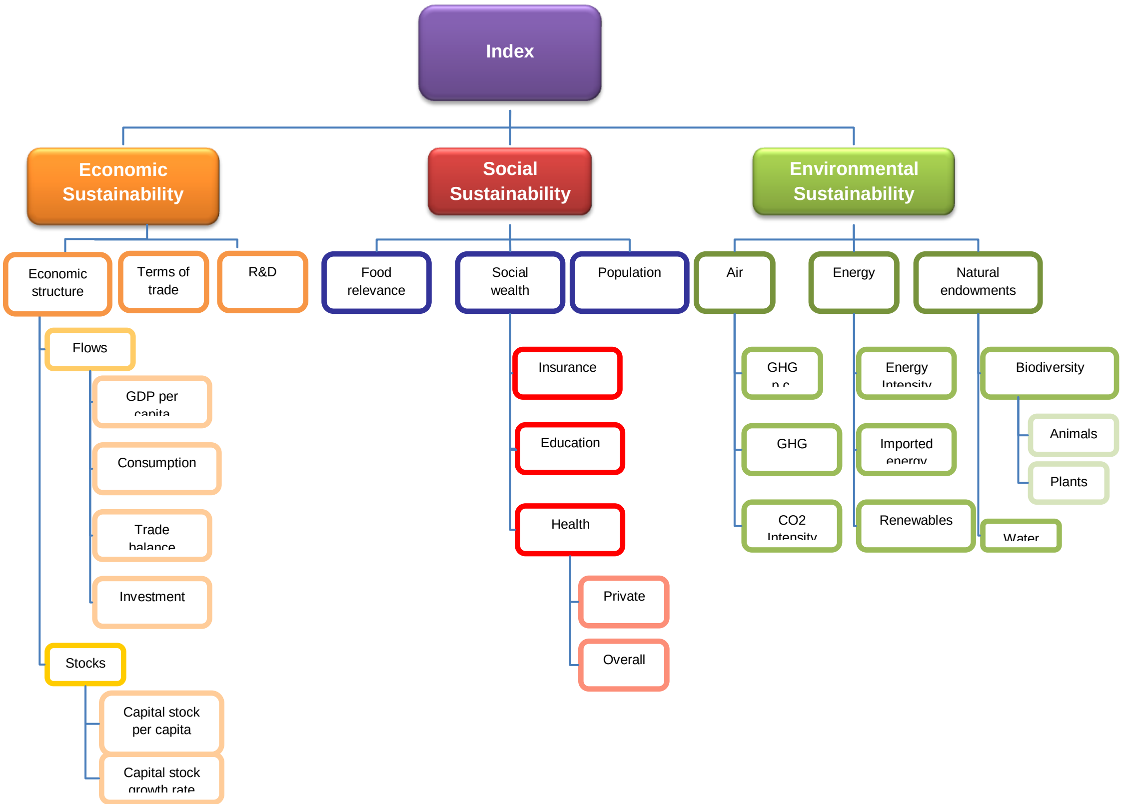
This higher complexity allows the more realistic "violation" of the linearity implicitly assumed in EWA. In fact, the "weight" of a coalition of sub-criteria (group of indicators or nodes) can be greater or lower than the sum of the weights of each of the sub-criterion belonging to the coalition itself[6].
Subsequently, a simple algorithm, the so-called Choquet Integral (De Waegenaere, 2001 and Murofushi, 1994), computes the weighted average, taking the coalitions of criteria into account, instead of considering the single criteria.

[5] For a more thorough review of weighting and aggregation methodologies the reader should refer to the Handbook on composite indicators (Saisana *et al.*, 2005).
[6] Naturally, if for every coalition the weight (or the *importance*) of each coalition is formed by the sum of the weights of each sub-set of its criteria forming a partition, we obtain again the EWA. Conversely, if for a coalition its weight is *inferior* to such a sum, a redundant interaction exists among the included criteria, while if it is *greater* than the sum, a synergic interaction exists.

5.2 Weighting and aggregation procedure for the FSI

The criteria (indicators or nodes) that have to be evaluated for the construction of the FSI are represented in the aggregation tree (Figure 2) that thematically groups the 23 selected indicators into 12 nodes[7].

Figure 2: FSI aggregation tree



This decision tree requires attributing weights to all combination of criteria (sub-coalition performances) present at each node: i.e. 2^n, where *n* is the number of criteria in the considered node.

Weights have been obtained through a questionnaire involving 12 respondents (10 FEEM researchers and 2 external) within a focus meeting in which a common consensus on each answer has been reach through discussion.

Following (Despic, 2000), to facilitate the complex evaluation, respondents were asked to weight each possible combination of indicators, but considering just their best or worst level. These were

[7]This tree is slightly different from that in Bosello et al. (2010). The R&D indicator is not anymore considered among Social sustainability indicators, but as an Economic one. This is in favor of more compliance with the UN CSD set of indicators which considers R&D expenditure as measure of Economic development (UN, 2007). Furthermore a rearrangement in the Economic Sustainability pillar, in particular in Economic structure, has been performed in order to simplify preference elicitation through questionnaires.

not defined in quantitative terms in order to avoid bias deriving from respondents disagreeing with the judgment given. It is up to the respondent to "imagine" what "best" and "worst" mean in each case. A weight of "0" corresponds to the WORST case where all indicators in a node have the worst performance, and "100" to the BEST one where all indicators have the best performance. The questionnaire thus boils down to the list of all the combinations of BEST and WORST values.

A further requirement has been imposed to respondents: the respect of the **monotonicity criterion** in weight assignment. This principle requires that the weight of a coalition cannot be less than the lowest weight of each sub-coalition included. Box 1 gives an example to clarify this process.

As Table B1 (in the BOX 1) shows, evaluations are normalized in the interval [0, 1]. Then the Möbius transformation (Grabisch, 2003) is applied to take into account synergic and redundancy interactions among indicators. If the Möbius weight is null, no interaction exists, as in the WA case, if it is positive there is a synergy, if negative, a redundancy.

Finally the **Choquet integral,** can be directly calculated using the Möbius values as weights, according to the following formula:

$$C_m(x_1, ..., x_n) = \sum_{T \subseteq S} a(T) \cdot \min_{i \in T}(x_i)$$

where x_i are values of single indicators considered in the node and $a(T)$ are the Möbius weights for the coalition T. A more exhaustive explanation of this procedure can be fount in the Technical Appendix II.

BOX 1: CONSTRUCTION OF THE INDICATOR-COALITION MATRIX NORMALIZATION AND MOBIUS TRANSFORMATION

Assume to weight the top nest of the FSI tree. This implies weighting the three main components of sustainability: Economic, Social and Environmental, *and* their possible combinations, i.e. to express 2^3= 8 weights.

The respondent give a valuation "at the edge" (evaluating only the Worst or the Best performance) of each indicator and of all sub-coalition of the nest (criteria). We can write down a matrix containing the valuations for each possible combinations of indicators. The first (all Worst) and the last (all Best) must necessarily be valued respectively 0 and 100, whereas the other combinations only need to respect the monotonicity criterion and may lie between 0 and 100.

Table B1: evaluations and Mobius transformation

SUSTAINABILITY			Weights	Normalized weights	Möbius transformation
Economic	**Social**	**Environment**			
Worst	**Worst**	**Worst**	**0**	**0**	**0**
Best	**Worst**	**Worst**	20	0.2	0.2
Worst	**Best**	**Worst**	50	0.5	0.5
Worst	**Worst**	**Best**	30	0.3	0.3
Best	**Best**	**Worst**	x≥50 e.g.60	0.6	0.6-(0.2+0.5)=-0.1
Best	**Worst**	**Best**	x≥30 e.g.50	0.5	0.5-(0.2+0.3)=0
Worst	**Best**	**Best**	x≥50 e.g.90	0.9	0.9-(0.5+0.3)=0.1
Best	**Best**	**Best**	**100**	**1**	**1**

Thus, the questionnaire allows to collect respondents judgments on good performance of a single indicator (e.g. Economic sustainability), but also on mutual good performance of two of them (e.g. Economic sustainability and Environmental sustainability); clearly this evaluation must be higher than the highest score of the indicators considered alone (e.g. if the Best performance only in Economic sustainability has score 20 and in Environmental sustainability 30, the evaluation must be bigger than 30 when both these indicators have the Best performance).

5.3 Results of the weighting procedure: the Shapley value of indicators

To describe the importance of each indicator or indicator node in the FSI tree as results of the expert responses the Shapley index can be used. It represents the weighted marginal effect of including a given indicator to every coalition of indicators not including itself, belonging to a hierarchically superior node (see Appendix II for details, for a more formal definition, see Grabish, 1995 and 1996). If it is null, it means that adding the indicator to any indicator coalition does not change the score, and thus it can be excluded. Conversely, if it equals 1 (the maximum value), every coalition excluding it scores as zero, every coalition including it scores 1, and, therefore it turns out to be the only important criterion.

The Shapley values were computed first at each node of the INSTREAM tree (see Table 7) to determine the local relative importance of each criterion (indicator or node) to a hierarchically superior node

Table 7: Local Shapley values of criteria

Node	Criterion	Shapley value
Index	Economic Sustainability	28.33%
	Social Sustainability	33.33%
	Environmental Sustainability	38.33%
Economic Sustainability	Economic structure	47.50%
	Terms of trade	20.00%
	R&D	32.50%
Social Sustainability	Population	29.17%
	Social Wealth	34.17%
	Food relevance	36.67%
Environmental Sustainability	Air	26.67%
	Energy	34.17%
	Natural endowments	39.17%
Economic structure	Flows	60.00%
	Stocks	40.00%
Social wealth	Insurance	21.67%
	Education	36.67%
	Health	41.67%
Air	CO2 intensity	37.50%
	GHG per capita	30.00%
	GHG intensity	32.50%
Energy	Energy intensity	37.50%
	Imported energy	30.00%
	Renewables	32.50%
Natural endowments	Biodiversity	45.00%
	Water	55.00%

Flows	GDP p.c.	33.33%
	Consumption	28.33%
	Relative trade balance	14.67%
	Investment	23.67%
Stocks	Capital stock per capita	60.00%
	Capital stock growth rate	40.00%
Health	Private health	35.00%
	Overall health	65.00%
Biodiversity	Plants	52.50%
	Animals	47.50%

Thus, for instance, the average marginal gain of adding the node *Economic Sustainability* to the final index is about 28.3%, while *Social Sustainability* and *Environmental Sustainability* contribute for 33.3% and 38.3%, of the FSI final performance respectively. The results are rounded to the second decimal place and it can be checked that the Shapley values at any given node always sum to one.

These, however, are *local* Shapley values for each criterion at a given node. In order to estimate how much each indicator contributes to the final index, one could multiply the Shapley values of each criterion needed to reach the index from each indicator. For instance, given that the node *Social Sustainability* contributes 33.3% to the final index and the indicator *Food relevance* contributes 36,7% to *Social Sustainability*, then the indicator *Food relevance* contributes about 12.2% to the FSI. Similar results, ranked by size of contribution, are reported in Table 8.

Table 9: Contribution of single indicators to the FSI

Indicator	Contribution
Food relevance	12.22%
Population	9.72%
R&D	9.21%
Water	8.26%
Terms of trade	5.67%
Energy intensity	4.91%
Renewables	4.26%
Education	4.18%
Imported energy	3.93%
CO2 intensity	3.83%
Plants	3.55%
GHG intensity	3.32%
Capital stock per capita	3.23%
Animals	3.21%
Overall health	3.08%
GHG per capita	3.07%
GDP p.c.	2.69%
Insurance	2.47%
Consumption	2.29%
Capital stock growth rate	2.15%
Investment	1.91%
Private health	1.66%
Relative trade balance	1.18%
Sum	100%

Note: different colors correspond to the three areas of sustainability: economic (yellow), social (purple), environmental (green)

According to respondents' judgments the most important indicator is *Food relevance,* followed by *Population* and the ratio of *R&D expenditure* . The indicator with the lowest marginal contribution is *Relative trade balance*. A noteworthy feature of the contribution of indicators to the final index is that they are influenced both by the evaluation of decision makers (which determine the local Shapley values) and by the structure of the tree, in which some indicators contribute more directly than others to the final index.

6 Results: FSI calculation

Table 11 describes the FSI sustainability ranking in 2010 and in 2020 for the 40 ICES regions in the reference scenario. EU regions/countries are highlighted in yellow.

