

The Power of Global Performance Indicators

Global performance indicators (GPIs), such as ratings and rankings, permeate nearly every type of human activity, internationally and nationally, across public and private spheres. While some indicators aim to attract media readership or brand the creator's organization, others increasingly seek to influence political practices and policies. The Power of Global Performance Indicators goes beyond the basic questions of methodological validity explored by others to launch a fresh debate about power in the modern age, exploring the ultimate questions concerning real world consequences of GPIs, both intended and unintended. From business regulation to terrorism, education to foreign aid, Kelley and Simmons demonstrate how GPIs provoke bureaucracies, shape policy agendas, and influence outputs through their influence on third parties such as donors and market actors and, potentially, even broader global authority structures.

Judith G. Kelley is Terry Sanford Professor of public policy and political science and Dean of the Duke Sanford School of Public Policy. She is the author of Scorecard Diplomacy: Grading States to Influence Their Reputation and Behavior (Cambridge University Press) and Monitoring Democracy: When International Election Observation Works and Why It Often Fails (Princeton University Press) winner of the Chadwick F. Alger Prize.

Beth A. Simmons is Professor of Law, Political Science and Business Ethics at the University of Pennsylvania and an elected member of the National Academy of Sciences. She is the author of *Mobilizing for Human Rights: International Law in Domestic Politics* (Cambridge University Press) recognized by the American Political Science Association and American Society for International Law, among others, as the best book of 2010.





The Power of Global Performance Indicators

Edited by

JUDITH G. KELLEY

Duke University

BETH A. SIMMONS
University of Pennsylvania





CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org Information on this title: www.cambridge.org/9781108487207 DOI: 10.1017/9781108763493

© Cambridge University Press 2020

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2020

Printed in the United Kingdom by TJ International Ltd, Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Kelley, Judith Green, editor. | Simmons, Beth A., 1958- editor.
Title: The power of global performance indicators / edited by Judith
Kelley, Duke University, North Carolina, Beth Simmons, University of
Pennsylvania.

Description: New York: Cambridge University Press, 2020. | Includes bibliographical references and index.

Identifiers: LCCN 2019044097 | ISBN 9781108487207 (Hardback) | ISBN 9781108732741 (eBook)

Subjects: LCSH: Political planning--Evaluation. | Ranking and selection (Statistics)--Political aspects.

Classification: LCC JF1525.P6 P69 2020 | DDC 320.6--dc23 LC record available at https://lccn.loc.gov/2019044097

ISBN 978-1-108-48720-7 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.



Contents

Lis	t of Figures	page vii
Lis	t of Tables	xi
Lis	t of Contributors	xiii
Preface		XV
Ι	Introduction: The Power of Global Performance Indicators <i>Judith G. Kelley and Beth A. Simmons</i>	I
PAI	RT I RATINGS, RANKINGS, AND REGULATORY BEHAVIOR	
2	The Power of Ranking: The Ease of Doing Business Indicator and Global Regulatory Behavior Rush Doshi, Judith G. Kelley, and Beth A. Simmons	25
3	Blacklists, Market Enforcement, and the Global Regime to Combat Terrorist Financing Julia C. Morse	61
4	The Power of Indicators in Global Development Policy: The Millennium Development Goals Helena Hede Skagerlind	98
PAI	RT II THE NORMATIVE INFLUENCE OF RATINGS AND RANKINGS	
5	A Race to the Top? The Aid Transparency Index and the Social Power of Global Performance Indicators Dan Honig and Catherine Weaver	139
6	International Assessments and Education Policy: Evidence from an Elite Survey Rie Kijima and Phillip Y. Lipscy	174

