Routledge
Taylor & Francis Group

Innovation: The European Journal of Social Science Research

ISSN: 1351-1610 (Print) 1469-8412 (Online) Journal homepage: www.tandfonline.com/journals/ciej20

(De)politicizing good governance: the World Bank Institute, the OECD and the politics of governance indicators

Tero Erkkilä & Ossi Piironen

To cite this article: Tero Erkkilä & Ossi Piironen (2014) (De)politicizing good governance: the World Bank Institute, the OECD and the politics of governance indicators, Innovation: The European Journal of Social Science Research, 27:4, 344-360, DOI: 10.1080/13511610.2013.850020

To link to this article: https://doi.org/10.1080/13511610.2013.850020

Innovation: The European Journal of Social Science Research, 2014
Vol. 27, No. 4, 344–360, http://dx.doi.org/10.1080/13511610.2013.850020

Routledge
Taylor & Francis Group

(De)politicizing good governance: the World Bank Institute, the OECD and the politics of governance indicators

Tero Erkkilä* and Ossi Piironen

Department of Political and Economic Studies, University of Helsinki, Helsinki, Finland

(*Received 21 May 2013; final version received 17 July 2013*)

Since the early 1990s there has been a surge in international efforts to calculate the comparative performance of states in terms of various characteristics of governance. In this article we show how numerical objectification of social phenomena can function to depoliticize potentially political issues. As a case of example we examine the evolving field of measuring good governance through analyzing the documentation of the World Bank Institutes established Worldwide Governance Indicators and its recent contender, the OECD project "Government at a Glance", which argues to provide an alternative to the existing rankings. Although we observe certain methodological discontinuities in measurement practices of the OECD, these have hardly been serious enough to activate its potential in repoliticisizing the issue of "good governance". Moreover, the work of OECD further strengthens and legitimates the epistemic expert authority of global index producers.

Keywords: measuring; ranking; indicators; governance; expertise; World Bank; OECD

Introduction

Policy-making today is represented as a business that demands extensive technical understanding. Consequently, public policies are increasingly expected to be based on firm evidence and continuous evaluation. Since the early 1990s there has been a surge in international efforts to calculate the comparative performance of states in terms of various characteristics of governance (Arndt and Oman 2006). These aspects – accountability, effectiveness, legality, to mention only a few – are often presented as components of the more general discourse of good governance, which is seen as a cornerstone for a thriving economy and political stability. There have been manifold motivations for demanding and producing numeric expert-knowledge on governance, but the general idea has been to offer a tool for developing political and administrative processes in individual countries. Of course, the line between expertise and politics – knowledge production and policy processes, causal research and impact assessment, value neutrality and policy-relevancy – is extremely blurred.

Our contribution here is to show how production of policy-relevant numeric data serves to expand this ambiguity. In examining numerical data on good governance produced internationally we exemplify how numerical objectification of social phenomena functions often to *depoliticize* potentially political issues in naturalizing certain interpretations of reality at the expense of alternative visions.[1] To be sure, numbers do

*Corresponding author. Email: tero.erkkila@helsinki.fi

have also the potential to (*re)politicize* social phenomena, as, for example, the meager success of European universities in global university rankings have helped to bolster the status of science and innovation policy on the political agenda of the European Union and its member countries.

Also good governance indicators have caused political controversies. In 2007, representatives of countries such as China, Russia, Mexico and Argentina voiced their concerns about the World Bank Institute's Worldwide Governance Indicators (WGI) project, arguing that the Bank should not engage in such activity. One particular concern was Chinese low ranking in the voice and accountability component of the WGI (Guha and McGregor 2007). Recently, the Chinese criticism of the World Bank's Doing Business Report sparked a debate on the report's "liberal economic prescriptions" and use of rankings (Harding 2013; Zadek 2013; Vogl 2013). The above accounts show the indicators' potential for political conflict on the level of measurements.

Importantly, however, there is more to (de)politicization than public visibility of a political question. According to Palonen (2007, 41), depoliticization as a "movement towards closing a horizon" – as datasets may fix the parameters of the phenomena they seek to depict – can be actively and visibly effected and thus seen as "an extremely intensive form of politicking". Indeed, we have argued elsewhere that the short history of measuring governance has been one of depoliticization: more than anything else, indicators have strengthened the dominant economist visions of good governance by making descriptive data appear neutral and apolitical (Erkkilä and Piironen 2009). In this article we look at whether the argument still holds by looking at a new type of dataset that claims to break the universalizing logic of older measurements: can the new measurements repoliticize the debate on good governance?

The first broad comparisons of governance performance in 1990s were conducted by international organizations relatively free from governmental steering. At first, there was a trend toward creating composite indicators and reporting the results in the form of simplistic league table format – all countries were rendered commensurable in terms of pregiven performance standards. Recently, however, new players have entered the ranking arena, most notably intergovernmental actors such as the Organisation for Economic Co-operation and Development (OECD), as the critique of previous rankings has lent increased credibility to datasets with more openness, political acceptability and less inclination to create rankings through composite indicators (Knack, Kugler, and Manning 2003; Arndt and Oman 2006). There is, however, no guarantee that such a competition between datasets would challenge the underlying ideas of measuring governance.

Drawing primarily on their official publications, we first analyze the efforts of the constructors of the most prominent governance measurement, the WGI, to tone down their initially uncritical – albeit evolving – way of describing the potential uses and abuses of their data, after which we go on to examine a more recent attempt to quantify good governance, the OECD project on "Government at a Glance" launched in 2005 "with the aim of providing governments with high quality comparative information on the public sector". In January 2012, we conducted six background interviews in Washington, DC with experts representing the World Bank, World Bank Institute and Millennium Challenge Corporation. Based on the analysis we conclude that, although we can observe certain *methodological* discontinuities in measurement practices, these have hardly been serious enough to truly challenge the underlying ideas behind the current understanding of good governance. The potential for repoliticization of good governance is still in waiting.

Furthermore, based on our observations, we enlarge the scope of our primary analysis concerning (re)politicization to allow a hypothesis about the epistemic authority of legitimated experts.[2] It is still another facet of depoliticization, since it sets controls on who possesses and is able to provide legitimate information about a defined subject matter such as good governance. The fact that numbers must be fought with other numbers may well indicate that the practice of quantification serves merely as a criterion for inclusion and exclusion of expertise and relevant knowledge (Gieryn 1983): one's credibility as an expert authority rests largely on the capability to produce numerical data and/or willingness to accept the validity of the method. In effect, the normative (value-based) notion of good governance comes to be looked upon as a fixed set of numerical parameters differentiating universal good from universal bad.

