Constructing the FEEM Sustainability Index: a Choquet-integral application

Cruciani, C.°*, Giove, S.°*, Pinar, M.° and Sostero, M.°*

°Fondazione Eni Enrico Mattei, *Ca' Foscari University of Venice

5th Atlantic Workshop on Energy and Environmental Economics

A Toxa, Galizia, Spain. 25-26 June 2012

Outline

- **Overview of sustainability indices**
- **Conceptualizing sustainability**
  - **Sustainability indicators and FEEM SI tree**
- **Multi-attribute aggregation**
  - Choquet-integral aggregation
- **Expert elicitation**
  - Features of the questionnaire
  - Decision maker behaviour
  - Representative Decision maker
- **FEEM Sustainability Index**
  - FEEM SI ranking (Economic, Social and Environmental)
  - Robustness analysis
  - Dominance of sustainability and rankings
- **Concluding remarks and future research**

Overview of sustainability

- **Overview of sustainability**
- Brundtland report : "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED, 1987)
- Need to find ways to measure sustainability translated into a plethora of approaches and sustainability indicators that have been aggregated in different ways to obtain composite indices.
- Methodological review on the sustainability, see Bossell (1999), OECD JRC (2008), and Singh et al. (2009); and for list of sustainability indicators refer to the EU core set of indicators (EEA, 2005), and the UN Commission on Sustainable Development (2005).

- **Why a composite index?**
- A composite index allows for a quick assessment of sustainability performance across different countries and at different times. Moreover, sustainability indices conveys a straightforward message to stakeholders and policy makers, and also are able to highlight best practices and weaknesses of sustainability strategies (Ness et al. 2007).
- **What is out there as a sustainability index?**
- Singh *et al*. (2009) summarizes 41 sustainability indicators used in the literature.

1) Majority of those indices are either aggregated through equal weight assignment (e.g., Environmental Sustainability Index, Human Development Index, Sustainability Performance Index, etc.) or
2) weights given by experts (e.g., Index of Environmental Friendliness) to each sustainability indicator.

- None of those indices allow to capture the interrelations among different sustainability indicators. No synergies or redundancies across indicators are considered.

Objectives of the paper

What can be done further to obtain a sustainability index?

- Sustainability is characterized by many different aspect that are somewhat linked one to the other, which rules out the possibility of using simple aggregation techniques.

Objectives:

- 1) To construct a composite sustainability index by applying a *non-linear* aggregation methodology (i.e., the Choquet integral) which accounts for the interactions among sustainability indicators.
- **Whose sustainability? By whom the interactions and redundancies across sustainability indicators are determined?**
- 2) Expert elicitation and the derivation of weights for each sustainability indicator and their coalitions among each other.
- The aggregated index strongly depends on the *subjective relative importance of coalitions between different sustainability indicators, which may be different for* each expert.
- 3) We derive a consensus measures on sustainability indicators from many expert elicitations by using a metric distance (i.e., if the evaluation of an expert is in agreement with other experts, then this expert's valuation gets higher weight.
- **4)** Robustness analysis

