

U.S. Department of Labor
Bureau of International Labor Affairs

Office of Trade and Labor Affairs
Contract Research Program

Download this and other papers at
<http://www.dol.gov/ilab/media/reports/otla/>

The views expressed here are those of the author(s) and do not necessarily represent the views or official positions of the U.S. Government or the U.S. Department of Labor.

Formulating and Aggregating Indicators of Labor Rights Compliance

for
Research project: Refining the NAS Matrix

DOL099RP20744

Submitted to

Bureau of International Labor Affairs
Department of Labor

March 10, 2011

Mark Barenberg
Professor of Law
Columbia University
New York

- Preface..... 5**
- Introduction: Road Map, Eleven Proposals, and
Flow-Chart for Indicator Formulation 7**
 - 1. The Policy Goal of the Paper, and ILAB’s Role in Serving
that Policy 7**
 - 2. Road Map of the Paper 12**
 - 3. How to Formulate an Indicator – A Step-by Step Textual
Explanation..... 19**
 - 4. How to Formulate an Indicator – A Flow Chart 24**
- 1. The NAS Methodology and the Michigan Pilot-Testing of that
Methodology 27**
 - 1.1. The NAS Indicators and Matrix Methodology 27**
 - 1.2. The Michigan Pilot Testing of the NAS Indicator
Methodology 32**
- 2. The Traditional Compliance Model 40**
 - 2.1. International Labor Rights are Legal Concepts 41**
 - 2.2. Legal Concepts Call for Sustained Interpretive and
Regulatory Analysis 43**
 - 2.3. Traditional Legal Interpretation and Sources of Law 49**
 - 2.4. Traditional Compliance Methods 51**
 - 2.5. Conclusion 63**
- 3. Challenges to the Traditional Compliance Strategy:
Four Types of Legal Norms 65**
 - 3.1. Weaknesses in the Traditional Compliance Strategy 67**
 - 3.2 From Rules to Standards 71**
 - 3.3. From Fixed Norms to Revisable Norms 78**
 - 3.4. Conclusion: Four Types of Indicators
and the Circumstances to Which They are Best Suited..... 80**
- 4. Capacity-Building Indicators 85**
 - 4.1. The Three Categories of NAS Indicators vs. the Three
Categories of U.N. Indicators..... 86**
 - 4.2 A Fourth Category: Capacity-Building Indicators 90**

4.2.1. Requirements of Capacity-Building in International and Domestic Law	91
4.2.2. Conceptual Rationales for Applying Capacity-Building Indicators.....	97
4.2.3. Why Not Include Capacity-Building Indicators in the Category of Enforcement Indicators?	103
5. Input or Output Indicators? – Conceptual Problems	108
6. Normative Yardsticks: Longitudinal, Comparative, and other Numerical Indicators	118
6.1. The Case for Longitudinal and Comparative Indicators.....	119
6.2. Are Longitudinal, Comparative, and Revisable Indicators Consistent with Human Rights?	125
6.3. The Case for the Normative Metric of “Non-Trivial Number of Violations”	131
7. Binary Indicators.....	136
8. Control Variables: Hard and Soft Constraints.....	141
9. Data Problems, and Two Responses: Twin Indicators and Capacity-Building Indicators	146
9.1. The Problems of Unavailable, Unreliable, or Inconsistent Data	146
9.2. Twin Indicators	149
9.3. Capacity-Building Indicators	153
10. Probative Indicators, Diagnostic Indicators, and Assessment Indicators	154
10.1. Why Three Bodies of Indicators?.....	154
10.2. What Types of Indicators are Best Suited for Probative Indicators, for Diagnostic Indicators, and for Assessment Indicators?	155
10.3. Which Individual Indicators Should Serve as Probative and Diagnostic Indicators?	165
11. Alternative Strategies to Weight and Aggregate Indicators.....	175

11.1. Equal Weighting and Simple Summation of Indicators	177
11.2. Weighting Indicators Based on Hierarchies in Authoritative Legal Sources	183
11.3. Principal Component Analysis	187
11.4. <i>Ex Ante</i> Multivariate Modeling.....	188
11.5. <i>Ex Post</i> Multivariate Modeling: Learning by Monitoring.....	194
12. A Process Based on Iterative Dialogue? – Some Conjectures.....	199
13. Next Steps	208
ACRONYMS	211
BIBLIOGRAPHY	212
FIGURES	
Figure 1. How to Formulate an Indicator: A Flow Chart.....	24
Figure 2. The NAS Matrix	31
Figure 3. Converting Standards into Rules.....	76
Figure 4. Four Categories of Norm-Types for Particular Indicators	80
 APPENDIX A: Assessment Indicators	
 APPENDIX B: Diagnostic Indicators	
 APPENDIX C: Probative Indicators	

Preface



In 2004, the National Academy of Sciences (NAS), in cooperation with the Bureau of International Labor Affairs (ILAB) of the United States Department of Labor (DOL), published a methodology for assessing labor rights compliance and other labor market conditions of trading partners of the United States. The methodology includes batteries of Indicators, a matrix instrument, and a database of information sources for applying the Indicators and matrix to particular countries. Social scientists at the University of Michigan pilot-tested and evaluated the NAS methodology, and submitted their findings on February 23, 2009. The Michigan evaluation revealed several areas in which the Indicators, matrix, and database might be refined.

On September 10, 2009, ILAB and this author entered a contract for a research project on Refining the NAS Matrix. The subject of the research proposal is (1) to apply legal and regulatory analysis to formulate a refined body of Indicators dedicated to making determinations whether trading partners are compliant with United States trade legislation and trade agreements, and (2) to canvass alternative methodologies for weighting and aggregating the Indicators.¹ These topics require extended conceptual analysis of various types of indicators, and of the relation between indicators and composite indices measuring compliance with each overall labor right.

¹ The latter proposal is to analyze the conceptual pros and cons of the alternative strategies, but not to identify an optimal strategy or statistically validate the strategies.