Despite the OECD framework's potential for bringing a certain amount of *politics* back into the governance debate, disputes hitherto have been mild and methodological, not substantial, and the contestation has served more to reinforce the present boundaries of expertise than to challenge them. If some of the new datasets are based on openly normative conceptualizations and incorporate non-numerical information, the production of governance knowledge is still filtered through a relatively coherent community of experts.

It is not our intention to devaluate quantification as a method for producing knowledge as such. While we consider it often problematic to simplify and commensurate on value-based notions, we understand not only that rankings and indexes are here to stay, but also that they can be put into useful service. Although our starting point is "epistemic relativity" – a view "that all beliefs are socially produced, so that all knowledge is transient, and neither truth-values nor criteria of rationality exist outside historical time" (Bhaskar 1998, 62–63) – we advocate neither normative relativism – all forms of governance are as good as any other – nor scientific relativism – all measures of governance are as good as any other. Nevertheless, it is not our task to provide reconciliation or to specify validity criteria for scientific knowledge in general. In the present context we simply want to illustrate datasets' potential to depoliticize important issues, primarily, by fixing the parameters of discussion, and secondarily, by fixing the parameters of legitimate authorities.

Our statements here are somewhat indifferent to strict juxtaposition between postpositivist foundationalism and poststructuralist constructivism, because our interest lies in measurements as implications of the value-based concept of *good governance* – not as providers of empirical descriptions of real instances of governance. Nevertheless, in a broad agreement with Foucauldian-inspired intersubjectivism, we do believe that shared ideas, vocabularies and discourses are of importance in making events, institutions and practices meaningful, and that numeric objectifications are powerful devices in this process regardless of whether an external reference was posited or not. If political decision-making and administrative steering are to be based on numeric objectifications, it is recommended that their premises are visible, accessible and open for public deliberation.

Politics, numbers, and expert knowledge

The trend for governing through numbers was strengthened in 1999 when the Blair government published a White Paper, *Modernizing Government*, which institutionalized the discourse and practice of *evidence-based policy-making*, first in the UK and soon after in foreign and international arenas. Since 2001 the European Commission has been

committed to an evidence-based impact assessment of all major legislative proposals (European Commission 2001; Lee and Kirkpatrick 2006). The most sought-after type of information for purposes of policy planning, monitoring and evaluation is quantitative time-series data, which often allow international comparison and benchmarking (Arndt and Oman 2008). Comparative knowledge is now commonly recognized as a useful tool for improving policy outcomes and a resource for public communication – whether for purposes of justifying reform, collecting the plaudits or pointing to a scapegoat.

It is assumed that international policy coordination – for mitigating problems that individual countries are not able to deal with alone – has amplified the demand for internationally oriented knowledge (Haas 1992) and is helping to carve out political spaces for multilevel governance (see Hooghe and Marks 2003). Within the EU, for example, the increased use of the Open Method of Co-ordination – brought about especially to coordinate employment and social, education and culture policies – has applied indicator data in its benchmark type of steering (Lelieveldt and Princen 2011). In a similar fashion, the enormous databases of various international organizations such as the World Bank and OECD can be justified as vital tools for international management. However, if international cooperation and coordination play a role in the increased demand for indicator data, so does international competition. Globalization is often identified as a significant cause of accelerated competition between various *economic* entities. With the presumed competition comes, in turn, the need to enhance economic performance, acquire best practices and – simply – to give an appearance of being successful. Inherent in this thinking is the need for comparison; to benchmark one's position, quality, quantity and performance in relation to others.

Numbers allow those who make or possess the figures to "grasp" abstract phenomena and see their scope and limits (Miller and Rose 1990). Miller and Rose (1990, 168) call these technologies of government "ways of entering reality into calculations of government by means of inscription techniques and rendering it amenable to interventions". In other words, statistics play a role in defining the scope of governing. What we make statistics out of, how and why is a highly political choice, since this constructs abstract entities upon which we can politicize, debate and make decisions (Porter 1995). Producing a numerical description of certain social phenomena objectifies it (Desrosières 1998), entailing – or hiding – a potential struggle for definition.

Thus, while the numerical objectifications of governance can be used for politicking within a polity, and although they can potentially function as instruments for politicization, they are also powerful instruments of *depoliticization*. According to Palonen (2007, 41) depoliticization can be either active, "a movement towards closing a [political] horizon", or it can be passive, "based on exhaustion or on a diminishing interest in the horizon of politicking". Numbers often have the tendency to make the issues they describe appear as facts (Desrosières 1998), which is also why they are so often referred to in contested matters (Rose 1999, Chapter 6): the figures themselves often appear innocent in terms of politics, and they seem to belong to a sphere of expert knowledge where the rules of scientific verification apply (Porter 1995; Hummel 2006).

Numbers such as governance indices make new concepts of governing operable, bringing new issues to the fore and allowing them to become desired goals, guidelines or deficiencies of governing, but they also narrow the room for debate (Miller and Rose 1990; Hopwood and Miller 1994; Rose 1999). Indeed, it is our contention that more than anything else numerical objectifications of governance are instruments of constructing and maintaining a normative understanding of *good governance*. The depoliticization of

governance occurs when a set of indicators giving normative content to an attribute succeed in representing it in a neutral and non-political guise.

Statistics are increasingly being produced in the international context for the purposes of supranational governance (compare Mahon and McBride 2009). Such knowledge is not new but a strengthened manifestation of global governmentality; the production of knowledge about entities thus made governable (Larner and Walters 2004). It is easy to concur with Oded Löwenheim (2008) that an important function of statistical comparisons is to reproduce hierarchical structures of international system not only by subjecting states to (self-)evaluation of their politico-administrative conduct by standards set in the industrial West, but also by constructing a representation of states as ethical actors capable of enacting responsible policies. As such, unit-level comparisons help to sever the discursive linkages between powerful international actors and a wide variety of political, social and economic problems, which come to be treated as "domestic", and responsibility for their alleviation is foisted on national governments.

