

The politics of numbers: the normative agendas of global benchmarking

ANDRÉ BROOME AND JOEL QUIRK*

Abstract. Global benchmarks have grown exponentially over the last two decades, having been both applied to and developed by states, international organisations, corporations, and non-governmental organisations. As a consequence, global benchmarking is now firmly established as a distinct mode of transnational governance. Benchmarking chiefly involves the development of comparative metrics of performance, which typically take the form of highly stylised comparisons which are generated by translating complex phenomena into numerical values via simplification and extrapolation, commensuration, reification, and symbolic judgements. This process of translation takes what might otherwise be highly contentious normative agendas and converts them into formats that gain credibility through rhetorical claims to neutral and technocratic assessment. This politics of numbers has far-reaching ramifications for transnational governance, including the dimensions and effects of indirect power, expertise and agenda-setting, coordination, regulation and certification, and norm contestation and activism. This Special Issue draws upon an emerging literature to explore how and why benchmarks both align with and expand upon established models of International Relations theory and scholarship. It does so by critically examining the role of global benchmarks in key areas such as state ‘failure’, global supply chains, disaster management, economic governance, corporate social responsibility, and human development.

André Broome is Associate Professor of International Political Economy in the Department of Politics and International Studies and is Director of the Centre for the Study of Globalisation and Regionalisation, University of Warwick. His publications include *The Currency of Power* (Palgrave, 2010) and *Issues and Actors in the Global Political Economy* (Palgrave, 2014).

Joel Quirk is an Associate Professor in the Department of Political Studies, University of the Witwatersrand. His publications include *The Anti-Slavery Project* (Penn, 2011) and *Mobility Makes States* (Penn, 2015). Joel is a current member of the International Scientific Committee of the UNESCO Slave Route Project, and is also an *openDemocracy* editor for Beyond Trafficking and Slavery.

The political appeal of numbers

In November 2014, the Walk Free Foundation launched the second edition of their flagship Global Slavery Index. The press release for the launch declared the index to be ‘the most accurate and comprehensive measure of the extent and risk of modern slavery’, building upon an ‘improved methodology’ which included random-sample

* As editors of this Special Issue, we are grateful for feedback on earlier versions of all of the articles from participants at the ‘Benchmarking in Global Governance’ Research Workshop, University of Warwick, 12–14 March 2014. We are also grateful for financial support from GR:EEN, European Commission Project Number: 266809, the Global Research Priority in Global Governance, the Department of Politics and International Studies, University of Warwick, and a Warwick International Partnership Award with the University of the Witwatersrand on ‘Benchmarking in Global Governance’.

surveys from 19 countries.¹ This was significant because the inaugural 2013 index had been heavily criticised for using unreliable, incomplete, and inappropriate data.² Despite this emphasis on improved methods, the second version of the index once again attracted sustained criticism. One of the sharpest challenges came from Anne Gallagher, a noted expert on human trafficking, who reported that the index contained ‘critical errors of fact and logic’.³

The errors which Gallagher and others have identified chiefly relate to the steps which are required in order to generate a large volume of numerical information describing practices through the globe. Complex social, economic and political phenomenon, such as state stability and discrimination, are made both easily accessible and globally commensurable through a process of radical simplification and ‘guesstimation’. Distinctive qualities are converted into numerical quantities, which are then compared and assessed in terms of orders of magnitude. Challenging and contested concepts, such as slavery, acquire fixed and unproblematic meanings, which are presumed to be universally applicable irrespective of cultural context. Once it has been crudely converted into numerical values, the Global Slavery Index then goes on to assign countries a ranking, with a ranking of 1 being the worst and a ranking of 167 being the best. This also extends to a ‘Top Ten’ list, which is a popular device used to draw attention to the ‘best’ or ‘worst’ performers. In a world where school and university league tables, university rankings and other metrics have become pervasive, the idea of ranking countries based on performance is a concept we are already very familiar with.

The Global Slavery Index is not a new innovation, but can instead be best understood as the extension of an already well-established template to a new topic. Over the last two decades, there has been a remarkable explosion in the prevalence of national, regional, and global benchmarks. Some notable examples from a much larger trend include the Fragile/Failed States Index (Fund for Peace, from 2005), the Corruption Perceptions Index (Transparency International, 1995), the Climate Change Performance Index (Germanwatch & Climate Action Network Europe, 2006), and the more venerable Freedom in the World report (Freedom House, 1972), the Trafficking in Persons Report (Department of State, United States, 2001), the Human Development Index (United Nations Development Programme, 1990), the Foreign Direct Investment Confidence Index (A. T. Kearney, 1998), and the Bloomberg Innovation Index (Bloomberg, 2013). It is hard to think of any area of International Relations, from international security to global political economy, grand strategy, climate change, human rights, international development, and global public policy, which has not been pulled into this politics of numbers. As this Special Issue demonstrates, this recent popularity can be chiefly traced to the capacity of global benchmarks to cloak normative agendas in languages of neutral and technocratic assessment.