The Statement of Work enumerates eight Tasks. This document is submitted in fulfillment of Task 7. The Statement of Work defines that Task, and the relevant earlier Tasks to which Task 7 refers, as follows:

Task 4: Conduct original research as described in contract.

Task 5: Produce a draft research paper and submit it to ILAB researchers for peer review and to serve as a basis of consultation with ILAB policy and program officials. ILAB staff must be given at least one month to review and comment on draft papers.

The draft research paper will include:

1. Introduction, overview of key issues, and motivation for study
2. Literature review
3. Conceptual framework
4. Analytic strategy, and a description of data collection and data, as appropriate to the contract
5. Results
6. Policy implications
7. Caveats and limitations
8. Bibliography
9. Appendices, as necessary.

Task 6: Conduct virtual or in-person consultations with ILAB policy and program officials to identify recommendations that follow from the research and will aid ILAB in its work.

Task 7: The Contractor will prepare a final research paper taking on board relevant comments from ILAB peer review and consultations. The final paper must be of a caliber equal to an article in a peer-reviewed economic journal such as those listed in *The Journal of Economic Literature Abstracts*. In addition to hardcopy submission of draft and final products, the researcher will provide an electronic copy of the research paper and any data used in the analysis to USDOL. See Tasks 4 and 5 for final research paper requirements.

Introduction: Road Map and Overview of Eleven Proposals



This Introduction has four sub-parts. The first sub-part sets out the policy goal of the paper, the role of ILAB in serving that policy goal, and the major conceptual issues raised by the policy goal and by ILAB’s role. The second sub-part gives the customary road map of the thirteen Parts of the paper. The bulk of the paper provides conceptual analysis of several categories and types of Indicators for measuring compliance with internationally recognized labor rights. The third sub-part, as both a preview and overview, explains how the proposed methodology constructs individual Indicators – that is, how each Indicator set out in Appendices A, B, and C incorporates the various categories and types of indicators discussed in this paper. The fourth sub-part then presents the same explanation in the form of a Flow Chart.

1. The Policy Goal of this Paper, and ILAB’s Role in Serving that Policy

This paper is motivated by a policy goal and by the conceptual questions underlying that goal. The broad policy goal is to “refine the NAS Indicators and Matrix” for assessing the labor rights compliance of United States’ trading partners and the U.S. itself.² There are four

² Throughout this paper, “Indicators” with a capital “I” refers to the NAS Indicators and the revised Indicators formulated in Appendices A, B, and C, whereas generic references to “indicators” use the lower case “i.” The distinction is not always hard and fast.

sub-goals. The first is to develop conceptual foundations on which an effective Indicator methodology should rest. The second is to propose bodies of Indicators for (a) workers' freedom of association, rights to organize, and rights to bargain collectively; (b) rights against employment discrimination; and (c) acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health. The third is to examine the pros and cons of alternative strategies for weighting and aggregating the Indicators and for creating composite indicators for each of the three overall rights. The fourth is to examine the process by which ILAB may apply and potentially revise the bodies of Indicators over time.

One could imagine innumerable ways to operationalize the rights listed above, and to undertake pertinent conceptual analysis for formulating indicators for such operationalized rights, depending on the practical purpose for doing so.

The practical analysis and proposals of this paper are bounded by the institutional role and resources of ILAB. Congress sets the broad obligations of trading partners to enforce worker rights. The touchstone in defining those obligations is Congress's intent in enacting the worker rights provisions in trade statutes and in legislation that implements trade agreements. Congress also plays a critical role in determining the level of resources available to the Executive Branch – including the United States Trade Representative (USTR), the State Department, and the Department of Labor – for purposes of assessing the compliance of U.S. trading partners with their Congressionally intended obligations. The responsibility to enforce U.S. trade legislation and agreements rests in the first instance with

the USTR and ultimately with the President, who has the constitutional obligation to “faithfully execute” the laws.

In this constitutional and political context, the role of ILAB – using the resources appropriated by Congress – is to formulate a methodology and to gather and analyze data

ELEVEN PROPOSALS

- 1. THREE BODIES OF INDICATORS:** Probative Indicators for triage, Diagnostic Indicators for more intensive screening, and Assessment Indicators for evaluation. *See Parts 1 and 10, and Appendices A, B, and C.*
- 2. NEW CATEGORY OF INDICATORS:** Add “Capacity-Building Indicators” to existing categories of “Substantive Law Indicators,” “Enforcement Indicators,” and “Outcome Indicators.” *See Part 4.*
- 3. FRAME INDICATORS IN BINARY FORM:** Instead of NAS’s 3 x 3 matrix. *See Part 7.*
- 4. USE FOUR QUALITATIVE TYPES OF INDICATORS SUITED TO DIFFERENT ASPECTS OF EACH RIGHT:** Bright-line Rules vs. Standards; Fixed vs. Revisable. *See Part 3.*
- 5. USE DROP-DOWN VIZUALIZATION:** To convert more general Indicators into detailed country-specific sub-indicators, to revise Indicators over time, and to store and display information sources. *See Part 3.*
- 6. PRIMARILY USE “INPUT” INDICATORS, NOT “OUTPUT” INDICATORS:** But use outcome measures for defining input Indicators, for ILAB research purposes, and for probative and diagnostic purposes. *See Part 5.*
- 7. USE LONGITUDINAL AND COMPARATIVE INDICATORS:** Where international standards provide no clear yardstick. *See Part 6.*
- 8. START WITH ONLY ONE CONTROL VARIABLE:** Income per capita. *See Part 8.*
- 9. INCENTIVIZE DATA COLLECTION:** Use “Twin Indicators” to measure data collection for specific Indicators, and Capacity-Building Indicators to measure overall systems for data collection. *See Part 9.*
- 10. WEIGHTING AND AGGREGATING INDICATORS:** Pros and cons of five alternative strategies. *See Part 11.*
- 11. THE OVERARCHING PROCESS: ITERATIVE DIALOGUE:** Exchange of data and explanations among U.S. analysts, trading partners, and stakeholders. *See Part 12.*

that can inform the USTR's and President's ultimate judgments about whether trading partners comply with the obligations imposed by legislation and agreements.