While we are not suggesting that the use of numerical techniques is merely aimed at the promotion of private interests, we nevertheless believe that there is reason to put more emphasis on tactical considerations when looking at the production of governance data. A need for new actors to establish themselves as experts on the governance field seems to be an important supply-side incentive, as many actors wish to engage in producing quantitative data (Erkkilä and Kauppi 2010; Espeland and Stevens 2008; Arndt and Oman 2008, 10–11). Governance indices form a fast-evolving field of expert knowledge where international governmental organizations and non-governmental organizations (NGOs) are most active. While some are more established than others – the World Bank relies less on public visibility than Transparency International – all urge to be recognized as experts on their field.

One means for attaining such credibility is to produce seemingly neutral numerical knowledge, which helps to legitimate their existence and resourcing (cf. Marcussen 2002, also Gieryn 1999, 23). At the same time, however, quantified assessments represent a type of information that is costly to collect, effectively making the circle of experts to engage in this activity somewhat exclusive. This restrictive effect is furthermore strengthened because it is "difficult for new initiatives – to gain attention, because the most-widely used indicators are well-established and dominate the market" (Arndt and Oman 2008, 11).

The numerical form of the new knowledge products tends to captivate the actors by the same ontology: numbers are best debated with other numbers (Erkkilä and Kauppi 2010). Developers of indices often correlate their own data against competitors – not only to prove the superiority of their numbers – but primarily to validate theirs belonging to the field of expertise. In effect, rather than disqualifying the others, such mild contrasting serves to lend credibility and thus legitimacy to the referents also. The numbers themselves become a policy discourse that coordinates the consensus formation of the actors involved (cf. Schmidt 2008). Interestingly, however, while the production of expert knowledge and the boundary-work to protect their epistemic authority in Gieryn's terms (1999) builds on consensus, our examination on the production of governance indicators suggests also potential, albeit usually limited, internal tensions.

In the following we explore empirically the tensions around the production of comparative data on governance performance: are these tensions enough to unsettle the current understanding of "good governance" or the epistemic authority of legitimated experts? We do this by a critical reading of key documents explaining the theoretical and methodological bases of the measurement along with documents providing information

on measurement experiments and application of the data. Our assessment of the Worldwide Governance Indicators documentation by the World Bank Institute shows how the producers of an established measurement have responded to challenges from both within the World Bank group (see Knack, Kugler, and Manning 2003) and outside it (Thomas 2010; Arndt and Oman 2008) by not needing to alter either the conceptual or the methodological core of their instrument.

We pay particular attention to the activities of the OECD where interest toward creating its own – competing – dataset on governance has yielded results in the latter half of the 2000s. While the new approach – looser datasets, increased openness and wider stakeholder participation in the development – adopted by the OECD implicates a new potential for re-politicizing knowledge on good governance, it is our conclusion that the work for redefining the boundaries of expertise remains limited. Indeed, it can be hypothesized that some of the fundamental ideas – a set of shared normative and causal beliefs – behind the existing figures are immune to almost any challenge that a set of countering numbers can put forth.

Although producers of governance data constantly compete against each other, they are still to a great extent dependent on mutual acknowledgment in reciprocally upholding their expert status. Critics are still dependent on the very same community of experts in arguing for the scientific validity of their own activities (cf. Gieryn 1983).

Worldwide Governance Indicators: establishing the knowledge production in governance assessments

In political science, "governance" as a concept has come to define the redefinition of the changing role of state since the early 1990s. Here the term is generally being used to describe the outward shifts of state powers toward international organizations, private companies and local government through decentralization (Pierre and Peters 2000, 77–91). In addition, a normative concept of "good governance" was coined by the World Bank wanting to overcome the legal constraints preventing it from interfering in member countries' internal affairs through its lending criteria in the late 1980s (Thomas 2007).[3]

The original idea of good governance proxied that of "good market regulation" (Argyriades 2006, 158–160; Doig, Mcivor, and Theobald 2006, 241; Drechsler 2004; Seppänen 2003; Zanotti 2005, 470). Thus, many of the standards of good governance are identical to the policy prescriptions of the New Public Management initiatives that were launched in the West at about the same time and whose emphasis was on increasing public sector efficiency through market-led steering (Drechsler 2004). To be sure, the "good" in good governance can refer to whatever attributes an analyst sees fit. Nevertheless, we argue that the most important measurements of governance attributes have tended to be in line with the economist understanding of good governance (see below). By economism we do not refer to the science of economics but to all definitions that primarily and fundamentally focus on economic aspects of governance – often at the cost of alternative values (e.g. democracy, environment, social cohesion).

One of the most important and ambitious attempts to measure governance performance is the World Bank Institute's WGI project, which has produced data since 1996 that covers at present 212 countries and territories and is based on a total of 441 individual variables from 35 different sources produced by 33 different organizations (Kaufmann et al. 2009). According the producers, the purpose is to generate knowledge concerning the quality of governance defined as "the traditions and institutions by which [public] authority in a country is exercised" (Kaufmann et al. 2009, 5). The WGI consists

of six aggregate indicators that are designed to capture political, economic and institutional dimensions of governance: (1) voice and accountability; (2) political stability and absence of violence; (3) government effectiveness; (4) regulatory quality; (5) rule of law; and (6) control of corruption. If not before, assignment of interval scale values to the performance variables effectively turns governance indicators into normative good governance indicators. As a consequence, the norm of good governance comes to be defined in terms of the indicators included and their relative weights.

A look at the attributes and the indicators[4] reveals that democracy is not emphasized too heavily in the conception of good governance promoted by the WGI: of its six components only "voice and accountability" measures, clearly democratic aspects of governance, others – "political stability" excluded – revolve around the business friendliness of public policies and infrastructure, protection of private property and controls for public authority exercised for private gain. The selection of relevant attributes and their respective indicators, and aggregating them into index-values depicting each of the six attributes, is done by the scholars at the World Bank Institute.[5] It is the six dimensions that provide the form of presentation and analysis. While the WGI publications do not offer an overall league table on countries' performance, such data is present for ranking country performance according to the six mid-level aggregates.

Although we do not engage in a full-blown analysis of any particular measurement, we want to stress certain characteristics of the WGI, which we hope will suffice to highlight both the paradoxical dynamics of (de-)politicization inherent in most measurements of construct concepts. On the one hand, index-production is necessarily normatively loaded, because of both the multiple judgmental choices that developing a dataset demands and the inherent logic of measurement. Conceptual judgments deal with decisions of what to measure and how to define the object of measurement (see also Thomas 2010). To define "regulatory quality" – as the WGI does – in terms of "the ability of the government to formulate and implement sound policies and regulations that permit private sector development" should hardly be seen as the only possible way of looking at qualities of regulation.