Table 7: FSI Regional Ranking 2010 – 2020

	Regions	2010 baseline	Rank	2020 baseline	Regions	
1	SWE	0.684	=	0.685	SWE	1
2	CHE	0.632	-1	0.649	AUT	2
3	AUT	0.628	1	0.635	CHE	3
4	FIN	0.620	=	0.631	FIN	4
5	GBR	0.583	=	0.618	GBR	5
6	FRA	0.574	=	0.597	FRA	6
7	DNK	0.561	-2	0.579	CAN	7
8	CAN	0.560	1	0.575	GER	8
9	JPN	0.557	-1	0.567	DNK	9
10	GER	0.549	2	0.567	JPN	10
11	NOR	0.519	=	0.543	NOR	11
12	ITA	0.508	=	0.535	ITA	12
13	NZL	0.491	=	0.514	NZL	13
14	ESP	0.479	-1	0.510	BNLX	14
15	BNLX	0.470	1	0.501	ESP	15
16	RUS	0.466	=	0.496	RUS	16
17	Baltic	0.442	-1	0.469	PRT	17
18	RoEU	0.440	-1	0.460	Baltic	18
19	PRT	0.438	2	0.454	RoEU	19
20	ARG	0.394	-8	0.415	SEA	20
21	USA	0.389	-2	0.408	RoE	21
22	RoE	0.375	1	0.397	GCM	22
23	MEX	0.370	-2	0.395	USA	23
24	ZAF	0.362	-2	0.387	BRA	24
25	BRA	0.353	1	0.385	MEX	25
26	SEA	0.348	6	0.383	ZAF	26
27	GCM	0.340	5	0.380	AUS	27
28	AUS	0.332	1	0.370	ARG	28
29	IDN	0.318	=	0.358	IDN	29
30	RoAsia	0.315	-3	0.352	TUR	30
31	POL	0.313	=	0.349	POL	31
32	BUL	0.294	=	0.329	BUL	32
33	RoLA	0.282	-1	0.318	RoAsia	33
34	FSU	0.274	-2	0.313	RoLA	34
35	TUR	0.266	5	0.276	CHN	35
36	CHN	0.257	1	0.244	FSU	36
37	MEast	0.231	=	0.239	MEast	37
38	RoAfrica	0.211	=	0.218	RoAfrica	38
39	NorthAfr	0.158	=	0.176	NorthAfr	39
40	IND	0.143	=	0.154	IND	40

EU and other developed countries are currently among the best performers both in 2010 and 2020. North European countries (Sweden and Finland), Austria, United Kingdom and Switzerland rank in the first 5 places in 2010, while Mediterranean and Eastern European countries are between the 10^{th} and 20^{th} place. Developing countries, especially African and Asian score very low.

Along the simulation period there is only a modest re-ranking; it is marginal in the top and bottom ten. Considering that in the time frame analyzed there are not huge economic and environmental transformations this homogeneous behavior should not surprise.

It is however interesting to analyze the information the FSI provides on regions' pathway toward full sustainability, associated to a score of 1. In the 10 years total sustainability improves on average by 5%. South East Asia, Greece, Cyprus and Malta (GCM), Turkey highlights improvements above average, whereas Argentina, and Former Soviet Union countries worsen. Moreover Sweden, the top performer, in 2020 still scores 0.7, rather far from 1.

All this signals that even a scenario of generalized economic growth, without excessive pressure on resources, and embedding the assumption of economic convergence between developed and developing countries may not necessarily imply improvements in sustainability, not to mention its rapid reach. Respect to this, the worsening of sustainability in economies like those in transition, that cannot be considered "poor", is particularly enlightening.

In order to understand the reasons underling the ranking one could examine in detail the performance of each single indicator, however the analysis of the three top pillars of sustainability (the economic, the environmental and social one) is sufficiently informative to convey the main message (Table 12).

Table 8: Economic – Social and Environmental Regional Ranking 2010 – 2020

Economic Sustainability				
Regions	**2010**		**2020**	**Regions**
FIN	0.814	=	0.828	FIN
SWE	0.774	=	0.799	SWE
JPN	0.766	-3	0.772	DNK
DNK	0.754	1	0.754	CHE
CHE	0.746	1	0.725	CAN
CAN	0.683	1	0.724	JPN
USA	0.641	-1	0.710	GER
GER	0.638	1	0.674	USA
AUT	0.636	-1	0.664	FRA
BNLX	0.635	-1	0.659	AUT
FRA	0.613	2	0.652	BNLX
AUS	0.572	=	0.624	AUS
GBR	0.545	=	0.603	GBR
NZL	0.544	-1	0.578	RUS
RUS	0.525	1	0.560	NZL
NOR	0.503	-1	0.543	ITA
ITA	0.493	1	0.533	NOR
ESP	0.487	=	0.520	ESP
RoAsia	0.475	-1	0.511	CHN
CHN	0.462	1	0.483	RoAsia
RoEU	0.432	=	0.471	RoEU
PRT	0.415	=	0.461	PRT
MEast	0.405	=	0.426	MEast
SEA	0.386	-1	0.419	Baltic
Baltic	0.384	1	0.406	SEA
ZAF	0.341	-1	0.376	TUR
RoE	0.318	-2	0.372	ZAF
IND	0.311	=	0.356	IND
FSU	0.304	-2	0.353	RoE
MEX	0.290	-2	0.349	GCM
TUR	0.263	5	0.328	FSU
GCM	0.257	2	0.303	MEX
ARG	0.251	-5	0.303	POL
IDN	0.248	-1	0.298	BUL
BRA	0.244	-1	0.279	IDN
POL	0.240	3	0.263	BRA
BUL	0.223	3	0.208	RoLA
RoLA	0.191	1	0.188	ARG
NorthAfr	0.149	=	0.165	NorthAfr
RoAfrica	0.129	=	0.141	RoAfrica

Social Sustainability				
Regions	**2010**		**2020**	**Regions**
DNK	0.679	=	0.680	DNK
GER	0.665	=	0.677	GER
SWE	0.651	-	0.662	AUT
GBR	0.644	-	0.653	JPN
JPN	0.641	1	0.645	GBR
CHE	0.636	-	0.639	SWE
AUT	0.626	4	0.635	CHE
USA	0.599	=	0.614	USA
BNLX	0.591	-	0.605	CAN
POL	0.587	-	0.599	BNLX
ITA	0.586	-	0.585	NZL
CAN	0.574	3	0.583	ITA
FIN	0.567	-	0.583	POL
NZL	0.563	3	0.583	FRA
FRA	0.561	1	0.571	FIN
RoEU	0.528	-	0.551	ESP
ESP	0.525	1	0.529	RoEU
PRT	0.513	-	0.521	ARG
Baltic	0.506	-	0.517	PRT
ARG	0.491	2	0.498	ZAF
GCM	0.472	-	0.489	Baltic
AUS	0.456	-	0.485	MEX
RUS	0.456	-	0.482	GCM
ZAF	0.455	4	0.482	RUS
BUL	0.455	-	0.482	AUS
MEX	0.450	4	0.468	NOR
NOR	0.450	1	0.460	BUL
RoE	0.396	-	0.458	BRA
BRA	0.361	1	0.439	IDN
IDN	0.308	1	0.389	RoE
SEA	0.279	=	0.376	SEA
FSU	0.267	-	0.324	RoLA
RoLA	0.265	1	0.299	TUR
RoAsia	0.248	-	0.274	FSU
CHN	0.200	-	0.245	RoAsia
MEast	0.190	-	0.239	CHN
TUR	0.189	4	0.197	MEast
RoAfrica	0.178	=	0.195	RoAfrica
NorthAfr	0.098	=	0.117	NorthAfr
IND	0.009	=	0.037	IND

Environmental Sustainability				
Regions	**2010**		**2020**	**Regions**
NOR	0.759	=	0.795	NOR
SWE	0.726	=	0.752	SWE
ARG	0.669	=	0.686	ARG
FIN	0.648	=	0.681	FIN
BRA	0.646	-1	0.644	AUT
AUT	0.630	1	0.642	BRA
GBR	0.623	=	0.633	GBR
CHE	0.608	=	0.611	CHE
FRA	0.593	=	0.598	FRA
RoLA	0.558	=	0.590	RoLA
SEA	0.547	-1	0.587	RoE
CAN	0.531	-1	0.548	SEA
Baltic	0.526	-1	0.543	CAN
IDN	0.526	-10	0.538	Baltic
TUR	0.523	-3	0.525	GER
RoE	0.510	5	0.524	ITA
GER	0.506	2	0.516	JPN
ITA	0.499	2	0.508	TUR
DNK	0.495	=	0.500	DNK
JPN	0.495	3	0.490	NZL
MEX	0.493	-1	0.486	ESP
RoAfrica	0.482	-5	0.486	MEX
ESP	0.468	2	0.486	RUS
NZL	0.466	4	0.475	IDN
RUS	0.460	2	0.462	PRT
IND	0.457	-5	0.461	BNLX
PRT	0.443	2	0.458	RoAfrica
GCM	0.427	=	0.442	GCM
RoEU	0.423	=	0.435	RoEU
RoAsia	0.418	=	0.433	RoAsia
BNLX	0.409	5	0.383	IND
NorthAfr	0.363	=	0.376	NorthAfr
ZAF	0.355	=	0.362	ZAF
BUL	0.345	=	0.324	BUL
POL	0.313	=	0.318	POL
USA	0.291	-1	0.307	AUS
CHN	0.291	-2	0.289	USA
FSU	0.274	-2	0.238	MEast
AUS	0.254	3	0.234	CHN
MEast	0.235	2	0.221	FSU

It can be noted that countries performing well in terms of FSI not only have "higher scores", but also perform well in all the three sustainability dimensions. On the contrary, countries lower in the rank show a bad performance in at least one of the "sustainability pillars". Thus for instance top countries like Sweden, Finland and UK (GBR) maintain or improve their high environmental social and economic sustainability in the period 2010-2020. Conversely, countries like India, China and Russia, but also the USA, perform well economically, but badly environmentally or like Argentina well environmentally, but badly economically.

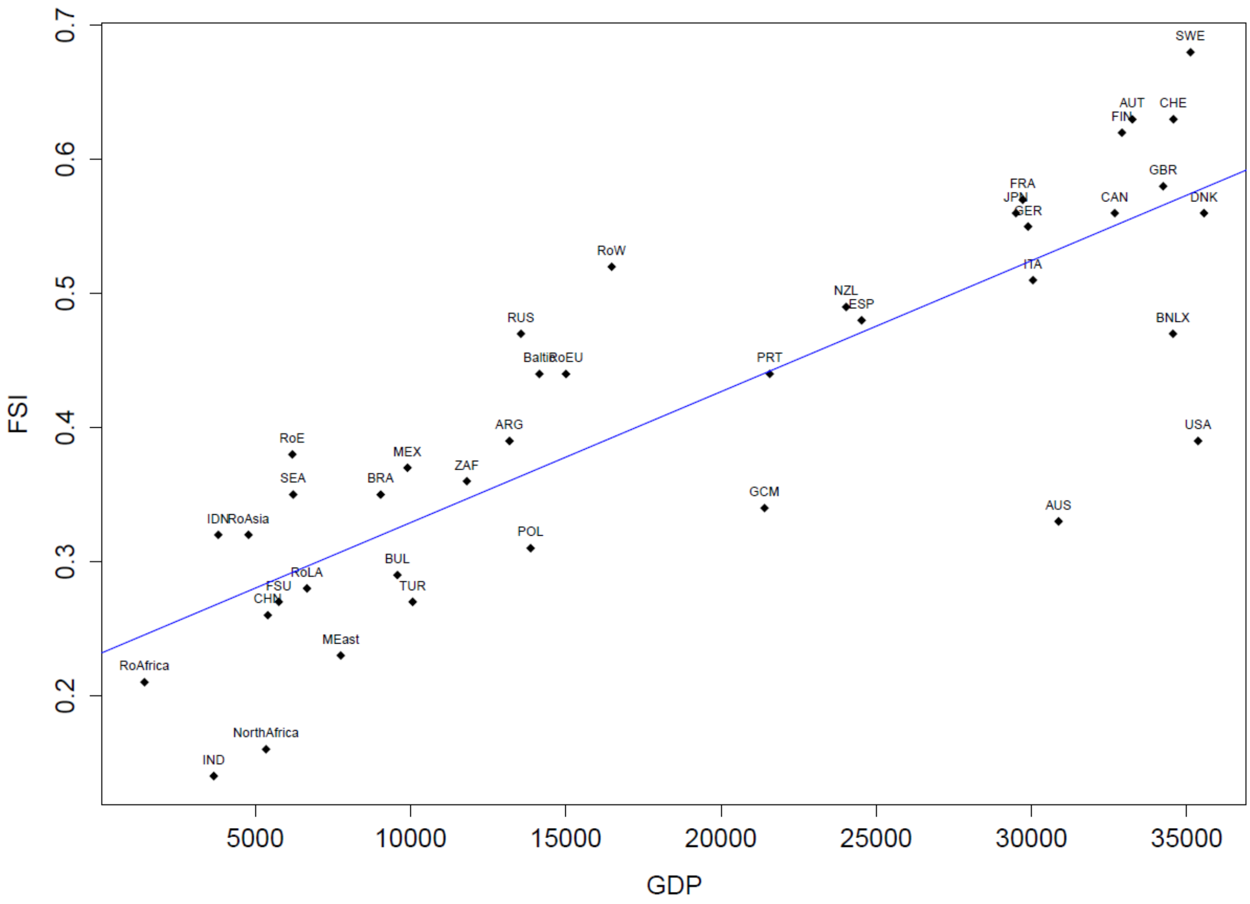
This expresses clearly the idea of sustainability implicitly endorsed by the experts and replicated by FSI: that of the complementarity of the different pillars (recalling section 5.1., in the DM an "AND" attitude prevails). Therefore a good economic, social or environmental performance alone is not sufficient to produce high scores in the FSI. This is biased downward by the worst result among the three dimensions. As a consequence, it also points out that there is not one "dominant" (or at least clearly dominant) sustainability component over the other able to summarize all the informative content of the FCI. This is particularly evident when the FSI ranking is compared with the one produced by GDP (Table 13).