The practical purpose animating this paper, then, is to produce constructive proposals for ILAB's performance of that role. This paper offers eleven such proposals. As a byproduct, the proposals may prove useful to other actors – such as other Executive Branch officials, parties filing petitions under U.S. trade legislation, human rights organizations monitoring compliance with labor rights, and of course governments seeking to comply with U.S. legislation and treaties – to assess whether particular countries comply with the labor rights and obligations contained in U.S. trade legislation and treaties. The proposals may also prove useful to ILAB officials in making decisions about allocating resources for technical assistance and research.

To serve the practical purposes of ILAB, this paper formulates three bodies of Indicators, rather than simply refining the single body of NAS Indicators. This strategy is premised on ILAB's resource constraints and its various functions apart from monitoring compliance with trade legislation and agreements. The first body – Probative Indicators – serves an initial screening function.³ The purpose of the Probative Indicators is to identify those countries that are most likely to have especially poor compliance records, warranting more intensive diagnosis, comprehensive assessment, technical assistance, or targeted research. The second – Diagnostic Indicators – serves the purpose of more intensive screening of countries to more confidently identify those with actual or likely compliance

³ See Appendix C.

problems.⁴ (Probative Indicators are essentially a short list of Diagnostic Indicators.) The third – Assessment Indicators – serves the purpose of comprehensive evaluation of a trading partner’s compliance with its statutory and treaty obligations.⁵

To lay the groundwork for formulating the three bodies of Indicators, this paper carries out the conceptual analysis required by the current state-of-the-art in developing composite indices generated by bodies of indicators. There are four foundational conceptual problems. First, indicators measure one or more underlying or “ultimate” concept(s). We must precisely define those ultimate concepts – for our purposes, the ultimate concepts of “denial of internationally recognized worker rights,” “taking steps to afford internationally recognized worker rights,” and “effective enforcement of labor law.”⁶ Second, the Indicators themselves

⁴ See Appendix B.

⁵ See Appendix A.

⁶ In fact, the relevant statutes and treaties use variations of this language. Section 301 of the Trade Act speaks of a trading partner’s “deni[al]” of internationally recognized rights rather than “fail[ure] to effectively enforce,” the language used in treaties. 19 U.S.C. § 2411(d)(3)(B)(iii)(I),(II). Section 301 also speaks of “standards for minimum wages, hours of work, and occupational safety and health,” rather than “acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health,” the language used in the other statutes and treaties. Cf. 19 U.S.C. § 2411(d)(3)(B)(iii)(V) with 19 U.S.C. § 2467(4)(E). As to the concept of “taking steps,” the Generalized System of Preferences asks whether the country “has taken or is taking steps to afford” internationally recognized worker rights. 19 U.S.C. § 2462(c)(7), while Section 301 asks whether the country “has taken, or is taking, actions that demonstrate a significant and tangible overall advancement in providing throughout the foreign country” internationally recognized worker rights. 19 U.S.C. § 2411(d)(3)(C)(i)(I). The Caribbean Basin Economic Recovery Act contains the same “taking steps” language. See 19 U.S.C. § 2702(b)(7). The African Growth and Opportunity Act states the concept of progressive compliance as “making continual progress toward establishing...protection of internationally recognized worker rights...” 19 U.S.C. § 3703 (a)(1)(F). The conceptual implications of these variations are discussed at relevant points in the paper.

Note, however, that all the statutes and treaties use the language of “internationally recognized worker rights” or the equivalent “internationally recognized labor rights” as the key concept for the substantive rights that the parties are required to effectively enforce, to not deny, or to take steps to afford or advance. It is true that the U.S.-Peru template explicitly refers to ILO law on freedom of association, collective bargaining, and nondiscrimination, and requires the parties to incorporate ILO rights in domestic law. See Article 17.2(1), United States-Peru Trade Promotion Agreement, April 12, 2006, at www.ustr.gov. But note that even in that treaty, the core obligation of the parties, stated in Article 17.3(1)(a), is to “effectively enforce its labor laws,” and Article 17.8 defines “labor laws” as “a Party’s statutes and regulations, or provisions thereof, that are directly related to the following internationally recognized labor rights,” followed by an enumeration that includes freedom of

may take alternative conceptual forms. That is, there are several distinctive types of indicators, each with its own conceptual advantages and disadvantages. Third, there are conceptual challenges in specifying the relationships among the Indicators and between the Indicators and the ultimate concepts. Fourth, we may conceptualize the process through which ILAB will gather data; will apply Indicators to the data; will revise Indicators and appraisals in light of those iterative inquiries; and will communicate with trading partners about ILAB's provisional appraisals, about additional pertinent data, and about possible justifications for poor compliance.

2. Road Map of the Paper

Before taking up the conceptualization, refinement, and revision of the NAS Indicators, we must know what it is we are refining. Part 1 of this paper summarizes the NAS methodology and the Michigan pilot testing of that methodology.

Part 2 makes the straightforward point that U.S. trade legislation and treaties are legal instruments. The formulation of Indicators should therefore be based on sustained legal and regulatory analysis. Part 2 therefore presents a thought experiment: How might a body of labor rights Indicators look if it were based systematically on traditional compliance methods?