The second series of judgments concern methods for dealing with missing data, rescaling, weighing and aggregation, all of which have the potential to affect the results – composite indicators and rankings – immensely (OECD 2008; Arndt and Oman 2008). For example, while the WGI involves a rather complicated weighing procedure, it is still informative, for the sake of example, to observe that, from a total of 40 representative "rule of law" indicators, 10–15 measure protection of property rights, the exact count depending on the interpretation.[6]

Finally, practical judgments concerning the selection of data sources and the form of presentation play a role in the design and interpretation of a measurement. Indeed, inclusion of indicators seems to be determined more in terms of data availability than theoretical considerations about the exact composition of the aggregates. However, more fundamentally, *the very act of measuring is sufficient to render the measured "phenomena" normative*. In assigning interval scale values to observed qualities of governance – possibly a neutral or empirical concept as such – the WGI (and any other similar measurement) necessarily comes to classify aspects of governance as wanted or unwanted, good or bad. Governance is thus politicized.

Yet perhaps paradoxically, the very practice of aggregation as a numerical reduction tends to hide the inherent normativity – even to the extent that the domain in question becomes depoliticized. Numerical index data partly functions as a mechanism through which room for debate is narrowed by framing the meaning of good governance.

The WGI, for example, represents the concept of governance as something fixed and agreed upon. Despite the name – Worldwide Governance Indicators – the focus is not on the concept of governance but on countries' performance in relation to a very particular economist objectification not present in the description of the dataset.

That said, compared with some of its counterparts – Transparency International's Corruption Perception Index, Bertelsmann Transformation Index, the Economist Intelligence Unit's Index of Democracy, the Ibrahim Index of African Governance, just to mention a small sample – the WGI carries the potential for diverging interpretations of governance. Whereas the direct substantive critique against the "good governance" discourse seems to have reached only a fraction of the community of knowledge producers, there has been much debate on the measurement methodology. A close reading of the World Bank Institute's *Governance Matters* series reveals that the producers of the WGI are increasingly aware of problems related to reductionist measurements striving for extreme parsimony: the numerical objectification (WGI) is treated all the more carefully as criticism against it hardens with the potential of undermining its seeming neutrality. External criticism and internal adjustment could potentially weaken the depoliticizing elements of the WGI. However, as we show below, this has not been the case.

In their early published documents the tone of WGI developers when describing their objectives and achievements is optimistic and matter-of-fact (see e.g. Kaufmann et al. 1999a, 1999b). The technical problems of measurement are discussed, but they are either represented as the necessary evil of all social scientific quantification – and thus deemed acceptable – or as problems in competing measurements that are minimized in the WGI framework. Nevertheless, the tone of discussion has become more cautious in each succeeding *Governance Matters* publications (published almost yearly between 1999 and 2009) as problematic issues are brought into consideration by the developer team. Possible problems ranging from subjective measurement methodology to consequences of alternative weighing rules are discussed and, more or less convincingly, treated (see e.g. Kaufmann et al. 2007). In early 2000s the stakes were raised when certain components of WGI were announced to be applied in the Millennium Challenge Corporation's (MCC) funding criteria. Warnings against reading the WGI scores without paying due respect to the margins of error were sharpened and emphasized (Kaufmann et al. 1999a; cf. Kaufmann et al. 2003, 25–26).[7]

Although the developer team still holds that the WGI can and should be applied in the MCC and similar funding devices, they, from 2003 onwards, have warned against trusting the WGI too much in assessing governance in individual countries (Kaufmann et al. 2003, 40). In the 2008 edition of *Governance Matters* (Kaufmann et al. 2008, 5), the ranking of countries in a league table manner is definitively denounced: "More generally, recognizing the importance of margins of error and the imprecision of country rankings, we do not follow the popular practice of producing precisely ranked 'top ten' or 'bottom ten' lists of countries according to their performance on the WGI, recognizing that such seemingly precise 'horse races' are of dubious relevance and reliability". In fact, starting from 2012, the MCC's financing criteria were revised and the "voice and accountability" component based on WGI data was changed with a component of transparency using data from Fringe and Open Net Initiative, both small NGOs providing "second generation" datasets.

On the one hand, this echoes the more general distrust of country rankings of any kind that has gained the upper hand during the latter part of the decade. On the other, it positions the WGI against its competitors on the index-market: Transparency International, although

making levels of confidence public, keeps receiving far superior media attention annually by publishing the results of its corruption measurements in the form of a name-and-shame league table. The WGI wants to be taken as serious science and is aimed less at the general public than at academic and administrative audiences. While the importance of the six aggregate constructs is still emphasized, there is a parallel tendency to provide more detailed data to the public. From 2006 onwards, almost all disaggregated data underlying the six aggregates have been published and shared on the WGI homepage (Kaufmann et al. 2006).

In fact, in the interviews conducted at the World Bank Institutes, the developers of the WGI noted that the decision to use aggregation to begin with was largely due to the data available at the time. Arguably there was much noise in the data, meaning that the results of different measures were rather inconsistent. Aggregation provided a way around this. By making an aggregate figure, the "noise" could be eliminated. Nevertheless, this decision further legitimated the use of aggregation in the "first generation" governance rankings, owing to the prominence of the WGI.

The WGI is not only important because of its high status ("produced by the World Bank Institute"), empirical scope and methodological sophistication, but being a "composite of composite" it contains data from multiple sources. In doing so it literally binds a community of knowledge producers together and squeezes various perspectives into a unified whole – the WGI. While it resists reducing "quality of governance" into one aggregate – instead having six of them – it still plays an important role in the institutionalization of certain causal beliefs and normative goals into "self-evidences" or "truths"; it is an instrument for constructing consensus on standards and – in Palonen's terms – an active attempt to depoliticize domains previously (or potentially) open to politicking.

Governance at glance: politicization or steepening expertise?

