Measuring Good Governance: Kaufmann-Kray
Outline

1. Assumption, definition, and data sources of KK
2. Constructing aggregate indicators
3. Interpreting i) Levels; ii) Changes; and iii) Trends:
4. Uses and Limitations of Governance indicators
   • Why *subjective* data?
   • Margins of error for *objective indicators*?
   • *Ideological biases* in expert assessments?
   • Margins of error and *aid allocation rules*?
5. Summary, implications & discussion exercise
Discussion Exercise

1. What is meant by the concept of ‘good governance’ and what are its central components?

2. How would you **evaluate the pros and cons** of using expert or representative surveys to assess perceptions of good governance?

3. Which of the **evaluative criteria** selected by Kaufmann-Kray would you prioritize in assessing the performance of good governance, and why? What criteria would you add or delete?

4. If you were to carry out an expert survey, how would you design a **sampling frame** to collect the perceptions in each country? Who would you include? How could you insure that your sample would be ‘representative’?

5. How will you use KK in your reports?
Evaluating the good governance indicators

Recap: Munck and Verkuilen’s criteria
- Valid? – Measures the underlying concept
- Reliable? – Can be repeated consistently
- Comprehensive? – Covers all major dimensions of concept
- Equivalence across societies? – Measures like-with-like
- Replicable? – Can be recreated using the same steps
- Balanced? – Not skewed towards only partial aspect
- Robust? – Generates same results in analysis
- Value bias? – Generates consensus
Measuring Good Governance

Governance Indicators

Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi

The World Bank Institute

For details see: http://info.worldbank.org/governance/

Data is in the QoG dataset
Governance Matters 2007
Worldwide Governance Indicators, 1996-2006

"Bravo to the World Bank for providing reformers with such a valuable tool for their work on governance. Now we can benchmark performance in various governance areas and measure progress!"

Ngizzi Okonjo-Iweala,
former Minister of Finance and of Foreign Affairs of Nigeria

The Worldwide Governance Indicators (WGI) project reports aggregate and individual governance indicators for 212 countries and territories over the period 1996-2006, for six dimensions of governance:

- Voice and Accountability
- Political Stability and Absence of Violence
- Government Effectiveness
- Regulatory Quality
- Rule of Law
- Control of Corruption

The aggregate indicators combine the views of a large number of enterprise, citizen and expert survey respondents in industrial and developing countries. The individual data sources underlying the aggregate indicators are drawn from a diverse variety of survey institutes, think tanks, non-governmental organizations, and international organizations.

The six aggregate indicators and the underlying data sources can be viewed interactively on the Governance Indicators webpage of this site. Documentation of the latest update of the WGI can be found in "Governance Matters VII: Governance Indicators for 1996-2006." Further documentation and research using the WGI is available on the Resources page of this website.

Disclaimer: The data and research reported here do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. The WGI are not used by the World Bank Group to allocate resources or for any other official purpose.

The WGI are produced by:

Daniela Kaufmann, World Bank Institute
Aarik Kayay, World Bank Development Economics Research Group
Massimo Mastruzzi, World Bank Institute

Provide feedback

Back to top
Good governance assumptions

• Instrumental goals:
  • Targeting aid towards countries with good institutions and policies makes sense for efficient investment of limited resources

• Transparency:
  • Using publicly-available eligibility criteria encourages monitoring, accountability, progress

• Implementation:
  • Therefore need effective, comprehensive, reliable and accurate indicators of ‘good governance’
Good governance & democracy
**Governance: the traditions & institutions by which authority is exercised.** This includes:

1. The process by which those in authority are selected and replaced
   - VOICE & ACCOUNTABILITY;
   - POLITICAL STABILITY & ABSENCE OF VIOLENCE

2. The capacity of government to formulate and implement policies
   - GOVERNMENT EFFECTIVENESS
   - REGULATORY QUALITY

3. The respect of citizens and state for institutions that govern interactions among them
   - RULE OF LAW,
   - CONTROL OF CORRUPTION
Six indicators of good governance

1. Voice and accountability
   Citizen participation, independent media

2. Political instability and violence
   Threat of state coup

3. Government effectiveness
   Quality of civil service

4. Regulatory burden
   “Market-unfriendly” policies

5. Rule of law
   Perceptions of crime, effective judiciary, enforceable contracts

6. Corruption
   - Perceptions of corruption
“Perceptions” (subjective) data on governance from 25 different sources constructed by 18 different organizations.

Data sources include cross-country surveys of firms, commercial risk-rating agencies, think-tanks, government agencies, international organizations, etc.

Over 200 proxies for various dimensions of governance.