V



vi		Contents
7	Reporting Matters: Performance Indicators and Compliance in the International Labor Organization (ILO) Faradj Koliev, Thomas Sommerer, and Jonas Tallberg	203
8	Freedom House's Scarlet Letter: Assessment Power through Transnational Pressure Jordan Roberts and Juan Tellez	242
	RT III BEYOND AND WITHIN STATE: INFLUENCES AND IMPACTS OF N-STATE ACTORS	N
9	Can Blacklisting Reduce Terrorist Attacks? The Case of the US Foreign Terrorist Organization (FTO) List Hyeran Jo, Brian J. Phillips, and Joshua Alley	271
10	Assessing International Organizations: Competition, Collaboration, and Politics of Funding <i>Ranjit Lall</i>	300
	RT IV SKEPTICAL VOICES: NULL RESULTS, UNINTENDED NSEQUENCES	
II	Third-Party Policymakers and the Limits of the Influence of Indicators Melissa M. Lee and Aila M. Matanock	341
12	The Millennium Development Goals and Education: Accountability and Substitution in Global Indicators James H. Bisbee, James R. Hollyer, B. Peter Rosendorff, and James Raymond Vreeland	375
13	Global Performance Indicators: Themes, Findings, and an Agenda for Future Research <i>Judith G. Kelley and Beth A. Simmons</i>	409
Inc	lex	425



Figures

I.I	GPI contestation across issues, time, and creators	page 4
1.2	From GPIs to policy change: pathways, policy response,	
	and reassessment	10
2.1	Theory of the social influence of the Ease of Doing	
	Business ranking	28
2.2	Doing Business website visits, annually (2003–2016)	32
_	Days to start a business	41
	Procedures to start a business	41
3.1	The figure shows the number of non-listed countries	
	(dark grey, bottom of stacked bars) and listed countries	
	(light grey, top of stacked bars) by the total number of	
	failing ratings	69
3.2	The figure shows the percent of nonlisted countries (solid line)	
	and listed post-2009 countries (dashed line) that have adopted	
	FATF-compliant laws on terrorist financing. The dotted vertical	
	line indicates the 2009 announcement of the revamped FATF	
	noncomplier process that issued its first noncomplier list in	
	February 2010	76
3.3	The figure shows the change in the cumulative probability of	
	criminalizing terrorist financing in line with FATF guidelines,	
	comparing listed countries with high market integration	
	(one standard deviation above the mean) with listed countries	
	with low market integration (one standard deviation below	
	the mean). First difference calculations estimated using the	
	results of a cox proportional hazards model (Model 4 in	
	Table 3.1). Dotted lines show the 95 percent confidence interval,	
	calculated using a Monte Carlo simulation sampling over	
	500 iterations	83

vii



viii List of Figures

3.4	The figure shows cross-border liabilities (billions of US dollars)	
	from 2009 to 2015. The FATF listed Thailand in February 2010	
	and in February 2012, placed Thailand on its "black" list for	
	failing to improve its laws in a timely fashion. Following	
	significant legal changes, the FATF removed Thailand from its	
	monitoring process in June 2013	92
	MDG 3 policy goal adoption, mean for 15 countries	104
	MDG 3 policy goal adoption, by goal	105
	Adoption of implementation plans, girls' education	106
	Adoption of implementation plans, women's wage employment	106
4.5	Adoption of implementation plans, women's political	
	participation	107
4.6	Adoption of policy goals and implementation plans, mean for	
	15 countries	108
	Adoption of policy goals and implementation plans, Kenya	117
-	Adoption of policy goals and implementation plans, Ethiopia	123
-	Histogram of overall ATI scores, 2011–2016	150
	Net change by agency on ATI over coverage period	151
5.3	Marginal effect of treatment by independence for non-aid	
	agencies	155
	Marginal effect of treatment by independence for aid agencies	156
6.1	The number of participants in cross-national assessments in	
	education (1959–2012)	178
	Marginal effects on CIRI improvements (Model 1)	227
7.2	Conditional marginal effects of ILO reporting. (a) Democracy	
	and (b) Capacity	230
	Verbal conflict events by CAMEO category	257
8.2	The Freedom House classification discontinuity.	
	(a) Discontinuity: Freedom score and average verbal conflict	
	from democracies, (b) edge cases: Verbal conflict from	
	democracies and non-democracies, and (c) distribution of	
0 .	edge cases over time	258
_	Scarlet letter effect across sources of verbal conflict	260
	Temporal decay of the Scarlet Letter effect	261
_	Placebo tests for the Scarlet Letter effect	262
9.1	Estimated impact of FTO status on group terrorist attacks.	
	Estimates from Model 5 in Table 9.1, with other variables at	- 00
	their means Townsiet attacks and funding types Al Ocada McV and	288
9.2	Terrorist attacks and funding types. Al-Qaeda, MeK, and FARC. Vertical dashed line indicates the year of FTO	
	designation. Of the three groups, only MeK was heavily and	
	singularly dependent on private funding such as donations	202
TO 7	Relationship between performance ratings and changes in	292
10.1	contributions	307
	V01111110H110110	1 ∨ /



