7
Seven *

[Tentative Tenets on Governance Indicators]

Daniel Kaufmann, Brookings Institution
dkaufmann@brookings.edu
http://www.brookings.edu/experts/kaufmannd.aspx

Presentation at APSA Conference on Democracy Audits and Governmental Indicators, at UC Berkeley, October 29-30, 2009

* At one level, this numerical (7) as a heading strictly reflects the number of tenets raised in this presentation and in the accompanying brief paper (at http://indicatorsinfo.pbworks.com/f/Kaufmann.pdf). At another, the title signals the need to increase the ratio of data points per text words in assessments and critiques of governance indicators.
1. The ‘True’ Quality of Governance is Inherently Unobservable.

- What is the ‘quality of democracy’ in a country?
- Or its actual extent of corruption?
- Or the effectiveness of its government and of its rule of law?
- The ‘true’ level of governance cannot be directly observed, due to: i) complex and multi-dimensional nature; ii) incentive to hide actual governance, iii) lack of definitional consensus about what constitutes governance and its key components, and, iv) measurement challenges
- Corollaries of Unobservability: i) Trivial: fallacy in ‘my’ index being real thing vs. everybody else’s,
  ii) Non-Trivial: Margins of Error...
2. Transparency and Precision about the Imprecision in Governance Indicators

- Taking margins of error seriously in all Indicators.
- Observed governance is a proxy for ‘true’ governance.
- This means that each proxy is an imprecise approximation of the true value.
- Any measured governance variable has a margin of error around its observed estimate.
- Such margin of error can be estimated, and disclosed.
- Disclosure and Estimation of Margins of Error rarely done.
- Yet extremely important--for appropriate use and interpretation by researchers, advocacy groups, policy analysts and policy-makers.
- Margins of error (and thus uncertainty/confidence intervals) do inform: what inferences can and cannot be made; which cross-country and over time comparisons are statistically meaningful, and which ones are not.
- Exhibits
Taking Margins of Error Seriously (1), Control of Corruption in WGI, selected countries, 2008

Source for data: 'Governance Matters VIII: Governance Indicators for 1996-2008', by D. Kaufmann, A. Kraay and M. Mastruzzi, June 2009, www.govindicators.org. Colors are assigned according to the following criteria: Dark Red: country is in the bottom 10th percentile rank ('governance crisis'); Light Red: between 10th and 25th percentile rank; Orange: between 25th and 50th percentile rank; Yellow, between 50th and 75th; Light Green between 75th and 90th percentile rank; and Dark Green: between 90th and 100th percentile (exemplary governance). Estimates subject to margins of error.

Comparison between 2008, 2003, 1998 (top-bottom order)

Source for data: "Governance Matters VIII: Governance Indicators for 1996-2008", D. Kaufmann, A. Kraay and M. Mastruzzi, June 2009 (http://www.govindicators.org). Colors are assigned according to the following criteria: Dark Red, bottom 10\textsuperscript{th} percentile rank; Light Red between 10\textsuperscript{th} and 25\textsuperscript{th}; Orange, between 25\textsuperscript{th} and 50\textsuperscript{th}; Yellow, between 50\textsuperscript{th} and 75\textsuperscript{th}; Light Green between 75\textsuperscript{th} and 90\textsuperscript{th}; Dark Green above 90\textsuperscript{th}. 
Uncertainty in Ibrahim Index of African Governance (IIAG), Safety & Rule of Law Component -- Margin of Errors from Missingness and Measurement (M&M) 2007/08

Horizontal lines indicate terciles and median (56.0):
- Top tercile (62.4): 18 countries included; 4 countries with high confidence
- Middle Tercile (50.8-62.4): 17 countries included; no countries with high confidence
- Bottom tercile (50.8): 18 countries included; 6 countries with high confidence
Uncertainty in IIAG Human Development Margin of Errors (M&M) Graph, 2007/08 Period

Horizontal lines indicate terciles and median (49.9):
• Top tercile (53.1): 18 countries included; 5 countries with high confidence
• Middle Tercile (45.9-53.1): 17 countries included; no countries with high confidence
• Bottom tercile (45.9): 18 countries included; no countries with high confidence
When is a Change Significant, when it is not?
Changes over Time and Uncertainty/Margins of Error in IIAG Composite, 2000/01-2007/08 Periods

*Margins of errors (M&M) provided for countries experiencing greater than 6 point changes in governance over time

*M&M stands for Measurement and Missing Variable Uncertainty
3. Benefits of Aggregation in Composite Indicators.