There have been claims for a paradigm shift in the production of governance indices, a shift toward second-generation governance indices (Knack, Kugler, and Manning 2003). At a minimum this would mean that (1) the indicators are replicable through documentation and use politically acceptable sources, (2) they have broad country coverage, (3) they are consistent conceptually and in measuring different countries, and (4) they are specific about what is being measured (Knack, Kugler, and Manning 2003, 350). Along with the above attempt to improve the governance measurements, there now is a new debate concerning the presentation of the results (OECD 2008). Many previous measurements of good governance are under attack for creating aggregate numbers of their assessments, often presented in a ranking order. Thus there is a new demand for looser datasets and open access to all the raw data. This has also opened an opportunity for new knowledge products to enter the policy field. Most notably, there has been a recent attempt to measure aspects of governance by OECD's Public Governance and Territorial Development Directorate.

Although the OECD is a central organization of transnational governance and has a long history of collecting statistical hard-data on macroeconomic performance (Mahon and McBride 2009), it has been somewhat absent in the sphere of applied governance measurements. Since 2005, however, the OECD has prepared its own governance measurements, now launched under the name "Government at a Glance" (GG). What separates the OECD project from the previous numerical assessments of governance is its explicit criticism of accumulated numbers and single rankings. However, this criticism has grown during the evolution of the OECD's measurements, becoming a justification

for its participation in the activity of governance measurement. Moreover, although the newcomer GG is clearly a manifestation of the second-generation approach, it fails to challenge the boundaries of established expert authority on governance studies as will be explained in the following.

The scope of the GG dataset is the 30 OECD member countries. The OECD has identified the topics of its core data as follows (see also Lonti and Woods 2008): (1) institutions and administrative structures; (2) revenues; (3) budgeting and public expenditures; (4) e-government; (5) fighting corruption; (6) public employment and management; (7) and regulatory reform and management. There are also changing "hot topics" that are assessed, but the above classification perhaps best represents the understanding of "good government" assessed in the measurements. The OECD's "government at a glance" is tied to the contemporary ideas of performance management. It has an underlying input–output model, assessing also the funding with which the measured outcomes are achieved. In the OECD, the lack of comparative information on productivity between the OECD countries was identified already in 1999 (OECD 1999, 7; see also OECD 2005a, 5).

Launching the production of the GG dataset, only in 2005 however, the OECD conducted a detailed feasibility study that covered its previous activities on gathering comparative data on the member states but also on the international governance indices that were produced at the time. Acknowledging the growing interest in the measurements of government, OECD refers to experiences from the Eurostat and certain of its member states (OECD 2005a). Before the launching of the GG, the OECD had already published five major "at a glance" series: Society at a Glance, Health at a Glance, Education at a Glance, Pensions at a Glance, and Agricultural Policies in OECD Countries. Initially, the feasibility study identified the first three as less evaluative and hence more like-minded to the planned GG publication (OECD 2005b, 62). In particular, the Education at a Glance Indicators, first produced in 2001, are mentioned as the forerunner for the coming GG indicator set (OECD 2005b, 3).

In justifying its participation in the new venture for measuring governance, the 2005 feasibility report refers to the methodological problems of the existing indices and rankings, denting their reliability. The indices in question were the World Bank Governance Indicators, the European Central Bank's Public Sector Efficiency Study, the World Economic Forum's Public Institutions Index in the Global Competitiveness Report, and the "Government Efficiency" Indicator developed by the International Institute for Management Development in the World Competitiveness Yearbook:

> There is a significant growth in broad measures of "governance", including some comparative data concerning public sector bureaucratic quality. However, most of these data are based on subjective assessments, and were not initially collected with comparative analysis of public management as a principal aim. … Reviews of these data note that these indicators incorporate significant methodological problems. The data often do not adequately measure what they claim to measure, and can aggregate many diverse indicators, achieving statistical quality at the price of significant loss of conceptual precision. Often data amount to broad subjective evaluations combined with service-specific performance indicators. The former can be excessively impressionistic and the latter cannot be aggregated in any meaningful way. (OECD 2005a, 6)

This criticism is based on general notions of the insufficient validity and reliability of the data. At this point, the attempt at forming a single accumulated ranking is not criticized.

In fact, the low quality of the data used is seen as a problem *for* aggregation. Aggregation *per se* is not perceived to be a problem. The feasibility study refers to the construction of aggregate indicators, but only in reference to the At a Glance publications on Pensions systems and Agricultural Policies. This is not explicitly criticized but it is merely noted that such "more evaluative" assessments are less suitable examples for the planned GG. Aggregate indicators and rankings of other producers are also referred in the assessment, but the "problem" of aggregation is not explicitly mentioned in the lengthy assessment of existing figures.

The same occurred in 2006 when a document on the use of outcome measurements in GG idealized an effort of ranking over 50 countries according to their educational performance in the PISA study (OECD Programme for International Student Assessment; OECD 2006a, 32). The aggregation of the results is not presented as a problem. Instead, in the 2006 memo, aggregation or presenting the national indicators in a broader portfolio is seen as preferable (OECD 2006a, 7). Moreover, in the preparation for the GG the possibility of engaging in collaboration with other data producers is proposed (OECD 2006a, 32–33).[8] In a more technically inclined OECD document of 2006, aggregation merely appears as a means for improving output quality (OECD 2006b, 34).

In 2006, the OECD's critique primarily concerned the composite indicators, meaning that the data is derived from various sources (such as the WGI), although the OECD uses "aggregate indicator" and "composite indicator" interchangeably (OECD 2006c, 7). An OECD technical paper claims that there are risks involved in the use of composite indicators. First of all, there is a political risk, understood by the OECD as a risk of political debate not being analytical enough to learn from the measurements. The political risk is in fact somewhat similar to the effect of depoliticization that we have examined in this article. However, the OECD's concern is different, pointing to a missed opportunity to make informed reforms, rather than of being uncritical about the numerical information as such. Secondly, another risk regarding composite indicators, according to the OECD, stems from the imprecision in inter-country rankings that might be misleading (OECD 2006c, 45).

In assessing the "the contribution of aggregate governance indicators to an OECD debate", these are dismissed as expressing only slight variations of the same themes (OECD 2006c, 60). Moreover, the OECD paper notes that there is a "sense that each [development] agency needs its own signature index" (OECD 2006c, 60). The concept of "governance" is criticized for being vague: "The absence of a well-accepted theoretical framework for governance ensures that any composite indicators are largely devices for communication – for crystallizing concerns about corruption etc. into a single short and pithy summary" (OECD 2006c, 60).