List	of Figures	ix
10.2	Funding trends consistent with indicator influence	309
10.3	Funding trends inconsistent with indicator influence	310
10.4	Marginal effects of average ratings on changes in	
	contributions at different levels of institutional competition	
	and partnership robustness	325
II.I	Number of countries in the Corruption Perceptions Index	350
11.2	Media coverage of the Corruption Perceptions Index	
	(bottom ten countries compared to the next ten countries)	352
11.3	Media coverage of the WGI's control of corruption	
	(bottom ten percent of all countries compared to the next	
	ten percent)	353
11.4	CPI: First differences results. (a) Δ CPI rank and	
	(b) Δ CPI score	360
11.5	CPI: Levels results. (a) CPI rank and (b) CPI score	360
11.6	CPI: Sectoral aid results. (a) Aid prone to capture,	
	CPI rank, (b) Aid prone to capture, CPI score, (c) Aid not prone	
	to capture, CPI rank, and (d) Aid not prone to capture,	
	CPI score	362
	Histogram of change in ranks	363
	CPI rank change results	365
	CPI big improvers results	365
12.1	The pathways of GPIs to policy change. Our chapter	
	focuses on the unintended consequences reprioritization can	
	have on state behavior (highlighted in black)	379
12.2	Linear marginal effects plots characterizing heterogeneity in	
	effect of MDG on primary enrollment rates over values of	
	democracy (a) and transparency index (b)	396
12.3	First stage (a) and overall (b) results from mediation analysis	
	on the mechanism of foreign aid. The first stage results show	
	that foreign aid for primary education increases following	
	MDG adoption, particularly for countries in the bottom	
	quartile of primary enrollment rates. However, the overall	
	results are only marginally significant (pseudo p-values	
	presented in the legend) and represent a small fraction of the	
	overall relationship, suggesting that this mechanism is inactive	403





Tables

I.I	Chapter overview of mechanisms and topics	page 14
2.1	Market share of the ease of doing business index	32
2.2	Countries with reform committees directly using the EDB data	37
2.3	The efficiency of bureaucratic reforms for ascending the	
	rankings	38
2.4	Overview of de jure reform measures (dependent variables)	40
2.5	The effect of ranking on reductions in time, cost, and	
	procedures for starting a business	43
2.6	Experimental results of ranking differences on investment	
	likelihood	47
3.I	Listing, market enforcement, and criminalization:	
-	Cox proportional hazards models for full sample – Hazards	
	ratios for cox proportional hazards models	81
3.2	Listing, market enforcement, and criminalization: Cox	
-	proportional hazards models for matched sample and	
	eligible-for-listing sample – Hazards ratios for cox proportional	
	hazards models	84
3.3	Listing, market enforcement, and criminalization: Cox	
,	proportional hazards models for US power alternatives –	
	Hazards ratios for cox proportional hazards models	85
3.4	The effect of listing on cross-border liabilities – Dependent	,
	variable is logged cross-border liabilities	88
4.I	Observable implications of causal mechanisms	113
	2016 ATI indicators and weights	145
	Summary statistics of selected variables	149
-	Performance on the ATI	151
	ATI as treatment	153
	Perceptions of ATI influence on donor transparency behavior	158
	1 /	,