Table 9: GDP per capita and FSI ranking

		FSI 2010		GDP pc 2010		
1	SWE	0.684	=	1.000	SWE	1
2	CHE	0.633	-2	1.000	DNK	2
3	AUT	0.629	-4	1.000	USA	3
4	FIN	0.620	-4	0.979	CHE	4
5	GBR	0.583	-1	0.978	BNLX	5
6	FRA	0.575	-7	0.962	GBR	6
7	DNK	0.561	5	0.912	AUT	7
8	CAN	0.560	-1	0.896	FIN	8
9	JPN	0.557	-5	0.884	CAN	9
10	GER	0.549	-2	0.794	AUS	10
11	NOR	0.520	-8	0.752	ITA	11
12	ITA	0.509	1	0.747	GER	12
13	NZL	0.492	-3	0.743	FRA	13
14	ESP	0.479	-1	0.738	JPN	14
15	BNLX	0.470	10	0.613	ESP	15
16	RUS	0.467	-7	0.601	NZL	16
17	Baltic	0.443	-4	0.539	PRT	17
18	RoEU	0.440	-2	0.535	GCM	18
19	PRT	0.438	2	0.425	NOR	19
20	ARG	0.394	-4	0.394	RoEU	20
21	USA	0.390	18	0.376	Baltic	21
22	RoE	0.376	-11	0.370	POL	22
23	MEX	0.371	-4	0.363	RUS	23
24	ZAF	0.362	-1	0.355	ARG	24
25	BRA	0.354	-4	0.326	ZAF	25
26	SEA	0.349	-6	0.289	TUR	26
27	GCM	0.340	9	0.285	MEX	27
28	AUS	0.332	18	0.278	BUL	28
29	IDN	0.318	-9	0.267	BRA	29
30	RoAsia	0.316	-7	0.232	MEast	30
31	POL	0.314	9	0.191	RoLA	31
32	BUL	0.294	4	0.175	SEA	32
33	RoLA	0.283	2	0.174	RoE	33
34	FSU	0.274	=	0.158	FSU	34
35	TUR	0.267	9	0.145	CHN	35
36	CHN	0.257	1	0.143	NorthAfr	36
37	MEast	0.231	7	0.122	RoAsia	37
38	RoAfrica	0.212	-2	0.086	IDN	38
39	NorthAfr	0.159	3	0.080	IND	39
40	IND	0.144	1	0.000	RoAfrica	40

The two classifications are very different with 57% of countries showing a displacement of at least four positions. The case of USA is striking. According to GDP per capita they rank first, but their relatively low environmental performance (high GHG emissions per capita and energy intensity) place them at the 21st place in FCI. Its relation with GDP is thus mild as visualized also by Figure 3.

Figure 3. FSI and GDP p.c. (US$ 2010) correlation in 2010



To investigate further the value added of the overall procedure, the section concludes comparing the FSI results with those produced by different "equal weighting" or additive aggregation methodologies. Among the many possible two are chosen (Table 14): "I_23", starting from the same indicator values, equally weights all indicators (each indicator thus weights 1/23 on the final value of the index); "I_3" equally weights the three pillars of sustainability (each pillar thus contributes 1/3 to the final index performance).

In Table 14, the top two regions (Sweden and Switzerland) maintain the first ranking in the three aggregation methodologies, however several differences are registered in the mid-lower level of the table which are substantively different. In total, 25% of countries changes their ranks of at least four positions across the different methodologies, while only 5 or 6 remain unchanged moving from FSI to I_3 and I_23 respectively.

Table 10: FSI comparison to FSI_23 and FSI_3 in 2010

		FSI		I_23	
1	SWE	0.684	=	0.739	SWE
2	CHE	0.632	=	0.698	CHE
3	AUT	0.628	-3	0.680	FIN
4	FIN	0.620	1	0.679	DNK
5	GBR	0.583	-3	0.661	NOR
6	FRA	0.574	-4	0.654	AUT
7	DNK	0.561	3	0.637	CAN
8	CAN	0.560	1	0.636	GBR
9	JPN	0.557	=	0.636	JPN
10	GER	0.549	-1	0.623	FRA
11	NOR	0.519	6	0.605	GER
12	ITA	0.508	-2	0.571	BNLX
13	NZL	0.491	=	0.569	NZL
14	ESP	0.479	-5	0.555	ITA
15	BNLX	0.470	3	0.541	ARG
16	RUS	0.466	-2	0.534	Baltic
17	Baltic	0.442	1	0.529	USA
18	RoEU	0.440	-3	0.528	RUS
19	PRT	0.438	-1	0.526	ESP
20	ARG	0.394	5	0.513	PRT
21	USA	0.390	4	0.509	RoEU
22	RoE	0.376	-7	0.476	BRA
23	MEX	0.371	-2	0.476	SEA
24	ZAF	0.362	-10	0.468	AUS
25	BRA	0.354	3	0.464	MEX
26	SEA	0.349	3	0.452	IDN
27	GCM	0.340	-4	0.443	RoAsia
28	AUS	0.332	4	0.441	TUR
29	IDN	0.318	3	0.439	RoE
30	RoAsia	0.316	3	0.429	RoLA
31	POL	0.314	-1	0.429	GCM
32	BUL	0.294	-1	0.420	POL
33	RoLA	0.283	3	0.418	BUL
34	FSU	0.274	-3	0.393	ZAF
35	TUR	0.267	7	0.388	CHN
36	CHN	0.257	1	0.385	MEast
37	MEast	0.231	1	0.368	FSU
38	RoAfrica	0.212	-1	0.355	NorthAfr
39	NorthAfr	0.159	1	0.348	RoAfrica
40	IND	0.144	=	0.346	IND

	FSI		I_3		
SWE	0.684	=	0.762	SWE	1
CHE	0.633	-1	0.721	FIN	2
AUT	0.629	-4	0.718	CHE	3
FIN	0.620	2	0.698	DNK	4
GBR	0.583	-6	0.685	JPN	5
FRA	0.575	-3	0.658	NOR	6
DNK	0.561	3	0.648	AUT	7
CAN	0.560	=	0.647	CAN	8
JPN	0.557	4	0.631	FRA	9
GER	0.549	=	0.629	GER	10
NOR	0.520	5	0.629	GBR	11
ITA	0.509	-2	0.589	NZL	12
NZL	0.492	1	0.568	BNLX	13
ESP	0.479	-3	0.563	ITA	14
BNLX	0.470	2	0.561	USA	15
RUS	0.467	=	0.552	RUS	16
Baltic	0.443	-1	0.540	ESP	17
RoEU	0.440	-2	0.535	Baltic	18
PRT	0.438	-2	0.525	ARG	19
ARG	0.394	1	0.503	RoEU	20
USA	0.390	6	0.496	PRT	21
RoE	0.376	-2	0.484	MEX	22
MEX	0.371	1	0.482	SEA	23
ZAF	0.362	-5	0.481	RoE	24
BRA	0.354	-3	0.465	RoAsia	25
SEA	0.349	3	0.461	AUS	26
GCM	0.340	-6	0.460	IDN	27
AUS	0.332	2	0.456	BRA	28
IDN	0.318	2	0.425	ZAF	29
RoAsia	0.316	5	0.419	CHN	30
POL	0.314	=	0.413	POL	31
BUL	0.294	-2	0.409	RoLA	32
RoLA	0.283	1	0.408	GCM	33
FSU	0.274	-3	0.402	BUL	34
TUR	0.267	-1	0.372	MEast	35
CHN	0.257	6	0.349	TUR	36
MEast	0.231	2	0.345	FSU	37
RoAfrica	0.212	=	0.308	RoAfrica	38
NorthAfr	0.159	-1	0.298	IND	39
IND	0.144	1	0.275	NorthAfr	40

7. Sensitivity analysis

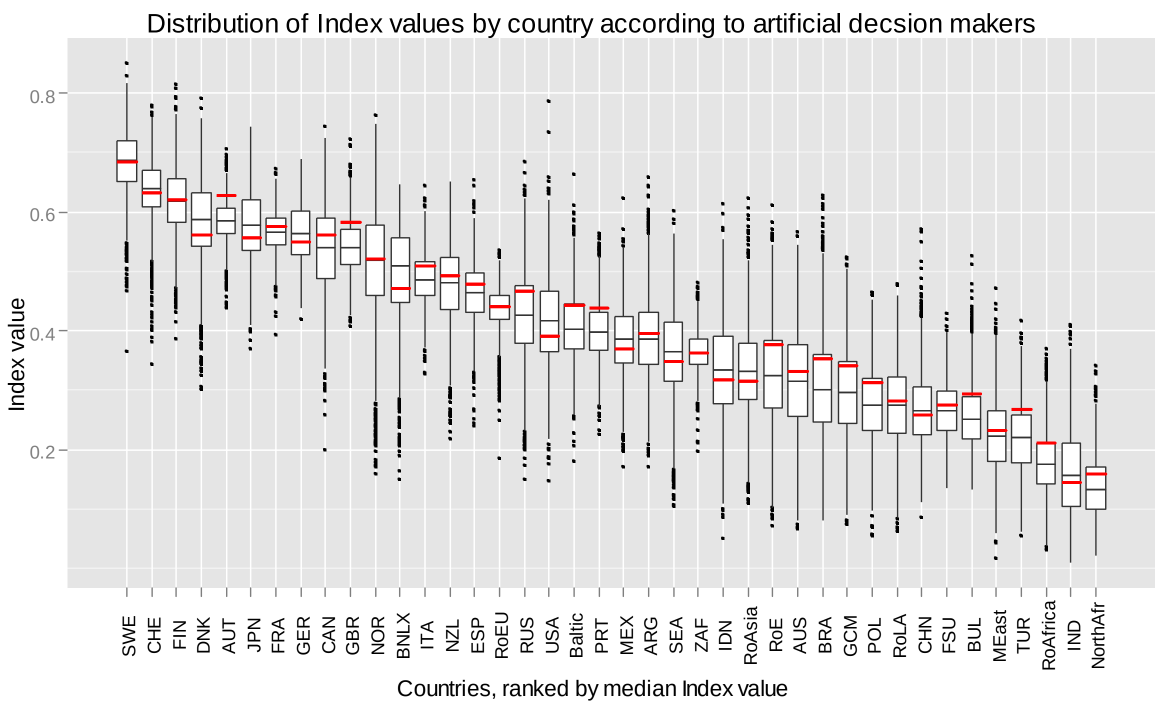
In a complex aggregation such as the one used for the FSI, the attitude of the decision maker is a key component of the process. It is thus important to analyse how robust/stable the ranking actually is to change in this attitude. This exercise is performed for the FSI in the baseline in 2010.

To introduce some variability in the determination of subjective weights, a set of 2000 artificial decision makers was generated. Each artificial decision maker provided a random, monotonic set of weights for every node of the tree. Since each set of decision maker weights produces a different FSI value, every country is associated to 2000 alternative FSI values depending on which set of weights was used. These sets were constructed varying within an interval of +/- 10% the "optimism" of the original set. This "optimism" is that expressed by the experts and reflect their attitude in considering the good performance of one indicator compensating the bad performance of another one. This attitude, (recalling section 5.1 the degree of preference for an AND or OR situation) can be readily measured by an appropriate index, that we call Or-ness varying between 0 and 1. Is the Or-ness that is allowed to vary +/- 10%.

The sensitivity analysis is then conducted with three different approaches: analyzing the distributions of FSI scores, analyzing the "agreement in ranking" and comparing the reference FSI with the median of the distribution of the generated FSI values.

The distribution of FSI scores is depicted in figure 4. The reference values of the index, for each country, are displayed as red lines in the graph.

Figure 3: Sensitivity analysis

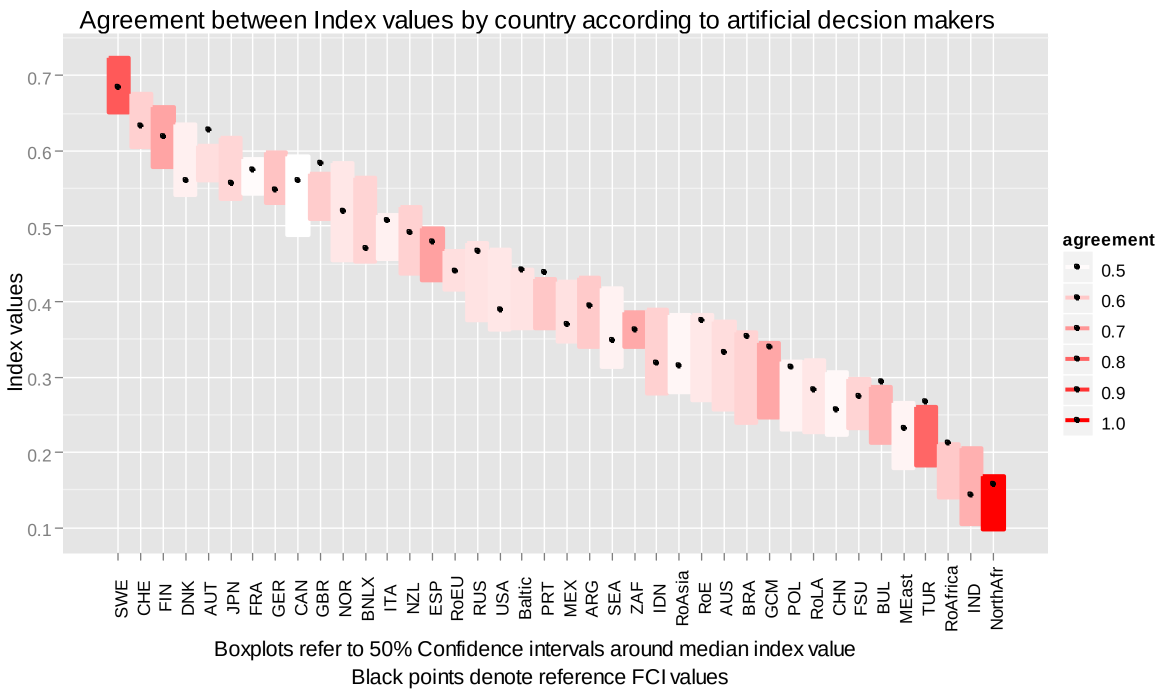


It is immediately appreciable that some countries, such as South Africa (ZAF), Indonesia (IDN), France or the Rest of Europe (RoEU) have a FSI index more robust to changes in decision maker weights than others, such as Norway or Australia (see also Appendix II for details). Nonetheless on a first approximation, it would seem that countries with neighboring positions in the ranking are quite likely to switch positions, since the distribution of FSI values often tend to overlap between consecutive countries.