¹ ‘35.8 Million People are Enslaved Across the World’, *Global Slavery Index* (5 November 2014), available at: {<http://www.globalslaveryindex.org/35-8-million-people-are-enslaved-across-the-world/>} accessed 29 July 2015.

² Andrew Guth, Robyn Anderson, Kasey Kinnard, and Hang Tran, ‘Proper methodology and methods of collecting and analyzing slavery data: an examination of the Global Slavery Index’, *Social Inclusion*, 2:4 (2014), pp. 14–22; Ronald Weitzer, ‘Miscounting human trafficking and slavery’, *openDemocracy*, available at: {<https://www.opendemocracy.net/beyondslavery/ronald-weitzer/miscounting-human-trafficking-and-slavery>}; Neil Howard, ‘Keeping Count: The Trouble with the Global Slavery Index’, *The Guardian* (13 January 2014), available at: {<http://www.theguardian.com/global-development-professionals-network/2014/jan/13/slavery-global-index-reports>} accessed 29 July 2015.

³ Anne Gallagher, ‘The Global Slavery Index is Based on Flawed Data – Why Does No One Say So?’, *The Guardian* (28 November 2014), available at: {<http://www.theguardian.com/global-development/poverty-matters/2014/nov/28/global-slavery-index-walk-free-human-trafficking-anne-gallagher>} accessed 29 July 2015.

The normative agendas of global benchmarking

Global benchmarking involves the classification of relative performance or value. In this Special Issue we use *global benchmarking* as an umbrella term for a wide range of comparative evaluation techniques that systematically assess the performance of actors, populations, or institutions. This can include techniques such as audits, rankings, indicators, indexes, baselines, and targets, all of which work on the basis of standardised measurements, metrics, and rankings. We define global benchmarking as a specifically ‘transnational’ practice that goes beyond the jurisdictions of individual states, although particular benchmarks may not be fully ‘universal’ in scope. Benchmarking is further understood as encompassing one or more of the following forms of comparative assessment: (1) quality of *conduct*, or how well actors have discharged their responsibilities in specific areas; (2) quality of *design*, or how well specific policies, laws, or institutions have been formulated and applied; and (3) quality of *outcomes*, or how well activities in specific areas align with defined goals (irrespective of who is actually responsible for the overall outcomes).

The main focus in the Special Issue is on benchmarking by external transnational actors, rather than internal self-benchmarking by states or corporations.⁴ As we discuss further in the lead article for the Issue,⁵ the recent proliferation of global benchmarks has been driven by the political appeal of numbers as information shortcuts that are frequently assumed to present unbiased facts.⁶ This is a recursive process whereby complex and contested normative values are translated into simplified numerical representations, which in turn enables global benchmarks to be represented as ‘evidence’ that can be used to establish a foundation for initiating particular kinds of political conservations as well as potentially influencing the design of policy interventions and reforms. This process of numerical translation helps to obscure the normative values that underlie global benchmarks, and also enables non-experts to make simplistic comparisons of relative performance regarding complex phenomena at a transnational level.

We suggest that this process of numerical translation is common to all forms of global benchmarking, and can be divided into four (non-sequential) components: (1) simplification and extrapolation; (2) commensuration; (3) reification; and (4) symbolic judgement.⁷ There is a distinct form of the ‘politics of numbers’ at work here. In the short term, translating qualities into quantities can potentially expand the political traction that global benchmarks achieve, by increasing their salience at the same time as it simplifies the message they communicate. In the longer

⁴ See, for example, Sakiko Fukuda-Parr, ‘Millennium development Goal 8: Indicators for international human rights obligations?’, *Human Rights Quarterly*, 28:4 (2006), pp. 966–7; Tore Fougner, ‘Neoliberal governance of states: the role of competitiveness indexing and country benchmarking’, *Millennium: Journal of International Studies*, 37:2 (2008), pp. 303–26; Peter Larmour, ‘Civilizing techniques: Transparency international and the spread of anti-corruption’, in Brett Bowden and Leonard Seabrooke (eds), *Global Standards of Market Civilization* (Abingdon: Routledge, 2006), pp. 95–106; Laura Langbein and Stephen Knack, ‘The worldwide governance indicators: Six, one, or none?’, *Journal of Development Studies*, 46:2 (2010), pp. 350–70; Nehal Bhuta, ‘Governmentalizing sovereignty: Indexes of state fragility and the calculability of political order’, in Kevin E. Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry (eds), *Governance by Indicators: Global Power through Quantification* (Oxford: Oxford University Press, 2012), pp. 132–62.