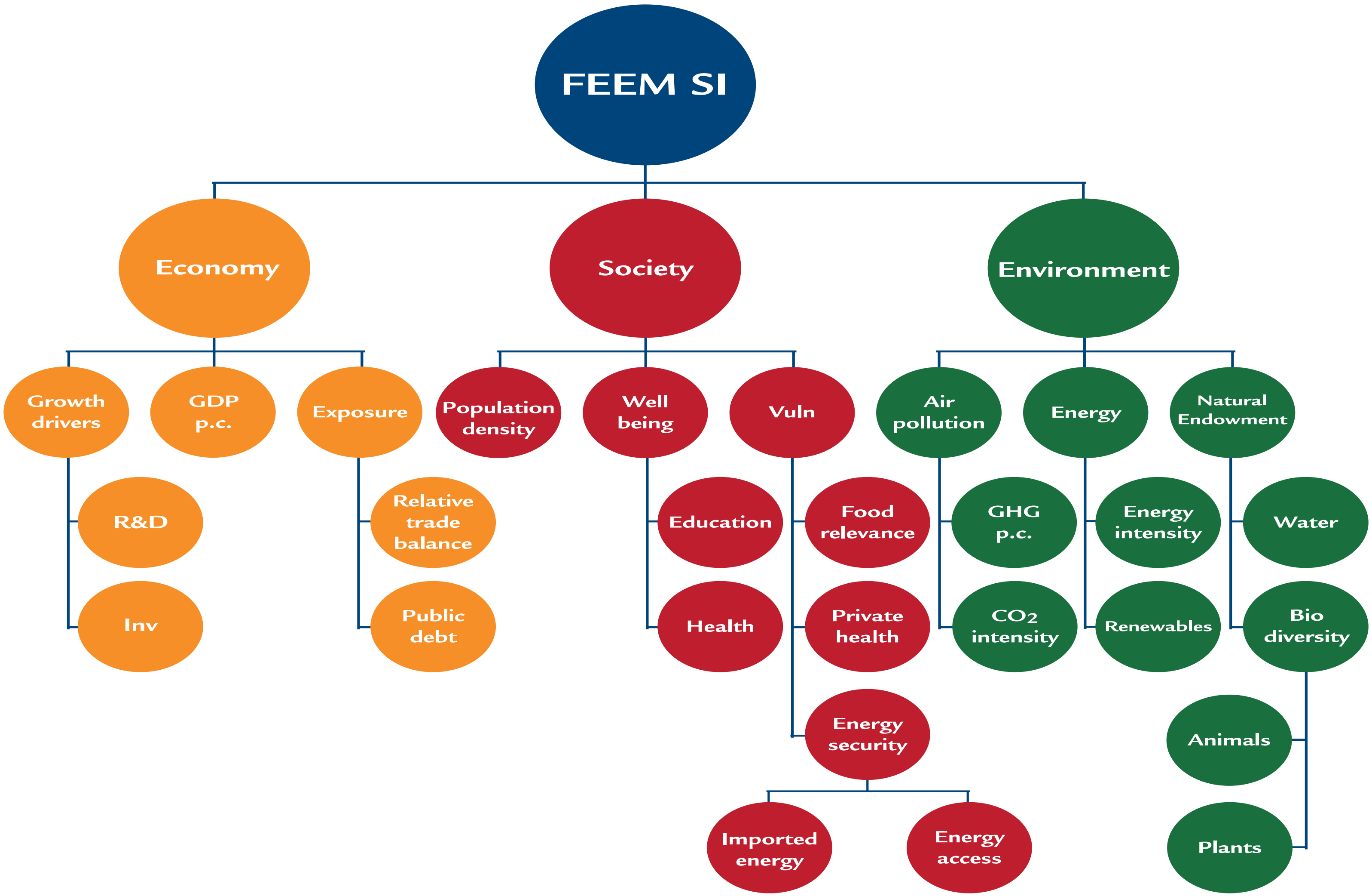
The FEEM Sustainability Index

- 19 indicators belonging to the three pillars of sustainability (i.e., economic, social and environmental).
- The indicators are constructed within a recursive-dynamic general equilibrium model ICES-SI (Carraro et al., 2012) which allows producing future projections of all indicators in the time frame 2011-2020 that can be used in comparative static policy analysis.
- Carraro, C., Campagnolo, L., Eboli, F., Lanzi, E. Parrado, R., Portale, E. (2012). Beyond GDP: A New Tool for Sustainability Assessment. Fondazione Eni Enrico Mattei, *mimeo.*
- The indicators are then normalized using a policy-oriented benchmarking technique developed ad hoc for the FEEM SI before proceeding to the aggregation stage.

  See FEEM Sustainability Index Methodological Report 2011, Section 3 for the normalization procedure and detailed indicator benchmarks

Conceptualizing sustainability

The FEEM Sustainability Index Aggregation Tree



Sustainability indicators

Sustainability indicators

Economic pillar (5 variables):

GDP per capita (+), Relative trade balance (+), Public debt (as percentage of GDP) (-), Investment (Investment as percentage of GDP) (+), R&D Expenditure (as percentage of GDP) (+)

Social pillar (7 variables):

Population density (-), Public education expenditure (+), Total health expenditure (+), Food consumption (as a share of all household consumption (-), Private health expenditure (as a share of all health consumption) (-), Energy imported (-), Energy access (electrification) (+)

Environmental pillar: (7 variables):

Co2 intensity (Co2 emission per energy consumed) (-), GHG emission per capita (-), Energy intensity (gross inland consumption of energy/GDP) (-), Renewable energy (share of renewable energy over energy consumption) (+), Plant biodiversity (the percentage of endangered species /total species) (-), Animal biodiversity (-), Water pressure (percentage of the total freshwater abstracted annually compared to the total available renewable resources) (-)

Multi-attribute aggregation

Definition of NAM and the Choquet integral.