This perception, however, ignores the fact that decision-maker weights tend to have a somewhat homogeneous effect across countries on the FSI score. It can be shown that, on average, the weights provided by a more "optimistic" decision maker (i.e., one giving set of weights with a higher-than-average or-ness score to all the nodes in the tree) tend to produce higher FSI score for most countries.

Taking this into account is possible to determine a measure of agreement of decision makers with respect to the median ranking. This amounts to comparing successive countries in the ranking pair wise and counting how many artificial decision makers, out of 2000, provided measures such that the former countries has a higher index than the latter. The measure of agreement among artificial decision makers can then be summarized in a single number for each country. This measure would take a value of 1 if there is perfect agreement among all 2000 artificial decision makers that a given country is more sustainable than the next one in the ranking; would take a value of 0.5 if only 50% of them agree etc… These measures are displayed in the figure 5 below, where the shade of their box plot indicates the extent of agreement among artificial decision makers.

Figure 4: FSI sensitivity



According to this measure of agreement, some positions, particularly those at the top and at the bottom of the ranking are more stable than others. When individual countries are concerned nearly 82% of artificial decision makers agree that Sweden has a higher FSI value than Switzerland and 79%

seem to prefer Turkey to Rest of Africa. These countries are likely to be ranked in the same way with respect to one another irrespective of the measures provided by decision makers. By contrast, only 52% of artificial decision makers provide measures such that Italy is preferred to New Zealand. In general an agreement lower than the 60% concerns the position of France, Canada, South East Asia (SEA), Rest of Asia (RoAsia), Poland, China and Middle East. Therefore this part of the ranking is more brittle. Table 15 finally compares the reference with the simulated FSI values

Table 11: Reference and simulated median FSI values by country for the year 2010.

		FSI 2010		Median simulated FSI 2010		
1	SWE	0.684	=	0.687	SWE	1
2	CHE	0.633	=	0.640	CHE	2
3	AUT	0.629	-2	0.619	FIN	3
4	FIN	0.620	1	0.588	DNK	4
5	GBR	0.583	-5	0.584	AUT	5
6	FRA	0.575	-1	0.577	JPN	6
7	DNK	0.561	3	0.566	FRA	7
8	CAN	0.560	-1	0.564	GER	8
9	JPN	0.557	3	0.541	CAN	9
10	GER	0.550	2	0.540	GBR	10
11	NOR	0.520	=	0.519	NOR	11
12	ITA	0.509	-1	0.509	BNLX	12
13	NZL	0.492	-1	0.485	ITA	13
14	ESP	0.480	-1	0.481	NZL	14
15	BNLX	0.470	3	0.464	ESP	15
16	RUS	0.467	-1	0.441	RoEU	16
17	Baltic	0.443	-2	0.427	RUS	17
18	RoEU	0.440	2	0.417	USA	18
19	PRT	0.438	-1	0.404	Baltic	19
20	ARG	0.395	-2	0.397	PRT	20
21	USA	0.390	3	0.386	MEX	21
22	RoE	0.376	-5	0.386	ARG	22
23	MEX	0.371	2	0.365	SEA	23
24	ZAF	0.362	=	0.363	ZAF	24
25	BRA	0.354	-4	0.334	IDN	25
26	SEA	0.349	3	0.331	RoAsia	26
27	GCM	0.340	-3	0.326	RoE	27
28	AUS	0.332	=	0.315	AUS	28
29	IDN	0.318	4	0.300	BRA	29
30	RoAsia	0.316	4	0.295	GCM	30
31	POL	0.314	=	0.276	POL	31
32	BUL	0.294	-3	0.275	RoLA	32
33	RoLA	0.283	1	0.265	CHN	33
34	FSU	0.274	=	0.265	FSU	34
35	TUR	0.267	-2	0.250	BUL	35
36	CHN	0.257	3	0.222	MEast	36
37	MEast	0.231	1	0.221	TUR	37
38	RoAfrica	0.212	=	0.175	RoAfrica	38
39	NorthAfr	0.159	-1	0.155	IND	39
40	IND	0.144	1	0.133	NorthAfr	40

It should be noted that there is generally close agreement between the reference FSI values and the central interval of the simulated FSI distribution. For most countries, the reference value of the FSI tends to be relatively close to the median value of the simulation.

It is therefore misleading to assume that overlapping distributions of FSI values necessarily imply high variability in ranking positions, although these do vary to some extent in practice.

8. Policy application

In order to show the further potential of FSI coupled with CGE analysis a policy scenario is introduced and compared to the baseline. The idea is to assess the policy impact on sustainability within and outside the EU highlighting potential tradeoffs between the different sustainability pillars and across countries. The effects of the European mitigation policy on sustainability in the EU are represented by Figure 6a and b.

Figure 6a. FSI: Baseline vs Policy Scenario in EU

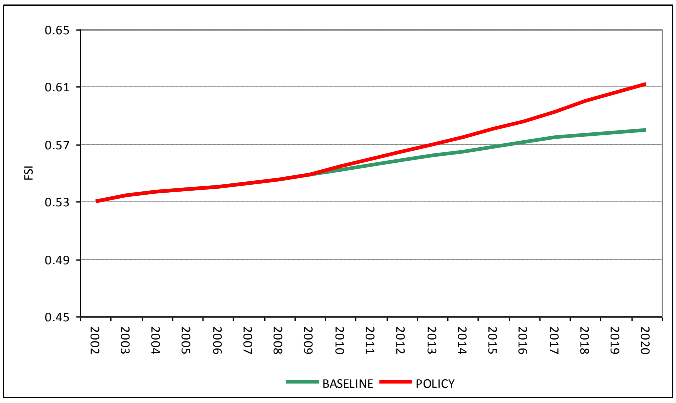
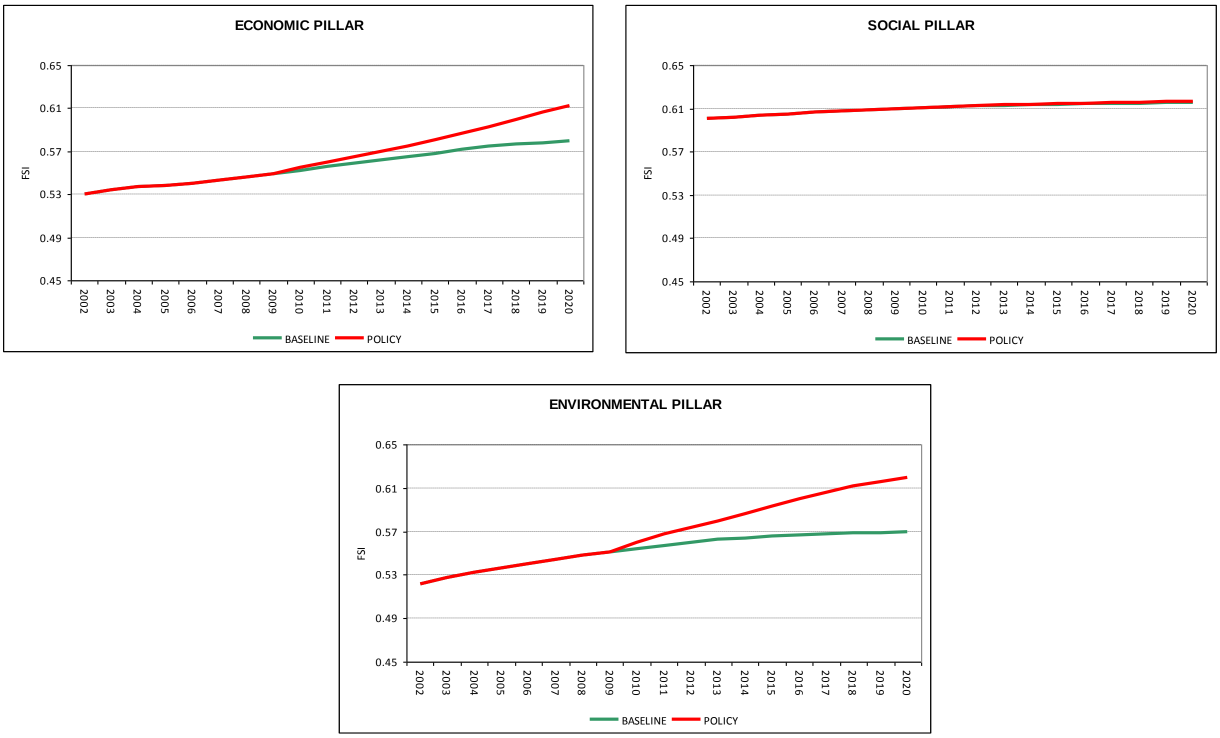


Figure 6b. Economic, social and environmental pillars: baseline vs policy scenario. EU



The EU mitigation policy improves the overall sustainability performance of the EU (+5.62% in 2020). This is driven by the amelioration of the environmental pillar (+8.8% in 2020) and of the economic pillar (+2.51% in 2020). The social pillar remains basically unaffected (+0.16% in 2020). Interestingly the GDP per capita performance of the EU worsens (-0.84% of the GDP per capita-related indicator in 2020). GDP declines because the mitigation policy imposes additional costs to the EU production systems in the form of higher prices of energy/carbon intensive goods. However this is more than compensated in the economic pillar by the improvement in the EU terms of trade (+12.26% in 2020). According to experts' weighting (see table 8) these are more important than GDP in determining the final economic performance. Terms of trade improvement depends on two factors: an increase in the EU export prices, and a decrease in energy intensive import prices. These decline as the mitigation policy fosters carbon/energy efficiency and thus lowers the EU demand of energy intensive foreign goods.

Therefore, within the EU context, the EU climate policy does not trigger conflicts between the three main dimensions of sustainability. Nonetheless, the EU policy impacts sustainability also outside its borders. In fact, in the Non-EU region total sustainability worsens slightly (-0.2% in 2020, Figure 7). This is the net effect of a worsening of the social and the environmental pillars (-0.31% and -0.26% in 2020 respectively) and of a slight improvement in the economic pillar (+0.07% in 2020) (see Figure 8). The social component is adversely affected as indicators generally decrease: education expenditure on GDP (-0.8% in 2020) and public health expenditure on GDP (-0.65% in 2020). This is caused by the increasing in competitiveness of the energy intensive sector which captures economic resources and affects expenditure in the welfare sector (health, pension and education). Environmental sustainability worsens because of the carbon leakage effect. Energy and carbon intensive goods of Non-EU countries, not burdened by a climate policy, become cheaper than the EU homologues, experience an increase in demand (notwithstanding the demand decline in the EU) with a subsequent increase in emissions. The carbon leakage also explains the improvement in Non-EU economic performance. In summary, the EU mitigation policy can partially conflict with sustainability, especially social and environmental, but outside the EU.

FIGURE 7 –Baseline vs. Policy Scenario In Non-EU Countries

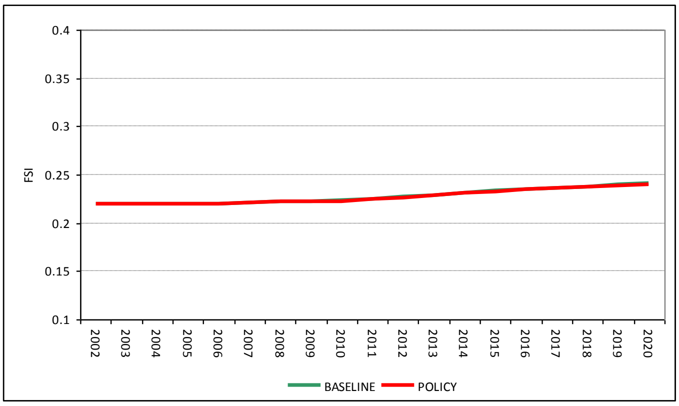
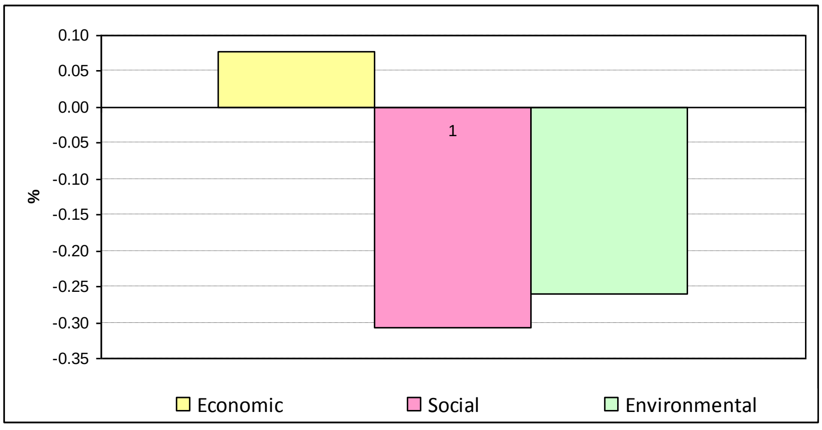


Figure 8: % change wrt baseline in Non-EU countries of the three pillars after climate policy in 2020



To conclude the section, Table 16 and 17 report the re-ranking induced by the policy considering the FSI and the three dimension of sustainability respectively.
When the FSI is concerned, position shifts are minimal. Basically the EU mitigation policy affects quite uniformly EU countries (positively) and Non-EU countries (negatively).
As can be expected, larger changes are observable, focusing on environmental sustainability. In this case East European Countries, particularly Hungary and the RoEU aggregate boost their environmental performance climbing respectively 16 and 22 places. This happens at the expenses of economic growth. Indeed Hungary looses three positions and RoEU one in the economic dimension. In both cases positive terms of trade effects cannot compensate the higher costs. As a net result, Hungary does not change its FSI pre-policy ranking, whereas that of the RoEU improves by three positions.

