

SUMMARY NOTES

IFI Roundtable, 23 March 2010

“Measuring Law: How to do it Right in Real-World Circumstances”

The following notes reflect selected remarks from discussions between senior officials from legal departments of international financial institutions and social scientists who specialize in measuring law. The excerpts reflect a variety of viewpoints. There is no attempt to reconcile differences in views. None of the comments should be understood as reflecting the positions of any given participant.

Sections to Jump to:

1. [Purposes and Mandates of IFIs](#)
2. [Global Norms](#)
3. [Diagnostic Instruments of IFIs](#)
4. [Organizational Factors](#)
5. [Theory and Concepts](#)
6. [Assumptions](#)
7. [Indicators](#)
8. [Methodological Approaches](#)
9. [Measurement Issues](#)
10. [Data Analysis](#)
11. [Inferences](#)
12. [Reform and Implementation Issues](#)

Purposes and mandates of IFIs

- Crisis reaction or crisis prevention
 - This requires assessment of some sort to anticipate crises and to remedy faults before the onset of a crisis
- Technical assistance in reform efforts
- IFIs measure in three broad ways:
 - In crisis situations
 - Deep analyses within countries in non-crisis situations
 - Periodic or routine cross-national comparisons

[Top](#)

Global Norms

- as a general matter, it is useful for countries to be exposed to global norms – more difficult is to know how they can be made useful or relevant for a given country and its particular situation.

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 2 of 9

- there are two contrasting models of orientation between global norms and countries: one is to start with the global norms and carry them to countries; the other is to start with national policy issues and look for solutions, which may sometimes be found at the global level.
- rather than formulate global norms, IFIs might identify policy choices that legislators will face and show what options are available
- there is a danger that the norms for one country or legal family (e.g., common law countries) get elevated to global best practices regardless of their fit in particular environments.
- in some areas of law reform, e.g., tax, the substantive law is so complex, the national models are so varied, the policy preferences are so different that it makes no sense to create external standards—rather, evaluation and subsequent reforms can be tailored for each country.

[Top](#)

Diagnostic Instruments of IFIs

- ROSCs – Reports on Observance of Standards and Codes (IMF/World Bank)
- Country-specific diagnoses (using what protocol for analysis?)
- Perception surveys of lawyers (law-on-books, law-in-action)
- Surveys of experts (e.g., a law firm, professors) that evaluate law-on-books country by country against an external standard
- Surveys which ask informants to report on observed behaviors (e.g., how long does it take to file or enforce a pledge)

[Top](#)

Organizational Factors

- The extent of money available for measuring law affects what instruments can be used and thus what validity and reliability there is in inferences on the bases of data that are collected.
- ‘Tied’ donor funding, which demands that evaluations be undertaken by experts from the country that provided the funding, provides IFIs less opportunities to do it right than “untied” funding.
- Core institutional funding for measuring law is more desirable than donor funding that comes with various strings attached. Core funding allows more methodical and consistent evaluations.
- If IFI legal departments are evaluated internally by the number of new statutes they put on the books, or legal institutions they construct, this will create perverse incentives that lead to superficial evaluations and ineffective legal change
- Cross-country comparisons may be very difficult to make validly or reliably, but IFIs are often under pressure for pragmatic reasons to provide comparative rankings or produce cross-sectional data.

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 3 of 9

- In globalized markets, e.g., capital markets, IFIs may be charged with cross-country comparisons in order to induce convergence.
- If an IFI legal department had a research officer (i.e., a well-trained, experienced social science researcher), and research panels (e.g., of social science or socio-legal advisors), then more sophisticated measurement, analysis and conclusions would be possible.
- How can IFIs get reliable field data if they are mandated to work with state authorities who do not necessarily want the IFIs to hear alternative voices to those of current policy-makers?