Organize these measures into six clusters corresponding to definition of governance, for four periods: 1996, 1998, 2000, and 2002, covering up to 199 countries.
Why Subjective Governance Data?

1. For some dimensions (e.g. corruption), no cross-country objective data exist
   - Limited quantitative measures of corruption focus on differences in procurement costs relative to materials purchased

2. Subjective data can pick up crucial distinction between *de jure* and *de facto* institutional arrangements –
   - Most countries in the world now have elections, anti-corruption commissions, and decent anticorruption laws in the books

3. Perceptions do matter
Methodological issues

Polls of experts reliable and valid?
- Small N.
- Colored by economic outcomes?
- Ideological agenda of rating organizations?

Representative surveys reliable?
- Larger N.
- Interpreted in culturally-specific ways?
- Often in smaller range of countries
• Cross-Country Surveys of Firms:
  – Global Competitiveness Survey, World Business Environment Survey, World Competitiveness Yearbook, BEEPS

• Cross-Country Surveys of Individuals:
  – Gallup International, Latinobarometro, Afrobarometer

• Expert Assessments from Commercial Risk Rating Agencies:
  – DRI, PRS, EIU, World Markets Online,

• Expert Assessments from NGOs, Think Tanks:
  – Reporters Without Borders, Heritage Foundation, Freedom House, Amnesty International

• Expert Assessments from Governments, Multilaterals:
## Inputs for Governance Indicators 2002

<table>
<thead>
<tr>
<th>Publisher</th>
<th>Publication Source</th>
<th>Country Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wefa’s</td>
<td>DRI/McGraw-Hill Country Risk Review Poll</td>
<td>117 developed and developing</td>
</tr>
<tr>
<td>Business Env. Risk Intelligence</td>
<td>BERI Survey</td>
<td>50/115 developed and developing</td>
</tr>
<tr>
<td>Columbia University Columbia U.</td>
<td>State Failure Poll</td>
<td>84 developed and developing</td>
</tr>
<tr>
<td>World Bank</td>
<td>Country Policy &amp; Institution Assessment Poll</td>
<td>136 developing</td>
</tr>
<tr>
<td>Gallup International</td>
<td>Voice of the People Survey</td>
<td>47 developed and developing</td>
</tr>
<tr>
<td>Business Env. Risk Intelligence</td>
<td>BERI Survey</td>
<td>50/115 developed and developing</td>
</tr>
<tr>
<td>EBRD</td>
<td>Transition Report Poll</td>
<td>27 transition economies</td>
</tr>
<tr>
<td>Economist Intelligence Unit Country</td>
<td>Indicators Poll</td>
<td>115 developed and developing</td>
</tr>
<tr>
<td>Freedom House</td>
<td>Freedom in the World Poll</td>
<td>192 developed and developing</td>
</tr>
<tr>
<td>Freedom House</td>
<td>Nations in Transit Poll</td>
<td>27 transition economies</td>
</tr>
<tr>
<td>World Economic Forum/CID</td>
<td>Global Competitiveness Survey</td>
<td>80 developed and developing</td>
</tr>
<tr>
<td>Heritage Foundation</td>
<td>Economic Freedom Index Poll</td>
<td>156 developed and developing</td>
</tr>
<tr>
<td>Latino-barometro</td>
<td>LBO Survey</td>
<td>17 developing</td>
</tr>
<tr>
<td>Political Risk Services</td>
<td>International Country Risk Guide Poll</td>
<td>140 developed and developing</td>
</tr>
<tr>
<td>Reporters Without Borders</td>
<td>Reporters sans frontieres (RSF) Survey</td>
<td>138 developed and developing</td>
</tr>
<tr>
<td>World Bank/EBRD</td>
<td>BEEPS Survey</td>
<td>27 transition economies</td>
</tr>
<tr>
<td>IMD, Lausanne</td>
<td>World Competitiveness Yearbook Survey</td>
<td>49 developed and developing</td>
</tr>
<tr>
<td>Binghamton Univ.</td>
<td>Human Rights Violations Research Survey</td>
<td>140 developed and developing</td>
</tr>
</tbody>
</table>
Ingredients for Rule of Law Indicator

Surveys of Firms

- **BEEPS** *Courts Honest? Crime? Property rights protected?*
- **Global Competitiveness Survey** *Crime, money laundering, judicial independence, protection of financial assets*
- **World Competitiveness Yearbook** *Justice fairly administered, personal security and private property protected*

Surveys of Individuals

- **Gallup** *Trust in legal system*

Risk Rating Agencies

- **BERI** *Contract enforcement*
- **DRI** *Costs of crime, enforceability of contracts*
- **EIU** *Costs of crime, enforceability of contracts, property rights protection*
- **PRS** *Law and order*
- **World Markets Observer** *Judicial independence, crime*

Think Tanks

- **Freedom House** *Rule of law*
- **Heritage Foundation** *Property rights, black market activity*

Governments

- **State Dept Human Rights Report** *Judicial independence*
2. Building Aggregate Indicators

- **Use Unobserved Components Model** (UCM) to construct composite governance indicators, and margins of error for each country.

