Quantitative storytelling in the making of a composite indicator

Article

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

ORIGINAL RESEARCH

Quantitative Storytelling in the Making of a Composite Indicator

Marta Kuc-Czarnecka[1] · Samuele Lo Piano[2,3] · Andrea Saltelli[3,4]

Abstract

The reasons for and against composite indicators are briefly reviewed, as well as the available theories for their construction. After noting the strong normative dimension of these measures—which ultimately aim to 'tell a story', e.g. to promote the social discovery of a particular phenomenon, we inquire whether a less partisan use of a composite indicator can be proposed by allowing more latitude in the framing of its construction. We thus explore whether a composite indicator can be built to tell 'more than one story' and test this in practical contexts. These include measures used in convergence analysis in the field of cohesion policies and a recent case involving the World Bank's Doing Business Index. Our experiments are built to imagine different constituencies and stakeholders who agree on the use of evidence and of statistical information while differing on the interpretation of what is relevant and vital.

✉ Marta Kuc-Czarnecka
   marta.kuc@pg.edu.pl

   Samuele Lo Piano
   s.lopiano@reading.ac.uk

   Andrea Saltelli
   andrea.saltelli@uib.no

[1] Faculty of Management and Economics, Gdansk University of Technology, Traugutta 79, 80-233 Gdańsk, Poland

[2] School of the Built Environment, University of Reading, JJ Thomson Building, Whiteknights Campus, Reading RG6 6AF, UK

[3] Open Evidence Research, Universitat Oberta de Catalunya, Carrer dels Almogàvers 165, 08018 Barcelona, Spain

[4] Centre for the Study of the Sciences and the Humanities, University of Bergen, Postboks 7805, N-5020 Bergen, Norway

1 Scope

Composite indicators have seen dramatic growth in use as well as in impact, and are increasingly used at face value by a plurality of actors. At the same time, academia maintains an honest scepticism about these measures, tempered by methodological developments that aim to remedy their most evident shortcomings.

True to its nature, a composite indicator is usually built to 'tell a story'. It is thus ideally suited to identify and bring attention to a possibly latent phenomenon. For this, it is appreciated by 'value entrepreneurs'.[1]

When used for policy, the unidirectionality of composite indicators is less desirable. In the context of policy analysis and negotiation, in which different options, as well as different 'ends in sight', are relevant, composite indicators may fall short. Ideally, different stakeholders could confront one another, armed with different measures and indicators. Could the concept of composite indicators be stretched so as to accommodate these settings? Could one force it to tell 'more than one story'?

In this contribution, we give this concept a try based on examples.

2 The Fortune of Composite Indicators

A recent search on Scopus performed in the summer of 2019 indicates growth in interest in composite indicators, proxied by their mention in the scientific literature. Figure 1 suggests that not only is the use of composite indicators increasing but so is the pace of their growth.

The 'Report by the Commission on the Measurement of Economic Performance and Social Progress', prepared by Joseph Stiglitz, Amartya Sen, and Jean-Paul Fitoussi for the French president Nicolas Sarkozy (Stiglitz et al. 2009), says that the growth in the number of statistical indicators reflects different concurring trends: improvements in the level of education, increases in the complexity of modern economies, and the widespread use of information technology.

As discussed in Becker et al. (2017), human well-being and progress are areas in which composite indicators are popular, covering themes from happiness-adjusted income to environmentally adjusted income, from child development to information and communication technology. They are also used in the analysis of innovation (Balcerzak and Pietrzak 2017a; Dutta, Lanvin and Wunsch-Vincent 2018; Hausken and Moxnes 2019; Żelazny and Pietrucha 2017), analysis of real estate markets (Małkowska and Głuszak 2016), countries' competitiveness (Cheba and Szopik-Depczynska 2017; World Bank 2019; Kruk and Waśniewska 2017; Schwab 2019), socio-economic development (Bartkowiak-Bakun 2017; Mazziotta and Pareto 2016), the quality of institutions (Balcerzak and Pietrzak 2017b), sustainable development (Balcerzak and Pietrzak 2017c; Luzzati and Gucciardi 2015; Semenenko et al. 2019), the standard of living (Greyling and Tregenna 2017; Kuc 2017), well-being (Barrington-Leigh and Escande 2018; Chaaban et al. 2016; Peiro-Palomino and Picazo-Tadeo 2018) and many others (Aparicio and Kapelko 2019; Capecchi and Simone

---

[1] Value entrepreneur is a term coined by Emily Barman's in 'Caring Capitalism', cited in Popp Berman and Hirschman 2018). Those are experts whose job is to measure social value of different measures for the purpose of investment, and whose choices are driven by 'communicative goals'; a value entrepreneur might want to establish legitimacy, show conformity, change behaviour or justify a field.
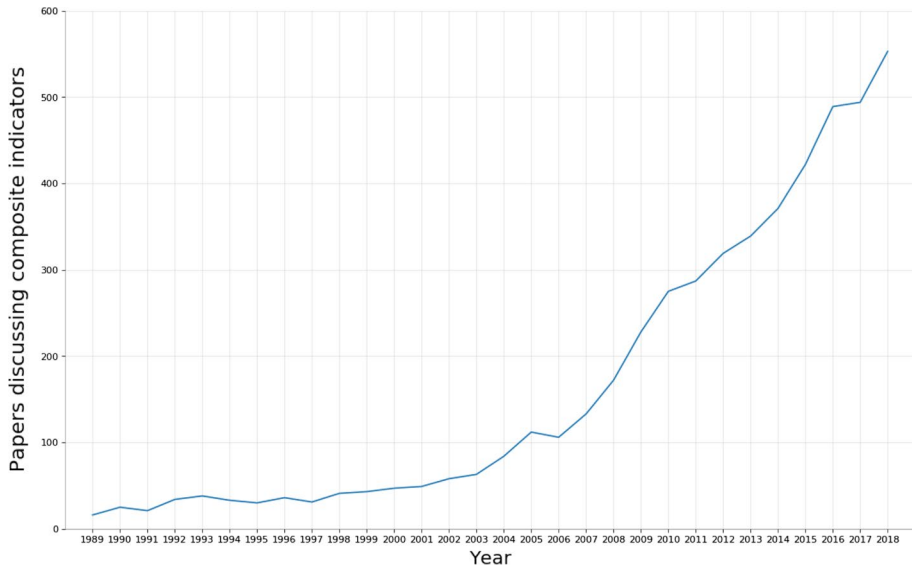
Fig. 1 Search on www.scopus.com using as search string: TITLE-ABS-KEY ("composite indicator*") OR TITLE-ABS-KEY ("composite index") OR TITLE-ABS-KEY ("composite indices")

2019; Dinis et al. 2019; Marozzi 2015; Mann and Shideler 2015; Michener 2015; Miro and Piffaut 2019; Rogalska 2018). Many composite indicators target issues of sustainability. In this respect, an ongoing line of inquiry is to understand why these measures seem to have little traction, e.g. in relation to their capacity to displace established indicators such as gross domestic product (GDP) as measures of progress (Boulanger 2018; Popp Berman and Hirschman 2018). We return to these themes later in the paper.

3 Pros and Cons of Composite Indicators

An existing handbook on composite indicators (OECD-JRC 2008) lists several pros and cons of these measures (Table 1).

The report by Stiglitz et al. (2009) considers composite indexes (CIs) problematic because even when the weighting procedure is presented transparently, the normative implications of the weights are seldom spelt out or justified. Along the same lines, for Popp Berman and Hirschman "successful quantification projects tend to hide their assumptions". The critique of CIs also feeds into the present reflection on the abuse of metrics (Muller 2018), with the example of university ranking and similar league tables standing out as controversial (Saisana et al. 2011; Wilsdon 2016).

For mainstream economics—for the purpose of this paper exemplified by the World Bank's economist Martin Ravallion—CIs are guilty of not being constructed on sound economics. Ravallion identifies two types of indices: those built on economic theory, either

Table 1 Pros and cons of composite indicators Source: OECD-JRC 2008

Pros	Cons
1. Can summarise complex, multi-dimensional realities to support decision-makers	1. May send misleading policy messages if poorly constructed or misinterpreted
2. Are easier to interpret than a battery of many separate indicators	2. May invite simplistic policy conclusions
3. Can assess the progress of countries over time	3. May be misused, e.g. to support the desired policy, if the construction process is not transparent and/or lacks sound statistical or conceptual principles
4. Reduce the visible size of a set of indicators without dropping the underlying information base	4. The selection of indicators and weights could be the subject of political dispute
5. Thus, make it possible to include more information within the existing size limit	5. May disguise severe failings in some dimensions and increase the difficulty of identifying proper remedial action if the construction process is not transparent
6. Place issues of country performance and progress at the centre of the policy arena	6. May lead to inappropriate policies if dimensions of performance that are difficult to measure are ignored
7. Facilitate communication with the general public (i.e. citizens, media) and promote accountability	
8. Help to construct/underpin narratives for lay and literate audiences	
9. Enable users to compare complex dimensions effectively	

direct monetary aggregates or based on shadow prices[2]; and all others, which are dismissively termed 'mashup indices'—a definition that Ravallion applies to existing measures, such as the Human Development Index (HDI) and the Multidimensional Poverty Index (MPI).

More, in general, the use of Economics as a master discipline to adjudicate the soundness of any form of knowledge has suffered, after the internal discussion in economics on the role of mathematics and modelling. Initiated after the models' inability to predict the last recession (Mirowski 2013), this discussion gave birth to a new term, 'mathiness'—coined by the economist Paul Romer (2015), defined as the use of mathematics to veil positions which are in fact normative. The use of prices—whether real or shadow—is in itself fraught with the crucial normative assumption that socio-ecological outcomes can be represented in monetary terms (Funtowicz and Ravetz 1994). Finally, doubts remain as to what constitutes a 'sound economic theory', given the present controversy in the profession against the prevailing economic paradigm (Reinert 2008; Rethinking Economics 2017).