[Top](#)

Theory and Concepts

- Measurement requires theory: Why are you measuring? For what societal outcomes?
- Every indicator is predicated on a theory, implicit or not. Make the theory explicit. Examine it.
- If it is conceded that institutions matter for economic development, it is difficult to know which institutions. There is no general and/or empirically robust theory of institutions available.
- Often there is a lack of scholarly consensus on issues of fundamental importance to IFIs, e.g., what kind of political or administrative or legal system is optimal or even facilitative to make economic policy effective.
- IFIs engage in massive reform efforts often on the basis of (economic) theory that has not been rigorously critiqued or empirically verified, however plausible it seems on its face (cf. IFI efforts on secured transactions without any corroborative evidence that the existence of security or registries actually makes more or cheaper credit available).
- A concept that is to be evaluated or measured should not be confused with a *proxy* for it. There is a constant danger of elevating the proxy to some kind of reality. The measure becomes the thing rather than being one of several ways to get at it.
- To measure law requires (a) starting with the ultimate behavior that law is supposed to produce, (b) mapping the social norms or normative regime that surrounds this activity, and (c) disaggregating the various elements around it.
- Take a concept like judicial independence: if it is an end, rather than a means to an end, then in some contexts it would have little impact and even lead to greater levels of corruption because of a lack of accountability mechanisms.
- Labeling of behaviors often as a policy prescription built into it, e.g., the ease of workers being fired may be a cost of doing business or policy failure. In any event, the cost of firing workers may be non-linear.

[Top](#)

Summary Notes – IFI Roundtable on Measuring Law
March 23, 2010
Page 4 of 9

Assumptions

- That good government is necessary for economic development, economic stability
- That government should be minimalist
- That the elements of good government are well-known (e.g., independent Central Bank, independent judiciary)
- That better harmonized laws across countries will facilitate trade among them and economic development within them
- That law matters for development: but this may not be true, the relationship may be spurious, or the inverse may be the case, i.e., development produces law that serves it; some kind of reciprocal relationship is likely
- That the principal motivation for action by economic entities or even states is financial, i.e., that economic incentives are the principal drivers of change rather than other powerful motivators such as tradition, religion, status, power, political ideology, etc.
- That legal origin matters – and, even if it seems to matter, this may have to do with other factors associated with legal origin, not legal origin per se
- That state law is all that matters, whereas informal or non-state law is often pervasive and influential in many countries; Legal pluralism cannot be ignored
- That passing a law has an impact – it may not, or the impact may be perverse
- That all effects of law (so far as they can be teased out) on economies operate in a linear function. But what about non-linear or curvilinear relationships, i.e., those impacts that are positive up to a point, then stop, or even turn negative

[Top](#)

Indicators

- Indicators should be designed to make theories testable
- Indicators in some areas have become increasingly more numerous and more complex; however, it is a difficult problem to decide how to combine them, most especially how to weight various elements against others as these vary in importance by country situations and over time within a country
- Over-simplified indicators, or summary indicators, can foster gaming by governments as they try to rank higher on rating scales
- Avoid situations where countries use indicators not as a tool but as a goal, e.g., they reverse engineer indicators and change policies to target them

[Top](#)

Methodological approaches

- Law on books only
- Law in action/effectiveness of law
- Triangulation, or a multiplicity of methods, or hybrid assessments, are much to be preferred to a single method, as they can overcome biases in particular methods

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 5 of 9

- Comparative methods, e.g., country comparisons
 - This is exceptionally difficult. Even if you have indicators that can measure variation across countries, it is often difficult to know what the results mean (e.g., does a higher arrest rate mean the government is cracking down on crime or becoming more repressive?)
 - Quantitative comparisons of legal or governance systems are very difficult, so much so that it is doubtful they are valid, reliable or meaningful
 - A principal purpose of comparisons of countries on a criterion (e.g., comprehensiveness/effectiveness of law) is to induce low-scoring countries to reform
- Time-series or historical methods, i.e., comparisons within the same country over time
 - These are more defensible than cross-national comparisons – at least the contexts are kept relatively constant
 - It is valuable to measure the same thing, year after year
 - Going back repeatedly to stakeholders, key informants, etc., as reforms roll out is crucial in order to discover “frictions” in the system
 - Return periodically to measure whether a legal reform begins to bite in the near term, intermediate term, longer term, i.e., need to know medium and long-term effects
- Functional holistic analyses
- Media analysis, e.g., of news media across countries and time can provide a new kind of information on many factors related to diagnosis and reform programs
- Qualitative data:
 - Narrative accounts are valuable
 - We need to know *why* economic and legal actors decide to act in one or another way—what is going in their heads?
- Panel studies – for each country/area of law create panels of informants who represent key constituencies for enactment and implementation and return to them year after year for over-time quantitative and qualitative data
 - It may be difficult to get a representative group together, even in one country