Let $N = \{1,2,3,....,n\}$ be the set of attribute for a given node in the tree.

A *non additive (monotonic) measure* is a set function $m : S \subseteq N \to [0,1]$

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

The two first constraints are two border condition, while the second represents a monotonicity constraint, a rational property.

A NAM is ***additive*** *if* $m(S \cup T) = m(S) + m(T),\ S \cap T = \emptyset$

$m(S \cup T) < m(S) + m(T),\ S \cap T = \emptyset$ the measure is called *sub-additive*, implying a ***redundancy*** effect

$m(S \cup T) > m(S) + m(T),\ S \cap T = \emptyset$ *super-additive* (a ***synergic*** effect).

Choquet integral

Let $(x_1,\ldots,x_n)$ be the values of the (normalized) criteria, obtained from the benchmark filtering.

Let $(x_{(1)},\ldots,x_{(n)})$ the ordered vector of the vector $(x_1,\ldots,x_n)$ obtained by a suitable permutation of indices, so that $0 = x_{(0)} \le x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$

The Choquet integral of the vector $(x_1,\ldots,x_n)$ with $x_i \in [0,1]$ with respect to the (non additive) measure $m : S \subseteq N \to [0,1]$ is given by:

$$C_m(x_1,\ldots,x_n) = \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \cdot m(A_{(i)}) \quad A_{(i)} = \{i, i+1, i+2, \ldots, n\} \quad A_{(n+1)} = \varnothing \quad x_{(0)} = 0$$

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

$$C_m(x_1, x_2, \ldots, x_n) = \sum_{T \subseteq N} \alpha_m(T) \cdot \min_{i \in T} \{x_i\}$$

$$\alpha_m(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|}\, m(T), \quad \forall S \subseteq N$$

For example n=3, then we have 8 capacities. Then the Mobuis for each measure can be calculated
a(1)=m(1)-m(0), …, a(12)=m(12)-m1-m2,…,a(23)=m(23)-m(2)-m(3)
Finally, a(8)=m(8)-a(1)-a(2)-a(3)-a(4)-a(5)-a(6)-a(7)

Shapley value (Importance factor)

- The Shapley value characterizes the "relative importance" of each criterion and can be derived directly by the NAM values.
- The Shapley value can be computed for each criterion at every node of the hierarchy tree.
- It is obtained by averaging all the marginal gains obtained by adding the criterion to every coalition not including itself (Grabisch, 1995 and 1996).
- For the *i-th* criterion, the Shapley value is calculated as follows:

$$v(i) = \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\,t!}{n!} \left[m(T \cup i) - m(T) \right] \quad \text{where} \quad t = card(T)$$

$$\sum_{i=1}^{n} v(i) = 1$$

Shapley values vary between 0 and 1, higher value representing higher importance of that criterion.

ORNESS and ANDNESS degrees

- It is clear that the measure values, $m(A_{(i)})$,are close to (0,1,1,…,1), i.e. the maximum operator, the DM behaviour tends to be optimistic.

- Whereas the measure values, $m(A_{(i)})$, are close to (0,1,1,…,1), i.e. the maximum operator, the DM behaviour tends to be pessimistic.

Using the Möbius values of the measure, the ORNESS is computed as follows:

$$ORNESS_m(i) = \frac{1}{n-1}\sum_{T \subseteq N}\frac{n-t}{t+1}a(T)$$

If ORNESS=1, then the DM is fully optimistic, implicitly using the maximum operator (logical disjunction).

If ORNESS=0 (i.e. ANDNESS=1 since ORNESS+ANDNESS=1), then the DM is extremely pessimistic, corresponding to the minimum operator (logical conjunction)

If ORNESS=0.5 the DM is additive on average

Interaction indices

The intuition behind the interaction index is very similar to Shapley index but considering two indicators' contribution together rather than only one indicator's.

Let's consider two indicators, i and j

If the m(i,j)>m(i)+m(j), then it shows a ***complementary effect*** between i and j

Similarly, m(i,j) < m(i)+m(j) suggests that i and j interact in a ***redundant*** (substitutive) way.

if m(i j) = m(i) + m(j), it can be considered that the indicators i and j do not interact, i.e., that they have ***independent*** roles.