Indeed, we suggest that composite indicators is a field better addressed by social scientists in general, than it is by the field of economics in particular, as we go on to discuss in the next session.

4 Is a Theory for Composite Indicators Possible?

Without being exhaustive, we mention here some important ingredients of a possible theory of CIs. To start with, the OECD-JRC handbook (2008) offers ten recursive steps to build CIs, from building a theoretical framework to how to present the results, and includes

---

[2] Economists resort to various strategies to compute costs when they are not available. For example, a shadow price can be estimated based on the estimated 'willingness to pay'; in the absence of a market price, the value of a good or service is taken to be what people would be willing to pay for it.

advice on how to tackle technical choices on data selection, imputation, normalization, and aggregation. Even a cursory glance at this list is sufficient to grasp that numerous modelling assumptions are needed in these processes, not just the assignment of weights; the idea that a composite indicator can be made 'objective' should be put to rest. As mentioned above, 'telling a story' is precisely what a 'value entrepreneur' wishes to do. In this context, the problem may not be with the non-neutrality of a measure but, rather, with its purported neutrality. The production of apparently neutral scientific fact in the context of policy is known as 'stealth advocacy' (Pielke 2007), meaning that this form of advocacy is hidden behind a veneer of objectivity. Following Pielke, we suggest that scientists and CI developers be clear about their normative stances, rather than courting an impossible neutrality.

It is also important to note that most existing CIs suffer from technical shortcomings, especially when they are built using a linear combination of variables. For example, most indicators are built in such that the weights attached to the different variables (i) add up to one and (ii) reflect the importance of the variables. In fact, both (i) and (ii) are highly questionable, if not outright wrong. It can be proven that according to mathematical theory it is the sum of the squared weights which should be one (Becker et al. 2017) and that the actual importance of variables in a CI may deviate from its weights considerably (Paruolo et al. 2013). Both results owe to the theory of sensitivity analysis (Saltelli et al. 2008), a tool used in mathematical modelling. This theory proves that the importance of a given variable is given by the square of its weight divided by the sum of all squared weights, for the ideal case of uncorrelated and standardised variables (Saltelli et al. 2008, p. 47). For non-standardised and non-independent variables, the importance depends on the interplay of the relative variances and on correlation among the variables.

For Paul-Marie Boulanger (2014), CI can be situated at the intersection of three conceptual movements. The first is associated with the democratisation of expertise (Carrozza 2014), the idea that more knowledge needs to be brought to bear than that provided by experts. This is close to the concept of extended peer community advocated in post-normal science (Funtowicz and Ravetz 1993), in which the soundness of scientific practice is not judged simply by the peers of a given discipline but by several disciplines, because each discipline offers a different viewpoint. Laypeople with a direct stake, knowledge, or interest in the matter at hand are also involved in the process. The OECD-JRC handbook insists on the need for an indicator to be co-developed, possibly and ideally involving the community of those actors (individuals, institutions, countries, or regions), which are affected by the measure. This process should extend to all the stages in a CI's construction (OECD-JRC 2008).

The second constitutive element identified by Boulanger (2014) is CI as instrumental in the creation of a new public through a process of social discovery (Dewey 1938). Why are 'social discoveries' needed? For Dewey, there are 'publics' affected by a transaction taking place somewhere else, who need to be made aware of a problem situation—e.g. the pollution of an aquifer, a case of air contamination, the unintended consequences of technology, and so on. In his view, the "machine age has so enormously expanded, multiplied, intensified and complicated the scope of the indirect consequences […] that the resultant public cannot identify and distinguish". The German sociologist Ulrich Beck later called a society thus affected by a myriad of mostly invisible threats a 'risk society' (Beck 1992 [1986]).

In general, developing an index of, say, air quality, sustainability, the rule of law, corruption, or university performance can generate and mobilise new publics, at times producing important policy outcomes. Several recent books on the emerging field of sociology of quantification reviewed in Popp Berman and Hirschman (2018) address this issue.

The third element which can be used to shed light on CI comes from Charles Sanders Peirce, the father of semiotics, and his triadic conception of the sign as a structure connecting three elements: the sign properly said (S), an object (O), and an 'interpretant' (I). The example reported by Boulanger (2014) is that of the African vervet monkey, which possesses a sophisticated repertory of vocal signs for signalling the presence of a predator, distinguishing a terrestrial stalker such as a leopard, an aerial raptor such as an eagle, or a ground predator such as a snake. In this case, the 'object' is the predator, the 'sign' is the cry emitted by the vervet, and the 'interpretant' is the collective behaviour of the monkeys in reaction to the cry, e.g. climbing up a tree if the predator is terrestrial. Thus a CI is not just a sign (a number) pointing to an issue; it also entails an interpretant, understood as the policy or social change desired or suggested by the proponents of the measure. The same Dewey (1938) made the point that any social indicator is meaningful only in the context of a desired end in sight.

Boulanger enriched the analysis of indicators of sustainability and their interpretants in a subsequent work (2018) drawing from the German sociologist and systems theorist Niklas Luhmann. For Boulanger, mobilising publics with a CI is not just a matter of providing the 'facts'. According to Luhmann, modern societies should be seen as functionally differentiated, i.e. composed of several sub-systems, relatively independent (law, science, the economy, the mass media, etc.). These systems may either actively develop/use indicators to influence their evolution or be influenced by them when one system is observing another system with an indicator. The problem with sustainability is thus which system is observing or 'irritating' (the word used by Luhmann) which system. Science cannot easily 'irritate' e.g. the economy developing a measure of progress to replace GDP, as science and economy follow separate codes of communication, *true–false* for science and *gain–loss* for the economy.

There are instead cases where the strength of indicators, their capacity to appear objective, and their addictive character, manage to impact social systems, as evidenced by university league tables mentioned earlier (Saisana et al. 2011; Wilsdon 2016).

5 Frames and Quantitative Storytelling (QST)

It is a tautology that every measure of society corresponds to a frame and—according to Dewey—to 'an end in sight' as well. The use of frames in the policy discourse is a hot topic. For example, in the US, George Lakoff (2004, 2010) laments the liberals' cultural subjugation to frames developed by conservatives. In economics, Akerlof and Shiller (2015)discuss how economic actors operating in a market are forced to exploit the frames of their consumers in order to survive. This view reverses economics' (and Adam Smith's) cherished paradigm that the composition of individual selfishness in a market produces a common good. The market as the best arbiter of all negotiations is at the heart of the prevailing orthodoxy, neoclassical economics. As mentioned, this theory is at present contested (Mirowski 2013; Reinert 2008; Rethinking Economics 2017). The French legal scholar Alan Supiot (2015) sees the neoliberal creed in markets as particularly reliant on quantification, in which numbers replace laws to create a world of dystopian injustice and dysfunction.

Frames are very much at the core of discussion on the use of evidence for policy (known as evidence-based policy [EBP]). Dan Kahan, an expert in cultural cognition theory, studies how we process new knowledge in order to reinforce our existing worldviews and

political orientations (Kahan et al. 2011), so an actor's deeper knowledge reinforces—rather than cures—his or her polarisation.

It is thus a current refrain in EBP that facts and values cannot be always be separated, e.g. when social facts are brought to bear on the design of a policy (Gluckman 2017). The present—in our view disingenuous—brouhaha about the end of facts and the post-truth society (Flood 2016) mixes somewhat uncritically conscious strategies of confusion and manipulation—as witnessed in the recent US elections—with the existence of a plurality of legitimate frames around what constitutes a problem and hence a 'fact' in relation to the problems (Saltelli and Funtowicz 2017).

In other words, one should not equate the US administration's 'alternate facts' with the concept of 'extended facts', defined in the theory of post-normal science as the product of an extended peer community (Funtowicz and Ravetz 1993). An extended fact may be a loss to a constituency which has been overlooked, a fact that is not part of the set of facts brought to bear by a regulator or by the proponent of a policy.

Quantitative storytelling (QST; Giampietro et al. 2014; Saltelli and Giampietro 2017) posits that more legitimate frames and worldviews are upheld by different social actors. Thus, QST draws attention to the fact that economic and mathematical models used in EBP are often in the form of risk analyses or cost–benefit analyses. These models focus on a single framing of the issue under consideration. A classic example is when a monetary equivalent is assigned to a social or environmental good. As discussed above, in relation to the critique of Ravallion (2010), this implies a clear normative stance.

In the logic of QST, the deepening of the analysis corresponding to a single view of what the issue is—achieved by 'mathematizing' the problem – distracts from what could be alternative readings.

For Ravetz (1987) and Rayner (2012), alternative frames may represent 'uncomfortable knowledge', which is removed from the policy discourse. Lakoff (2004, 2010) suggests that frames may be used to generate 'hypo-cognition'.

Under this critical viewpoint, mathematical models—or a CI in the present context—can be seen as a tool for 'displacement' (Rayner 2012). This occurs when a model or a ranking becomes the end instead of the means, e.g. when universities monitor and manage the outcome of a ranking, rather than what happens within their walls (Saisana et al. 2011). After stakeholders realise they are on the receiving end of a strategy of hypo-cognition, their trust in the actors and institutions involved may be diminished (Saltelli and Giampietro 2017). A fundamental problem with EBP is that stronger players have access to better evidence, i.e. more data and indicators in the present context, and can use it strategically (Saltelli 2018). Thus, EBP is based on a power asymmetry (Boden and Epstein 2006; Strassheim and Kettunen 2014).

In QST, one drops the hope that neutral, impersonal facts will prescribe a policy and suggests instead acknowledging ignorance, as to identify 'clumsy solutions' (Rayner 2012), which may accommodate unshared epistemological or ethical principles. Similarly, post-normal science suggests 'working deliberatively within imperfections' (van der Sluijs et al. 2008). The solution offered by QST is to use quantification 'via negativa', by testing which of the available frames runs afoul of a quantitative or qualitative analytic check, as shown in the PISA example below. QST borrows from system ecology and attempts to refute the frames if they are based on unsound inference or violate the constraints of (Giampietro et al. 2014): (i) feasibility (can one afford a given policy in terms of external constraints, e.g. existing biophysical resources?), (ii) viability (can one afford it in the context of the internal constraints, governance, socio-economic, and technological arrangements?), and (iii) desirability (will the relevant constituency accept it?). For example, in examining the

transition to a carbon-free economy, one can test the availability of natural resources (lithium, cobalt, and other minerals needed for energy storage), whether legislation promoting transition is viable, and whether the transition is compatible with existing lifestyles. An instructive test case of QST exploring the deployment of intermittent electrical energy supply in Germany and Spain is reported in Renner and Giampietro (2019).