Part 3 then examines the weaknesses of any traditional compliance strategy that

association, the right to collective bargaining, the elimination of employment discrimination, and acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health. Hence, while the parties are obligated to incorporate ILO law in their substantive domestic law, their enforcement obligation still attaches to the concept of internationally recognized labor rights.

formulates detailed, static rules to serve as yardsticks for the realization of multidimensional rights in complex, fluid social settings. The conceptual weaknesses are pinpointed by recent practice and theory in indicator-driven regulation. The weaknesses lie in the overly specific and static nature of the Indicators – making at least some of the Indicators insufficiently adaptable to local context and economic change and excessively susceptible to opportunistic gaming by ill-intentioned governments. Part 3 examines indicators with the opposite features – generality and revisability. It therefore proposes four conceptually different types of legal indicators: (1) fixed bright-line rules, (2) revisable bright-line rules, (3) fixed standards, and (4) revisable standards. Part 3 concludes with a discussion of the subject domains to which each of the four types of legal indicators is best suited.

Part 4 discusses the three categories of Indicators formulated by the NAS – Legal Framework Indicators, Government Performance Indicators, and Overall Outcome Indicators. Part 4 proposes the addition of a fourth category – Capacity-Building Indicators. (For clarity, Part 4 also re-labels the other three NAS categories as Substantive Law Indicators, Enforcement Indicators, and Outcome Indicators.) The Capacity-Building Indicators are mandated by international and domestic law, and measure functionally advantageous features of sustainable compliance.

The Capacity-Building Indicators measure the government's system for collecting data about compliance, its setting of targets and policy plans for improved compliance, and its evaluation of success in achieving targets and implementing plans. Capacity-Building Indicators also measure the government's guarantee of stakeholder participation and

transparency in defining worker rights, enforcing those rights, setting targets for improvement, and evaluating improvements. Finally, Capacity-Building Indicators measure the government's participation in disciplined comparisons of its performance relative to its peer countries.

Part 5 discusses the conceptual relationships among "input" indicators (Substantive Law Indicators, Enforcement Indicators, and Capacity-Building Indicators) and "output" indicators (Outcome Indicators). Part 5 concludes that most Assessment Indicators should be framed as input indicators. The rationale is threefold: First, if the body of Indicators includes both input and output indicators, there is a risk of double-counting the cause (increased enforcement efforts) and its effect (improved outcomes resulting from those increased efforts). Second, U.S. legislation and trade agreements are directed primarily at incentivizing changes in policy instruments under the control of trading partners' governments. Third, and related, outcomes are influenced by innumerable variables other than government performance.

Part 5 concludes, nonetheless, that output measures remain relevant in three ways: First, outcome measures may be useful for ILAB research on the causal impact of changes in enforcement inputs. Second, outcome measures may be useful as Probative Indicators and Diagnostic Indicators that do not serve a high-stakes evaluative function. That outcome measures may yield false positives is not so costly, for purposes of triage and screening. Third, many input indicators should be *defined* in relation to outputs. For example, an input indicator might ask about the ratio of increased outputs to increased inputs.

Part 6 argues that many individual Indicators should incorporate quantitative yardsticks, including longitudinal and comparative measures. Such measures are appropriate where the Executive Branch, in fulfilling its role of reasonably interpreting Congressional intent, chooses not to stipulate absolute qualitative or quantitative yardsticks for normative evaluation of country performance. By asking whether the country is improving relative to previous performance, we need not stipulate an arbitrary metric. And, by asking whether a country is performing well relative to its peers, we are using a metric that has three advantages: It embodies a standard that is demonstrably feasible. It represents a prevailing international standard. And, it creates incentives for each government to continuously improve its compliance in order to out-perform its peers. Part 6 further argues that, if carefully constructed, longitudinal, comparative, and revisable Indicators are not inconsistent with basic conceptions of universal human rights.

Part 7 proposes that the ILAB methodology rest on Indicators taking the binary form, rather than on the 9-cell matrix proposed by the NAS. The pilot analysts had great difficulty in applying the NAS methodology, in part because the gradated thresholds on both the vertical and horizontal axis have no clear, settled meaning. Moreover, many Indicators pertaining to compliance “naturally” take the binary form, as evidenced by the fact that many NAS Indicators are themselves binary and therefore cannot be scored using the 3 by 3 matrix. Most important, the binary form has the advantage of simplicity – simplicity in formulating, applying, and explaining the methodology.

Part 8 addresses the question of whether the definition or weighting of Indicators

should be “adjusted” based on such country-specific characteristics as level of development, type of political regime, government capacity, labor relations model, religion, colonial history, or legal origins. Part 8 concludes that the methodology should start with only a single control variable: real income per capita. That variable is relevant to U.S. trade legislation, which authorizes the USTR and the President to consider a country’s level of economic development or living standards when granting or denying trade benefits. Other control variables, such as those just noted, should not be built into the Indicators until such time as ILAB is convinced – after iterative rounds of diagnosis and assessment – that they are indeed hard constraints rather than soft constraints that the policy instruments of trading partners can and should loosen.

Part 9 then turns to the urgent problem of data availability and reliability. That Part argues that we should be concerned more with incentivizing future production of data than with jerry-rigging the methodology to deal with current gaps in data. Two solutions to data insufficiency are proposed. First, many of the Indicators should be framed as twins – the first calling on analysts to apply the Indicator to the available data and the second asking whether governments themselves have verifiably demonstrated their compliance with the Indicator. The Twin Indicators therefore measure governments’ collection of data about *specific* aspects of worker rights. Second, the Capacity-Building Indicators proposed in Part 4 measure governments’ overall *systems* for collecting and publishing data. Both the Twin Indicators and the Capacity-Building Indicators create incentives for governments to come forward with data and with justifications for their comparative performance levels and

therefore *to produce* data, performance targets, and self-assessments.

Part 10 then defends the premise, posited from the start, that we should construct three bodies of Indicators: Probative Indicators, Diagnostic Indicators, and Assessment Indicators. That Part analyzes the types of Indicators that are best suited for each body. The actual Indicators are set forth in Appendices A, B, and C. Appendix A formulates proposed Indicators for the comprehensive body of Assessment Indicators. The body of Assessment Indicators is quite lengthy. It is, at this stage, presented as a set of candidate Indicators that ILAB may abridge, perhaps in the iterative process of applying and revising the Indicators (while maintaining valid weighting and balance among categories).