Table 12: FSI ranking 2020 baseline and policy

	Regions	2020 base		2020 policy	Regions	
1	SWE	0.685	=	0.689	SWE	1
2	AUT	0.649	=	0.665	AUT	2
3	CHE	0.636	-1	0.643	FIN	3
4	FIN	0.632	1	0.635	CHE	4
5	GBR	0.618	=	0.629	GBR	5
6	FRA	0.598	-1	0.604	GER	6
7	CAN	0.580	-2	0.602	FRA	7
8	GER	0.575	2	0.596	DNK	8
9	DNK	0.568	1	0.578	CAN	9
10	JPN	0.568	=	0.564	JPN	10
11	NOR	0.543	-1	0.547	ITA	11
12	ITA	0.536	1	0.543	NOR	12
13	NZL	0.515	-2	0.542	BNLX	13
14	BNLX	0.511	1	0.530	ESP	14
15	ESP	0.501	1	0.514	NZL	15
16	RUS	0.497	-1	0.512	RoEU	16
17	PRT	0.470	-1	0.495	RUS	17
18	Baltic	0.460	-1	0.493	PRT	18
19	RoEU	0.454	3	0.458	Baltic	19
20	SEA	0.415	-1	0.425	GCM	20
21	RoE	0.409	-1	0.414	SEA	21
22	GCM	0.397	2	0.401	RoE	22
23	USA	0.395	=	0.394	USA	23
24	BRA	0.388	=	0.388	BRA	24
25	MEX	0.385	=	0.385	MEX	25
26	ZAF	0.384	-2	0.383	POL	26
27	AUS	0.381	=	0.377	AUS	27
28	ARG	0.371	-1	0.373	ZAF	28
29	IDN	0.358	-1	0.364	ARG	29
30	TUR	0.353	-1	0.358	IDN	30
31	POL	0.350	5	0.351	TUR	31
32	BUL	0.330	=	0.336	BUL	32
33	RoAsia	0.318	=	0.317	RoAsia	33
34	RoLA	0.313	=	0.313	RoLA	34
35	CHN	0.276	=	0.277	CHN	35
36	FSU	0.245	=	0.246	FSU	36
37	MEast	0.239	=	0.239	MEast	37
38	RoAfrica	0.218	=	0.218	RoAfrica	38
39	NorthAfr	0.176	=	0.175	NorthAfr	39
40	IND	0.154	=	0.153	IND	40

Table 13: Three pillars baseline vs. policy in 2020

Economic Sustainability					Social Sustainability					Environmental Sustainability				
Regions	2020 baseline		2020 policy	Regions	Regions	2020 baseline		2020 policy	Regions	Regions	2020 baseline		2020 policy	Regions
FIN	0.828	=	0.841	FIN	DNK	0.680	=	0.680	DNK	NOR	0.795	=	0.793	NOR
SWE	0.799	=	0.802	SWE	GER	0.677	=	0.679	GER	SWE	0.752	=	0.772	SWE
DNK	0.772	=	0.769	DNK	AUT	0.662	=	0.663	AUT	ARG	0.686	-1	0.726	FIN
CHE	0.754	=	0.753	CHE	JPN	0.653	=	0.652	JPN	FIN	0.681	1	0.686	ARG
CAN	0.725	=	0.725	CAN	GBR	0.645	=	0.646	GBR	AUT	0.644	=	0.680	AUT
JPN	0.724	=	0.715	JPN	SWE	0.639	=	0.638	SWE	BRA	0.642	-2	0.677	GBR
GER	0.710	=	0.707	GER	CHE	0.635	=	0.632	CHE	GBR	0.633	1	0.653	RoEU
USA	0.674	-1	0.671	FRA	USA	0.614	=	0.613	USA	CHE	0.611	-2	0.641	BRA
FRA	0.664	1	0.669	USA	CAN	0.605	=	0.603	CAN	FRA	0.598	=	0.618	FRA
AUT	0.659	=	0.661	AUT	BNLX	0.599	=	0.601	BNLX	RoLA	0.590	-1	0.611	CHE
BNLX	0.652	=	0.661	BNLX	NZL	0.585	-2	0.590	POL	RoE	0.587	-1	0.588	RoLA
AUS	0.624	=	0.622	AUS	ITA	0.583	=	0.583	ITA	SEA	0.548	-2	0.571	RoE
GBR	0.603	=	0.605	GBR	POL	0.583	2	0.583	NZL	CAN	0.543	-2	0.566	GER
RUS	0.578	=	0.574	RUS	FRA	0.583	=	0.582	FRA	Baltic	0.538	-5	0.546	SEA
NZL	0.560	=	0.560	NZL	FIN	0.571	=	0.572	FIN	GER	0.525	2	0.541	CAN
ITA	0.543	=	0.549	ITA	ESP	0.551	=	0.551	ESP	ITA	0.524	-1	0.541	DNK
NOR	0.533	=	0.536	NOR	RoEU	0.529	=	0.535	RoEU	JPN	0.516	-5	0.539	ITA
ESP	0.520	=	0.528	ESP	ARG	0.521	=	0.520	ARG	TUR	0.508	-6	0.535	BUL
CHN	0.511	=	0.511	CHN	PRT	0.517	=	0.517	PRT	DNK	0.500	3	0.530	Baltic
RoAsia	0.483	=	0.483	RoAsia	ZAF	0.498	=	0.497	ZAF	NZL	0.490	-6	0.526	ESP
RoEU	0.471	-1	0.470	PRT	Baltic	0.489	=	0.488	Baltic	ESP	0.486	1	0.525	PRT
PRT	0.461	1	0.451	RoEU	MEX	0.485	=	0.484	MEX	MEX	0.486	-5	0.513	JPN
MEast	0.426	=	0.425	MEast	GCM	0.482	=	0.482	GCM	RUS	0.486	-5	0.506	GCM
Baltic	0.419	=	0.418	Baltic	RUS	0.482	-1	0.480	AUS	IDN	0.475	-5	0.505	TUR
SEA	0.406	=	0.407	SEA	AUS	0.482	1	0.480	RUS	PRT	0.461	4	0.504	BNLX
TUR	0.376	-1	0.373	ZAF	NOR	0.468	-1	0.469	BUL	BNLX	0.461	1	0.489	NZL
ZAF	0.372	1	0.371	TUR	BUL	0.460	1	0.468	NOR	RoAfri	0.458	-4	0.485	MEX
IND	0.356	-1	0.369	GCM	BRA	0.458	=	0.457	BRA	GCM	0.442	5	0.484	RUS
RoE	0.353	-1	0.352	IND	IDN	0.439	=	0.439	IDN	RoEU	0.435	22	0.474	IDN
GCM	0.349	2	0.345	RoE	RoE	0.389	=	0.387	RoE	RoAsia	0.433	-2	0.460	POL
FSU	0.328	=	0.328	FSU	SEA	0.376	=	0.376	SEA	IND	0.383	-2	0.459	RoAfri
MEX	0.303	=	0.303	MEX	RoLA	0.324	=	0.323	RoLA	North	0.376	-2	0.430	RoAsia
POL	0.303	=	0.285	POL	TUR	0.299	=	0.299	TUR	ZAF	0.362	-2	0.383	IND
BUL	0.298	-3	0.280	IDN	FSU	0.274	=	0.277	FSU	BUL	0.324	16	0.373	North
IDN	0.279	1	0.265	BRA	RoAsia	0.245	=	0.244	RoAsia	POL	0.318	5	0.346	ZAF
BRA	0.263	1	0.208	RoLA	CHN	0.239	=	0.238	CHN	AUS	0.307	=	0.303	AUS
RoLA	0.208	1	0.203	BUL	MEast	0.197	=	0.197	MEast	USA	0.289	=	0.289	USA
ARG	0.188	=	0.176	ARG	RoAfri	0.195	=	0.195	RoAfri	MEast	0.238	=	0.237	MEast
North	0.165	=	0.157	North	North	0.117	=	0.118	North	CHN	0.234	=	0.235	CHN
RoAfri	0.141	=	0.141	RoAfri	IND	0.037	=	0.037	IND	FSU	0.221	=	0.221	FSU

9. Conclusions

The present work aimed to highlight:

(a) if and how modeling tools in general and CGE models in particular could be used also to investigate sustainability in its wider dimensions; the focus on CGE is determined by their widespread use in policy analyses especially environmental policies.

(b) to test the potential of composite indicators in providing synthetic measures of sustainability and deliver additional information compared to those conveyed by the "simple" GDP.

These issues have been investigated applying the recursive-dynamic general equilibrium model ICES to a reference and an emission reduction scenarios for the mid term (2010-2020). 23 sustainable development indicators belonging to the three pillars of sustainability (economic, environmental and social) have been extracted from the model output and compounded into an innovative sustainability index: the FSI. The major novelties in the approach followed are the weighting and aggregation procedures in building the FSI which take into account complementarity or substitutability of performances among different indicators. Weights have been elicited interviewing experts in focus groups.

With this exercise we were able to show that, *subjected to given experts' opinions*:

There is a group of countries composed by more developed economic systems (but not by all of them), where economic, environmental and social sustainability move together. There is another group of countries, mostly, but not necessarily only, developing countries, where at least one dimension of sustainability diverges from the other. This seems to partly support an "environmental Kuznets curve" (EKC) idea: when a given level of economic development is reached, good economic, environmental and social performances could not be in opposition, but below that level the contrast can be stronger.

The first group of country is also that of the top-performers while the second performs low in terms of FSI. This expresses clearly the idea of sustainability replicated by the composite sustainability indicator: the different dimensions of sustainability are complement, therefore a bad performance in any of them greatly lowers the final score.

A direct consequence of this is that one "dominant" (or at least clearly dominant) sustainability component over the other cannot be identified. Thus none of them singularly taken is able to summarize all the informative content of the FSI. This applies particularly to GDP. Its country ranking is very different from that of the FSI. The case of USA is striking. They rank first as per capita GDP, but the FSI places them 21^{st} due to their relatively high GHG emissions per capita and energy intensity.

The EU pursuit of an improvement in environmental sustainability, represented by the implementation of a unilateral 20% emission reduction policy, apparently does not originate conflicts across the different sustainability dimensions within the EU. This is potential good news for policy makers. In fact all the three sustainability pillars improve. However, the EU policy can trigger potential conflicts with sustainability, especially environmental and social, in the non EU countries.

Both are induced by the well known phenomenon of carbon leakages which on the one hand foster the economic performance of carbon intensive sectors in non EU countries on the other worsen their environmental performance and drain resources from health and education investment.

The following general conclusions can then be drawn:

Sustainability can be analyzed with a consistent quantitative modeling framework and this methodology is particularly useful to get insights on, and measure the relations between its different components. Moreover, inter-temporal modeling exercises can provide important informative support to anticipate possible trends in sustainability and its components in given (business as usual or policy) scenarios. This can be appealing to a decision maker.

Against this background the use of CGE models presents two specific advantages: their large database makes it possible to calculate the indicators for several regions and sectors; their explicit modeling of market interactions and international trade is ideal to capture how potential tradeoffs in sustainability originates and propagates through the economic system. A CGE approach presents also specific limitations that should not be hidden. The majors are: the full equilibrium view of the economic system, the assumed instantaneous often costless adjustments to that equilibrium, the crucial dependence of results on the calibration process; the simplified dynamics, the difficulty to deal with non market values. This said, the use of a modelling approach is a useful enrichment to the standard analysis of sustainability, particularly important to capture quantitatively and explicit the relations between very different domains. As such it can be a powerful communication device. Due to the multifaceted nature of sustainability, it must not be considered the “only” or the “best” approach to analyze sustainability though, but an additional instrument in an ampler toolbox.

Also, we totally agree that it is neither possible to summarize sustainability in just one figure, nor that subjectivity can be ruled out; and this no matter how comprehensive, complex and innovative its generation process is. Nonetheless, we strongly support the use of composite indicators. As shown in this exercise, they can be invaluable communication devices to explicit the preference structure and value judgments originating a given synthetic sustainability assessment. They can also offer the opportunity to investigate in depth if and how this assessment can change when those preferences and values change. These information can be very interesting for policy decision makers and, on our view, are as if not more important than the synthesis provided.

References

Benassi C., Malagoli F., The sum of squared distances under a diameter constraint, in arbitrary dimension, *Arch. Math.,* 90, 471-480, 2008.