хi



xii	I	List of Tables
6.1	Survey analysis – Elite response (Ordered probit)	194
6.2	Survey analysis – Domestic politics (Ordered probit)	195
6.3	Survey analysis – Transnational pressure (Ordered probit)	196
7.1	CIRI worker rights coding	216
7.2	Logistic regression of CIRI improvements, 1991–2011	220
7.3	Logistic regression of CIRI improvements, length of report	ing,
	1991–2011	222
7.4	Logistic regression of CIRI improvements, interaction effect	ets,
	1991–2011	224
7.5	Logistic regression of ILO targeting	232
9.1	Statistical analysis of terrorist attacks and FTO designation	n 287
9.2	Listing vs. sanctioning effect Analysis of Pre and	
	Post-9/11 Subsamples	291
10.1	Summary of IO performance assessments	304
10.2	Effect of performance ratings on changes in contributions	
	as mediated by institutional competition (Equation 10.1)	322
10.3	Effect of performance ratings on changes in contributions a	ıs
	mediated by governance partnerships (Equation 10.2)	326
10.4	Effect of pooled British ratings on changes in contributions	
	as mediated by institutional competition and governance	
	partnerships (Equations 10.3 and 10.4)	330
	Media coverage of the CPI, 1995–2014	351
	Naming and shaming results	356
	Millennium Development Goals	380
12.2	Summary statistics for pre- and post-millennium declaration	on
	by sample	393
12.3	The relationship between the Millennium Declaration and	the
	primary school enrollment rate	395
12.4	The relationship between the Millennium Declaration and	
	the secondary school enrollment rate	397
12.5	Coefficients relating the Millennium Declaration to the	
	difference between primary and secondary enrollment rates	S
	across countries	399
12.6	$\boldsymbol{\beta}$ substitution tests for equilibrium levels of secondary and	
	primary education	401



Contributors

Joshua Alley

Texas A&M University, USA

James H. Bisbee

New York University, USA

Rush Doshi

Brookings Institution, Washington, DC, USA

James R. Hollyer

University of Minnesota, USA

Dan Honig

Johns Hopkins School of Advanced International Studies, Washington, DC

Hyeran Jo

Texas A&M University, USA

Judith G. Kelley

Duke Sanford School of Public Policy, USA

Rie Kijima

University of Toronto, Canada

Faradj Koliev

Stockholm University, Sweden

Ranjit Lall

London School of Economics, UK

Melissa M. Lee

Princeton University, USA

xiii



xiv

List of Contributors

Phillip Y. Lipscy

University of Toronto, Canada

Aila M. Matanock

University of California, USA

Julia C. Morse

University of California, Santa Barbara

Brian J. Phillips

University of Essex, UK

Jordan Roberts

Duke University, USA

B. Peter Rosendorff

New York University, USA

Beth A. Simmons

University of Pennsylvania, USA

Helena Hede Skagerlind

Stockholm University, Sweden

Thomas Sommerer

Stockholm University, Sweden

Jonas Tallberg

Stockholm University, Sweden

Juan Tellez

Duke University, USA

James Raymond Vreeland

Princeton University, USA

Catherine Weaver

University of Texas at Austin Lyndon B. Johnson School of Public Affairs, USA



Preface

The seeds of this volume were planted early in 2011 when we were working on a project about the US efforts to combat human trafficking. That project was focused on the rating system of the US Department of State, which assesses government efforts to combat human trafficking. We wondered, does that "tier" system have any influence on a country's propensity to criminalize human trafficking? Halfway through the project, in the fall of 2011, we were skyping across the Atlantic in a series of intense revision efforts, when we began to realize that we might be onto something larger than this specific example. There was something about the process of rating that went beyond information gathering; it seemed to constitute social pressure, maybe even an effort at governance. While we had worked before on the effects of monitoring and global norms on state behavior and studied causal mechanisms from reputational concerns to domestic mobilization around global norms, this phenomenon seemed to bring together many of these past insights into a unique combination of causal mechanisms.

Soon thereafter, we were both grateful for the opportunity to attend a conference on the construction of performance indices at New York University in 2014 that was organized by Sally Merry and Benedict Kingsbury. The more we started to think about it, the more we were intrigued about the phenomenon of ratings and rankings.