The average interaction between two indicators i and j is calculated with the following interaction index (see Murofushi and Soneda, 1993):

$$I_m(ij) = \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\,t!}{(n-1)!} \left[m(T \cup ij) - m(T \cup i) - m(T \cup j) + m(T) \right]$$

$$I_m(ij) \in [-1,1]$$

The interaction index being 1 (respectively –1) represents to full complementarity (resp. substitutivity) between i and j (see Grabisch, 1997).

Expert elicitation

Questionnaire and evaluation of sustainability indicators

- The aggregation methodology prepared for the FEEM SI builds on the capacities of the multi attribute aggregation methodology.
- Indicators can take in theory a wide range of values, but in this procedure only two extreme qualitative values were identified and labelled “best” and “worst”.
- All the indicators of the FEEM SI are ***quantitative*** in nature, yet the aggregation methodology has been constructed using “***qualitative***” evaluations in order to make the procedure more comprehensible.
- Experts in the questionnaire have been asked to make reference to their own ideal “best” and “worst” for each indicator. Moreover, avoiding numerical evaluations of best and worst levels rules out bias deriving from respondents disagreeing with the judgement given.

Expert elicitation

Questionnaire and evaluation of sustainability indicators

Economic	Social	Environmental	Weights
Worst	Worst	Worst	0
Best	Worst	Worst	20
Worst	Best	Worst	50
Worst	Worst	Best	30
Best	Best	Worst	$X \geq 50$
Best	Worst	Best	$X \geq 30$
Worst	Best	Best	$X \geq 50$
Best	Best	Best	100

Example questionnaire

Sample questionnaire of the FEEM SI final node

qualtrics.com

Overall sustainability

You will have to rate the **sustainability** of some hypothetical scenarios.

In these scenarios, the following indicators can have a **GOOD or BAD performance**:

- **ECON**: Economic sustainability
- **SOC**: Social sustainability
- **ENV**: Environmental sustainability

Please **rate on a 0-100 scale** all of the following hypothetical situations.

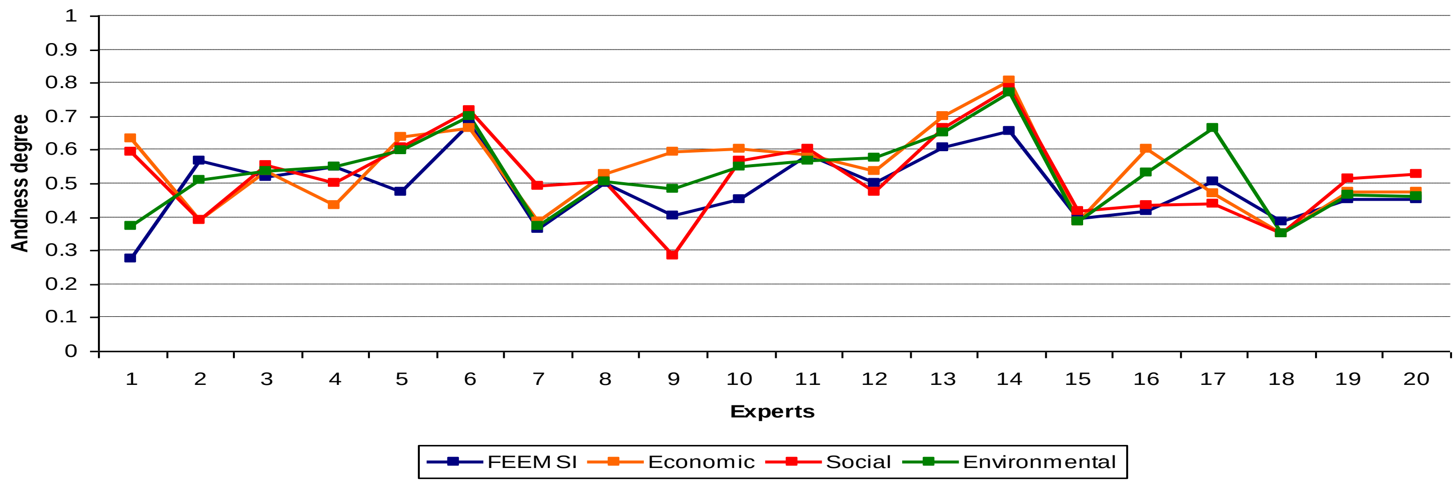
Remember: a scenario where TWO indicators have a "GOOD PERFORMANCE" must be rated at least as the maximum score assigned to a scenario where only ONE of these TWO indicators have a "GOOD PERFORMANCE"