- **Estimate of governance**: *weighted average* of observed scores for each country, re-scaled to common units.
  - Weights are proportional to *precision* of underlying data sources.
  - Precision depends on how strongly individual sources are correlated with each other.

- **Margins of error** reflects:
  - (a) *number of sources* in which a country appears, and
  - (b) the *precision of those sources*.
Unobserved Components Model

- Observed indicator $k$ of governance in country $j$, $y(j,k)$, is noisy indicator of true governance in country $j$, $g(j)$:

$$y(j,k) = \alpha(k) + \beta(k) \cdot (g(j) + \varepsilon(j,k))$$

- Variance in measurement errors is same across countries for each source, but different across sources:

$$E[\varepsilon(j,k)^2] = \sigma^2_\varepsilon(k)$$

- *Identifying assumption:* Measurement errors are uncorrelated across sources? highly correlated sources measure governance with more precision
Estimates of Governance from UCM

- UCM allows us to infer the distribution of governance in a country conditional on the observed data for that country.

- Best estimate of governance is the *mean* of this conditional distribution:

\[
E[g(j) \mid y(j,1), \ldots, y(j,K(j))] = \sum_{k=1}^{K(j)} w(k) \cdot \frac{y(j,k) - \alpha(k)}{\beta(k)}
\]

- So estimate of governance is *weighted average of re-scaled scores*, with weights proportional to precision of each source:

\[
w(k) = \frac{\sigma_{\varepsilon(k)}^{-2}}{1 + \sum_{k=1}^{K(j)} \sigma_{\varepsilon(k)}^{-2}}
\]
Precision of Estimates from UCM

- Reliability or precision of estimate of governance for each country is the standard deviation of this conditional distribution:

\[
SD[g(j) \mid y(j,1), \ldots, y(j,K(j))] = \left(1 + \sum_{k=1}^{K(j)} \sigma^2_{\varepsilon}(k)^{-2}\right)^{-1/2}
\]

- These standard errors are smaller for countries that (a) appear in more sources, and/or (b) appear in more reliable sources
Precision and Number of Sources: Rule of Law Indicator, 2002

The graph shows the relationship between the standard error of governance estimate and the number of sources. As the number of sources increases, the standard error decreases, indicating improved precision.
Levels of Corruption Across Countries, 2002

Note: This graph shows estimates of the indicated dimension of governance (on the vertical axis) for all countries graphed against each country's percentile rank (on the horizontal axis) for 2002. The vertical bars show the statistically-likely range of values of governance for each country, with the midpoint of each bar corresponding to the best single estimate. Selected countries are labeled. As emphasized in the text, the ranking of countries along the horizontal axis is subject to significant margins of error, and this ordering in no way reflects the official view of the World Bank, its Executive Directors, or the countries they represent.
Assigning Countries to Governance Categories: Margins of Error Matter

Note: Confidence Interval: 90%
Voice and Accountability (2006)


Note: The governance indicators presented here aggregate the views on the quality of governance provided by a large number of enterprise, citizen and expert survey respondents in industrial and developing countries. These data are gathered from a number of survey institutes, think tanks, non-governmental organizations, and international organizations. The aggregate indicators do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. Countries’ relative positions on these indicators are subject to indicated margins of error that should be taken into consideration when making comparisons across countries and over time.
Interpreting Cross-Country Differences in Governance

- Governance scores are normally distributed; they run from about –2.5 to 2.5 (i.e. denominated in S.D. units)
- Global average is zero for all indicators and periods- so indicators measure countries’ relative positions
- Margins of error are substantial:
  - Small differences in estimates in governance unlikely to be statistically significant
  - Relatively few countries have 90% confidence intervals entirely within a single quartile
- More countries do have confidence intervals entirely within their half, and even within their tercile
- Differences in governance between best and worst group of countries in the world are unambiguous
Margins of Error Are Not Unique to Subjective Indicators

Many potential objective/quantitative indicators of governance:

- **Regulatory Quality**: Days to start a business

- **Rule of Law**: Contract-intensive money (share of M2 held in banking system, confidence in property rights protection)

- **Government Effectiveness**: Stability of budgetary revenue and expenditure shares (policy instability), share of trade taxes in revenue (narrow tax base)

- All indicators provide imperfect proxies for broader notions of governance – and so have implicit margins of error relative to these broader concepts
Do Expert Poll Assessments Reflect Ideological Biases of Rating Institutions?