Perhaps the best application of the concept of QST *ante litteram* is an old study of how European citizens perceive the existing narratives and conflicts in the adoption of genetically modified (GMO) food and products (Marris et al. 2001). Although the prevailing narrative is that 'GMO food is safe' and that consumer reluctance is rooted in anti-technology or anti-science prejudice, Marris and co-authors showed that the people interviewed did not care about the safety of GMO food but, rather, expressed concern about a totally different set of issues, including:

- who would benefit from these technologies
- why they were introduced in the first place
- whether existing regulatory authorities would be up to the task of resisting regulatory capture from powerful industrial incumbents.

Quantitative storytelling, like other tools for evidence appraisal, such as sensitivity analysis (Saltelli et al. 2008), NUSAP (Funtowicz and Ravetz 1990; van der Sluijs et al. 2005), and sensitivity auditing (Saltelli et al. 2013), can be useful for gauging and possibly deconstructing existing measures. Thus, proposers of new CI should factor this danger in if they wish to anticipate criticism. They should be the first to test the relevance and robustness of their constructs, following the well-known Mertonian principle of 'organized scepticism' (Merton 1973), in which scientists strive to falsify their own results and invite fellow scientists to attempt such a deconstruction.

In the present work, we apply the QST approach to the construction of a composite indicator. We consider that social convergence, with its dense web of interconnected interests, policies, and outcomes, offers an ideal environment for such an experiment. The Cohesion Policy and convergence issues are still being discussed in the policy arena (e.g. the 7th Cohesion Forum, held in Brussels 26–27 June 2017; European Commission 2017a, b, c, d, e), as well as in the academy (Anagnostou et al. 2015; Baddeley 2006; Balcerzak and Rogalska 2016; Cosci and Mirra 2017; Furkowa and Chocholata 2017; Horridge and Rokicki 2017; Pietrzak and Balcerzak 2017; Próchniak and Witkowski 2016; Scheurer and Haase 2017; Stanickova 2017). The role of narratives linked to the Cohesion Policy is likely to be perceived as increasingly important for mitigating the present difficulties of the European project (Applebaum 2017).

6 A Previous Example of Quantitative Storytelling

Quantitative storytelling has been used in relation to the ranking of the OECD-PISA study (Araújo et al. 2017; Saltelli 2017). We describe this work here, as it shows how the methodology can be used to deconstruct a measure. In the test cases of the following sections a constructive use of quantitative storytelling is demonstrated.

Since the publication of its first results in 2000, the Programme for International Student Assessment (PISA) implemented by the Organisation for Economic Co-operation and Development (OECD) has been a subject of controversy. PISA has been presented

by some as a measure of a country's innovation and growth potential, while others found these metrics—published every 3 years with considerable media amplification—irrelevant and potentially counter-productive. Noticeably, the PISA dispute was the subject of a letter published in *The Guardian* newspaper and signed by several educationalists and scholars and of a subsequent exchange with the OECD (Meyer and Zahedi 2014; for additional references, see Araújo et al. 2017).

OECD-PISA is a convenient example for discussing the importance of the issue of frames in policy, as well as some limitations in the concept of EBP.

For advocates of 'evidence-based' or 'informed' policy, PISA incarnates the dispassionate, objective facts which nourish the formulation of sound policies by allowing for comparison across countries and possibly for the identification of good practices worth emulating. For opponents of this survey, the relation between PISA and economic growth represents a neoliberal framing of education policies within a context of globalisation which is perceived as unacceptable. QST showed that—while international comparability is desirable—more tends to be read into these ranking than the quality of evidence allows (Araújo et al. 2017).

According to the analysis in Araújo et al. (2017), a number of issues emerged.

(1) Over-interpretation of PISA results: According to PISA supporters (Woessmann 2014), "If every EU Member State achieved an improvement of 25 points in its PISA score (which is what for example Germany and Poland achieved over the last decade), the GDP of the whole EU would increase by between 4% and 6% by 2090; such a 6% increase would correspond to 35 trillion Euro."

(2) PISA scoring strongly depends upon the modelling assumption, the design of the sample, the choice of the items (questions) included or excluded, and the number and typology of students sampled. Previous works reviewed in Araújo et al. (2017) showed that shifts in the relative position of a country were attributed to the success or failure of educational policies when, in fact, they were due to different compositions of the share of students excluded from the test.

(3) PISA ranking lacks uncertainty and sensitivity analysis. PISA just offers a summary and non-conservative measure of the error of a country score.

(4) The non-availability of the full data hampers a full analysis of the sensitivity of PISA scores to modelling assumptions.

(5) PISA embeds strong normative stances, foremost the fact that education is investigated as an input to growth.

(6) PISA may adversely affect what is taught and might run counter to our desires concerning what education should be about. It encourages focusing on the subset of educational topics being selected at the expense of others.

(7) In measuring what it considers 'life skills', PISA assumes that these skills are the same across countries and cultures. All societies are bound to become 'knowledge' societies. However, diversity in the curriculum being taught might be a source of country-specific creativity and well-being.

We recognise in this list many of the 'flags' from QST (and from sensitivity auditing as well; see Araújo et al. 2017), e.g. in technical shortcomings in the interpretation of analysis, its non-transparency, the non-desirability of the adopted narrative, and the institutional conflicts on whether countries or a supra-national organisation such as the OECD should dictate curricula.

While in the example just given QST was used to deconstruct a frame, in the present work it will be used to enrich the spectrum of frames in order to test a new style of use for CI.

7 First Case: Analysis of Convergence

As discussed above, social convergence offers an ideal arena for testing QST. With the implementation of the European Pillar of Social Rights, a stronger focus is placed on social performance and employment, in which what Europe needs is less division and more cohesion, especially now, when the European Union is struggling with Brexit, a refugee crisis, a multispeed union and a populist upsurge of euro-scepticism.

8 Second Case: Doing Business Index (DBI)

The World Bank's Doing Business Index—also known as the ease of doing business score—is an extremely popular CI, constructed by aggregating forty-one component indicators over ten thematic areas (World Bank 2019):

1. Starting a business.
2. Dealing with construction permits.
3. Getting electricity.
4. Registering property.
5. Getting credit.
6. Protecting minority investors.
7. Paying taxes.
8. Trading across borders.
9. Enforcing contracts.
10. Resolving insolvency.

The forty-one component indicators are first normalised according to a min–max scheme and then aggregated through a simple average within the thematic areas to which they belong and finally into the overall index. Each of the ten areas has the same weight, and so do the individual CIs in an area.

One-hundred-ninety countries are then ranked from the highest index value to the lowest.

A Google search on 'world bank' and 'doing business index' in July 2019 yielded as many as 5290 hits, while a search on Scopus with the search strings (TITLE-ABS-KEY 'world bank' AND 'doing business index') resulted in fifteen documents.

9 Methodology for the Convergence Analysis

9.1 Composite Indicators

The classical approach to constructing CIs implies the assignment of variables to a given pillar (based on researchers' knowledge or experts' opinion), then the aggregation of variables within the pillar, and finally aggregation into a holistic CI. We follow here this popular approach.

Variables used in the analysis have a different impact on social performance. Stimulants are factors that have a positive impact on the phenomenon analysed (e.g. the employment rate), while destimulants have a negative impact (e.g. the infant mortality rate). In regional research, destimulants are often transformed to stimulants based on the inversion formula:

$$x_{ijt}^s = \frac{1}{x_{ijt}}(i = 1, \dots, n; j = 1, \dots m; t = 1, \dots, k) \tag{1}$$

where x_{ijt}^s is the value obtained as stimulant j in country (region) i in year t obtained from transforming the original destimulant x_{ijt} (the superscript s stands for stimulant), x_{ijt} is the value of destimulant j in country (region) i in year t.

The inversion formula is the simplest transformation method, and it gives all the diagnostic variables the same interpretation in the sense of their impact on the phenomenon analysed, i.e. the higher the value, the better in the optic of the index.

After this transformation all variables are normalised according to the formula:

$$x_{ijt}' = \frac{x_{ijt}^s - min\ x_{ij2005}}{max\ x_{ij2005} - min\ x_{ij2005}}(i = 1, \dots, n; j = 1, \dots m; t = 1, \dots, k) \tag{2}$$

where $min\ x_{ij2005}$ is the minimum value of variable j in 2005. $max\ x_{ij2005}$ is the maximum value of variable j in 2005.

This normalisation method enables the results to be compared and their dynamics to be analysed by providing a fixed reference point (Pawełek 2008). In this paper, we assume that each dimension is equally important, so the CI is calculated as:

$$CI_{it} = \frac{1}{p}\sum_{q=1}^{p} z_{iqt}(i = 1, \dots, n; t = 1, \dots, k) \tag{3}$$

where CI_{it} is the composite indicator describing social performance in country (region) i in year t, z_{iqt} is the composite indicator in country (region) i calculated for variables included in group q in year t, p is the number of groups.

The value of z_{iqt} is calculated as the mean average of all variables in each dimension. In this case, the higher the CI value is, the better for the phenomenon analysed.

9.2 Measuring Convergence

In a convergence analysis of EU Cohesion Policy, several measures are customarily employed: sigma convergence, delta convergence, gamma convergence, and beta convergence.

The sigma convergence concept measures gaps among time series by examining whether cross-sectional variation (measured by either standard deviation, coefficient of variation, Gini index, or any other dispersion measure) decreases over time, as would be anticipated if two

series converged (Kong et al. 2019). To investigate the existence of a sigma-convergence trend, regression is usually used:

$$V_t = \alpha_0 + \alpha_1 t + \varepsilon_t \tag{4}$$

where V_t is the coefficient of variation in the year t.