- Appreciated: virtues of disaggregated indicators
- Less appreciated: virtues of aggregate indicators
- Aggregation of many independent proxies for the same concept (say, corruption) reduce margins of error (and less nonsensical outliers).
- Composite of 4 independent sources: standard error one-half that of a single source.
- Also, serves as organizing framework for disparate array of data.
- Particularly useful when broad concept, such as democracy or rule of law (vs. formal elections or trust in the police).
- Aggregation can also make sense if notion is specific
- Complementarities: Aggregate Indicators are composites of Disaggregate, both available
4. To Weight or Not to Weight variables in aggregating?: That is [Not] the Question. 

Instead, the Question is: What is In, and what is Out?

- Reality: In aggregating, weights always assigned to individual variables
- Issue is whether such weights are explicit, or implicit
- If latter, equal weights (most often)
- More important than weights of included variables: what was left out altogether, vs. what was included?
- Errors of Omission: less focus than errors of commission
- But if omission: implicit zero weight, making large difference vs. inclusion (eg CPI vs WGI vs IIAG)
- Unintended consequences of ‘Zombie’ variables: subverting weights
5. **Expert Polls vs. Citizen/ Surveys vs. ‘Objective’ Data: A Moot Contest?**

- Each method to obtain data has merits
- And their own margins of error (and biases)
- Large Variation in imprecision *within* each category (experts, surveys, and ‘objective’ data)
- It dwarfs differences in imprecision *across* the 3 categories.
- Within each category, there are noisy and less noisy indicators
- This is not surprising: each type has ‘pros’ and ‘cons’
  - e.g. expert polls benchmarking vs. high N of surveys
  - ‘Objective’ data: element of judgment; de jure (vs. de facto, which captures informality)
- Exhibits
Subjective and Objective Measures of Ease of Business Entry: Developing Country Sample

Difficulty of Starting a Business (EOS)

Low

High

Number of Days to start a Business (DB)

Low

High

$\text{r} = 0.24$
On Elections: De Jure and De Facto Measures

De Jure

\[ y = -3.14x + 85.69 \]
\[ R^2 = 0.00 \]

De Facto

\[ y = 23.09x + 55.30 \]
\[ R^2 = 0.19 \]
Capturing Informality: Bribery vs. ‘Legal Corruption’ (& State & Regulatory Capture), 2004

% Firms report ‘corruption’

Source: Author’s calculations based on EOS 2004.
6. On the Conceptual: Avoiding Obese or Anorexic Definitions helps convergence

- Pitfalls of both Obese and of Anorexic Definitions of Governance, Rule of Law
- Obese definitions: problems of falsifiability, tautologies, attribution, impairs research on links
- Anorexic definitions: excluding important components
- Opting for definitions within mid-range: facilitates continuing progress on empirical areas
- More progress in conceptualization has taken place than acknowledged
- Existing indicators draw from such progress
- Alternative theoretical approaches: small or large difference empirically?
- Continue to progress: conceptual & empirical tracks
7. Power of Statistics to Address Challenges in Indicators

- Embracing Statistics to construct robust indices and carry out empirical analysis
- *Statistics* to estimate of margins of error
- Econometrics and Statistics: tools to deal w/ noisy indicators, missing data, non-random samples, causality, endogeneity.
- Statistics: help address critiques of indicators
- Power of data for the policy-maker
- Concluding onwards:  i) Margins of Error--seriously;
  ii) what ingredients crucial in notions of democracy, ‘voice & democratic accountability’, government effectiveness, rule of law, and corruption;
  iii) what is useful, and what is amiss, and can be improved in existing individual and composite measures in these dimensions, and,
  iv) capturing informal rules of the game/institutions
A modicum of perspective...

‘Not everything that counts, can be counted, and, not everything that can be counted, counts’

Albert Einstein
Selected Links to papers associated with this presentation

• Brief paper associated with this presentation for this conference:
  http://indicatorsinfo.pbworks.com/f/Kaufmann.pdf

• Survey paper on governance indicators:

• Access to WGI data, resources and papers:
  www.govindicators.org