- This “popular” critique subject to empirical scrutiny
- Look at difference between (i) country rankings based on expert assessments (potentially prone to ideological biases), and (ii) country rankings based on firm surveys (not prone to biases)
- Regress this difference in assessments on a variable measuring ideology of government in power:
  
  \[
  \text{IDEOLOGY} = 1 \text{ (Left-Wing)}, \quad =2 \text{ (Center)}, \quad =3 \text{ (Right-Wing)}
  \]
## Ideological Bias?

### Results of Ideology Regressions: Rule of Law 2002

<table>
<thead>
<tr>
<th></th>
<th>PRS</th>
<th>PIA</th>
<th>EIU</th>
<th>DRI</th>
<th>CDU</th>
<th>BRI</th>
<th>QLM</th>
<th>HUM</th>
<th>HER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideology</td>
<td>1.52</td>
<td>3.39</td>
<td>5.61</td>
<td>5.67</td>
<td>4.68</td>
<td>7.32</td>
<td>6.47</td>
<td>5.32</td>
<td>7.42</td>
</tr>
<tr>
<td>t-stat</td>
<td>0.41</td>
<td>0.73</td>
<td>1.65</td>
<td>1.46</td>
<td>1.21</td>
<td>1.65</td>
<td>1.63</td>
<td>1.19</td>
<td>1.91*</td>
</tr>
<tr>
<td>n</td>
<td>52</td>
<td>47</td>
<td>51</td>
<td>46</td>
<td>42</td>
<td>25</td>
<td>49</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>Adj R-Sq</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.05</td>
<td>0.03</td>
<td>0.01</td>
<td>0.05</td>
</tr>
</tbody>
</table>

- Positive estimates mean source rates right-wing governments higher than corresponding survey
- Little evidence of significant ideological biases
- Only Heritage tends to rate right-wing governments higher, magnitude of difference is small (less than ten percentile rank points)
- No overall difference between polls & surveys
Example of U.S. Millennium Challenge Account

To be eligible for MCA funds, potentially-eligible IDA countries with per capita GDP less than $1435 must score above median in half the indicators in three categories:

1. **Ruling Justly**:
   - Six indicators, including Voice, Government Effectiveness, Rule of Law, Control of Corruption

2. **Investing in People**:
   - Four indicators covering health and education spending and outcomes

3. **Promoting Economic Freedom**:
   - Six economic policy indicators including Regulatory Quality

   *and must score above median on Control of Corruption*
Recommendations for MCA Eligibility Rules

- Important to take margins of error seriously (for all indicators) – non-trivial risk of misclassifying countries
- Using multiple indicators reduces misclassification risk, but it remains substantial for “hard” corruption hurdle
- Rely on additional sources of data, especially for borderline cases just above or below the cutoff – complement with diagnostics
- Measuring progress over time is difficult but important
- Maximizing country coverage for all indicators is key
Summary & Conclusions

• Targeting aid towards countries with good institutions and policies makes sense

• Transparent publicly-available eligibility criteria encourages monitoring, accountability, progress

  at the same time...

• Have to consider margins of error, especially with “hard” in-or-out rules like corruption hurdle: focus on ‘yellow light’ group just below the median

• Need to gather more information, country diagnostics, etc.

• Aggregate indicators advantage on margins of error

• Margins of error a major challenge for all other indicators as well -- which also need to address issue of country coverage gaps and timeliness

• How would you use for your policy reports??
Pros and cons

Critique by Merilee Grindle
- Overwhelming demands on governments in developing countries
- ‘Shopping lists’ of demands
- ‘Good enough’ governance

Response
- Greater transparency in evaluations
- Evidence-based policy analysis
- Benchmarks for performance-based management
Discussion Exercise

1. What is meant by the concept of ‘good governance’ and what are its central components?

2. How would you evaluate the pros and cons of using expert or representative surveys to assess perceptions of good governance?

3. Which of the evaluative criteria selected by Kaufmann-Kray would you prioritize in assessing the performance of good governance, and why? What criteria would you add or delete?

4. If you were to carry out an expert survey, how would you design a sampling frame to collect the perceptions in each country? Who would you include? How could you insure that your sample would be ‘representative’?

5. How will you use KK in your reports?
Next class

Lab sessions on QoG database and reports