drag to determine your evaluation

	0	10	20	30	40	50	60	70	80	90	100
ECON is GOOD SOC is BAD ENV is BAD											
ECON is BAD SOC is GOOD ENV is BAD											
ECON is BAD SOC is BAD ENV is GOOD											
ECON is GOOD SOC is GOOD ENV is BAD											
ECON is GOOD SOC is BAD ENV is GOOD											
ECON is BAD SOC is GOOD ENV is GOOD											

Decision maker behaviour

ANDNESS degree



1) The evaluations of sustainability indicators at a given node do vary among different decision makers.
2) For example, Expert 6 and Expert 14 have more *non-compensative* while Expert 15 have more *compensative* attitude towards all three final pillars and the FEEM SI node

Aggregation of decision makers (experts)

Derivation of representative DM

- Let is the valuation (i.e. judgement) of *k-th* DM for *i-th* coalition at a given sub-node.
- Each sub-node have possible coalitions that Decision Makers may have different judgements where *n* is the number of indicators at a given sub-node.
- Let us denote , as the total absolute distance of *k-th* DM's judgements to all other DMs' judgements.

$$D_k = \sum_{i=1}^{i=2^n} \sum_{l=1}^{l=m} \left| v_{ki} - v_{li} \right|$$

where , $l \neq k$, $l = 1,2,\ldots,m$ are the DMs, *i* is the valuation of a criterion or any possible combinations of criteria and *n* is the number of indicators at a given sub-node.

Then, the sum of absolute distances of all DMs calculated as:

$$\overline{D} = \sum_{k=1}^{k=m} D_k$$

Aggregation of decision makers (experts)

Derivation of representative DM

Each DM's weight is inversely related to the ratio of DM's absolute distance to the sum of all absolute distances. For instance, if a DM has the lowest absolute distance, that DM's valuations should be weighted more than the other DMs.

$$W_k = \left(\frac{D_k}{\overline{D}} \right)^{-1}$$

Weights given to each DM's evaluation can be further normalized so that the weights are bounded between 0-1.

$$w_k = \frac{W_k}{\sum_{k=1}^{k=m} W_k} \quad \text{where} \quad \sum_{k=1}^{k=m} w_k = 1$$

Since we have the normalized weights for each DM, one can obtain the "representative" DM valuations, v_i^r, for all possible coalitions at a given sub-node

$$v_i^r = \sum_{k=1}^{k=m} w_k v_{ki} \quad \text{for} \quad \forall i \quad \text{where } i = 1,2,\ldots,2^n$$

Decision maker behaviour

Interaction and ANDNESS among sustainability indicators

Node	Interaction indices			ANDNESS degree
FEEM SI	Economic	Social	Environmental	0.493
Economic	NA	-0.024	0.020	
Social		NA	-0.019	
Environmental			NA	
Economic	Growth drivers	GDP pc	Exposure	0.538
Growth drivers	NA	0.047	0.026	
GDP pc		NA	0.041	
Exposure			NA	
Social	Pop. Density	Well being	Vulnerability	0.525
Pop. Density	NA	0.016	0.041	
Well being		NA	0.020	
Vulnerability			NA	
Environmental	Air pollution	Energy	Endowments	0.532
Air pollution	NA	0.021	0.037	
Energy		NA	0.037	
Endowments			NA	
Vulnerability	Food	Private Health	Energy Security	0.528
Food	NA	0.040	0.022	
Private Health		NA	0.022	
Energy Security			NA	

Decision maker behaviour

Interaction and ANDNESS among sustainability indicators

Node	Indicators	Interaction index	ANDNESS degree
Growth drivers	R&D, Capital Accumulation	0.058	0.529
Exposure	Relative trade, Public debt	0.187	0.5935
Well being	Education, Health	0.029	0.5145
Energy security	Imp. energy, Energy access	0.000	0.5
Air pollution	GHG p.c., CO2 Intensity	0.183	0.5915
Energy Use	Energy Intensity, Renewables	0.053	0.5265
Endowments	Biodiversity, Water	0.058	0.529
Biodiversity	Animal, Plant	0.171	0.5855

- Majority of the sub-nodes, the representative DM features an ANDNESS index that is greater than 0.5 (i.e. more *non-compensative* attitude towards the nodes) and a positive interaction index value among two indicators at a given node (i.e. two indicator being more *complementary*).