[Top](#)

Measurement issues

- Measurement must always begin with intended effects. They must be specified, then evaluate against those
- The only meaningful measurement takes into account local circumstances. Standardized measures can be inefficient and terribly misleading

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 6 of 9

- What is to be measured is an expression of power. This implies that designers of measurement tools and their consumers be aware of what other things might have been measured, of who made decisions about what to measure and what to ignore
- Measuring outcomes should be the priority for IFIs
- IFIs should publish their methodologies so that (a) the appropriate inferences can be drawn from results, and (b) observers can see what is being measured—and how—and what is not
- To make commensurable (to measure) the same concept in different places, it is necessary to take a higher level concept, then operationalize it or specify it in terms of behaviors appropriate for that place. E.g., skipping by children is a good measure of child development only in those countries where children skip
- It is necessary to measure separately (a) the operation of the law, and (b) the independent variables which are thought to influence law in operation
- Be aware of the danger of measuring what is *easy* to measure rather than what is *important* to measure. E.g., it is easier to measure performance of formal arrangements than informal arrangements, hence the former becomes the object of inquiry and the latter is neglected—despite the fact that the latter may be more important in dispute resolution, for instance
- Perception surveys (e.g. of lawyers) make it difficult to compare across countries because perceptions themselves may reflect cultural attitudes (e.g., something that is reported to work well in one country may actually work worse than in another country that reports it works less well)
- Opinion surveys make cross-national comparisons difficult because people are answering against implicit normative standards that differ from country to country
- Sampling:
 - Be constantly vigilant about bias in data collection and measurement
Always ask—who is *not* being asked? Biased sources have a powerful effect on the probability of successful implementation.
 - Evaluating a country on the basis of a single report by a single person or organization, even if detailed, is probably unreliable.
 - Basing a country diagnosis on evidence from the national capital may tell little about what happens in provincial capitals, towns in the hinterland, or the countryside
 - Basing diagnoses on a biased sample of stakeholders (e.g., interviewing creditors, but not debtors, in a corporate insolvency appraisal), may lead to wrong inferences, bad policy choices, and eventually resistance and non-implementation
 - Talking only to people who are part of formal structures misses much of the norm-driven business activity that goes on in a society, much of the conflict, much of the dispute settlement outside of courts
 - Talking only to legal actors—not business people, workers, others—risks significant bias and faulty inferences from the results

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 7 of 9

- It is imperative to obtain knowledge from local communities
- Data collection and legitimacy:
 - if IFIs listen to a broad range of salient interests in a country, then their recommendations for reform are likely to be more legitimate, and implementation more probable
 - if nation-states are compelled through financial pressure from IFIs to adopt law reforms without having had their voices fully heard, the more they are likely to disregard or subvert a law
 - obtaining local knowledge is critical
- Countries or individuals that respond to efforts at measurement may produce ‘false’ returns if they know what answer the IFI expects, or if there is a conflict of interest in which an IFI has a financial interest in the answer
- Assessors of a country’s law or practice may confound the results if they are not seen to be neutral, e.g., if they have partisan interests in the results and the reforms prescribed on the basis of the results. This can apply to professionals, for instance, who diagnose and recommend reforms that benefit their profession
- To measure law and legal institutions intelligently requires:
 - Understanding the contexts of institutions, e.g, to focus on courts alone will tell little unless we also know what social structures make courts work or not
 - Understanding of the distribution of power in a society: measure the distribution of power that is relevant to enactment or implementation. We need effective ways to ascertain a nation’s political economy as it relates to particular law-related outcomes.
 - Understanding informal institutions, because they always affect how and why formal institutions will work or not.

