Methodological issues in constructing league tables

Jouni Kuha and George Gaskell
London School of Economics

International indicators and rankings

- Lots of them, on many topics
- Used for a wide variety of political and other purposes
- Apparently here to stay – “the measurement culture” (House of Commons Public Admin Select Committee)
- But numbers and metrics have a special status

Human Development Index (HDI)

- Based on three concepts and three indicators:
  - "A long and healthy life": Life expectancy at birth
  - "Knowledge": Adult literacy rate and enrolment ratio of 1-2-3 level education
  - "A decent standard of living": GDP per capita
- Rescaled and aggregated with equal weights
- Data from relevant national organisations

Corruption Perceptions Index (CPI)

- 16 studies on extent of corruption
  - Surveys and panels of resident and non-resident experts
- 3-14 of these available for each country, aggregated (fairly elaborately) with equal weights

Types of indices

- ‘Objective’ (official statistics etc.) vs. ‘subjective’ (surveys and panels of ‘experts’)
- Aggregated (combinations of several indicators) vs. disaggregated (single indicator)
- Different combinations raise different methodological issues

Quality of measurement

- *Validity* – how well does the scale capture what it purports to measure?
  - is it theoretically well-motivated (conceptual)
  - is it plausible (intuitive)
  - does it correlate with other indicators (empirical)
- *Reliability* – consistency of measurement
- *Replicability* – transparency and procedural clarity
- *Measurement error* – systematic and random

Other properties of useful performance indicators

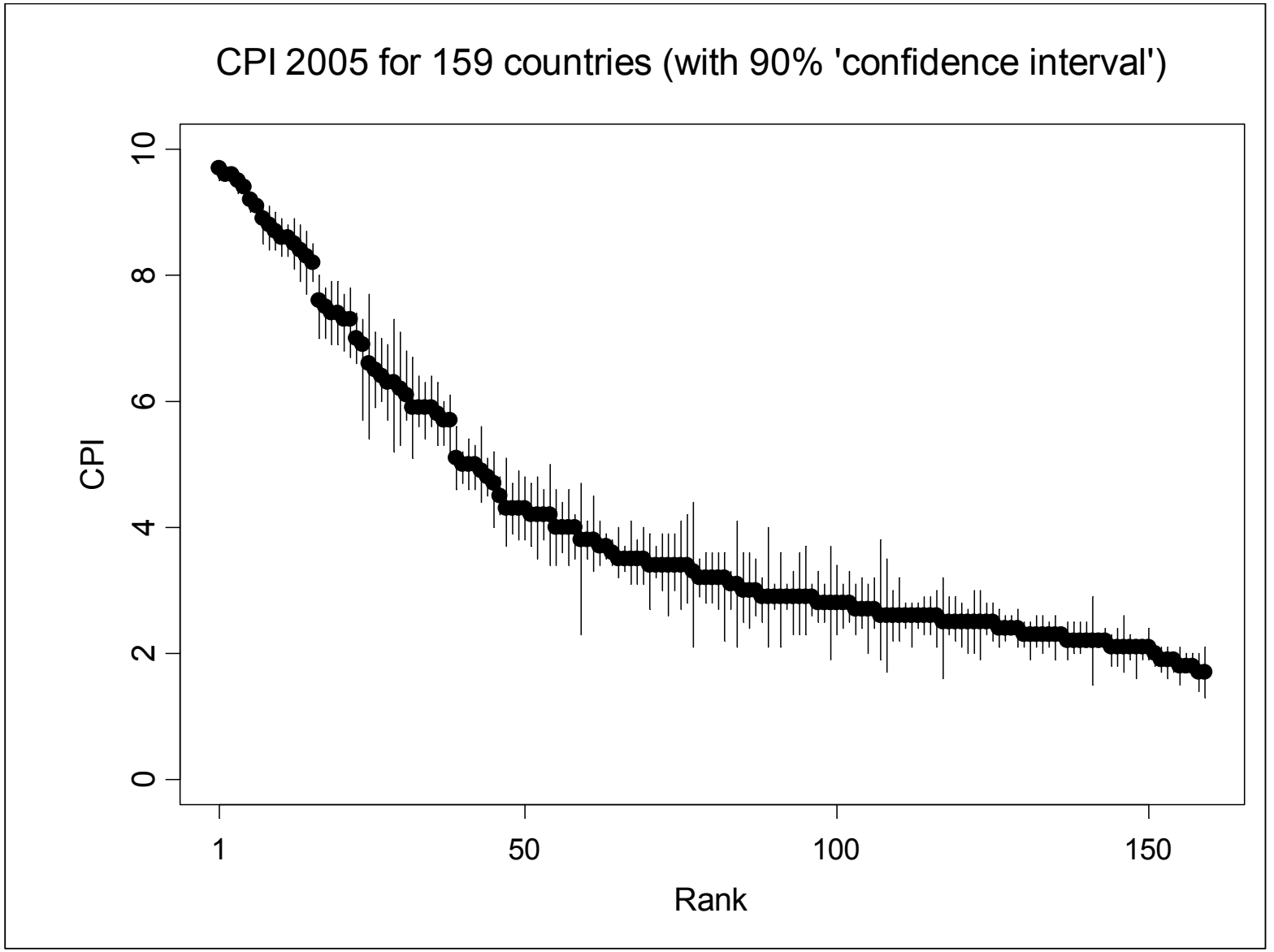
- Transparency
  - Indicator needs also to be *perceived* to be valid and legitimate
  - Need to know what to do to change a score.
  - However, greater transparency may induce perverse incentives
- Comparability across cases and across time

Challenges for different types on indices

- 'Objective' measures: variation in definitions, data availability and measurement error across countries
  - Need for standardisation of official statistics
- 'Subjective' measures: a host of transparency issues
- Aggregated measures: choice of components, choice of weights, apples and oranges
- Disaggregated measures: measure only what the single indicator does, and only as well as it does

Acknowledging uncertainty

- The values of most indicators are uncertain
  - Sampling variability of survey-based measures
  - Other kinds harder to quantify, but still there
- Translates into often substantial uncertainty about *rankings* of countries, i.e. low diagnostic precision.



*Reflects the variability in scores for each country, underestimating total uncertainty

Conclusions and implications

- Error in measurement (very broadly defined) is pervasive and a serious threat to information quality
- Careful consideration of methodological issues should be an essential part of metric construction and interpretation
- The “take it at face value” approach could lead to unjustified and unwise policy responses