The following set of hypotheses was tested:

H_0 $V_1 = V_2 = V_t$ no sigma convergence or divergence,
H_1 $V_1 > V_2$ the existence of sigma convergence,
H_{1a} $V_1 < V_2$ the existence of sigma divergence

where V_1 is the coefficient of variation in a given year, V_2 is the coefficient of variation in the next year.

If the estimated value of parameter α_1 turns out to be negative and statistically significant, then sigma convergence is taking place, and diversity among the objects analysed is decreasing. In the case of a positive sign, sigma divergence occurs, i.e. diversity among entries is accelerating (Barro and Sala-i-Martin 1999). In some case studies, a simple plot which shows a tendency of cross-sectional variance to decrease over time is taken as evidence in favour of sigma convergence (Tsionas 2002). This concept is widely used in policy literature (i.e. European Commission 2014, 2016).

Gamma convergence is a concept proposed by Boyle and McCarthy (1997). It requires an examination of the change in the ranking of countries. A simple measure that captures the change in rankings is Kendall's index of rank concordance, calculated as:

$$\tau = \frac{C - D}{n(n - 1)} \tag{5}$$

where C is the number of concordant pairs of countries, D is the number of discordant pairs of countries, n is the number of observations (countries).

Two observations are called concording if the two members of one observation are larger than the respective members of the other observation. Two observations are said to be discording if the two members of one observation are in opposite order to the respective members of the other observation (Kendall 1938).

If τ is closer to zero, than changes within the distribution are higher, and gamma convergence occurs, in the so-called overtaking effect. The advantage of this approach is the ability to capture dynamics and mobility among objects (Boyle and McCarthy 1997; Holzinger et al. 2011). Gamma convergence is usually based on a comparison of the linear ordering of analysing observations (countries, regions) based on the CI's vale. Usually it is the first and the last period of the analysis, that is taken into consideration. If Kendall tau is statistically insignificant or negative, one can say that gamma convergence occurs and the overtake effect can be observed.

A less known concept, which is nonetheless important policy-wise, is delta convergence. This concept was proposed by Heichel et al. (2005) and focused on decreasing distance towards an exemplary model, or a frontrunner object. Delta convergence can be measured by the Euclidean distance from the top performer:

$$d_i = \sqrt{\sum \left(x_{ijt} - max\ x_{ijt} \right)^2} \tag{6}$$

where d_i is the distance of country i from the frontrunner, $max\ x_{ijt}$ is the frontrunner.

If the sum of distances from the frontrunner is decreasing, that suggest that objects are converging. Otherwise, divergence patterns can be observed.

$$D_f = \alpha_0 + \alpha_1 t + \varepsilon_t \qquad (7)$$

where D_f is the sum of distances.

The following set of hypotheses was tested:

H_0 $D_{f1} = D_{f2} = D_f$ no delta convergence or divergence,
H_1 $D_{f1} > D_{f2}$ the existence of delta convergence,
H_{1a} $D_{f1} < D_{f2}$ the existence of delta divergence.

In the European Union, cohesion policy sigma and delta convergence are more desirable than beta convergence, as policymakers and the general public are also interested in reducing the disparities, not pure growth per se (Eurofound 2018; European Commission 2015). Because of this fact, beta convergence is not discussed in this paper.

10 Quantitative Storytelling on the Convergence Test Case

We focused on convergence at a national scale, allowing the composition of the index to vary. Different narratives are associated with different measures of convergence. As mentioned in the previous section, sigma convergence concerns a reduction in the disparities among countries, gamma convergence seeks changes in the distribution, and delta convergence corresponds to reducing the distance from the frontrunner.

Here, we test QST in the context of CI to investigate the existence of social convergence among EU countries in 2005–2017. Therefore, for the sake of illustration we introduce a set of new CIs, constructed by using the 24 variables that are in the European Pillar of Social Rights. Six variables describe governance and fairness, and six variables are related to health care. The data come from the European Pillar of Social Rights, Eurostat, the World Health Organization, Eurostat, and national statistical offices. In the case of missing data, we used an imputation procedure based on multiple regression (James et al. 2017).

In our research, we assumed the existence of four different stakeholder groups. Each group has a different point of view about which dimensions should be included in the CI. The starting point in our analysis is the set of variables in the European Pillar of Social Rights, grouped by category (Table 2). These variables are the choice of stakeholder no. 1 (see Table 3). Stakeholder no. 2 agrees with the first that those three categories are important, but from her point of view, a 'social Europe' should include measures of governance and fairness. A third stakeholder thinks that governance is not relevant in social convergence analysis and that the functioning of health care should be investigated instead. Finally, a fourth stakeholder argues that for an exhaustive social convergence analysis, and all previously mentioned dimensions should be included (see Table 2).

Table 3 lists the variables included in each dimension. CI values were calculated using Eq. (3). Those values were the basis for estimating sigma convergence from Eq. (4) using an ordinary least squares (OLS) method. The results are in Table 4.

Table 4 shows that sigma convergence occurs for stakeholders nos. 1 and 2. By contrast, stakeholder no. 3 sees more variation among countries. In addition, the divergence among the member states is increasing (positive α_1) for stakeholders no. 3 and 4;

Table 2 Stakeholders and the dimensions they recommend including in the composite indicator

Stakeholder no. 1	Stakeholder no. 2	Stakeholder no. 3	Stakeholder no. 4
1. Opportunities and access to the labour market	1. Opportunities and access to the labour market	1. Opportunities and access to the labour market	1. Opportunities and access to the labour market
2. Dynamic labour market and fair working conditions	2. Dynamic labour market and fair working conditions	2. Dynamic labour market and fair working conditions	2. Dynamic labour market and fair working conditions
3. Public support/Social protection and inclusion	3. Public support/Social protection and inclusion	3. Public support/Social protection and inclusion	3. Public support/Social protection and inclusion
	4. Governance/Fairness	4. The functioning of health care	4. Governance/Fairness
			5. The functioning of health care

Table 3 Stakeholders and their desired dimensions in the composite indicator

| Dimension	Variables	Main data source
Opportunities and access to the labour market	1. Early leavers from education and training 2. Adult participation in learning 3. Tertiary educational attainment, age group 30–34 4. Income quantile share ratio (S80/S20) 5. Population at risk of poverty or social exclusion 6. Severe material deprivation rate 7. Persons living in a household with very low work intensity 8. Young people neither in employment nor in education and training, age group 15–24 9. Gender gap in part-time employment 10. Gender employment gap	European Pillar of Social Rights
Dynamic labour market and fair working conditions	1. Employment rate 2. Unemployment rate 3. Activity rate 4. Youth unemployment rate 5. Share of long-term unemployment 6. In-work at-risk-of-poverty rate	European Pillar of Social Rights
Public support/Social protection and inclusion	1. Impact of social transfers (other than pensions) on poverty reduction 2. General government expenditure on social protection 3. General government expenditure on health 4. General government expenditure on education 5. Aggregate replacement ratio for pensions 6. Children age less than 3 years in formal child care 7. Healthy life years at the age of 65 (males) 8. Healthy life years at the age of 65 (females)	European Pillar of Social Rights
Governance/Fairness	1. Control of corruption 2. Government effectiveness 3. Political stability and absence of violence 4. Regulatory quality 5. Rule of law 6. Voice and accountability	World Bank database

Table 3 (continued)

Dimension	Variables	Main data source
The functioning of health care	1. Infant mortality rate 2. Diseases of the circulatory system per 100,000 3. Malignant neoplasm per 100,000 4. Tuberculosis per 100,000 5. Hospital beds per 100,000 6. Physicians per 100,000	World Health Organization database

Table 4 OLS estimations of sigma convergence (Eq. (4)) for different stakeholders

|	α_0	α_1	R^2
Stakeholder no. 1	0.3491 (<0.001)	-0.0055 (0.0255)	0.3773
Stakeholder no. 2	0.3305 (<0.001)	-0.0026 (0.0289)	0.3645
Stakeholder no. 3	0.2313 (<0.001)	0.0132 (0.0050)	0.5363
Stakeholder no. 4	0.2585 (<0.001)	0.0086 (0.0020)	0.5972

p-values are in brackets
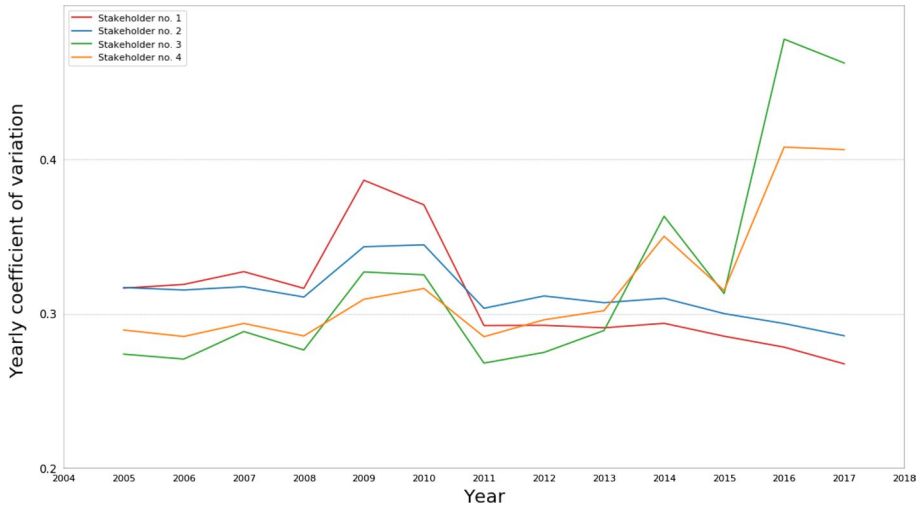
Fig. 2 The dynamic of the coefficient of variation

Table 5 Values of Kendall–tau coefficient and corresponding p-values

	τ	*P* value
Stakeholder no. 1	0.6182	<0.001
Stakeholder no. 2	0.6980	<0.001
Stakeholder no. 3	0.5556	<0.001
Stakeholder no. 4	0.5954	<0.001

therefore, it can be assumed that sigma divergence occurs. We recall that stakeholders 3 and 4 are those who included the functioning of health care dimensions.