[Top](#)

Data Analysis

- If IFIs would make public (e.g., in online databases) the data they collect from legal measurement, then scholars could extract additional value from the datasets, create models, test theories, make rebuttable instances, etc.
- There are a lot more data available from IFIs now than a few years ago
- Weighting measures across time or place presents many difficulties – how is it known how heavily to weight the import of a factor that has different significance across countries or even across a long time period?
- There is an aesthetics of presenting quantitative data in order to compete with other IFIs or to persuade governments—but in the inevitable simplifications of

Summary Notes – IFI Roundtable on Measuring Law
March 23, 2010
Page 8 of 9

- color-coding, charts, figures, etc., many of the complexities of real situations are lost
- Ordinal ranking is more likely than not to be perverse; any kind of ranking is likely to be harmful. Countries are likely to game the ranking system
 - Data analysis that meets standards of social science would (a) better report data in a range, and (b) would provide confidence intervals

[Top](#)

Inferences

- Too often correlation (e.g., good law, advanced economy) is treated as causation (good law causes advanced economy). Most cross-sectional analyses (e.g., LLSV) do not allow causal conclusions, but robust time-series data are scarce
- Cannot infer there has been a change of practice simply by looking at (new) structures or institutions. Formal compliance may mask non-compliance in practice
- Looking at high or low conviction rates, or rates of any kind, often tells little about whether or not law or institutions are working. It is necessary to know what the rates *mean*. What is really going on to raise or suppress rates of filings, convictions, outcomes, etc.?
- Countries often cut-and-paste global norms or other standards to satisfy IFIs, not to produce real change by adapting norms and standards to the local situation
- If countries are subject to cross-national comparisons and surveys, as in *Doing Business*, and they get low rankings, the possibility looms large that countries will undertake reforms for the wrong reasons and often will little result
- Consequences of reforms are often contingent:
 - The fact that a legal rule functions in a certain way in one situation is no guarantee it will function in a similar way in another situation
- Effects are often conditional:
 - Because a law or institution works (or does not) in one country is no guarantee it will work (or not) in another
 - Without knowing *why* a given law or legal institution works in a specific context, it is impossible to transport it elsewhere. I.e., it is necessary always to know not only how the target of measurement works (substantive laws, regulatory agencies), but *under what conditions* it will or will not work
- To rely on a single number to summarize the legal situation in an entire country is quite dangerous and probably very misleading

[Top](#)

Summary Notes – IFI Roundtable on Measuring Law

March 23, 2010

Page 9 of 9

Reform and Implementation Issues

- local knowledge is crucial:
 - this includes getting hard data that are meaningful; discovering the array of interests; identifying factions; knowing where resistance to reform is likely to arise
 - if countries have severely impaired institutional capacity, then how is it possible to obtain local knowledge that is key to reliable measurement or diagnosis?
 - Obtaining local knowledge is crucial for assessments and reform *but* in most countries it is exceptionally difficult to obtain (no hard data, refusal to share those data that are available, inability for institutions to appraise their own performance, limited civil society)
- the ultimate criterion for measurement of law or a law reform is its impact – whether we call this ‘effectiveness’ or ‘implementation’. Hence measurement techniques need to be developed to validly and reliably measure ‘behavior of law.’ These must grapple with counterfactuals—what would have happened had the law not changed
- to appraise law and to reform legal systems requires that we measure the social norms, or normative regime, or social structures that support the law—the private professions, or legal press, or academic experts, or informal norms that influence law-related behavior. To measure the ‘law’ or ‘legal institution’ without knowing its supports will provide illusory results that can lead to failed reform programs
- the more an IFI’s funding goes to the private sector, the less leverage it has through use of conditionality, for instance, and thus the more it must rely on persuasion or policy dialogue
- There is a huge gap between the theory behind reform (e.g., judicial independence) and what prospects there are for reform in practice
- It should never be assumed that judiciaries or Ministries of Justice are “embodiments of justice.” Quite the contrary – as often as not they are reactionary, tradition-bound, lacking in vision

[Top](#)