Bigano A., Bosello F., Roson R. and Tol, R.S.J. (2006), *Economy-Wide Estimates of the Implications of Climate Change: a Joint Analysis for Sea Level Rise and Tourism*, Fondazione Eni Enrico Mattei Working Paper N.135.2006.

Boehringer, C. and A. Loeschel (2004) *Measuring Sustainable Development: The Use of Computable General Equilibrium Models*, Center for European Economic Research (ZEW).

Bosetti, V., C. Carraro, M. Galeotti, E. Massetti and M. Tavoni (2006), "WITCH: A World Induced Technical Change Hybrid Model", *The Energy Journal*, Special Issue. Hybrid Modeling of Energy-Environment Policies: Reconciling Bottom-up and Top-down, 13-38.

Burniaux, J.-M. & Truong, T.P. (2002), *GTAP-E: An energy environmental version of the GTAP model*, GTAP Technical Paper n.16.

Carone, G., Denis, C., Mc MorROW, K., Mourre G., and W. Röger (2006) “Long-term labour productivity and GDP projections for the EU25 Member States : a production function framework” EC DG for Economic and Financial Affairs Economic Papers n° 253

Climate Analysis Indicators Tool (CAIT), version 3.0, Washington, DC: World Resources Institute, 2005. Available at http://cait.wri.org and http://earthtrends.wri.org.

Despic O., Simonovic S.P., Aggregation operators for decision making in water resources, *Fuzzy Sets and Systems*, 115, 11-33, 2000.

Dimaranan, B. V. (2006) *Global Trade, Assistance, and Production: The GTAP 6 Data Base*, Center for Global Trade Analysis, Purdue University.

Energy Information Administration (2008), *International Energy Outlook*.

Energy Information Administration (2009), *Annual Energy Outlook*.

European Communities Committee (ECC, 2002), *Investing in people: an imperative for Europe*, Bruxelles, 10.01.2003, Com (2002) 779 final

European Health strategy http://europa.eu/legislation_summaries/public_health/european_health_strategy/index_en. htm

EU SDS (2001) http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52001DC0264:EN:NOT

Food and Agriculture Organization of the United Nations (FAO), Aquastat Database: http://www.fao.org/nr/water/aquastat/main/index.stm.

Fujimoto K., Murofushi T., Some characterizations of the system represented by Choquet and multi-linear functional through the use of the Möbius inversion, *Int. Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, 5, 5, 547-562 1997.

Grabisch M., Fuzzy integral in multicriteria decision making. *Fuzzy Sets & Systems* 69 (1995), 279-298.

Grabisch M., The application of fuzzy integrals in multicriteria decision making. *European J. of Operational Research* 89 (1996), 445-456.

Grabish M., Lebreusche C., Vasnick J.C., On the extension of pseudo-boolean functions for the aggregation of interaction criteria, *European Journal of Operational Research*, 148, 1, 28-47, 2003.

Hertel, T.W. (1997), *Global Trade Analysis: Modeling and applications*, Cambridge University Press, Cambridge.

Horridge, M. (2005) *SplitCom: Programs to disaggregate a GTAP sector (Preliminary Draft),* centre of Policy Studies, Monash University, Melboune, Australia.

IMAGE (2001), *The IMAGE 2.2 Implementation of the SRES Scenarios*, RIVM CDROM Publication 481508018, Bilthoven, The Netherlands.

Intergovernmental Panel on Climate Change Agreements (IPCC), *Special Report on Emissions Scenarios*, SRES, 2000

International Energy Agency (2008), *World Energy Outlook*.

International Union for Conservation of Nature and Natural Resources (IUCN) Red List of Threatened Species, http://www.iucnredlist.org/.

Klaassen, G. and Miketa, A. (2003) *Defining Sustainable Development Objective and Common Policy Instruments: Methodologies for Integrating Impact Assessment in the Field of Sustainable Development*, International Institute for Applied Systems Analysis (IIASA).

Klement E.P., Mesiar R. and Pap E., 2000; *Triangular norms*, Kluwer Academic Publishers, Netherlands, 2000.

Lee, H. (2003) *The GTAP non-CO2 Emissions Data Base*, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University.

Mackay (2003) *Reducing Global Poverty Best Practice: Education and Distance Education*, Professor Lindsay, Deakin University

McKibbin, Warwick J. and Peter J. Wilcoxen (1998), The Theoretical and Empirical Structure of the G-Cubed Model, *Economic Modelling* , 16(1), pp. 123-148.

Murofushi T., Sugeno M., Machida M., Non-monotonic fuzzy measures and Choquet integral, *Fuzzy Sets and Systems*, 64, 73-86, 1994.

OECD Main Science and Technology Indicators, http://www.oecd.org/document/26/0,3343,en_2649_34451_1901082_1_1_1_1,00.html

OECD (2007) Health at a Glance 2007: Indicators http://lysander.sourceoecd.org/vl=3896448/cl=28/nw=1/rpsv/health2007/index.htm

OECD data, Education and Training http://www.oecd.org/topicstatsportal/0,2647,en_2825_495609_1_1_1_1_1,00.html

Pearson M.A. and Martin J.P (2005), *Should We Extend the Role of Private Social Expenditure*?, IZA Discussion Paper No. 1544 http://www.thepresidency.gov.za/learning/reference/factbook/11-01-04-g01.htm

Roson, R., (2003), *Modelling the Economic Impact of Climate Change*, EEE Programme Working Papers Series, International Centre for Theoretical Physics “Abdus Salam”, Trieste, Italy.

United Nations (2009), World Population Prospects: the 2008 Revision Population Database, http://esa.un.org/unpp/index.asp?panel=2

Von Altrock Constantin, 1995; *Fuzzy logic and neurofuzzy applications explained*, Prentice Hall PTR, Upper Saddle River, NJ, 1995.

World Bank (2008), World Development Indicators, New Delhi, OUP, http://web.worldbank.org/

World Health Organization (WHO), Global strategy for health for all by the year 2000. Geneva, WHO,

Yager R. R., 1993; Families of OWA operators, *Fuzzy Sets and Systems*, 59, 1 25-148, 1993.

Cagetti M. (2003), *Wealth Accumulation over the Life Cycle and Precautionary Savings*, in Journal of Business & Economic Statistics, Vol. 21.

Annex II: Non-Additive Measures for the Aggregation of the FEEM Sustainability Index

1. Non-Additive Measures and Aggregation Operators

This document briefly describes the methodological aspects necessary to implement the computation of the FEEM Sustainability Index (the root of the Decision Tree). The methodological framework belongs to the context of Multi Attribute Value Theory approach (MAVT), using an innovative and general approach. In the real world application, the most used aggregation functions is the Weighted Averaging approach (WA), which simply computes the weighted average of the numerical score of each criterion. Even if it is a very simple approach, nevertheless no interaction among the criteria can be considered, since it is a compensative method and requires the satisfaction of the Preferential Independent axiom, implicitly assumed but poorly verified in human decision processes. Trying to bypass such limitations, many other methods were proposed (the Geometric Averaging, the OWA operators (Ordered Weighted Averaging) (Yager, 1993), the compensation operator (Von Altrock, 1995). Here we do not address the theoretical framework required by these operators, see Klement (2000). But we remark the usefulness to deal with general operators which can be easily parameterized and tuned by the Decision Maker, and, at the same time, which need to be mathematically justified. This requirement avoids to fall down into traps which can strongly damage the results of the final computation. Nowadays, it is widely recognized that the non additive measures approaches (NAM, for brevity) satisfy these theoretical requirement, and at the same time are sufficiently general to cover a lot of preference structures. Moreover, the required parameters can be easily obtained by means of a simple questionnaire. The price to pay consists into an exponential increasing of the numerical complexity with the number of parameters. Nevertheless, in the current applications this is not a strong disadvantage, since the number of nodes is limited enough for each node of the Sustainability tree. A method based on a NAM is nothing else that the extension of the WA approach, but, despite than the assignment of a weight to every (single) element of the attribute set, it assigns a weight to every possible subset of the criteria which refer to the considered node in the tree. Than the algorithm simply aggregates every coalition computing the weighted averaging of all the subset. In so doing, the importance of a subset can be greater, equal, or less than the sum of the importance (the weights) of each criteria included in the subset. Synergic and redundant interactions among the criteria can be explicitly considered. Moreover, some indices can be computed, measuring the Decision Maker psychological behaviour. In fact, the Decision

Maker (DM, in what follows) can move to the pessimism (compensative) to the optimism, passing through the neutral behaviour as in the AW approach. An algebraic engine computes the score of each nodes, and the algorithm repeats the same computation for each node at each level of the tree, moving backward to the root, and finally computing the Sustainability Index. Let us now to briefly introduce the concept of NAM.

Definition 1. Let $N = \{1,2,3,....,n\}$ be the set of attribute for a given node in the tree. A non additive (monotonic) measure (NAM in what follows) is a set function $m: S \subseteq N \rightarrow [0,1]$, so that:

$$\forall S,T \subseteq N: \ m(\emptyset) = 0, \ \forall S,T \subseteq N: \ S \subseteq T \Rightarrow m(S) \leq m(T), \ \ m(N) = 1$$

(1.1)

It is remarkable that the set function assigns a weight to every subset of the criteria and not only to a single criterion only, as in the case of the WA algorithm. Besides the natural border conditions (the first and the last ones), the second constraint implies the monotonicity property, a quite intuitive constraint, even if, in rare cases, non monotonic measures could be applied too, (De Waegenaere, 2001 and Murofushi, 1994). Usually a NAM is additive if $m(S \cup T) = m(S) + m(T), \ S \cap T = \emptyset$, and this case corresponds to the WA operator. Conversely, if the "joint effect" is greater than the sum of the two effects considered separately, that is $m(S \cup T) < m(S) + m(T), \ S \cap T = \emptyset$, the measure is called sub-additive, and represents a redundancy effect. While if the contrary holds, $m(S \cup T) > m(S) + m(T), \ S \cap T = \emptyset$ it is super-additive, representing a synergic effect. If if n is the number of the criteria, despite the WA approach, which needs only n parameters, a NAM requires the specification of $2^n - 2$ parameters, i.e. the number of all the subsets of the n criteria, minus 2 (the border conditions force the value of the measure for the empty and for the universal sets).

As soon as that the NAM values are assigned, the (normalized) values of the criteria can be aggregated using a suitable extension of the WA approach, namely the so called Choquet integral or other methods derived by it, as the multi-linear approach (Grabisch, 1995 and 1996). Varying the values of the measure, the Choquet integral aggregator generalizes the WA, obtaining as sub-cases the Ordered Weighted Averaging (OWA, Yager, 1993), the Min and the Max operators, the k-order statistics, their combination, and other ones.

For practical applications, a first problem consists into the assignment of the measures, directly by data, if they are available, or by the preference structure of one or more DM(s). The latter is often used in social and economic sciences, given that the computation of an aggregated index, as in our case study, strongly depends on the subjective relative importance of a coalition with respect to another one. First of all, let us recall a theoretical result. To every set of NAMs, an alternative representation exists, based on the Möbius transform (Grabisch, 2003). This transformation assigns to every subset a value directly connected to the measures. If this value is null, no interaction exists among the element of the subset, as in the WA case, if it is positive there is a synergy, if negative, a redundancy. The Choquet integral can be directly calculated using the Möbius values, see 2). Moreover, using these values some possible extensions of the Choquet integral can be easily obtained, as the multi-linear algorithm, a smoother modification of the Choquet integral.

To obtain the values of the measure, a suitable questionnaire was developed by Despic (2000), where all the (0/1) combinations of criteria was considered (for a given node in the tree), meaning that “0” corresponds to the WORST case, and “1” to the BEST one. The questionnaire is a list of some possible scenarios, i.e. all the combinations of BEST and WORST values. If n is the number of sub-criteria for the considered node, 2^n is the number of possible questions. As soon as the questionnaire is fulfilled, the NAM is defined, and an algorithm computes the value of the multi-linear operator.

2. Aggregation by the Multi-Linear Operator

Given the values of the criteria, in the first step they are normalized in a common scale using a transform function which filters the sampled data, as usually done in MAVT methods, , used as benchmark for practical uses. The most common shape of such transform functions is piecewise linear, but also bell, quadratic, polynomial or spline can be used. We do not consider here the problem of the determination of the analytical form of such functions, usually obtained from statistical considerations, or, more correctly in the multi-criteria case, from the expertise. Let $(x_1, \ldots, x_n)$ be the values of the normalized criteria, obtained from the benchmark filtering. The next step includes the ordering phase, then the vector $(x_1, \ldots, x_n)$ is transformed in the vector $(x_{(1)}, \ldots, x_{(n)})$ in such a way that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.

Let us consider the computation of the Choquet integral defined as follows:

$$C_m(x_1, \ldots, x_n) = \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \cdot m(A_{(i)})$$

(2.1)

where $m(A_{(i)})$ is the measure of the set $A_{(i)} = \{i, i+1, i+2, \ldots, n\}$, and $A_{(n+1)} = \emptyset$, $x_{(0)} = 0$.