The coefficient of variation in 2005 for different stakeholders ranges from 0.27 (stakeholder no. 3) to 0.32 (Stakeholder no. 1 and 2; see Fig. 2). Hence, different stakeholders may perceive the overall spread among the member states differently, depending on the relevance of social performance dimensions. During the financial crisis (2008–2010), divergence patterns were observed, no matter which components were used to build the CI. This proves that country differences in challenging economic conditions are growing. In 2014 and afterwards, a significant increase in the value of the coefficient of variation can be observed for stakeholders nos. 3 and 4, which is once again connected
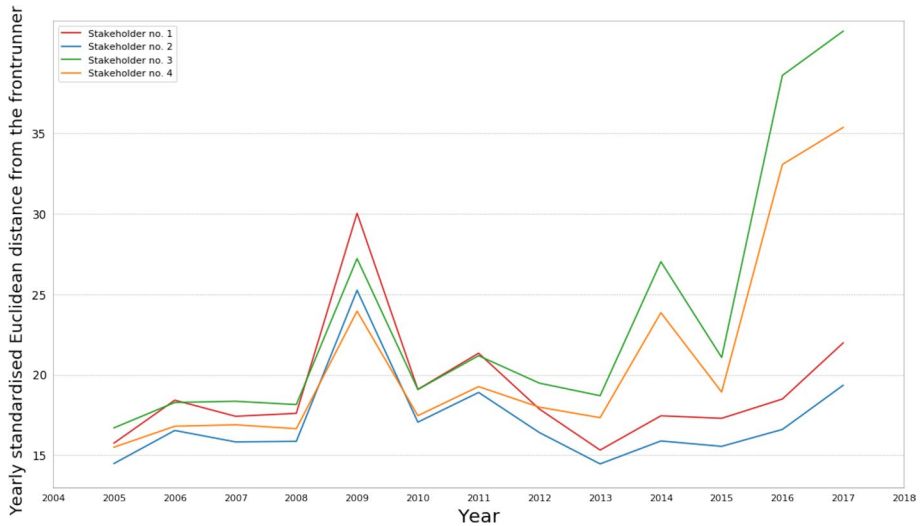
Fig. 3 The dynamic of standardised Euclidean distance from the frontrunner

Table 6 OLS estimation for Eq. (7) for 27 European Union countries

	α_0	α_1	R^2
Stakeholder no. 1	18.8463 (<0.001)	0.0335 (0.9100)	0.0012
Stakeholder no. 2	16.8488 (<0.001)	0.0342 (0.8791)	0.0021
Stakeholder no. 3	13.0741 (0.0028)	1.4868 (0.0054)	0.5206
Stakeholder no. 4	12.6301 (0.0007)	1.1961 (0.0051)	0.5255

The *p*-value is in brackets

with the unfavourable situation in the health-care system. Also, the perspective of stakeholder no. 2, including governance and fairness, indicates the occurrence of greater variation than stakeholder no. 1, who takes only social performance into consideration.

Table 5 presents the results on gamma convergence. For each stakeholder, the Kendall-tau measure is positive and statistically significant, which implies no evidence of gamma convergence. In other words, the ranking of countries is relatively stable, and no overlapping effects are observed.

Figure 3 presents the countries' aggregated distance from the best performer. As in the case of sigma convergence, a definite increase in the distance was also observed during the economic crisis. For all stakeholder, the sum of Euclidean distances was bigger at the end of the period analysed than in the initial year. The increase in distance for stakeholders nos. 1 and 2 was around 35%, whereas it was 125% for stakeholders 3 and 4. Thus Fig. 3 indicates that delta convergence did not occur over the analysed period. The findings on hypothesis testing for delta convergence, Eq. (7), are presented in the Table 6.

Analysing data presented in Table 6 we are not able to say whether the delta-convergence occurs or not. First of all, the parameter is statistically insignificant, and, more importantly, the coefficient of determination is extremely low. A comparatively well-fitted model can be obtained for stakeholders nos. 3 and 4, for whom the estimated parameter

was statistically significant and positive, indicating that delta divergence occurs. Therefore, it can be argued that adding variables related to health care affected the overall results. The significant and increasing differences among EU countries in health-care organisation, well-being and poverty, and disease prevention may have an impact on an already weakened European Union. Furthermore, they substantiate the notion of a union with many speeds.

While the European project experiences an objective moment of difficulty, it is noticeable that situation appears better when the 'official set' of convergence measure is used (stakeholder no.1) than when other sets of variables meeting different concerns (fairness, health) are included. The tension between the official set and the antagonist sets discussed here is artificial, but these situations exist in practice. One illuminating example is a controversy between the French representatives of the trade unions (and their militant experts) and the statistical office INSEE about how to measure poverty in France. Bernard Sujobert recounts this episode in the volume *Stat-activisme: comment lutter avec les nombres* (Bruno et al. 2014). Noticeable in this story – involving a new statistical measure known as BIP_{40}—is that the initial resistance of the official statisticians was successfully softened by a combination statistical activism, dialogue, and mediatic echo of the new proposed measure. What motivated the stat-activists was precisely the mismatch between what they perceived as a worsening of poverty for segments of the population and the reassuring message conveyed by the official measures of INSEE.

11 Quantitative Storytelling for the Doing Business Index

The DBI has been far from uncontroversial. Conceived as a measure of competitiveness, the methodology for its assessment has changed repeatedly over time. New component indicators have been introduced, some topics were removed, and other methodological changes were made. For instance, the 2016 version of the index introduced new component indicators, including the concept of *building quality index* (in the 'dealing with construction permits' topic), reliability of supply, and transparency of tariff index ('getting electricity' topic). Possibly the most prominent variation in the DBI assessment is the exclusion of the 'employing workers' thematic area from the 2011 edition onwards. As regards methodological changes, the component indicator 'total tax and contribution rate' (previously 'total tax rate') is aggregated in a non-linear fashion in the 'paying taxes' area from 2015. The quantity is elevated to the power of 0.8 before the min–max normalisation.

These changes in the DBI assessment resulted in variations in ranking for some countries from one edition to the next. For instance, Chile's performance ranking deteriorated over the two presidencies of Michelle Bachelet. It has been alleged that this was a deliberate manoeuvre to discredit the left-wing Bachelet against the mandates of the conservative Sebastián Piñera.

When questioned about this result, the then–World Bank chief economist, Paul Romer, objected that the index trend was not the result of any deteriorating performance by Chile; rather, it resulted from the introduction of new component indicators (Talley 2018). In a later blog post, Romer (2018b) clarified that this was not a deliberate move by the World Bank and provided an independent analysis of the data. Thus, Romer argued that the controversy was caused merely by insufficient clarity in the World Bank's communication. The loss of credibility caused by this episode, however, might be one of the reasons why Romer resigned from his duties as World Bank chief economist (Lawder and Wroughton 2018;

Fig. 4 Chile's world ranking according to its DBI score over the period 2014–2018, depending on the accounting methodology (DBI reports vs. Paul Romer's QST). The charts show four points for each method, rather than five because the figures for 2014 and 2015 overlap
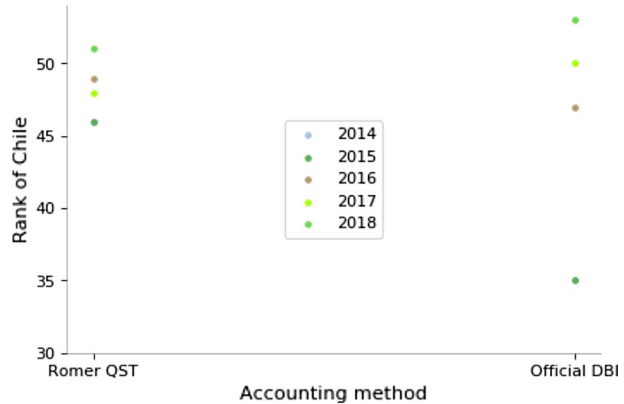
Zumbrun 2018). Romer is not an ordinary economist; in the past he demonstrated considerable intellectual openness by starting a discussion on the misuse of mathematical models in economics, coining the already mentioned neologism 'mathiness' (2015) to signify the use of mathematics to veil normative stances in growth models.

Romer (2018a) applied QST to DBI in an attempt to test the robustness of his collaborators' calculations. He did so by performing a new assessment, in which he included all the CIs available over the period of the study.

His objective was to remove the effect of introducing new variables in the computation of the DBI assessment methodology, thus producing more stable and comparable rankings over the years.

Romer implemented a *Jupyter* notebook[3] with his calculations, which has been made publicly available on his *GitHub* repository.[4] One of his main findings was that picking only the set of twenty-four component indicators available for the entire period 2014-2018 would have produced a less volatile ranking for Chile (Fig. 4). Romer ultimately encouraged his blog readers to repeat his analysis and evaluate longer trends, rather than ranking annual variations.

The prominence of the World Bank and the notoriety of the DBI led various countries to implement economic policies that would target an increase in their DBI score. The foreword to the 2019 edition states 'What gets measured gets done' and notes: 'Since its launch in 2003, Doing Business has inspired more than 3500 reforms in the 10 areas of business regulation measured by the report'.

Developing countries are particularly keen to pursue DBI-inspired reforms and to receive the approval of the World Bank, which conceived the DBI for this very purpose. Yet this did not happen without controversy: criticisms of multiple aspects of the DBI were raised in many quarters.

For opponents of so-called governance by numbers, the metric strategy pursued by international organizations has led to a deplorable erosion of the law and human condition (Supiot 2015). Berg and Cazes (2007) criticise the political perspective of the DBI in relation to the framework of labour laws, disputing the narrative whereby countries with less

---

[3] https://github.com/paulromer149/DB-Calcs/blob/master/DB-Calcs.ipynb.