The candidate Assessment Indicators are formulated by identifying aspects of NAS Indicators that call for refinement, and by revising, replacing, or supplementing them. Annotations in the body of Indicators explain how particular NAS Indicators can be strengthened by applying the conceptual criteria discussed in the main body of this paper, and how the proposed revised Indicators satisfy those criteria. The body of Assessment Indicators therefore comprises refinements of the three categories of NAS Indicators, together with the new category of Capacity-Building Indicators.

Appendices B and C then set out the more limited sets of Probative and Diagnostic Indicators. Although some of the Probative and Diagnostic Indicators are drawn from the Assessment Indicators, they are not just a stripped down version of selected Assessment Indicators. The Probative and Diagnostic Indicators are instead constructed as their respective names suggest. Probative Indicators probe for likely pathologies – for

governments that may be least committed to enforcement of worker rights or are most likely to actively defile those rights. Diagnostic Indicators more intensively screen countries, to identify countries that warrant comprehensive evaluation via the Assessment Indicators.

Part 11 canvasses alternative strategies for weighting and aggregating the Assessment Indicators: (1) equal weighting; (2) weighting based on hierarchies in the legal sources; (3) principal component analysis; (4) *ex ante* multivariate modeling; and (5) iterative, *ex post* multivariate weighting. Quantitative testing and validation of the alternative weighting and aggregation strategies are beyond the scope of this paper. As stipulated in the research proposal, Part 11 discusses the pros and cons of the alternative strategies, looking to the use of those strategies in other indices both in the labor field and in other domains of rights enforcement. Part 11 concludes that principal component analysis is unsuitable and that multivariate modeling is unfeasible, and discusses the relative advantages of the other two strategies.

Part 12 offers some conjectures about an overarching process for conducting diagnosis and evaluation of trading partners. That Part proposes that ILAB apply the three bodies of Indicators through an iterative dialogue with each trading partner and relevant stakeholders. In each round of dialogue, ILAB will apply the pertinent body of Indicators and reach a provisional appraisal; the trading partner will then have an opportunity to offer new data and argumentation to excuse the compliance problems provisionally found by ILAB analysts; and stakeholders will have an opportunity to respond to ILAB's provisional findings and the trading partner's arguments. ILAB analysts or other Executive Branch officials will

decide whether the new data and arguments warrant a revision of the provisional appraisal or a still deeper revision of the Indicators and sub-indicators applicable to each country.

Part 13 briefly discusses some possible future steps for this research program.

3. How to Formulate an Indicator: A Step-by-Step Textual Explanation

This paper proposes various categories and types of Indicators. As a preview and overview, this sub-part summarizes how the various categories of Indicators fit together, in the formulation of each Indicator. The subsequent sub-part then presents a flow chart depicting the same process for formulating each Indicator.

First, each Indicator fits into **one of three bodies of Indicators**: Probative Indicators to identify countries that may have the poorest compliance records; Diagnostic Indicators for more intensively screening countries for symptoms of compliance failures; and Assessment Indicators for comprehensive evaluation of countries' compliance.

Second, each Indicator measures compliance with **one major labor right or standard**. In this paper, we focus on the following three major rights and standards: (1) freedom of association, rights to organize, and rights to bargain collectively; (2) rights against employment discrimination; and (3) acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health.

Third, every Indicator is a **binary** (yes/no) Indicator. Almost all Assessment Indicators are drafted so that an answer of "yes" indicates positive performance by the government and "no" indicates negative performance. For purely grammatical reasons, there are some

Indicators for which an answer of “yes” indicates negative performance and “no” indicates positive performance. In either event, the Assessment Indicator is scored “1” for an answer that indicates positive performance by the government and “0” for an answer that indicates negative performance. The scores for individual Indicators may then be adjusted based on the differential weights attached to each Indicator, depending on the weighting and aggregating scheme that ILAB ultimately chooses to adopt. All Probative Indicators and Diagnostic Indicators are framed so that an answer of “yes” indicates negative performance.

Fourth, each Indicator addresses either one of **three phases of enforcement** or the **outcome of enforcement**: (1) Substantive Law Indicators measuring the government’s performance in putting laws on the books that satisfy standards for internationally recognized worker rights; (2) Enforcement Indicators measuring the institutions, resources, procedures, and practices that the government devotes to enforcing the laws on the books; (3) Capacity-Building Indicators that measure the government’s systems for collecting data, setting policy targets for improved enforcement, evaluating whether those targets are met, including stakeholders in all phases of defining and enforcing rights, ensuring transparency in compliance, and comparing performance with peer countries; and (4) Outcome Indicators that measure actual compliance by employers.

Fifth, the bulk of proposed Assessment Indicators are “**input**” indicators – that is, Substantive Law, Enforcement, and Capacity-Building Indicators, as distinguished from Outcome Indicators – since it is government effort, including effort in producing data, that the U.S. government wishes to incentivize, and since there is a danger of double-counting the

cause (enforcement effort) and the effect (employer compliance). Nonetheless, as mentioned in sub-part 2 of this Introduction, outcome measures are used in three ways. First, Outcome Indicators are included for purposes of ILAB research on the impact of government effort on employer compliance. Second, Outcome Indicators are used for probative and diagnostic purposes. Third, many input Indicators are *defined* in terms of the relationship between inputs and outcomes. For example: What is the rate of reduction of workplace fatalities for each additional dollar spent on occupational safety and health enforcement?

Sixth, each Indicator is framed as **one of four qualitative “norm types,”** depending on the degree of detail and revisability of the Indicator: (1) a fixed bright-line rule; (2) a revisable bright-line rule; (3) a fixed standard; or (4) a revisable standard. Each norm-type is best suited to different aspects of compliance, depending on the precision and fluidity of both the facts and the values that apply to the particular aspect of compliance.