Although a substantial body of work on such ratings and rankings – or indicators as they were often referred to – was developing, much of the work assumed that these indices mattered, and, therefore, they explored their construct validity, deconstructed their methodologies, and debated their normative content. Only a very few studies attempted to assess or theorize effects on state behaviors, and the ones that did typically did not test behavioral outcomes. We made this the goal of our research on the US tier ranking of countries for their efforts to combat human trafficking.

xv



xvi Preface

Our research led us to see just how widespread the phenomenon of rating and ranking states was. Indeed, the practice of arraying information to elicit explicit comparisons seemed to be spreading among states, international organizations, and non-state actors. In order to better understand this phenomenon, we conducted interviews in London, Washington, DC, and New York during the summer of 2014, and spoke with over forty individuals from twenty different organizations. These interviews were incredibly rich, allowing us to gain an appreciation for the creators' goals, and for the very intentional strategies to deploy indices to shape not only organizational brands and discourse but also actual policy outcomes. The raters explained how states reacted to being ranked, as well as to specific rankings. It started to appear plausible to us that countries cared about how they compared to others.

Our own analysis of the *Trafficking in Persons Report* revealed that the State Department's rating system had significantly impacted countries' criminalization of human trafficking through multiple causal mechanisms.¹ We were struck by several aspects of the State Department's exercise, which helped us detect and define a broader ranking phenomenon. We felt the flurry of ranking creation and contestation was increasingly going well beyond measurement; it represented a governance struggle that was broadly understood. First, ratings and rankings were *intentional* in their efforts to influence policies, by engaging their targets' sustained attention. This also meant such information had to be *public*. Second, many ratings and rankings schemes were *recurrent*, which appeared to be important for their effectiveness in invoking the enduring reputational concerns that motivated sometimes difficult change over time.

To explore the broader applicability of our ideas, we organized a call for papers for a conference on global performance indices, sponsored by the journal *International Organization*. For focus, we specifically asked that papers examine ratings and rankings as explanatory variables, and examine consequences, for example, on state policies, power relationships, rule-setting, or dominant discourse. We encouraged papers that used a variety of methodological approaches from single-case studies to large N studies, to experiments or statistical analyses. The call elicited nearly 100 proposals. The conference, "Assessment Power in World Politics," was held at the Harvard Weatherhead Center for International Affairs (WCFIA) in May 2016 and was followed in September by a mini-conference at the annual conference of the American Political Science Association (APSA) that same year. These conferences brought together scholars who worked on a wide range of assessments of state qualities (e.g. transparency), state policies (e.g. press freedom), or

¹ This work was eventually published as "Politics by Number: Indicators as Social Pressure in International Relations," in the *American Journal of Political Science*, Vol 59(1): 1146–1161, 2015, and as *Scorecard Diplomacy: Grading States to Influence Their Reputation and Behavior*, Cambridge University Press, 2017.



Preface xvii

prevalent social practices within a state's jurisdiction (e.g. corruption). Papers contributed theoretical and empirical insights about the use and consequences, or lack thereof, of indicators for domestic politics, transnational and/or international relations. While a subset of the contributions to this volume was published as a symposium with *International Organization*, we are delighted to be able to gather the much broader set of contributions in this edited volume. Together, these contributions present substantial evidence that global performance indicators – ratings and rankings – can influence the behavior of states. They also demonstrate the varied mechanisms of influence, and begin to probe the conditions for their effectiveness. As a set of studies, these chapters make a stronger argument than our sole focus on human trafficking ever could: global performance indicators are a tool of global governance, and rankings often do more than help "brand" organizations. They often alter the normative discourse around a subject, create competition around ranking status, and alter the incentives for specific policies.