[4] https://github.com/paulromer149.

protective labour laws had a higher ranking. According to the authors of the study, a country would be incentivised by the DBI to foster labour market deregulation, whereas the economic benefits on the ground of such a framework would be questionable. Other authors disputed the normative perspective of some of the component indicators on the country's employment performance. For instance, Benjamin et al. (2010) claim that the component indicators do not adequately map onto the state of labour regulation. These authors suggested integrating the assessment performed by the World Bank with other aspects, such as '*microlegislation, labour market institutions and juridical interpretation*'. The controversy around the deregulation of labour laws contributed to the removal of this thematic area from the DBI assessment from the 2011 edition. From this version onwards, the working employment area is discussed as a separate annexe, in which the dominant narrative is to seek a balance between worker protection and flexibility.

More technical criticism of the DBI comes from Høyland et al. (2012) and Pinheiro-Alves and Zambujal-Oliveira (2012). The former argue that the index completely neglects uncertainty and, with it, possible volatility in the country rankings. The latter argue that the selection of variables in the DBI may be misleading, as several of them do not contribute to variations in the score in the thematic area that they are part of. That is, they are 'silent' in the sense discussed in (Paruolo et al. 2013; Saisana et al. 2005). This may convey inadequate information to investors who are scrutinising the countries' DBI performance.

The effect of the different DBI component indicators is discussed by Schueth (2011) who analysed the performance of Georgia according to the DBI and the Global Competitiveness Index (GCI). Georgia's position in the DBI rank has been rising, moving from 100/155 in 2006 to 11/183 in 2010. In the 2019 edition of the DBI report, Georgia ranked sixth out of 190. By contrast, Georgia's GCI ranking languished: the country was reported to be 85th out of 125 in 2006, and it remained at 90/133 in 2009. Even in the most recent version of the GCI report, Georgia still ranks 66th out of 140. Schueth (2011) argues that this extreme discrepancy in performance can be ascribed to how the different sets of variables included in the indicators capture economic phenomena. This could also be seen as QST setting, whereby a Georgian policy maker who wishes to attract investment to the country would showcase the rapid improvement in Georgia's DBI ranking; an opposition leader might lament Georgia languishing GCI ranking as a proof of the ineffectiveness of the country's policies on competitiveness and business friendliness. Doing-business controversies lend themselves naturally to QST experiments.

12 Conclusions

The reflections given in this paper in the context of CIs are likely to apply to a much larger set of quantification practices. As Popp Berman and Hirschman (2018) inquire, in the age of algorithms and indicators, "what qualities are specific to rankings, or indicators, or models, or algorithms?" In particular, the misuse of metrics, statistical inference, mathematical modelling, and algorithms exhibit some common patterns (Saltelli 2019, 2020).

The solution offered here is by no means unique. For example, to address the predicament of using fragile mathematical instruments to measure soft concepts, some authors have suggested resorting to the theory of partially ordered sets. This approach offers a synthesis of multidimensional indicator systems in which the original variables are not aggregated, and the individuals being ranked (e.g. countries or regions or districts; see an example in Beycan et al. (2019); Carlsen and Bruggemann (2014, 2017)) are partially ordered

graphically. This procedure removes the design and modelling choices needed for a CI, such as weights, normalisation, and an aggregation scheme. Partial ordering is thus, by design, more robust than CIs.

We focus here on the present generation of CIs, briefly reviewing the existing debates and offering some constructive criticism. In particular, we modify the philosophy of CIs from 'analysis cum advocacy' to 'analysis with multiple storytelling'. In other words, we examine a situation in which different stakeholders agree on the importance of evidence and the need to use statistical data while disagreeing on what 'the end in sight' should be, as exemplified in the real world by the BIP_{40} story (Bruno et al. 2014).

Cohesion policy offers a convenient battleground for testing this methodology, as it is clear that multiple definitions of cohesion are possible and desirable, at a moment when a clear overall EU narrative seems elusive (Applebaum 2017).

Should measures of fairness or health be part of the portfolio of policies to be targeted by a cohesion policy? Clearly, depending on the answers to these questions, different diagnoses can be produced as to the state and the progress of cohesion.

Unsurprisingly, EU countries differ more upon a dimension which we loosely call fairness and which includes corruption, political functioning, stability and accountability, regulatory quality, and the rule of law.

EU countries become more equal when health care is included, but at the same time, this equality is eroded by the recent onset of a divergence trend.

The case of the Doing Business shows that, in practice, some sort of multiple-frame analysis—what we call quantitative storytelling—is already taking place under pressure from stakeholders. This contributed to variation in the structure of the underpinning thematic areas in the DBI over the years. The most prominent of them is the exclusion from the 2011 edition of the index of a controversial thematic area such as 'employing workers'. The primary role played by stakeholders is also reflected by the fact the DBI is simultaneously a measure and a target proposed to developing countries. For this reason, the danger of the Goodhart (or Campbell) Law—that when a measure becomes a target, it ceases to be a good measure as 'players' start adapting to it (Muller 2018), has been contrasted by scholars and stakeholders alike, signalling a mismatch between the measure and the desirability of the resulting policy.

13 Software and Data

Software and data used for the present work can be retrieved at the GitHub repository https ://github.com/Confareneoclassico/Quantitative_storytelling_making_composite_indicator

References

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for Phools*. New Jersey: Princeton University Press.

Anagnostou, A., Kallioras, D., & Kollias, Ch. (2015). Governance Convergence Among the EU28? *Social Indicator Research, 129*(1), 133–146. https://doi.org/10.1007/s11205-015-1095-2.

Aparicio, J., & Kapelko, M. (2019). Enhancing the measurement of composite indicators of corporate social performance. *Social Indicators Research, 144*(2), 807–826. https://doi.org/10.1007/s11205-018-02052-1.

Applebaum, A. (2017). A New European Narrative, New York Review of Books, 12 October.

Araújo, L., Saltelli, A., & Schnepf, S. V. (2017). Do PISA data justify PISA-based education policy? *International Journal of Comparative Education and Development, 19*(1), 1–17. https://doi.org/10.1108/IJCED-12-2016-0023.

Baddeley, M. (2006). Convergence or divergence? The impact of globalisation on growth and inequality in less developed countries. *International Review of Applied Econometrics, 20*(3), 391–410. https://doi.org/10.1080/02692170600736250.

Balcerzak, A. P., & Pietrzak, M. B. (2017a). Digital economy in Visegrad countries. Multiple-criteria decision analysis at regional level in the years 2012 and 2015. *Journal of Competitiveness, 9*(2), 5–18. https://doi.org/10.7441/joc.2017.02.01.

Balcerzak, A. P., & Pietrzak, M. B. (2017b). Human development and quality of institutions in highly developed countries. In M. H. Bilgin, H. Danis, E. Demir, & U. Can (Eds.), *Financial environment and business development. Proceedings of the 16th Eurasia Business and Economics Society* (pp. 231–241). Berlin: Springer. https://doi.org/10.1007/978-3-319-39919-5.

Balcerzak, A. P., & Pietrzak, M. B. (2017c). Sustainable Development in the European Union in the years 2004-2013. In M. H. Bilgin and H. Danis, E. Demir, & U. Can (Eds.). *Regional Studies on Economic Growth, Financial Economics and Management. Proceedings of the 19th Eurasia Business and Economics Society*. Vol. 7, Springer, Berlin pp. 193–213. https://doi.org/10.1007/978-3-319-54112-9_12.

Balcerzak, A. P. & Rogalska, E. (2016). Non-Keynesian Effects of Fiscal Consolidations in Central Europe in the Years 2000-2013. In M. H. Bilgin & H. Danis (Eds.), *Entrepreneurship, Business and Economics. Proceedings of the 15th Eurasia Business and Economics Society*, Vol. 2 (pp. 271–282). Berlin: Springer. https://doi.org/10.1007/978-3-319-27573-4_18.

Barrington-Leigh, C., & Escande, A. (2018). Measuring progress and well-being: A comparative review of indicators. *Social Indicators Research, 135*(3), 893–925. https://doi.org/10.1007/s11205-016-1505-0.

Barro, R. J., & Sala-i-Martin, X. (1999). *Economic Growth*. Cambridge: MIT Press.

Bartkowiak-Bakun, N. (2017). The Diversity of socioeconomic development of rural areas in poland in the western borderland and the problem of post-state farm localities. *Oeconomia Copernicana, 8*(3), 417–431. https://doi.org/10.24136/oc.v8i3.26.

Beck, P.U. (1992 [1986]). Risk society: Towards a new modernity. Sage Publications, Thousand Oaks.

Becker, W., Paruolo, P., Saisana, M., & Saltelli, A. (2017). Weights and importance in composite indicators: Mind the gap, In R. Ghanem, D. Higdon, H. Owhadi (Eds.), *Handbook of Uncertainty Quantification*, pp. 1187–1216, Berlin: Springer. https://doi.org/10.1007/978-3-319-12385-1.

Benjamin, P., Bhorat, H., & Cheadle, H. (2010). The cost of "doing business" and labour regulation: The case of South Africa. *International Labour Review, 149*(1), 73–91. https://doi.org/10.1111/j.1564-913X.2010.00076.x.

Berg, J., & Cazes, S. (2007). *The doing business indicators: Measurement issues and political implications*. Geneva: International Labour Office.

Beycan, T., Vani, B. P., & Bruggemann, R. (2019). Ranking Karnataka districts by the multidimensional poverty index (MPI) and by applying simple elements of partial order theory. *Social Indicators Research, 143*, 173–200. https://doi.org/10.1007/s11205-018-1966-4.

Boden, R., & Epstein, D. (2006). Managing the research imagination? *Globalisation and Research in Higher Education, Globalisation, Societies and Education, 4*(2), 223–236. https://doi.org/10.1080/14767720600752619.