Seventh, every Indicator is framed as a **specific type of normative metric** – that is, a values-based metric that *evaluates* what constitutes positive or negative performance:

(1) a **qualitative norm**, asking about the existence of some law, institution, or procedure; or

(2) a **quantitative norm**, including

(a) an **absolute quantitative metric**, such as a requirement of overtime pay for work weeks in excess of 48 hours, or

(b) a **zero tolerance norm** for certain quantifiable events, such as the number of prosecutions of known murderers of trade unionists, or

(c) a norm of a **“trivial” number of violations** for certain quantifiable events, such as the number of trials that are closed to the public in labor cases, or

- (d) a quantitative metric based on the country's **longitudinal** rate of improvement of compliance, or
- (e) a quantitative metric based on the country's performance **compared to some quantitative measure of peer countries' performance**, such as the average performance among countries in the same quintile of real income per capita.

Eighth, every Indicator will constitute the **heading of a drop-down window**. When applying the Indicator to a specific country, the analyst will identify compliance issues (sub-indicators) that are specific to that country and enter those issues in the window, alongside sub-indicators that are common to all countries. The sub-indicators will flesh out the Indicator. The analyst will also enter country-specific information sources. This process is illustrated in Figure 2 in Part 3.2 below. Hence, through the iterative process of probing, diagnosing, and assessing countries, ILAB analysts will accumulate country-specific sub-indicators and data sources.

4. Formulating Each Indicator: A Flow Chart

Figure 1 below is a flow chart showing the steps in formulating each Indicator, for all three bodies of Indicators set out in Appendices A, B, and C – that is, Probative Indicators, Diagnostic Indicators, and Assessment Indicators. For example, consider the following Indicator:

In the preceding two years, did the tribunals that hear cases on employment discrimination publish their final decisions in writing, in all but a trivial number of cases?

Since every Indicator takes a binary form, that element of the Indicator is not shown in the flow chart. Following the flow chart, this Indicator was constructed as follows:

Step One: This is an **Assessment Indicator**, since it is intended as one Indicator in the body of comprehensive Indicators used to evaluate whether the country is in full compliance with its obligations under U.S. legislation and treaties. The treaties specifically require tribunals to publish final decisions in writing.

Step Two: This is an Indicator that measures compliance with **rights against employment discrimination**.

Step Three: This is an **Enforcement Indicator**, since it measures the actual functioning of labor tribunals, which are enforcement institutions.

Step Four: This is an **input Indicator**, since it measures government effort.

Step Five: This is a **fixed bright-line rule**, since it defines the government's obligation with precision – it must publish final decisions in writing – and there is no reason to think that changes in facts or values will alter the international consensus that final decisions must be issued publicly and in writing.

Step Six: This is a **quantitative** norm, using the threshold of “**non-trivial number of violations**,” appropriate to rules or standards imposing requirements that are absolute but that nonetheless cannot reasonably be expected to be satisfied in every instance.

FIGURE 1: How to Formulate an Indicator: A Flow Chart

STEP ONE: Is the Indicator a **PROBATIVE INDICATOR**, a **DIAGNOSTIC INDICATOR**, or an **ASSESSMENT INDICATOR**?

1. If ILAB's purpose is the high-stakes **EVALUATION** of country compliance, then the Indicator is an **ASSESSMENT INDICATOR**.
2. If ILAB's purpose is to **IDENTIFY** countries that likely have **ESPECIALLY POOR COMPLIANCE**, calling for further research, technical assistance, or comprehensive assessment, then the Indicator falls in a short-list of **PROBATIVE INDICATORS** or a long-list of **DIAGNOSTIC INDICATORS**.

STEP TWO: Which **LABOR RIGHT** does the Indicator measure?

1. An Indicator for **FREEDOM OF ASSOCIATION, RIGHTS TO ORGANIZE, and RIGHTS TO BARGAIN COLLECTIVELY**.
2. An Indicator for **RIGHTS AGAINST EMPLOYMENT DISCRIMINATION**.
3. An Indicator for **ACCEPTABLE CONDITIONS** with respect to **MINIMUM WAGES, HOURS OF WORK, and OCCUPATIONAL SAFETY AND HEALTH**.

STEP THREE: Which **PHASE OF COMPLIANCE** does the Indicator measure?

1. **SUBSTANTIVE LAW INDICATOR** for measuring law on the books.
2. **ENFORCEMENT INDICATOR** for measuring institutions, resources, procedures, and the government's actual performance in enforcing law on the books.
3. **CAPACITY-BUILDING INDICATOR** for measuring data collection, policy targets, self-evaluation, stakeholder participation, transparency, and comparisons with peers.
4. **OUTCOME INDICATOR** for measuring actual employer compliance.

STEP FOUR: Make **LIMITED** use of **OUTPUT MEASURES** in formulating **ASSESSMENT INDICATORS**.

NB: **INPUT** Indicators include Substantive Law Indicators, Enforcement Indicators, and Capacity-Building Indicators. Outcome Indicators are **OUTPUT** measures. The primary goals of assessment are to measure inputs – that is, the government’s use of policy instruments within its control – and to incentivize collection of data about such inputs. Using Outcome Indicators as proxies for such inputs will reduce the incentive to collect data about inputs. These rationales are much weaker for Probative and Diagnostic Indicators.

1. Some **INPUT INDICATORS** can be **DEFINED** by the relation between inputs and outputs. For example: “In the preceding five years, what is the ratio of increased employer compliance with minimum wage laws (output) and increased number of labor inspectors (input)?”
2. Some **OUTCOME INDICATORS** are useful for **ILAB RESEARCH** on the actual impact of government enforcement measures.
3. Some **OUTCOME INDICATORS** are useful as **PROBATIVE** or **DIAGNOSTIC INDICATORS**.

STEP FIVE: Every Indicator is framed as one of the following **FOUR “NORM TYPES”**:

NB: An Indicator is a bright-line rule if it defines in detail the precise facts measured by the Indicator. An Indicator is a standard if it states a general principle or policy that the analyst must apply to complex facts that cannot be pre-specified.