- The representative DM evaluates indicators at those nodes as being more *complementary* and therefore, for a country to have a higher sustainability level, it needs to perform well in both indicators rather than simply having a satisfactory performance in only one of those.

FEEM SI Rankings

FEEM Sustainability Index Rankings in 2011

Rank	Country	FEEM SI
1	Norway	0.823
2	Sweden	0.774
3	Switzerland	0.700
4	Austria	0.691
5	Finland	0.661
6	Denmark	0.653
7	Canada	0.641
8	France	0.630
9	Ireland	0.620
10	New Zealand	0.609
11	USA	0.554
12	Australia	0.553
13	Brazil	0.546
14	UK	0.531
15	RoEurope	0.529
16	Germany	0.525
17	Portugal	0.522
18	RoLA	0.512
19	Spain	0.497
20	Benelux	0.495

Rank	Country	FEEMSI
21	Russia	0.493
22	RoEU	0.493
23	Mexico	0.492
24	Korea	0.477
25	Italy	0.472
26	Japan	0.456
27	Turkey	0.453
28	Middle East	0.450
29	Poland	0.430
30	South Africa	0.426
31	Greece	0.399
32	RoAfrica	0.398
33	RoWorld	0.385
34	SEastAsia	0.368
35	RoFSU	0.367
36	North Africa	0.342
37	RoAsia	0.325
38	Indonesia	0.299
39	China	0.287
40	India	0.240

Benelux: Belgium, Netherlands, and Luxembourg; **RoAfrica**: Rest of Africa; **RoAsia**: Rest of Asia; **RoEU**: Rest of European Union; **RoEurope**: Rest of Europe; **RoFSU**: Rest of Former Soviet Union; **RoLA**: Rest of Latin America; **RoWorld**: Rest of World; **SEastAsia**: Southeast Asia

Relative importance of indicators

Shapley Values
Relative importance of each sustainability indicator

Node	Criterion	Shapley value
FEEMSI	Economic	0.332
	Social	0.316
	Environmental	0.352
Economic	Growth drivers	0.378
	GDP per capita	0.355
	Exposure	0.267
Social	Population Density	0.254
	Well Being	0.415
	Vulnerability	0.331
Environment	Air pollution	0.351
	Energy	0.330
	Natural Endowment	0.319
Growth Drivers	R&D	0.522
	Investment	0.478
Exposure	Relative Trade Balance	0.554
	National Debt	0.446
Well Being	Education	0.508
	Health	0.492
Vulnerability	Food relevance	0.395
	Energy Security	0.275
	Private Health	0.330
Energy Security	Energy Imported	0.500
	Energy Access	0.500
Air pollution	GHG per capita	0.520
	CO_2 Intensity	0.480
Energy	Energy Intensity	0.458
	Renewables	0.542
Natural Endowment	Biodiversity	0.455
	Water	0.545
Biodiversity	Animals	0.516
	Plants	0.484

Global importance of indicators

Indicator's contribution to overall index

Indicator	Contribution to overall index
GDP per capita	0.1179
Population Density	0.0803
Education	0.0666
R&D	0.0655
Health	0.0645
GHG per capita	0.0642
Renewables	0.0630
Water	0.0612
Investment	0.0600
CO2 Intensity	0.0593
Energy Intensity	0.0532
Relative Trade Balance	0.0491
Food relevance	0.0413
National Debt	0.0395
Private Health	0.0345
Animals	0.0264
Plants	0.0247
Energy Imported	0.0144
Energy Access	0.0144

By multiplying the Shapley values of every hierarchically superior node of every indicator, from the bottom of the aggregation tree to the top (e.g., contribution of health is calculated by multiplying the Shapley values of health, well being and social pillar, since health indicator is under the node of well being which is a node of social pillar)

Robustness analysis

- In a complex aggregation such as the one used for the FEEM SI, the attitude of the representative decision maker is a key component of the process. Thus, it is important to check how robust the ranking is to a change in the representative decision maker's attitude.