One year later, in 2007, as the GG was starting to take shape, the OECD's work was contrasted – and legitimized – through its denouncing of "aggregate assessments", such as World Bank Worldwide Governance Indicators and Transparency International's Corruption Perception Index (OECD 2007). Referring to the heterogeneity of its member countries, the OECD aimed now at providing more "nuanced" picture of national governments than the above assessments:

> The World Bank Worldwide Governance Indicators and Transparency International Corruption Perceptions Index provide aggregate assessments of governance at the country level. By contrast, "Government at a Glance" will provide data with which a country can assess itself. … Like other OECD "At a Glance" publications, the data collected allows for some nuanced distinctions to be made between OECD countries, reflecting their distinctive

> administrative and social traditions. More aggregate indicators tend to show all OECD countries as being similar in most dimensions. (OECD 2007, 3)

Yet what credible alternatives or added value does the GG offer the existing figures? It is different from its content, resembling more traditional statistics produced by national statistical bureaus and Eurostat, with whom the OECD closely cooperates. There are also significantly fewer countries involved as only the OECD member states are assessed. Moreover, the data is provided by the national governments, which are also partially involved in the committees of the OECD, although the policy work is not explicitly intergovernmental. In fact, the GG is different enough to make an argument of its own right. It is still framed by references to existing index figures that during its planning have acted both as sources of ideational input and potential collaboration but which, finally, are argued to be plagued by shortcomings that the new OECD figure is to mend. It is based on a familiar technocratic management attitude and it seeks to complement figures by better figures. (OECD 2009, 10)

In the interviews the representatives of the World Bank pointed out that there was also keen collaboration between the Bank and the OECD in the data production. A central figure in the OECD GG project was actually a World Bank employee who was on a leave at the time, working for the OECD.

In practice, nevertheless, the OECD figure offers an "actionable" alternative to the previous measurements mainly technically, qualifying it as a second-generation governance measurement, the coming of which was anticipated by Knack, Kugler, and Manning already in 2003. Instead of providing a single ranking figure, the OECD's Government at a Glance aims at providing a multidimensional tool for assessing various aspects of governance. It does not dictate the balancing between different indicators or give premade combinations of these to the user. However, the numerical objectification aims at creating "social facts" over the phenomena that are being measured.

In a sense, the OECD's position on rankings makes it part of the critiques of governance indices and rankings, such as the WGI. Yet, as the organization chooses to formulate its critique in a set of numbers, it is itself becoming a player in the field (compare Erkkilä and Kauppi 2010). This is undoubtedly a deliberate choice and there exists a sincere wish to improve public management in OECD countries. Nevertheless, the decision to bandwagon has a significant function in reinforcing the expert status of the organization. Interpreted as boundary work, it "is strategic practical action" with objectives such as defending the autonomy of scientist – or experts for our purposes – and their material resources (Gieryn 1999, 23). Hence, the second generation of governance assessments need not mark a juncture in the numerical assessments of governance but rather leads to their further institutionalization by allowing sufficiently "qualified" actors to join in. Embracing new, critical members does not threaten the position of recognized members of the epistemic community. Neither does it in any way challenge the basis of credible epistemic authority, quite the contrary.

The growing number of actors in the field of governance measurements could potentially mark an opening for the politicization of the subject matter, (good) governance, but only if the organizations have a different set of premises on which they operate. The OECD has a somewhat different approach to quantification, true, but it nevertheless also tends to hold a view of governance favoring economic effectiveness. Moreover, it explicitly aims at providing the member states with information through which they can monitor their performance, leading altogether to the states' reflexivity

over the particular aspects of governing present in the indicators. Instead of breaking the instrumental nature of the governance measurements by exposing these to political debate, the OECD's alternative figure is likely to further mount to the streamlining of national administrations.

The OECD's "boundary work" is political in a limited sense: it wins the organization an edge in the newly established policy field. However, it shares most of the normative and causal beliefs, concepts and even the ontology of the previous assessments of good governance. While OECD's initiative does raise some concerns over measurability and methodology, it does not mark an opening for reconsidering the constituents of "good" governance.

Conclusions

The production of measurement knowledge on governance is a polymorphous act of power. Not only does measuring good governance significantly affect the actual policy-making of nation states, but it also produces and upholds power relations in demarcating the boundaries of expertise. Nevertheless, the political character of expert knowledge such as statistics often remains tacit.

As we have shown, the numerical objectifications of good governance tend to depoliticize the normative foundations of these assessments. The numbers make new domains of governance governable. Yet at the same time they tend to remove new issues of concern from the sphere of political debate, reframing them as expert knowledge based on scientific inquiry. To render the political nature of the numbers visible, they should be debated for their underlying normative and causal beliefs and, to some extent, for their technical composition.

Regarding the debate, the numerical representation of the above governance assessments excludes those who are not able to produce alternative figures. Moreover, comparative expert knowledge is costly and difficult to collect and produce, and the actors that are able to participate to the debate are limited in number. Nevertheless, since the 1990s, the number of actors in the field of governance indices has been growing steadily.

Although producers of governance data are constantly competing against each other with alternative rankings, they are still to a great extent dependent on mutual acknowledgment in upholding their expert status. Although there are increasingly critical voices in the debate on proper methodology for creating indices, this can nevertheless be understood in the realm of deepening expertise: critics are dependent on the very same community of experts to assess the scientific validity of their own activities. As has been argued in this article, the producers of indices on good governance incline to an expert community within which the political charges against other data producers can be likened to scientific boundary work, where also newcomers have to legitimize their knowledge products against the existing ones.

Nevertheless, we are witnessing a mild paradigm shift around the numerical assessments of good governance. Numerical assessments of governance have shifted from accumulated rankings and compound indices toward more detailed and varied aspects of the phenomena, rendering the results more actionable and suited for causal research. The underlying assumptions of governance, emphasizing economic virtues of public administration, have not however allowed for a significant reconsideration of the term.

The new techniques, second-generation governance measurements, abstain from evaluating administrative practices in terms of good and bad qualities, but they still set

the parameters within which governance is to be assessed. Potentially, second-generation datasets could imply alternative parameters for good governance, emphasizing, for instance, democratic aspects of governance. Yet even here the conflict is held within a domain of epistemic authority of experts, taking form in methodological disputes over aggregation.