1. **FIXED BRIGHT-LINE RULE**, for contexts where facts and values are static.
2. **FIXED STANDARD**, for contexts where facts are complex and values are static.
3. **REVISABLE BRIGHT-LINE RULE**, for contexts where facts are likely to change but values are fixed.
4. **REVISABLE STANDARD**, for contexts where facts and values are fluid.

STEP SIX: Every Indicator must contain one of these **NORMATIVE METRICS:**


NB: **NORMATIVE METRICS** tell analysts whether a “yes” score for an Indicator indicates positive or negative performance by the government.

1. **QUALITATIVE NORM**, asking about the existence of some rule of law, institution, procedure, or practice. For example: “Does the law prohibit the discharge of workers for anti-union reasons?”
2. **QUANTITATIVE NORM**, including:
 - a. **ABSOLUTE NUMBER**. For example: “Does the law require employers to pay at least 125 percent of the regular hourly wage for hours worked in excess of 48 per week?”
 - b. **ZERO TOLERANCE NORM**. For example: “In the preceding year, did the government fail to prosecute any known murderer(s) of trade union supporters?”
 - c. **“TRIVIAL” NON-COMPLIANCE**. For example: “In the preceding year, were all but a trivial number of evidentiary hearings in labor cases open to the public?”
 - d. **LONGITUDINAL METRIC**. For example: “Did the number of labor inspectors increase by at least 5 percent per year in the last five years?”
 - e. **COMPARATIVE METRIC**. For example: “Does the labor administration budget per worker exceed the average among countries in the same quintile of real income per capita?”

STEP SEVEN: Every Indicator is the **HEADING** of a **DROP-DOWN WINDOW** for **EACH COUNTRY**.

The window will show **SUB-INDICATORS**, some of which are country-specific, some of which are common to all countries. The window will also show **DATA SOURCES FOR EACH COUNTRY**. ILAB analysts will enter new sub-indicators and data sources as they apply Indicators to each country.

1. The NAS Methodology and the Michigan Pilot-Testing of that Methodology



The NAS methodology is the starting point for the revised methodology proposed in this paper. The NAS committees' work was assiduous and productive. Nonetheless, much of the analysis in this and subsequent Parts of the paper diverge from the NAS methodology – for two reasons: First, the point of this project is to refine the NAS methodology in light of problems found when the methodology was pilot-tested. The project necessarily focuses on the areas that call for refinement. Second, the NAS methodology rests on a somewhat different policy goal than does this paper; hence, this paper necessarily focuses on the conceptual underpinnings that distinguish the two projects. But at all stages of the analysis, the reader should bear in mind that the current research could not get off the ground without the invaluable, foundational work of the NAS committees.

1.1. The NAS Indicators and Matrix Methodology

In 2001, the ILAB engaged the National Academy of Sciences to formulate a methodology and information base for the assessment of countries' compliance with labor rights and acceptable conditions of work. The NAS research produced five volumes, published in 2003 and 2004, summarizing the content of workshops and fora in the areas of

quality of information, domestic regulation, national legal frameworks, international perspectives, and regional regulation. (National Research Council, 2004b; 2003a; 2003b; 2003c; 2003d).

The sixth and key volume, published in 2004, is entitled *International Labor Standards: Techniques and Sources of Information* (hereafter “*NAS Report*”). That *Report* generated the existing NAS body of Indicators, matrix methodology, and WebMILS database. (National Research Council, 2004). The content of that research and the source material on which it relies are systematically presented and summarized in that volume, are summarized again in the Michigan evaluation, and are discussed in this author’s *Summary and Analysis of Interviews with ILAB Staff and Literature Review and Bibliography*. (Barenberg, 2010).

The *NAS Report* concurrently surveys sources of information about both international and domestic labor rights and about social science data pertaining to labor market performance. In its survey of information sources, the *Report* looks primarily to reports produced by public international organizations, especially the International Labor Organization (ILO); by U.S. government agencies, particularly the State Department; by domestic statistical agencies; and by nongovernmental organizations such as the International Confederation of Free Trade Unions (now the International Trade Union Confederation), Freedom House, Human Rights Watch and the Fair Labor Association. (National Research Council, 2004, ch. 2-3.)

After surveying the sources of information, the *NAS Report* devotes individual chapters to each of the rights that the ILO has designated “fundamental” (freedom of

association, rights to organize, and collective bargaining rights; rights against forced labor; rights against child labor; and rights against employment discrimination) as well as acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health. The “internationally recognized worker rights” contained in U.S. legislation and agreements include all of these, except for nondiscrimination. The ILO fundamental rights, on the other hand, do not include acceptable conditions of work. (ILO, 1998b).⁷

Each chapter of the *NAS Report* begins with a brief discussion of the social problem (e.g., child labor) and the economic, historical, legal, and policy response to the problem. Each chapter ends with three lists of Indicators: one list for “Legal Framework,” another for “Government Performance,” and a third for “Overall Outcomes.” The Legal Framework Indicators ask primarily about substantive norms on the books. For example, one Indicator asks: “whether there are laws that prohibit discrimination in employment on the grounds of race, color, sex (including sexual harassment), religion, political opinion, national extraction, and social origin.” (National Research Council, 2004, p. 207).

The Government Performance Indicators ask largely about government institutions and processes, and the actual functioning of such institutions and processes in enforcing the substantive rules of law. For example, one Indicator states: “an agency to promote and enforce laws protecting occupational safety and health.” Another reads: “the effectiveness of

⁷ In its discussion of each right, the *NAS Report* does not distinguish between ILO conceptions of labor rights, other UN conceptions of labor rights, regional conceptions of labor rights, domestic conceptions of labor rights, and the concept of “internationally recognized worker rights.” The latter is an artifact of U.S. law, subject to the Executive Branch’s reasonable interpretation of Congress’s intent in enacting that term.

the complaint mechanism (in such terms as number of complaints brought compared with number of complaints heard, number of prosecutions, fines, or arrests, and length of time for complaint resolution).” (National Research Council, 2004, p. 240).