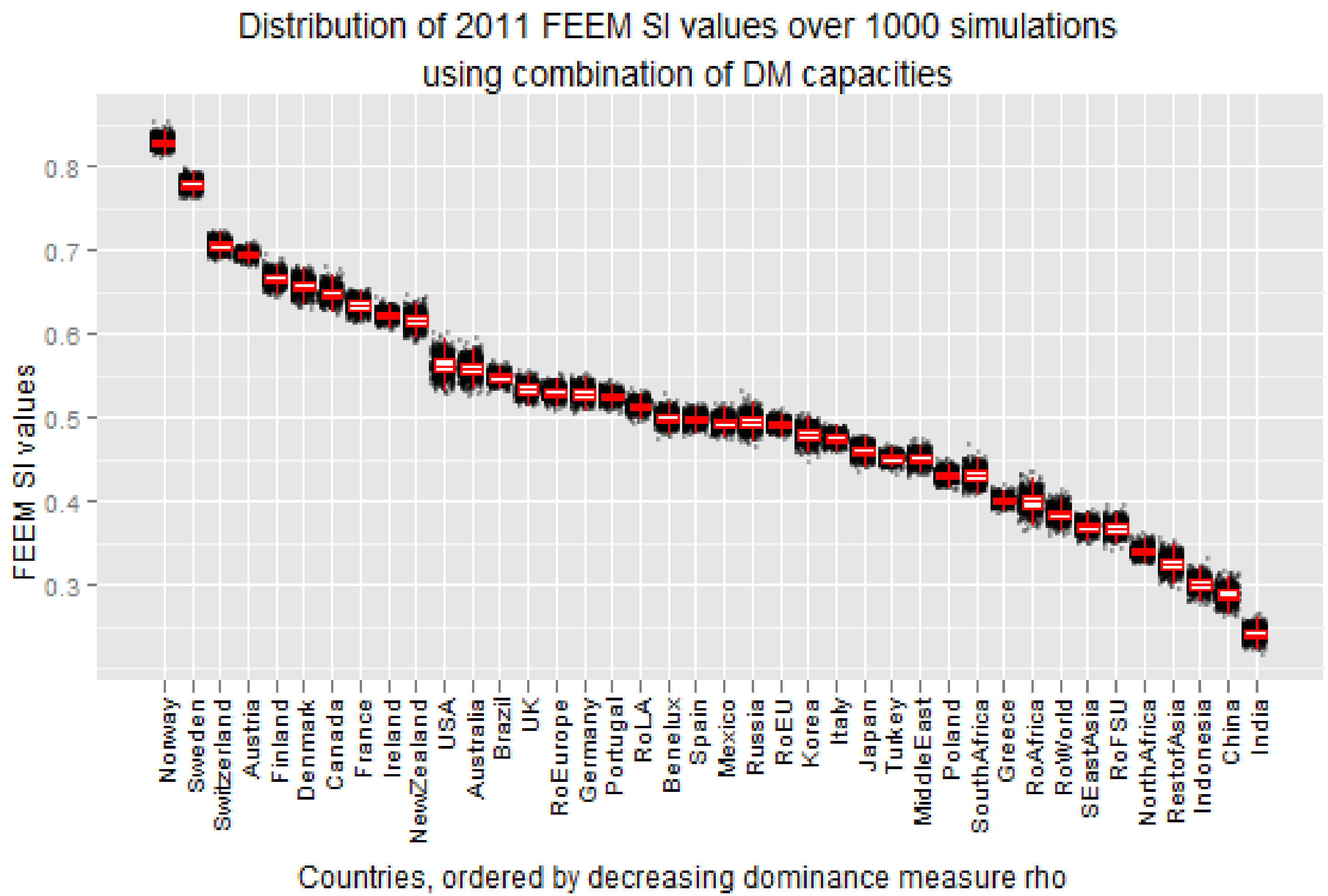
- A robustness analysis can be performed by building a linear convex combination of the values of the weights and run a significant number of simulations, as in a Monte Carlo approach.

- The robustness analysis has been carried out by generating 1000 sets of measures that are necessary to aggregate the indicators into the final FEEM SI.

- These sets have thus been called "artificial decision makers" (ADMs).

- The measures contained in the artificial decision makers has been used to aggregate, with the Choquet integral, the FEEM SI, using the same indicators as for the reference case.