⁵ See André Broome and Joel Quirk, ‘Governing the world at a distance: the practice of global benchmarking’, *Review of International Studies*, 41:5 (2015), pp. 819–41.

⁶ Lorenzo Fioramonti, *How Numbers Rule the World: The Use and Abuse of Statistics in Global Politics* (London: Zed Books, 2014), p. 192.

⁷ For a full explanation of these concepts, see Broome and Quirk, ‘Governing the world at a distance’.

term, global benchmarks can potentially create ‘anchoring effects’ by establishing referents that shape how different actors subsequently think about and ‘see’ specific issues.⁸ Different forms of global benchmarking may sometimes operate in competition, especially when benchmark producers are seeking to construct issue expertise or to otherwise compete for authority in a given area.⁹ Because individual global benchmarks often share a common normative orientation, however, they therefore contribute to the diffusion of overlapping normative agendas regarding what particular kinds of transnational actors should look like, what they should value, and how they should behave.

Towards a new IR research agenda on global benchmarking

Global benchmarking represents a new and distinctive application of authority in world politics. Benchmarking differs from earlier efforts to promulgate a ‘standard of civilization’ based on Western values,¹⁰ as well as more recent attempts to diffuse global ‘best practice’ policy norms through international regimes and organisations.¹¹ Because the quantification of particular norms and values relies primarily on the imprimatur of scientific expertise for establishing authority, and because this authority to make symbolic judgements is typically expressed through the language of numbers, global benchmarks are both easier to promulgate and, potentially, more difficult for target actors to ignore. It may be harder to argue with a numerical ranking, especially when this serves as a reference point for political debates about a given issue, than it is to resist technical advice from international actors on complex institutional reforms, except through deploying rival expertise and challenging the credibility of a global benchmark on the same ontological terrain.

Collectively, the nine articles in this Special Issue represent an important step in advancing the theorisation and empirical analysis of the politics of numbers and the dynamics of global benchmarking in IR. In the lead article for the Issue we seek to place global benchmarking on the IR agenda in accessible and general terms through developing a conceptual apparatus for the study of benchmarking, and by mapping out a fourfold typology for distinguishing between different types of benchmarking as a transnational practice.¹² This is followed by Alexandra Homolar’s article on human security benchmarks, which provides a robust critical examination of how the role of benchmarking within the wider human security agenda has served to reinforce the state as the main focal point of international security governance. Homolar’s analysis turns much of the conventional wisdom about human security in IR on its head, and shows how the emancipatory potential often associated with the shift towards the human security agenda has been undermined. As a consequence, global humanitarian

⁸ Peter Andreas and Kelly M. Greenhill, ‘Introduction: the politics of numbers’, in Peter Andreas and Kelly M. Greenhill (eds), *Sex, Drugs, and Body Counts: The Politics of Numbers in Global Crime and Conflict* (Ithaca, NY: Cornell University Press, 2010), p. 17; see also André Broome and Leonard Seabrooke, ‘Seeing like an international organisation’, *New Political Economy*, 17:1 (2012), pp. 1–16.

⁹ See Leonard Seabrooke, ‘Epistemic arbitrage: Transnational professional knowledge in action’, *Journal of Professions and Organizations*, 1:1 (2014), pp. 49–64; Ole Jacob Sending, *The Politics of Expertise: Competing for Authority in Global Governance* (Ann Arbor: University of Michigan Press, 2015).

¹⁰ See Shogo Suzuki, *Civilization and Empire: China and Japan’s Encounter with European International Society* (Abingdon: Routledge, 2009).

¹¹ See Broome and Seabrooke, ‘Seeing like an international organisation’.

¹² Using this typology, we have compiled a Global Benchmarking Database consisting of 205 benchmarks (as of June 2015), which is available at: {www.warwick.ac.uk/globalbenchmarking/database}.

governance has been inscribed with new power relations centred on the attempt to fix a particular conceptualisation of a 'secure' human life.

Tony Porter's examination of global benchmarking networks helps to unpack the organisational contexts of benchmarking. Using actor-network theory, Porter provides a systematic analysis of two key and underexplored features of global benchmarking: network embeddedness, and network publicness. The article demonstrates how the logic of benchmarking can create an impetus towards increased publicness, whereby benchmarking networks are characterised by expanding flows of information across widening circles of engagement. Rather than simply focussing on the power of benchmark producers, the relative weakness of target actors, or the formal design of benchmarking metrics, Porter makes an important and compelling argument of the need for shifting the focus of analysis towards the organisational contexts in which benchmarking networks operate.