Boulanger, P.-M. (2014). Elements for a comprehensive assessment of public indicators, Report EUR 26921 EN. Retrieved January 20, 2020 from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC92162/lbna26921enn.pdf.

Boulanger, P.-M. (2018). A systems-theoretical perspective on sustainable development and indicators. In S. Bell & S. Morse (Eds.), *The Routledge handbook of sustainability indicators*. London: Taylor & Francis. https://doi.org/10.4324/9781315561103.

Boyle, G. & McCarthy, T. (1997). Simple measures of convergence in per capita GDP: A Note on Some Further International Evidence, Economics, Finance and Accounting Department Working Paper

Series n751197, Department of Economics, Finance and Accounting, National University of Ireland - Maynooth.

Bruno, I., Didier, E. & Prévieux, J. (2014). Stat-activisme. Comment lutter avec des nombres. Paris: Zones, La Découverte.

Capecchi, S., & Simone, R. (2019). A Proposal for a model-based composite indicator: Experience on perceived discrimination in Europe. *Social Indicators Research, 141*(1), 95–110. https://doi.org/10.1007/s11205-018-1848-9.

Carlsen, L., & Bruggemann, R. (2014). The 'Failed State Index' offers more than just a simple ranking. *Social Indicators Research, 115,* 525–530. https://doi.org/10.1007/s11205-012-9999-6.

Carlsen, L., & Bruggemann, R. (2017). Fragile state index: Trends and developments. A partial order data analysis. *Social Indicators Research, 133,* 1–14. https://doi.org/10.1007/s11205-016-1353-y.

Carrozza, C. (2014). Democratizing expertise and environmental governance: Different approaches to the politics of science and their relevance for policy analysis. *Journal of Environmental Policy & Planning, 17,* 108–126. https://doi.org/10.1080/1523908X.2014.914894.

Chaaban, J., Irani, A., & Khoury, A. (2016). The composite global well-being index (CGWBI): A new multi-dimensional measure of human development. *Social Indicators Research, 129*(1), 465–487. https://doi.org/10.1007/s11205-015-1112-5.

Cheba, K., & Szopik-Depczyńska, K. (2017). Multidimensional comparative analysis of the competitive capacity of the European Union countries and geographical regions. *Oeconomia Copernicana, 8*(4), 487–504. https://doi.org/10.24136/oc.v8i4.30.

Commission, European. (2014). *Employment and social developments in Europe 2014*. Luxemburg: Publication Office of the European Union.

Commission, European. (2016). *Employment and social developments in Europe: Annual review 2016*. Luxemburg: Publication Office of the European Union.

Cosci, S., & Mirra, L. (2017). A spatial analysis of growth and convergence in Italian provinces: the role of infrastructure. *Regional Studies*. https://doi.org/10.1080/00343404.2017.1334117.

Dewey, J. (1938). *The public and its problems*. Redditch: Read Book Ltd. Edition.

Dinis, G., Costa, C., & Pacheco, O. (2019). Composite indicator for measuring the world interest by Portugal's tourism. *Journal of Spatial and Organizational Dynamics, 1*(7), 9–52.

Dutta S., Lanvin B., & Wunsch-Vincent S. (Eds.) (2018). Global innovation index 2018. Energizing the world with innovation 11th edition. Cornell University, INSEAD, and WIP. Ithaca, Fontainebleau, and Geneva.

Eurofound. (2018). *Upward convergence in the EU: Concepts, measurements and indicators*. Luxembourg: Publications Office of the European Union.

European Commission (2015). Speaking points by Employment, Social Affairs and Labour Mobility Commissioner Marianne Thyssen at the press conference to launch the 2016 European Semester, speech, Brussels, 26 November 2015.

European Commission (2017a). Assessment of the 2017 convergence programme for Czech Republic. European Commission. Retrieved October 1, 2017, from https://ec.europa.eu/info/sites/info/files/03_cz_cp_assessment.pdf.

European Commission (2017b). Assessment of the 2017 convergence programme for Hungary. European Commission. Retrieved October 1, 2017, from https://ec.europa.eu/info/sites/info/files/17_hu_cp_assessment.pdf.

European Commission (2017c). Assessment of the 2017 convergence programme for Poland. European Commission. Retrieved October 1, 2017, from https://ec.europa.eu/info/sites/info/files/21_pl_cp_assessment.pdf.

European Commission (2017d). Assessment of the 2017 convergence programme for The United Kingdom. European Commission. Retrieved October 1, 2017, from https://ec.europa.eu/info/sites/info/files/28_uk_cp_assessment.pdf.

European Commission (2017e). Commission Recommendation of 26.4.2017 on the European Pillar of Social Rights. Retrieved October 1, 2017, https://ec.europa.eu/commission/publications/commission-recommendation-establishing-european-pillar-social-rights_pl.

Flood, A. (2016). 'Post-truth' named word of the year by Oxford Dictionaries. *The Guardian*, 15.

Funtowicz, S. O., & Ravetz, J. R. (1990). *Uncertainty and quality in science for policy*. Dordrecht: Kluwer Academic.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the Post-Normal Age. *Futures, 25,* 739–755. https://doi.org/10.1016/0016-3287(93)90022-L.

Funtowicz, S. O., & Ravetz, J. R. (1994). The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics, 10,* 197–207.

Furkowa, A., & Chocholata, M. (2017). Interregional R and D spillovers and regional convergence: A spatial econometric evidence from the EU regions. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(1), 9–24. https://doi.org/10.24136/eq.v12i1.1.

Giampietro, M., Aspinall, R. J., Ramos-Martin, J., & Bukkens, S. G. F. (2014). *Resource accounting for sustainability assessment: The nexus between energy, food, water and land use*. Milton Park: Taylor & Francis.

Gluckman, P. (2017). Can science and science advice be effective bastions against the post-truth dynamic? Speech delivered at University College London. Retreived 1 October 8 from, www.pmcsa.org.nz/wp-content/uploads/17-10-18-UCL-speech.pdf.

Greyling, T., & Tregenna, F. (2017). Construction and analysis of a composite quality of life index for a region of South Africa. *Social Indicators Research, 131*(3), 887–930. https://doi.org/10.1007/s11205-016-1294-5.

Hausken, K., & Moxnes, J. F. (2019). Innovation, development and national indices. *Social Indicators Research, 141*(3), 1165–1188. https://doi.org/10.1007/s11205-018-1873-8.

Heichel, S., Pape, J., & Sommerer, T. (2005). Is there convergence in convergence research? An overview of empirical studies on policy convergence. *Journal of European Public Policy, 12*(5), 817–840. https://doi.org/10.1080/13501760500161431.

Holzinger, K., Knill, C., & Sommerer, T. (2011). Is there convergence of national environmental policies? An analysis of policy outputs in 24 OECD Countries. *Environmental Politics, 20*(1), 20–41. https://doi.org/10.1080/09644016.2011.538163.

Horridge, M., & Rokicki, B. (2017). The impact of European Union Accession on regional income convergence within the Visegrad countries. *Regional Studies, 52*(4), 1–13. https://doi.org/10.1080/00343404.2017.1333593.

Høyland, B., Moene, K., & Willumsen, F. (2012). The tyranny of international index rankings. *Journal of Development Economics, 97*(1), 1–14. https://doi.org/10.1016/j.jdeveco.2011.01.007.

James, G., Witten, G., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning*. Berlin: Springer.

Kahan, D. M., Wittlin, M., Peters, E., Slovic, P., Ouellette, L.L., Braman, D. & Mandel, G. N. (2011). The tragedy of the risk-perception commons: Culture conflict, rationality conflict, and climate change. Temple University Legal Studies Research Paper No. 2011-2026, DOI: 10.2139/ssrn.1871503.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika, 30*(1–2), 81–93. https://doi.org/10.1093/biomet/30.1-2.81.

Kong, J., Philips, C. B., & Sul, D. (2019). Weak sigma-convergence: Theory and applications. *Journal of Econometrics, 209*(2), 185–207. https://doi.org/10.1016/j.jeconom.2018.12.022.

Kruk, H., & Waśniewska, A. (2017). Application of the Perkal method for assessing competitiveness of the countries of Central and Eastern Europe. *Oeconomia Copernicana, 8*(3), 337–352. https://doi.org/10.24136/oc.v8i3.21.

Kuc, M. (2017). Social convergence in Nordic countries at regional level. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(1), 25–41. https://doi.org/10.24136/eq.v12i1.2.

Lakoff, G. (2004). *Don't think of an elephant: Know your values and frame the debate*. White River Junction: Chelsea Green Publishing.

Lakoff, G. (2010). Why it matters how we frame the environment. *Environmental Communication: A Journal of Nature and Culture, 4*(1), 70–81. https://doi.org/10.1080/17524030903529749.

Lawder, D., & Wroughton, L. (2018). World Bank economist Paul Romer quits after Chile comments. *Reuters*. Retrieved 20 Jan, 2020 from https://www.reuters.com/article/us-worldbank-economist-romer-idUSKBN1FD38Y.

Luzzati, T., & Gucciardi, G. (2015). A non-simplistic approach to composite indicators and rankings: An illustration by comparing the sustainability of the EU countries. *Ecological Economics, 113*, 25–38. https://doi.org/10.1016/j.ecolecon.2015.02.018.

Małkowska, A., & Głuszak, M. (2016). Pro-Investment local policies in the area of real estate economics: Similarities and differences in the strategies used by communes. *Oeconomia Copernicana, 7*(2), 269–283. https://doi.org/10.12775/OeC.2016.016.

Mann, J., & Shideler, D. (2015). Measuring Schumpeterian Activity Using a Composite Indicator. *Journal of Entrepreneurship and Public Policy, 4*(1), 57–84. https://doi.org/10.1108/JEPP-07-2013-0029.

Marozzi, M. (2015). Measuring trust in European Public Institutions. *Social Indicators Research, 123*(3), 879–895. https://doi.org/10.1007/s11205-014-0765-9.

Marris, C., Wynne, B., Simmons, P., & Weldon, S. (2001). *Public Perceptions of Agricultural Biotechnologies in Europe*. Final Report of the PABE research project funded by the Commission of European Communities Contract number: FAIR CT98-3844 (DG12 - SSMI).