While we observe OECD's dataset to be an alternative to the Worldwide Governance Indicators merely in methodological terms, a diversification of perspectives could finally prove to be the only viable solution to tackle the depoliticizing symptoms of numerical objectification of governance. This diversification would mean increased pluralism in ideological perspectives and premises behind the rankings. However, to provide an effective policy alternative, any index to challenge the existing paradigm of good governance would first have to surpass the institutional path dependencies and epistemic inertia of the field. The OECD's alternative ranking does not provide a real alternative, but celebrates the previous economic understanding of good governance with yet another set of numbers.

Acknowledgments

We would like to thank Thomas Pfister, Anna Horvath and our anonymous reviewers for their comments on previous versions of this article. Also, we would like to thank Max Eklund and Mark Waller for their assistance with editing and proofing the manuscript. Our work has received funding from Academy of Finland, Helsinki University Network for Higher Education and Innovation Research, and Helsinki University Network for European Studies.

Notes

1. Desrosières's (1998, 9) notion of numerical "objectification" is a fitting way of describing the processes in which social scientific quantification construct "objective phenomena".
2. Gieryn (1999, 1) defines "epistemic authority" as "the legitimate power to define, describe, and explain bounded domains of reality". While his analysis focuses on the modern meta-narrative – "science" – we seek to identify epistemic authority with bounded domains of expertise that, admittedly, draw much of their legitimacy from "scientific" conduct and contact.
3. Apparently, the World Bank – at the time presided by James Wolfensohn – used a somewhat corresponding strategy in attempting to redefine the "[previously unspeakable] 'C' word" as something non-political in order to make anti-corruption work fit its "neutrality" upholding statutes (Ivanov 2007, cited in Gephart 2009, 11).
4. World Wide Governance Indicators, http://info.worldbank.org/governance/wgi/index.asp
5. Laura Langbein and Stephen Knack (2008) argue that "the six indexes do not discriminate usefully among different aspects of governance" and that they merely reflect the quality of governance more generally.
6. Rule of law, http://info.worldbank.org/governance/wgi/pdf/rl.pdf
7. While it is reasonable to subscribe to the notion that corruption should be an important factor in allocating aid, it is important to emphasize that a simple "in-or-out" rule runs the risk of misclassifying some countries precisely because margins of error are not trivial … For the majority of countries there is a non-trivial probability that they could be mistakenly classified in the bottom half of the sample (Kaufmann et al. 2003, 25–26).
8. This was seen to be particularly relevant regarding trust in government as a diagnostic, where the OECD saw the institutions already making relevant surveys being prepared to deliberate collaboration (OECD 2006a, 33).

References

Argyriades, D. 2006. "Good Governance, Professionalism, Ethics and Responsibility." *International Review of Administrative Sciences* 72 (2): 155–170.

Arndt, C., and C. Oman. 2006. *Uses and Abuses of Governance Indicators*. Development Centre Studies. Paris: OECD.

Arndt, C., and C. Oman. 2008. "The Politics of Governance Ratings." *Working Paper MGSoG/2008/WP003*. Maastricht: Maastricht Gradute School of Governance.
Bhaskar, R. 1998. *The Possibility of Naturalism. A Philosophical Critique of the Contemporary Human Sciences*. 3rd ed. London: Routledge.
Desrosières, A. 1998. *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
Doig, A., S. McIvor, and R. Theobald. 2006. "Numbers, Nuances and Moving Targets: Converging the Use of Corruption Indicators or Descriptors in Assessing State Development." *International Review of Administrative Sciences* 72 (2): 239–252.
Drechsler, W. 2004. "Governance, Good Governance, and Government: The Case for Estonian Administrative Capacity." *Trames* 8 (4): 388–396.
Erkkilä, T., and N. Kauppi. 2010. "Alternatives to Existing International Rankings." In *World Social Science Report*, edited by F. Caillods, 239–241. Paris: UNESCO.
Erkkilä, T., and O. Piironen. 2009. "Politics and Numbers: The Iron Cage of Governance Indices." In *Ethics and Integrity in Public Administration: Concepts and Cases*, edited by R. W. Cox III, 125–145. Armonk, NY: M.E. Sharpe.
Espeland, W. N., and M. L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology* 49 (3): 401–436. doi:10.1017/S0003975609000150.
European Commission. 2001. *European Governance: A White Paper*. Brussels: Commission of the European Community. Accessed June 26, 2007. http://eur-lex.europa.eu/LexUriServ/site/en/com/2001/com2001_0428en01.pdf
Gephart, M. 2009. "Contextualizing Conceptions of Corruption: Challenges for International Anti-corruption Campaign." *GIGA Working Papers* 115. Hamburg: German Institute of Global and Area Studies.
Gieryn, T. F. 1983. "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists." *American Sociological Review* 48 (6): 781–795. doi:10.2307/2095325.
Gieryn, T. F. 1999. *Cultural Boundaries of Science. Credibility on the Line*. Chicago: The University of Chicago Press.
Guha, K., and R. McGregor. 2007. "World Bank Directors Test Zoellick." *Financial Times*, July 13. http://www.ft.com/intl/cms/s/0/fe1d7ece-30d8-11dc-9a81-0000779fd2ac.html#axzz2883nwxuM
Haas, P. M. 1992. "Introduction: Epistemic Communities and International Policy Coordination." *International Organization* 46 (1): 1–35. doi:10.1017/S0020818300001442.
Harding, R. 2013. "China Seeks to Water down Key World Bank Report." *Financial Times*, May 6. http://www.ft.com/intl/cms/s/0/a1ca36d2-b654-11e2-b1e5-00144feabdc0.html#axzz2SmcKhrYs
Hooghe, L., and G. Marks. 2003. "Unraveling the Central State, But How? – Types of Multi-Level Governance." *American Political Science Review* 97 (2): 233–243.
Hopwood, A., and P. Miller. 1994. *Accounting as Social and Institutional Practice*. Cambridge: Cambridge University Press.
Hummel, R. P. 2006. "The Triumph of Numbers: Knowledges and the Mismeasure of Management." *Administration & Society* 38 (1): 58–78. doi:10.1177/0095399705284202.
Kaufmann, D., A. Kraay, and M. Mastruzzi. 2003. "Governance Matters III: Governance Indicators for 1996–2002." *World Bank Policy Research Working Paper No.3106*. Washington, DC: World Bank.
Kaufmann, D., A. Kraay, and M. Mastruzzi. 2006. "Governance Matters V: Aggregate and Individual Governance Indicators for 1996–2005." *World Bank Policy Research Working Paper No.4012*. Washington, DC: World Bank.
Kaufmann, D., A. Kraay, and M. Mastruzzi. 2007. "Governance Matters VI: Governance Indicators for 1996–2006." *World Bank Policy Research Working Paper No.4149*. Washington, DC: World Bank.
Kaufmann, D., A. Kraay, and M. Mastruzzi. 2008. "Governance Matters VII: Aggregate and Individual Governance Indicators, 1996–2007." *World Bank Policy Research Working Paper No.4654*. Washington, DC: World Bank.
Kaufmann, D., A. Kraay, and M. Mastruzzi. 2009. "Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996–2008." *World Bank Policy Research Working Paper No.4978*. Washington, DC: World Bank.
Kaufmann, D., A. Kraay, and P. Zoido. 1999a. "Aggregating Governance Indicators." *World Bank Policy Research Working Paper No. 2195*. Washington, DC: World Bank.