For computational simplicity, let us consider a boolean string, so that a biunivocal correspondence is defined between each row in the Table. In particular, the string with all elements equals to 1, the string $(1,1,....,1)$, corresponds to the last row. Starting from this string, corresponding to $m(A_{(1)})$, the first element of $C_m(x_1,..., x_n)$ is computed multiplying the difference $(x_{(1)} - x_{(0)}) = x_{(1)}$ by $m(A_{(1)})$. Now the element corresponding to $x_{(1)}$, is set equal to 0 in the string that now refers to $m(A_{(2)})$. So the second element of $C_m(x_1,..., x_n)$ is computed in the same way, multiplying the difference $(x_{(2)} - x_{(1)})$ by $m(A_{(2)})$. Now the element corresponding to $x_{(2)}$ is set equal to 0, and so on.

As written above, the Choquet integral generalizes the AW approach, permitting, to compute many possible aggregation operators, varying the value of the corresponding NAM, including some logical combinations of the criteria values. It is remarkable that this cannot be done in AW approach, the most widely aggregation operator applied in sustainable indicator computation.

The Choquet integral is mathematically characterized by a set of properties and requirements that need to be satisfied by the preference structure of the DM, we limit to quote:

a) the Preferential Independence for comonotonic acts

b) idempotency: $C_m(x,..., x) = x$

c) monotonicity: $C_m(x_1,..., x_n) \geq C_m(y_1,..., y_n)$ if $x_1 \geq y_1,..., x_n \geq y_n$

d) border conditions: $C_m(0,...,) = 0$, $C_m(1,...1) = 1$

e) homogeneity extension property: $C_m(k1_A, 0_{A_C}) = kC_m(A),\ \forall k > 0, \forall A \subset N$

being $1_A = \begin{cases} 1, & i \in A \\ 0 & i \notin A \end{cases}$

Note that the homogeneity extension property implies the impossibility to represent preference structure characterized by rules like: "If all the criteria are MEDIUM, then the satisfaction degree is COMPLETE". In fact, if it were possible, from the homogeneity extension property it will follows that if all the criteria are HIGH the satisfaction should be more than COMPLETE, violating the second border condition.

3. The Möbius Transform

Any NAM $m(T), T \subseteq N$ can be alternatively represented by means of the Möbius transform as it follows:

$$m(T) = \sum_{S \subseteq T} a(S)$$

(3.1)

Where the coefficients $a(T)$ can be both positive, negative or null. If positive, it means that there is synergic interaction between the criteria belonging to the coalition T, if negative, there is redundancy interaction (somehow called also conflicting). If null, no interaction exist, and the measure becomes the sum of the singleton measures (the "weights") of each criterion belonging to T. The inverse transformation is given by:

$$a(S) = \sum_{T \subseteq S} (-1)^{|N|-|T|} m(T) \quad , \quad |N = card(S)| \tag{3.2}$$

Using the Möbius coefficients the Choquet integral is computable as:

$$C_m(x_1, ..., x_n) = \sum_{T \subseteq N} a(T) \cdot \min_{i \in T}(x_i) \tag{3.3}$$

Not any sequence of 2^n real numbers can be the Möbius transform of a NAM, see Fujimoto (1997) for the required conditions.

4. The Multi-Linear Aggregation Operator

In the Möbius space, the computation of the Choquet integral is given by formula (3.3). In this formula appears the computation of the minimum between all the criteria values belonging to a subset T of the criteria. The minimum operator belongs to an important class of aggregation operators, the T-norm based operator.

Definition

A T-norm is binary operator $TN(x, y)$ satisfying idempotency, monotonicity and border conditions: $TN(1,1) = 1$, $TN(0,1) = TN(x,0) = 0$.

The minimum operator is a T-norm, but also the product is a T-norm and many other ones, as the Lukasiewicz T-norm, and many other ones (Klement, 2000).

Using the Möbius transform, the Choquet integral is obtained applying the T-norm minimum, see (3.3), while the multi-linear operator is obtained applying the T-norm product, that is, substituting "min" with the product in formula (3.3), i.e.:

$$ML_m(x_1,..., x_n) = \sum_{T \subseteq N} a(T) \cdot \prod_{i \in T}(x_i) \qquad (3.4)$$

We used this formulation (3.4) since the product operator is smoother with respect to the min, even if such multi-linear operator suffers of theoretical properties with respect to the Choquet integral.

5. Implementation

For practical implementation, it can be useful to consider a pointer vector, i.e. a Boolean vector, like the one formed by the columns 2,3,4,5,6 of the following Table:

	C(1)	C(2)	C(3)	C(4)	C(5)	Val
1	0	0	0	0	0	m(0)
2	1	0	0	0	0	m(1)
3	0	1	0	0	0	…
4	0	0	1	0	0	…
5	0	0	0	1	0	…
6	0	0	0	0	1	m(5)
7	1	1	0	0	0	m(1,2)
8	1	0	1	0	0	m(1,3)
9	1	0	0	1	0	…

10	1	0	0	0	1	…
11	0	1	1	0	0	…
12	0	1	0	1	0	…
13	0	1	0	0	1	…
14	0	0	1	1	0	…
15	0	0	1	0	1	…
16	0	0	0	1	1	m(4,5)
17	1	1	1	0	0	m(1,2,3)
18	1	1	0	1	0	m(1,2,4)
19	1	1	0	0	1	…
20	1	0	1	1	0	…
21	1	0	1	0	1	…
22	1	0	0	1	1	…
23	0	1	1	1	0	…
24	0	1	1	0	1	…
25	0	1	0	1	1	…
26	0	0	1	1	1	m(3,4,5)
27	1	1	1	1	0	m(1,2,3,4)
28	1	1	1	0	1	m(1,2,3,5)
29	1	1	0	1	1	…
30	1	0	1	1	1	…
31	0	1	1	1	1	m(2,3,4,5)
32	1	1	1	1	1	m(1,2,3,4,5)

In this table, any 5-ple formed by the (0-1) elements in the columns 2,3,4,5,6, is the Boolean code of the progressive number in the first column. Each rows contains, in the last column, the value of the NAM corresponding to the coalition identified by the "1" elements in the 5 columns 2,3,4,5,6. Thus, for instance, the row number 28 corresponds to the measure of the coalition (1,2,3,5) since "28" is the coding of the Boolean vector (1,1,1,1,0). It corresponds to the case where the first 4 criteria are perfectly satisfied, and the last one completely unsatisfied. More in general, each row of the Table corresponds to a (fictious) scenario for which the "1" values mean "criterion totally satisfied", and the contrary for the "0" elements. They are nothing than the questions posed to each DM.

The following example will clarify the usefulness of such a table.

Let us consider 5 criteria, $n = 5$. Let $(x_1,..., x_5) = (0.3, 0.8, 0.1, 0.6, 0.2)$. We have $(x_{(1)},..., x_{(5)}) = (0.1, 0.2, 0.3, 0.6, 0.8)$, since $x_3 \le x_5 \le x_1 \le x_4 \le x_2$, and the corresponding string (or pointer vector) is $(3, 5, 1, 4, 2)$.

Given the initial value of the string $(1,1,1,1,1)$, we start by $x_{(3)} \cdot m(1,2,3,4,5) = 0.1 \cdot m(1,2,3,4,5) = 0.1$ since $m(1,2,3,4,5) = m(N) = 1$ for the border condition. Now the first element in the pointer vector is given by the third element in the string, set equal to 0, obtaining $(1,1,0,1,1)$, which points to the row n. 29, and the next element of $C_m(x_1,..., x_5)$ is computed as $(x_{(2)} - x_{(1)}) \cdot m(1,2,4,5) = (0.2 - 0.1) \cdot m(1,2,4,5) = 0.1 \cdot m(1,2,4,5)$. Next the last element is set equal to 0, and the string becomes $(1,1,0,1,0)$, and so on. So the steps are:

1. $(1,1,1,1,1)$: $x_{(3)} \cdot m(1,2,3,4,5) = 0.1 \cdot m(1,2,3,4,5) = 0.1$

2. $(1,1,0,1,1)$: $(x_{(2)} - x_{(1)}) \cdot m(1,2,4,5) = (0.2 - 0.1) \cdot m(1,2,4,5) = 0.1 \cdot m(1,2,4,5)$

3. $(1,1,0,1,0)$: $(x_{(3)} - x_{(2)}) \cdot m(1,2,4) = (0.3 - 0.2) \cdot m(1,2,4) = 0.1 \cdot m(1,2,4)$

4. $(0,1,0,1,0)$: $(x_{(4)} - x_{(3)}) \cdot m(2,4) = (0.6 - 0.3) \cdot m(2,4) = 0.3 \cdot m(2,4)$

5. $(0,1,0,0,0)$: $(x_{(5)} - x_{(4)}) \cdot m(2) = (0.8 - 0.6) \cdot m(2) = 0.2 \cdot m(2)$

To obtain , it is sufficient to sum up all the 5 terms.

6. ORNESS and ANDNESS Degrees

As for OWA operators, any non-additive measure is characterized by two extreme case, that are the minimum and the maximum aggregation operators. In the first case, the algorithm simply computes the minimum of the criteria values, and for this reason it is named as pessimistic, i.e. totally non compensative, given that the DM evaluation is guided only by the worst case, i.e. the logical AND operator (maximum prudence). In fact, using the minimum only the lowest values is considered, independently from the other ones. For instance, the two following scenarios are equally judged: (0.3,0.2,0.2) and (1,1,0.2)) with score equal to 0.2. It can easily be checked that the minimum operator can be implemented by a NAM with null value for all the coalition, m(A)=0, except for the universal set (the border condition): m(N)=1.

The opposite case is the maximum operator. In this case, the algorithm is totally compensative, corresponding to an optimistic DM behaviour, i.e. the logical OR operator, given that only the best case in considered. This happens when all the measure values equal 1, m(A)=1, for every subset A of N, except for the empty set (the other border condition): m(ø)=0.

The WA operator is in between these two extreme cases, when the measure of any coalition is the sum of the measures of the sub-sets measures, for any possible partition. Then it is clear that as soon as the measure values are close to (0,1,1,…,1), i.e. the maximum operator, the DM behaviour tends to me pessimistic, and the contrary holds in the opposite case.

To characterize the DM behaviour with respect to pessimism or to optimism, it is possible to compute an index, depending solely on the measure values, the ORNESS index, together with the ANDNESS index. The first one measures the tendency to optimism, while the second one, its complement to 1, measures the tendency to pessimism. Using the Möbius values of the measure, the ORNESS is computed as follows[8]:

$$\mathrm{Orness}_m(i) = \frac{1}{n-1}\sum_{T\subseteq N}\frac{n-t}{t+1}a_T$$

If Orness=1 the DM is optimistic, implicitly using the maximum operator (logical disjunction), if Orness=0, the DM is pessimistic, corresponding to the minimum operator (logical conjunction), if Orness=0.5 the DM is additive, and no interaction exist among the criteria.

[8] The ORNESS index can be computed also using the measure values, but the computation is more complicated, and it is not here reported.

The following Table furnishes the explicit formula for N=2,3,4,5:

N	Orness
2	$\frac{1}{2}(a_1 + a_2)$
3	$\frac{1}{2}(a_1 + a_2 + a_3) + \frac{1}{6}(a_{1,2} + a_{1,3} + a_{2,3})$
4	$\frac{1}{2}(a_1 + a_2 + a_3 + a_4) + \frac{2}{9}(a_{1,2} + a_{1,3} + a_{1,4} + a_{2,3} + a_{2,4} + a_3$ $+ \frac{1}{12}(a_{1,2} + a_{1,3} + a_{1,4} + a_{2,3} + a_{2,4} + a_{3,4})$
5	$\frac{1}{2}\sum_{i} a_i + \frac{1}{4}\sum_{i,j} a_{ij} + \frac{1}{8}\sum_{i,j,k} a_{ijk} + \frac{1}{20}\sum_{i,j,k,l} a_{ijkl}$

The ORNESS index is quite useful to characterize the DM behaviour, mainly in presence of more than one DMs, enhancing clusters or sub-groups of DM, if exist, with different way of reasoning, and even as a measure the consensus among them.