We have a lot of people to thank for the overall development of this volume. Sam Chase and Nadia Hajji helped us build the database of indicators on which the introduction is based. We are grateful to the following individuals whom we interviewed during the summer of 2014: Alexandra Gillies (PhD) with the Natural Resource Governance Institute which publishes *The* Resource Governance Index (RGI); Clare Doube with Amnesty International (AI), whom we interviewed about why AI does not have an index; Johan Gott, at A. T. Kearney, which publishes the FDI Confidence Index; David Roodman, who created the Commitment to Development Index for the Center for Global Development; and John Easton, Dan McGrath, and Dana Kelly from the National Center for Education Studies (NCES) who spoke with us about the US interaction with the PISA Index. We are grateful to Rita Ramalho, Nadine Shamounki Ghannam, and Betty Mensah at the World Bank, who spoke with us about the Ease of Doing Business (EDB), and to Jean-François Arvis and Christina Busch who explained the Bank's Logistics Performance Index to us. We are also grateful to Adriana Alberti, Richard Kerby, Zamira Dzhusupova, Anni Haataja, Deniz Susar, and Tyko Dyrksmeyer who helped produce the E-Governance Index for the United Nations Department of Economic and Social Affairs, and to Daniel Calingaert who spoke with us about the Freedom House's Freedom in the World Report. In addition, we thank the Hudson Institute's Kimberly Russell and Carol Adelman, who created the *Index of Global Philanthropy and Remittances*, and J. J. Messner and Nate Haken at The Fund for Peace, which publishes The Fragile States Index; Jeremy Tamanini with Dual Citizen LLC, which creates the Global Green Economy Index, and Nathaniel Heller and Hazel Feigenblatt, who help Global Integrity publish the Global Integrity Index. At the Heritage Foundation, which produces the *Index of Economic Freedom*, we are thankful to Anthony Kim, Terry Miller, and Kim Holmes. We also thank Klaus von Grebmer who works for the International Food Policy Research Institute



xviii Preface

and answered our questions about the Global Hunger Index; Leon Morse at International Research and Exchanges Board (IREX) who spoke with us about the Media Sustainability Index; and at Publish What You Fund, which produces the Aid Transparency Index, we thank Rachel Rank, Shreya Basu, and Mark Brough. We are grateful also to Adam Foldes at Transparency International, which produces the Corruption Perceptions Index, Milorad Kovacevic with the United Nations Development Program (UNDP) who talked to us about the Human Development Report Office, and Alejandro Ponce and Juan Botero at the World Justice Project, which publishes the Rule of Law Index. Finally, we thank Vivek Ramkumar and Michael Castro at the International Budget Partnership, which creates the Open Budget Survey, and Nathan Gamester and Novella Bottini at the Legatum Institute, which creates the Legatum Prosperity Index.

In addition to the many contributors to this volume, we also thank the other participants at the 2016 conferences at Harvard and APSA: Angelina Fisher, Nancy Green Saraisky, Laura A. Henry, Lisa Sundstrom, Steven Bernstein, David J. Gordon, Matthew Hoffman, Bradley C. Parks, and Takaaki Masaki, for their knowledge on ratings and rankings and their theoretical insights in early discussions. We also thank the paper discussants at these conferences, including Felipe Barrera-Osorio, Daniel Drezner, Martha Finnemore, Jeff Frieden, Thomas Gift, Alicia Harley, Connor Huff, Iain Johnston, Joshua Kertzer, Elise Le, Lucas Linsi, Christopher Lucas, John Marshall, Sally Merry, Richard Nielsen, Solé Prillaman, Karthik Ramanna, Anton Strezhnev, and Dustin Tingley. Robert Keohane and Judith Goldstein also gave us extensive advice on the project.

In addition to the support we received from *International Organization* to hold a 2016 workshop at Harvard, we are grateful to the Weatherhead Center for International Affairs (WCFIA) and to the Institute for Quantitative Social Science, both at Harvard, for supplemental funding. We are also grateful to WCFIA for the staff and logistical support to execute the event smoothly. Thanks also go to APSA for supporting a mini-conference that same year. Several of the contributions in this volume also benefitted from rich reviewer feedback from *International Organization*, and we are grateful to the editors and to the anonymous reviewers. We are grateful to Cambridge University Press for allowing us to include chapters in this volume originally printed in the symposium in *International Organization*, to our Cambridge University Press editor, Robert Dreesen, for shepherding this volume along, and for the production assistance of Jackie Grant at Cambridge University Press.

Beth A. Simmons would like to dedicate this book to her husband, Bruce Jackan. Judith G. Kelley would like to dedicate it to her husband, Michael Kelley.