Mazziotta, M., & Pareto, A. (2016). On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research, 127*(3), 983–1003. https://doi.org/10.1007/s11205-015-0998-2.

Merton, R. K. (1973[1942]). The normative structure of science. In R. K. Merton (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267–280). Chicago: University of Chicago Press.

Meyer, H.-D. & Zahedi, K. (2014) An open letter: To Andreas Schleicher, OECD, Paris, Global Policy Institute, 5 May, and Guardian, 6 May.

Michener, G. (2015). Policy Evaluation via composite indexes: qualitative lessons from International Transparency Policy Indexes. *World Development, 74,* 184–196. https://doi.org/10.1016/j.worlddev.2015.04.016.

Miro, D. R., & Piffaut, P. V. (2019). Financial quality index (ICF). *Cuadernos de Economia: Spain, 119*(42), 189–206. https://doi.org/10.32826/cude.v42i119.170.

Mirowski, P. (2013). Never Let a Serious Crisis Go Wasted. Verso Books.

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton: Princeton University Press.

OECD-JRC (2008). Handbook on constructing composite indicators: Methodology and user guide, OECD Statistics working paper JT00188147, STD/DOC(2005)3.

Paruolo, P., Saisana, A., & Saltelli, A. (2013). Ratings and rankings: Voodoo or science? *Journal Royal Statistical Society A, 176*(3), 609–634. https://doi.org/10.1111/j.1467-985X.2012.01059.x.

Pawełek, B. (2008). *Normalisation of variables methods in comparative research on complex economic phenomena*. Cracow: Zeszyty Naukowe Uniwersytet Ekonomicznego w Krakowie.

Peiro-Palomino, J., & Picazo-Tadeo, A. J. (2018). OECD: One or many? ranking countries with a composite well-being indicator. *Social Indicators Research, 139*(3), 847–869. https://doi.org/10.1007/s11205-017-1747-5.

Pielke, R. A., Jr. (2007). *The honest broker*. Cambridge: Cambridge University Press.

Pietrzak, M. B., & Balcerzak, A. P. (2017). A regional scale analysis of economic convergence in Poland in the Years 2004–2012. In M. H. Bilgin ה H. Danis, E. Demir, & U. Can (Eds.), *Regional studies on economic growth, financial economics and management. Proceedings of the 19th Eurasia Business and Economics Society*. Vol. 7 (pp. 257–268). Berlin: Springer. https://doi.org/10.1007/978-3-319-54112-9_16.

Pinheiro-Alves, R., & Zambujal-Oliveira, J. (2012). the ease of doing business index as a tool for investment location decisions. *Economics Letters, 117*(1), 66–70. https://doi.org/10.1016/j.econlet.2012.04.026.

Popp Berman, E., & Hirschman, D. (2018). The sociology of quantification: Where are we now? *Contemporary Sociology. A Journal of Reviews., 47*(3), 257–266. https://doi.org/10.1177/0094306118767649.

Próchniak, M., & Witkowski, B. (2016). On the use of panel stationarity tests in convergence analysis: empirical evidence for the EU Countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 11*(1), 77–96. https://doi.org/10.12775/equil.2016.004.

Ravallion, M. (2010). Mashup indices of development, Policy Research Working Paper 5432, World Bank Development Research Group. Retrieved 20 Jan, 2020 form http://documents.worldbank.org/curated/en/454791468329342000/pdf/WPS5432.pdf.

Ravetz, J. R. (1987). Usable knowledge, Usable Ignorance. *Knowledge, 9*(1), 87–116. https://doi.org/10.1177/107554708700900104.

Rayner, S. (2012). Uncomfortable knowledge: The social construction of ignorance in science and environmental policy discourses. *Economy and Society, 41*(1), 107–125. https://doi.org/10.1080/03085147.2011.637335.

Reinert, E.S. (2008). How Rich Countries Got Rich… and Why Poor Countries Stay Poor, Public Affairs.

Renner, A., & Giampietro, M. (2019). Socio-technical discourses of European electricity decarbonization: Contesting narrative credibility and legitimacy with quantitative story-telling. *Energy Research & Social Science*. https://doi.org/10.1016/j.erss.2019.101279.

Rethinking economics. (2017). 33 Theses for an Economics Reformation. Retrieved 20 Jan, 2020 from http://www.rethinkeconomics.org/journal/time-economics-reformation/.

Rogalska, E. (2018). Multiple-criteria analysis of regional entrepreneurship conditions in Poland. *Equilibrium Quarterly Journal of Economics and Economic Policy, 13*(4), 707–723. https://doi.org/10.24136/eq.2018.034.

Romer, P. M. (2015). Mathiness in the theory of economic growth. *American Economic Review, 105,* 89–93. https://doi.org/10.1257/aer.p20151066.

Romer, P. (2018a)., Doing Business. Retrieved 23 November 2019 from, https://paulromer.net/doing-business/.

Romer, P. (2018b). Comments about the doing business report. Retrieved 23 November 2019 from, https://paulromer.net/my-unclear-comments-about-the-doing-business-report/ .

Saisana, M., D'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research Policy, 40,* 165–177. https://doi.org/10.1016/j.respol.2010.09.003.

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society, A, 168*(2), 307–323. https://doi.org/10.1111/j.1467-985X.2005.00350.x.

Saltelli, A. (2017). International PISA tests show how evidence-based policy can go wrong. The Conversation, June 12.

Saltelli, A. (2018). Why science's crisis should not become a political battling ground. *Futures, 104,* 85–90. https://doi.org/10.1016/j.futures.2018.07.006.

Saltelli, A. (2019). Statistical versus mathematical modelling: a short comment. *Nature Communications, 10,* 1–3.

Saltelli, A. (2020). Ethics of quantification or quantification of ethics. *Futures.* https://doi.org/10.1016/j.futures.2019.102509.

Saltelli, A., & Funtowicz, S. O. (2017). To Tackle the Post-Truth World, Science Must Reform Itself, TheConversation.com, 27 January.

Saltelli, A., & Giampietro, M. (2017). What Is wrong with evidence based policy, and how can it be improved? *Futures, 91,* 62–71. https://doi.org/10.1016/j.futures.2016.11.012.

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J.P. & Funtowicz, S. O. (2013). What do i make of your latinorum? Sensitivity auditing of mathematical modelling. International Journal of Foresight and Innovation Policy, 9 (2/3/4), 213–234. https://doi.org/10.1504/ijfip.2013.058610.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). *Global sensitivity analysis: The Primer.* Hoboken: Wiley.

Scheurer, L., & Haase, A. (2017). Diversity and Social Cohesion in European Cities: Making Sense of Today's European Union-Urban Nexus within Cohesion Policy. *European Urban and Regional Studies.* https://doi.org/10.1177/0969776417736099.

Schueth, S. (2011). Assembling International Competitiveness: The Republic of Georgia, USAID, and the Doing Business Project: ECONOMIC GEOGRAPHY. *Economic Geography, 87*(1), 51–77. https://doi.org/10.1111/j.1944-8287.2010.01103.x.

Schwab, K. (Ed.). (2019). *The global competitiveness report 2019.* Geneva: World Economic Forum.

Semenenko, I., Halhash, R., & Sieriebriak, K. (2019). Sustainable development of regions in Ukraine: before and after the beginning of the conflict. *Equilibrium Quarterly Journal of Economics and Economic Policy, 14*(2), 317–339. https://doi.org/10.24136/eq.2019.015.

Stanickova, M. (2017). Can the implementation of the Europe 2020 strategy goals be efficient? The challenge for achieving social equality in the European Union. *Equilibrium: Quarterly Journal of Economics and Economic Policy, 12*(3), 383–398. https://doi.org/10.24136/eq.v12i3.20.

Stiglitz, J., Sen, A., & Fitoussi, J.-P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Retrieved June 2017 from, http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report.

Strassheim, H., & Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence & Policy: A Journal of Research, Debate and Practice, 10*(2), 259–277. https://doi.org/10.1332/174426514X13990433991320.

Supiot, A. (2015). *Governance by numbers: The making of a legal model of allegiance.* Oxford: Oxford University Press.

Talley, I. (2018). World Bank unfairly influenced its own competitiveness rankings, The Wall Street Journal. Retrieved 20 Jan, 2020 from https://www.wsj.com/articles/world-bank-unfairly-influenced-its-own-competitiveness-rankings-1515797620.

Tsionas, E. G. (2002). Another look at regional convergence in Greece. *Regional Studies, 36*(6), 603–609. https://doi.org/10.1080/00343400220146759.

van der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP system. *Risk Analysis, 25*(2), 481–492. https://doi.org/10.1111/j.1539-6924.2005.00604.x.

van der Sluijs, J., Petersen, A. C., Janssen, P. H. M., Risbey, J. S., & Ravetz, J. R. (2008). Exploring the quality of evidence for complex and contested policy decisions. *Environmental Research Letters, 3*(2), 024002. https://doi.org/10.1088/1748-9326/3/2/024008.

Wilsdon, J. (2016). *The metric tide: The independent review of the role of metrics in research assessment and management.* Thousand Oaks: Sage Publications.

Woessmann, L. (2014). The Economic Case for Education. Institute and University of Munich. European Expert Network on Economics of Education (EENEE). EENEE Analytical Report 20.

World Bank (2019). Doing Business. (2019). *Training for reform (2019).* Washington: International Bank for Reconstruction and Development.

Żelazny, R., & Pietrucha, J. (2017). Measuring innovation and institution: The creative economy index. *Equilibrium: Quarterly Journal of Economics and Economic Policy, 12*(1), 43–62. https://doi.org/10.24136/eq.v12i1.3.

Zumbrun, J. (2018). World Bank Chief Economist Paul Romer Resigns. Wall Street Journal. Retrieved 20 Jan, 2020 from https://www.wsj.com/articles/world-bank-chief-economist-paul-romer-resigns-1516823370.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.