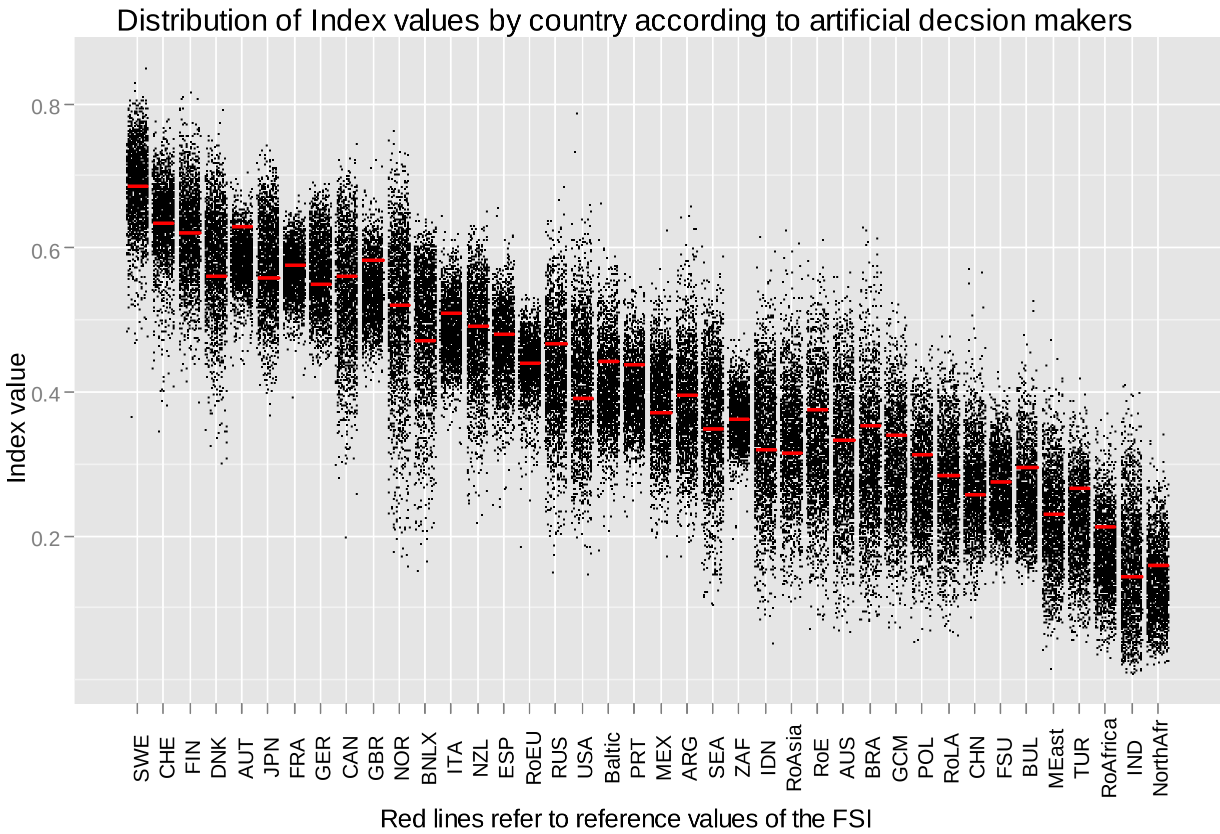
Robustness analysis

Given the reliance of the aggregation process on subjective measures provided by decision makers, it is important to understand how much specific decision makers' evaluations can influence the final value of the FSI for any given country and, therefore, the ranking. To this purpose, a valid approach to quantify the variability of index values involves changing the measures provided by the decision makers and aggregate the indicators accordingly, for the reference year 2010. This was performed by programming an R script and the kappalab package, to generate 2000 "artificial decision makers", each of providing a set of measures for every node of the FSI tree. Below is the code used to generate the measures for a node of `n` criteria (see Grabisch, Kojadinovic and Meyer, 2009).

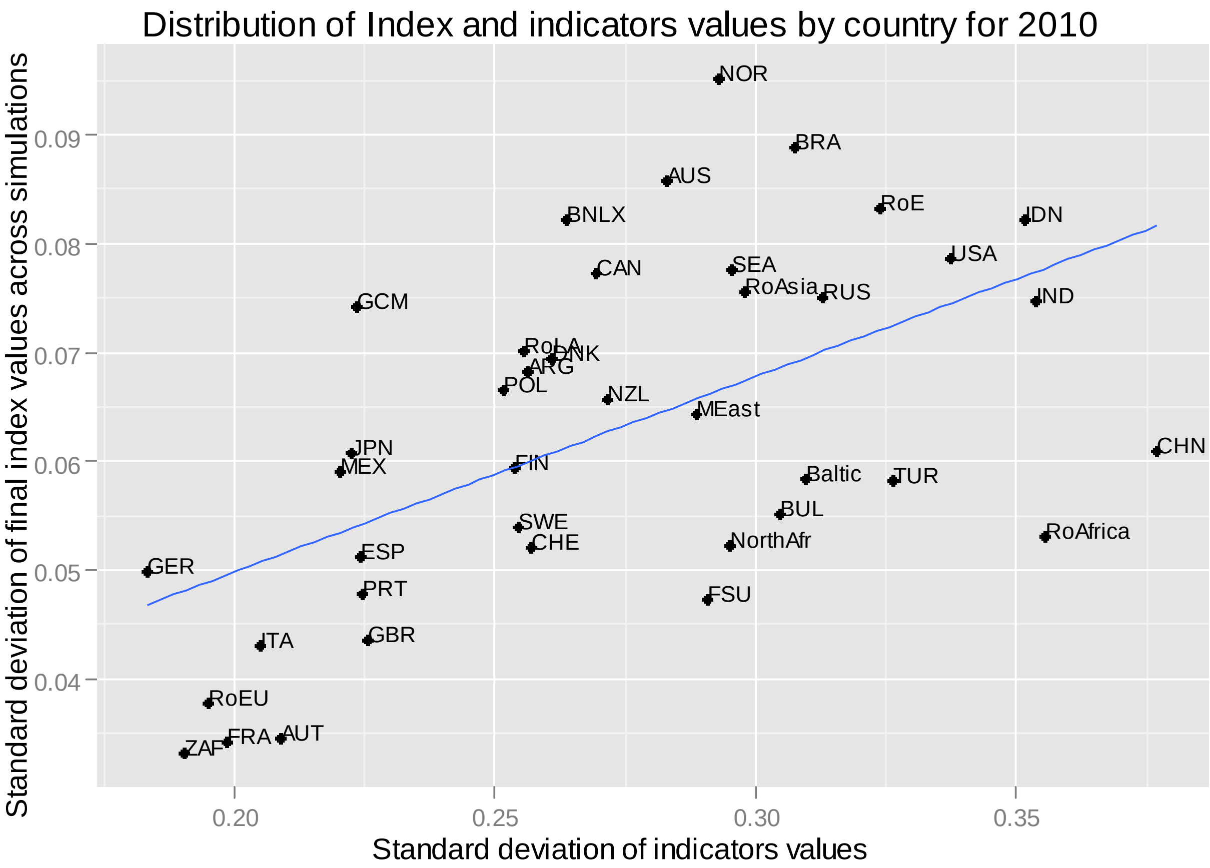
```
require(kappalab)

x <- runif(2^n-1)
for (i in 2:(2^n-1))
 x[i] <- x[i] + x[i-1]
mu <- normalize(capacity(c(0,x)))
```

In order to preserve the meaning of the aggregation process in the context of sustainability indicators, the non-additive measures were required to be monotonic and to feature an orness score within an interval of +/-10% with respect to the orness score of the measures used for the original aggregation. This amounts to ensuring that every artificial decision maker provides unbiased but meaningful set of measures, derived from a uniform random distribution. For every decision maker, the indicators were aggregated using the Choquet integral, as for the original values of the FSI. This allowed to obtain a distribution of 2000 FSI values for each country (pictured below), and a ranking based on the median value of these indices.



Among the factors that are likely to influence the variability of the simulated FSI index, the variance of the underlying indicators used to compute has a statistically significant correlation.



It can be checked empirically (correlation 0.5662, p-value=0.0001397) that countries with larger standard deviation of indicators, such as Brazil or Norway, tend to produce a more dispersed range of simulated indicators.

The distribution of final indices, however, provides only partial information concerning the vulnerability of the ranking to changes in the decision maker values. To understand how sensitive the ranking is to changes in decision maker measures, a pair wise comparison between successive countries in the ranking by median was performed. For very pair of countries, a measure of agreement among decision makers was derived. This is defined as the sum of all the artificial decision makers that provided measures such that the country ranking higher has a higher index than the subsequent one, divided by 2000. The values are reported in the table below.

	country	Median	Reference	St. dev.	St. dev.	agreement
1	SWE	0.687	0.684	0.054	0.255	0.822
2	CHE	0.640	0.633	0.052	0.257	0.585
3	FIN	0.619	0.620	0.059	0.254	0.672
4	DNK	0.588	0.561	0.069	0.261	0.5215
5	AUT	0.584	0.629	0.035	0.209	0.558
6	JPN	0.577	0.557	0.061	0.222	0.574
7	FRA	0.566	0.575	0.034	0.199	0.4985
8	GER	0.564	0.549	0.050	0.183	0.6115
9	CAN	0.541	0.560	0.077	0.269	0.4925
10	GBR	0.540	0.583	0.044	0.226	0.5935
11	NOR	0.519	0.520	0.095	0.293	0.539
12	BNLX	0.508	0.470	0.082	0.264	0.578
13	ITA	0.485	0.509	0.043	0.205	0.52
14	NZL	0.481	0.492	0.066	0.271	0.583
15	ESP	0.464	0.479	0.051	0.224	0.678
16	RoEU	0.441	0.440	0.038	0.195	0.556
17	RUS	0.427	0.467	0.075	0.313	0.548
18	USA	0.417	0.390	0.079	0.338	0.539
19	Baltic	0.404	0.443	0.058	0.309	0.547
20	PRT	0.397	0.438	0.048	0.225	0.6025
21	MEX	0.386	0.371	0.059	0.220	0.5515
22	ARG	0.386	0.394	0.068	0.256	0.5965
23	SEA	0.365	0.349	0.078	0.295	0.517
24	ZAF	0.363	0.362	0.033	0.191	0.6625
25	IDN	0.334	0.318	0.082	0.352	0.5835
26	RoAsia	0.331	0.316	0.076	0.298	0.5085
27	RoE	0.326	0.376	0.083	0.324	0.538
28	AUS	0.315	0.332	0.086	0.283	0.56
29	BRA	0.300	0.354	0.089	0.308	0.579
30	GCM	0.295	0.340	0.074	0.224	0.666
31	POL	0.275	0.314	0.066	0.252	0.5145
32	RoLA	0.275	0.283	0.070	0.255	0.541
33	CHN	0.265	0.257	0.061	0.377	0.506
34	FSU	0.265	0.274	0.047	0.291	0.577
35	BUL	0.250	0.294	0.055	0.305	0.6465
36	MEast	0.222	0.231	0.064	0.289	0.5125
37	TUR	0.221	0.267	0.058	0.326	0.797
38	RoAfrica	0.175	0.212	0.053	0.356	0.5985
39	IND	0.155	0.144	0.075	0.354	0.648
40	NorthAfr	0.133	0.159	0.052	0.295	1

The agreement measure is defined in such a way that, if there is unanimous agreement among decision makers for a given subsequent pair of countries in the ranking, the agreement measure takes value 1. Since NorthAfrica is the last country in the ranking, it has no successive country to compare to and takes agreement 1 by default.

This measure, together with the variability of simulated FSI values, allows to determine how robust the ranking is to changes in decision maker weights. It should be noted that the seemingly arbitrary requirement to artificial decision makers to supply weights with an orness score within an interval of +/- 10% actually provides a generous amount of freedom. In fact, it allows to compute a wide range of alternative sets of weights and thereby constitutes a thorough assessment of the robustness of the FSI values.

Determination of Shapley values for criteria

In order to determine the relative importance of criteria in the FCI decision tree, the Shapley values of every indicator or node in the tree has been computed. Grabisch (1996) defines the Shapley value of a criterion, in the context of multicriteria decision-making as follows:

let $X = \{x_1, \ldots, x_n\}$ *be the set of criteria* (in the case of the FCI tree, and indicators or a sub-node belonging to the same node*) and let* $\mu(A)$ *be the weight of importance of a set of criteria* A*. The Shapley value of criterion* x_i*, with respect to* μ *is defined by:*

$$\Lambda(x_i) = \sum_{A \subset X \setminus x_i} \frac{(|X| - |A| - 1)! \cdot |A|!}{|X|} \left[\mu(A \cup \{x_i\}) - \mu(A) \right]$$

Where $|A|$ *indicates the cardinality of A and 0! = 1. The Shapley values of all the indicators in a given node always sum to one:* $\sum_{i=1}^{n} \Lambda(x_i) = 1.$

In the context of the FCI tree, the Shapley values are derived from the sets of evaluations provided by the decision makers and denote how much they value the inclusion of criterion x_i to every coalition of criteria A not including it, weighted by a function of the sets' cardinalities.

References

Benassi C., Malagoli F., The sum of squared distances under a diameter constraint, in arbitrary dimension, Arch. Math., 90, 471-480, 2008.

Despic O., Simonovic S.P., Aggregation operators for decision making in water resources, Fuzzy Sets and Systems, 115, 11-33, 2000.

Fujimoto K., Murofushi T., Some characterizations of the system represented by Choquet and multi-linear functional through the use of the Möbius inversion, Int. Journal of Uncertainty, Fuzziness and Knowledge-based Systems, 5, 5, 547-562 1997.

Grabisch M., Fuzzy integral in multicriteria decision making. Fuzzy Sets & Systems 69 (1995), 279-298.

Grabisch, Michel. 1996. The application of fuzzy integrals in multicriteria decision making. *European journal of operational research* 89, no. 3: 445–456. http://linkinghub.elsevier.com/retrieve/pii/037722179500176X.

Grabisch M., The application of fuzzy integrals in multicriteria decision making. European J. of Operational Research 89 (1996), 445-456.

Grabisch, Michel, Ivan Kojadinovic, and Patrick Meyer. 2009. kappalab: Non-additive measure and integral manipulation functions.

Grabish M., Lebreusche C., Vasnick J.C., On the extension of pseudo-boolean functions for the aggregation of interaction criteria, European Journal of Operational Research, 148, 1, 28-47, 2003.

Klement E.P., Mesiar R. and Pap E., 2000; Triangular norms, Kluwer Academic Publishers, Netherlands, 2000.

Murofushi T., Sugeno M., Machida M., Non-monotonic fuzzy measures and Choquet integral, Fuzzy Sets and Systems, 64, 73-86, 1994.

R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. Vienna, Austria. Vienna, Austria. http://www.r-project.org.

Von Altrock Constantin, 1995; Fuzzy logic and neurofuzzy applications explained, Prentice Hall PTR, Upper Saddle River, NJ, 1995.

Yager R. R., 1993; Families of OWA operators, Fuzzy Sets and Systems, 59, 1 25-148, 1993.