Robustness analysis



Dominance analysis

Let's say the following measures have been implemented to compare any two countries *i* and *j* included in the ranking

$$\Delta(i, j) = \frac{1}{N}\sum_{k=1}^{k} F[R_k(i) - R_k(j)]$$

where *N* is the number of countries included in the ranking, $R(i)$ and $R(j)$ are the FEEM SI values for the i*th* and j*th* country respectively.

k is the number of simulations and F(x) takes the form:

$$F(x) = \begin{cases} 0 & if \quad x < 0 \\ x & if \quad x \geq 0 \end{cases}$$

$\Delta(i, j)$ represents the "average cardinal dominance" of country *i* on country *j*.

That is, the measurement expresses by how much, on average, the i*th* country dominates the j*th* across simulations.

Dominance analysis

The overall dominance measure of country *i* on every other country is given by:

$$\rho^{+}(i) = \frac{1}{N-1}\sum_{j=1}^{N}\Delta(i, j)$$

Whereas the degree to which country *i* is dominated by every other country is given by:

$$\rho^{-}(i) = \frac{1}{N-1}\sum_{j=1}^{N}\Delta(j, i)$$

Then, we can thus construct the following dominance ratio:

$$\rho(i) = \frac{\rho^{+}(i)}{\rho^{+}(i) + \rho^{-}(i)}$$

which indicates the extent of relative dominance of the i*th* country

This measures 1 if the country in question dominates any other across all simulations and measures 0 if country *i* is being dominated by all other countries.

β^+

Dominance analysis

Country	ρ^+	ρ^-	ρ
Norway	8.27	0	1
Sweden	7.02	0.03	0.995
Switzerland	5.24	0.13	0.977
Austria	4.99	0.15	0.972
Finland	4.32	0.22	0.951
Denmark	4.12	0.25	0.943
Canada	3.93	0.28	0.933
France	3.63	0.35	0.913
Ireland	3.38	0.41	0.892
NewZealand	3.26	0.44	0.880
USA	2.24	0.78	0.741
Australia	2.17	0.81	0.728
Brazil	1.98	0.89	0.689
UK	1.74	1.01	0.633
RoEurope	1.67	1.05	0.614
Germany	1.66	1.06	0.611
Portugal	1.59	1.10	0.592
RoLA	1.43	1.22	0.540
Benelux	1.22	1.39	0.467
Spain	1.21	1.41	0.462

Country	ρ^+	ρ^-	ρ
Mexico	1.164	1.451	0.4449
Russia	1.163	1.454	0.4445
RoEU	1.13	1.49	0.432
Korea	0.99	1.67	0.374
Italy	0.95	1.73	0.354
Japan	0.80	1.98	0.288
Turkey	0.73	2.12	0.255
MiddleEast	0.72	2.13	0.253
Poland	0.572	2.48	0.188
SouthAfrica	0.568	2.49	0.186
Greece	0.38	3.05	0.112
RoAfrica	0.37	3.09	0.108
RoWorld	0.29	3.39	0.080
SEastAsia	0.23	3.72	0.057
RoFSU	0.22	3.75	0.056
NorthAfrica	0.13	4.36	0.029
RestofAsia	0.09	4.70	0.020
Indonesia	0.05	5.29	0.009
China	0.03	5.61	0.005
India	0	6.77	0

Concluding remarks

- This paper aimed at proposing an application of non-linear aggregation methods to sustainability literature, extending the current work in this field to address the intrinsic complexity underlying the sustainability concept
- Computation of the Shapley index, it has been possible to address the relative importance of different indicators, which could also be used in the future to refine the current sustainability tree.
- The representative DM evaluates majority of the sustainability indicators as being more *complementary* and therefore, for a country to have a higher sustainability level, it needs to perform well in all indicators rather than simply having a satisfactory performance in only one of those.
- Robustness analysis has provided a measure of the subjectivity implied by the artificial decision makers developed, confirming the validity of the overall method in evaluating overall sustainability.

Future research

- Extend the pool of decision makers involved in the determination of the measures used in the aggregation operator.
- Extension of the pool of decision makers will not only allow for obtaining a more representative ‘consensus’ measures, but also for the evaluation of sustainability perceptions from different parts of the world.
- Given the heterogeneity of the current conditions (economic, social and environmental condition in general) in different countries (macro-regions), the need for future generations will vary and therefore importance given to sustainability indicators and their interactions may differ.
- If this is the case, a toll like the FEEM SI can offer different policy implications in different regions considered in the analysis