In their investigation of how activists use benchmarking initiatives, Leonard Seabrooke and Duncan Wigan characterise benchmarks as a form of 'symbolic violence' that NGOs utilise as weapons to exert pressure on a range of transnational actors, such as firms, states, and international organisations. Drawing on the work of Pierre Bourdieu, Seabrooke and Wigan's analysis of the role of benchmarking in transnational advocacy strategies helps to explain three constituent elements for successful benchmarking by NGOs: salience, will, and expertise. Differentiating between two types of benchmarking cycles, 'reformist benchmarking' and 'revolutionary benchmarking', Seabrooke and Wigan's contribution to the Issue helps to shed new light on the conditions under which activist organisations' campaigns on complex issues can be mounted and sustained using benchmarking practices.

Genevieve LeBaron and Jane Lister's article critically evaluates the power relations inherent in the 'ethical' compliance audit regime, focusing on the effectiveness of supply chain benchmarking as a mechanism for promoting environmental and social improvements in global retail supply chains. As LeBaron and Lister's research reveals, while benchmarks and audits have become increasingly trusted forms of corporate governance by governments and the public, they can function to obscure serious problems in global supply chains due to selective coverage and the challenges inherent in regulating labour rights and safeguards through numbers compared with capital and product quality. Within the context of a broader trend towards increasing reliance on global supply chain benchmarking as a form of transnational corporate governance, LeBaron and Lister's article makes an important intervention that highlights the potential for benchmarking to be used to expand corporate power over consumers as well as national policymakers, and helps to focus attention on the dangers inherent in corporate 'self-regulation' at the transnational level.

James Harrison and Sharifah Sekalala's article critically examines how global benchmarking initiatives by the United Nations (UN) help to promote compliance with international human rights norms. Through a comparative analysis of benchmarks produced by the UN for states and those produced for corporations, Harrison and Sekalala reveal the limits of global benchmarking as an effective mode of transnational governance. In particular, they explore how global benchmarking initiatives can be used to provide 'superficial legitimation' for the human rights performance of corporations. Rather than helping to close 'the compliance gap' identified in the implementation of international human rights norms, Harrison and Sekalala show how benchmarks based on self-reporting initiatives can instead create opportunities for powerful actors to evade their obligations by gaming the system.

Liam Clegg investigates the political contestation of how the success of the Millennium Development Goal (MDG) framework was benchmarked in UN reviews over the last fifteen years. Clegg shows that reviews by the UN of the progress achieved under the MDG framework was riven by a series of 'blame games' over how responsibility for poor performance was attributed. He identifies a clear split between developed and developing countries, with the former blaming internal governance failings in developing countries for poor performance, while the latter blamed aid shortfalls for a lack of success in achieving MDG targets. As Clegg's analysis shows, the problems caused by the imbrication of benchmarking and blame games over how responsibility is attributed for the success or failure of global development initiatives are likely to be carried over into the design of the new Sustainable Development Goals.

Caroline Kuzemko's analysis of European Union (EU) climate change benchmarking practices reveals the competing motivations that lie behind the development of climate governance, and shows how climate change benchmarking practices function to increase the legibility of certain political and economic practices, while obscuring others. Through a comparative analysis of two cases of 'best practice' countries – the United Kingdom and Germany – Kuzemko demonstrates how the quantification processes and short time horizons inherent in climate change benchmarking can give a distorted picture of domestic compliance with EU climate change initiatives. As a consequence, Kuzemko argues that climate change benchmarking often serves to gloss over important political issues for long-term climate mitigation, while simultaneously side-lining more radical policy alternatives.

The final article in the Issue, by Ole Jacob Sending and Jon Harald Sande Lie, critically examines how the World Bank benchmarks African economies through its Country Policy and Institutional Assessment ratings. Sending and Lie show how benchmarking connects policy and practice. In particular, their analysis shows that while benchmarks may be used to police and shape target actors' behaviour in some cases, in others they are used to manage and produce existing relations of power, thereby constituting particular forms of authority between actors based on their respective identities.

The key conceptual themes developed throughout the Special Issue therefore centre on the role of global benchmarking practices in attributing responsibility for good or bad performance, constructing issue expertise by both state and non-state actors, altering the terms of political rhetoric, shaping formal policy agendas, and defining the concepts used to quantify international norms. Many of the contributions also provide new insights into how global benchmarking practices may constitute new sets of relations between different types of transnational actors, and explore the role of 'third parties' who use benchmarks produced by others for a variety of political purposes. Covering a wide range of rich empirical terrain, the contributions to this Issue help to expand our understanding of how global benchmarking practices have emerged as a critical mode of transnational governance in the twenty-first century.