Kaufmann, D., A. Kraay, and P. Zoido. 1999b. "Governance Matters." *World Bank Policy Research Working Paper No. 2196*. Washington, DC: World Bank.
Knack, S., M. Kugler, and N. Manning. 2003. "Second Generation Governance Indicators." *International Review of Administrative Sciences* 69 (3): 345–364. doi:10.1177/0020852303693004.
Langbein, L., and S. Knack. 2008. "The Worldwide Governance Indicators and Tautology: Causally Related Separable Concepts, Indicators of a Common Cause, or Both?" *Policy Research Working Paper 4669*. Development and Research Group. Washington, DC: World Bank.
Larner, W., and W. Walters. 2004. "Globalization as Governmentality." *Alternatives* 29: 495–514.
Lee, N., and C. Kirkpatrick. 2006. "Evidence-Based Policy-Making in Europe: An Evaluation of European Commission Integrated Impact Assessments." *Impact Assessment and Project Appraisal* 24 (1): 23–33. doi:10.3152/147154606781765327.
Lelieveldt, H., and S. Princen. 2011. *The Politics of the European Union*. Cambridge: Cambridge University Press.
Lonti, Z., and M. Woods. 2008. "Towards Government at a Glance: Identification of Core Data and Issues Related to Public Sector Efficiency." *OECD Working Papers on Public Governance*, No. 7, Paris: OECD.
Löwenheim, O., 2008. "Examining the State: A Foucauldian Perspective on International 'Governance Indicators'." *Third World Quarterly* 29 (2): 255–274. doi:10.1080/01436590701806814.
Mahon, R., and S. McBride. 2009. "Standardizing and Disseminating Knowledge: The Role of the OECD in Global Governance." *European Political Science Review* 1 (1): 83–101. doi:10.1017/S1755773909000058.
Marcussen, M., 2002. *OECD og idéspillet. Game Over?* [OECD and the Diffusion of Ideas: Game Over?] Kopenhagen: Hans Reitzels Forlag.
Miller, P., and N. Rose. 1990. "Political Rationalities and Technologies of Government." In *Texts, Contexts, Concepts. Studies on Politics and Power in Language*, edited by S. Hänninen and K. Palonen, 166–183. Helsinki: The Finnish Political Science Association.
OECD. 1999. "Report of the Puma Expert Group Meeting on Measuring Productivity in the Government Sector." STD/NA/RD(99)1. Paris: OECD.
OECD. 2005a. "Management in Government: Feasibility Report on the Development of Comparative Data." GOV/PGC(2005)10/ANN. Paris: OECD.
OECD. 2005b. "Management In Government: Feasibility Report on the Development of Comparative Country Data Technical Annexes." GOV/PGC(2005)10/ANN. Paris: OECD.
OECD. 2006a. "Issues in Outcome Measurement for 'Government at a Glance'." OECD GOV Technical Paper 3. GOV/PGC(2006)10/ANN3. Paris: OECD.
OECD. 2006b. "Issues in Output Measurement for 'Government at a Glance'." OECD GOV Technical Paper 2. GOV/PGC(2006)10/ANN2. Paris: OECD.
OECD. 2006c. "How and Why Should Government Activity Be Measured in 'Government at a Glance'?" OECD GOV Technical Paper 1. GOV/PGC(2006)10/ANN1. Paris: OECD.
OECD. 2007. "Towards Better Measurement of Government." *OECD Working Papers on Public Governance*, 2007/1, Paris: OECD.
OECD. 2008. *Handbook on Constructing Composite Indicators. Methodology and Used Guide*. Paris: OECD. http://browse.oecdbookshop.org/oecd/pdfs/browseit/3008251E.PDF
OECD. 2009. *Government at a Glance 2009*. Paris: OECD.
Palonen, K. 2007. *Re-thinking Politics: Essays from a Quarter-Century*. The Finnish Political Science Association. Jyväskylä: University Printing House.
Pierre, J., and G. B. Peters. 2000. *Governance, Politics and the State*. New York: St. Martin's Press.
Porter, T. 1995. *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Rose, N. 1999. *Powers of Freedom. Reframing Political Thought*. Cambridge: Cambridge University Press.
Schmidt, V. A. 2008. "Discursive Institutionalism: The Explanatory Power of Ideas and Discourse." *Annual Review of Political Science* 11 (1): 303–326. doi:10.1146/annurev.polisci.11.060606.135342.
Seppänen, S. 2003. "Good Governance in International Law." *The Erik Castrén Institute Research Reports* 13/2003. Helsinki: The Erik Castrén Institute.
Thomas, M. A. 2007. "The Governance Bank." *International Affairs* 83 (4): 720–745. doi:10.1111/j.1468-2346.2007.00649.x.

Thomas, M. A. 2010. "What Do the Worldwide Governance Indicators Measure." *European Journal of Development Research* 22 (1): 31–54. doi:10.1057/ejdr.2009.2.
Vogl, F. 2013. "Challenge Is Not Without Precedent." *Financial Times*, May 8. http://www.ft.com/intl/cms/s/0/66999938-b715-11e2-841e-00144feabdc0.html#axzz2SmcKhrYs
Zadek, S. 2013. "China's Challenge to the World Bank." *Financial Times*, May 8. http://www.ft.com/intl/cms/s/0/661e93d2-b715-11e2-841e-00144feabdc0.html#axzz2SmcKhrYs
Zanotti, L. 2005. "Governmentalizing the Post-Cold War International Regime: The UN Debate on Democratization and Good Governance." *Alternatives* 30: 461–487.