The Overall Outcomes Indicators ask both about immediate compliance outcomes (e.g., “incidents of discrimination against union organizers, unions or employer associations”) and about labor market indicia (e.g., “distribution of women and other potentially vulnerable groups by sector”). (National Resource Council, 2004, pp. 119, 211).

The NAS noted that some of its Indicators – such as all the Overall Outcome Indicators for freedom of association and collective bargaining – need additional “interpretation to be used to measure compliance.” (National Resource Council, 2004, p. 119). The chapters on forced labor, child labor, and employment discrimination also include a section on “Associated Factors” – essentially additional or optional Indicators that “can be helpful” in assessing a government’s compliance. (National Research Council, 2004, p. 211).

The second component of the NAS methodology is a 3 by 3 matrix, shown in Figure 2 below. When assessing a government, the analyst must choose which of the nine boxes applies to each Indicator. Along one axis is the degree of the “problem” in a government’s compliance with the Indicator: “some problems,” “more extensive problems,” and “severe problems.” Along the other axis is the direction of change in compliance: “improving,” “steady state,” and “worsening.” The NAS encouraged analysts to take a flexible approach to the categories, based on country context:

[The Committee on Monitoring International Labor Standards (CMILS)] did not attempt to delineate the thresholds along either axis that would place a country above or below

the line separating the categories because we do not believe that it is possible to set thresholds that are appropriate for every country and every situation. Judgments about government performance have to reflect the resources that a government has to work with: the question is not solely whether a government is spending more or less than a given percent of its gross domestic product, but whether it is doing the best it can with what it has. And... many of the outcome indicators can mean different things under different circumstances. The burden of demonstration, argumentation, and persuasion necessarily falls on assessors and their critics to defend or challenge any given assessment.

(WebMILS at <http://webapps.dol.gov/webmils/matrix-assessment-tool.aspx>, visited 10-26-10).

Nonetheless, CMILS offered four factors for defining the degree of compliance problems for any given Indicator: numerical frequency, breadth in the workforce or country; qualitative severity; and the impact of the first three factors on the difficulty of coming into compliance.

Id.

FIGURE 2: The NAS Matrix

	Improving	Steady State	Worsening
Some Problems			
More Extensive Problems			
Severe Problems			

The third component of the NAS methodology is a website called WebMILS. The site allows searches by country, labor standard, Indicator type (Legal Framework, Government

Performance, and Overall Outcome), and information source (nongovernmental organizations, international organizations, national government sources, etc.) The methodology calls on analysts to use WebMILS as their information sources when scoring the matrices for particular Indicators applied to particular countries.

The NAS did not develop a methodology for calculating a composite score or comparative ranking of countries. The WebMILS website emphasizes that “CMILS recommended against precise scoring or ranking of countries against each other.” (WebMILS at <http://webapps.dol.gov/webmils/assessing-compliance.aspx>, visited 10-26-10).

1.2. The Michigan Pilot Testing of the NAS Indicator Methodology

In order to test and evaluate the NAS Indicators and methodology, ILAB engaged social scientists of the Institute for Research on Labor, Employment and the Economy of the University of Michigan. The results of the evaluation were published in the 2009 *Michigan Report*. (Root and Verloren, 2009).

The Michigan researchers assembled three-person panels to apply the Indicators to three countries. The panelists were experts and consultants, with varying expertise and experience in labor relations, labor policy, labor conditions, and labor rights. All three panels applied the Indicators for freedom of association, rights to organize, and rights to bargain collectively. In addition, one panel applied the Indicators for acceptable conditions of work, another panel applied the Indicators for forced or compulsory labor, and the third panel

applied the Indicators for rights against employment discrimination. Each panel member independently applied the Indicators, drawing on the WebMILS database. After convening to discuss the variation in their individual assessments the panelists, again independently, revised their assessments. In their written assessments, the panelists used the 3 by 3 matrix described above. In practice, the panelists used an improvised 4 by 4 matrix, adding a “no assessment” option along each axis.

The Michigan evaluation made several findings. Rates of “non-assessment” were high, indicating either confusion in the meaning of Indicators or lack of available information in the WebMILS database or other sources. Indicators with at least one non-assessment (per panel) ranged up to 81.3 percent, with a rate of non-assessment less than 50 percent for only one Indicator. The rates of initial agreement among all three panelists on “levels of compliance” were low – ranging from 0 percent to 34.2 percent, and in four out of six cases falling below 15.8 percent per Indicator. Rates of initial agreement among all three panelists on “direction of change” were somewhat higher, but still low – from 13.5 percent to 73.7 percent, with all but one falling below 31.6 percent. Pair-wise agreements among panelists (that is, agreement between two panelists) ranged from 30 to 70 percent. In other words, if ILAB analysts use the NAS methodology, the outcome of assessments will more often than not depend on the individual staff person assigned to the task – at least, if the Michigan test is a good predictor of ILAB assessments and if ILAB assessments are conducted by individual analysts without group deliberation.

The *Michigan Report* concluded that a significant problem is the lack of clarity in the

Indicators: “[M]any, if not most, of the individual Indicators are themselves complex and subject to interpretation....” (Root and Verloren, 2009, p. 16). The *Report* concluded that instances in which panelists rested their judgment on differing information were “[t]he easiest to resolve,” since the same information could be shared in the second round of assessments. (Root and Verloren, 2009, p. 18). But “differences in the interpretation of an Indicator” were “less likely to be resolved.” (Root and Verloren, 2009, p. 18). Likewise for panelists’ differing views about whether Indicators should be assessed relative to an absolute baseline or instead relative to similarly situated countries. (Root and Verloren, 2009, p.18).⁸

According to the Michigan panelists, the most prevalent problems in the clarity of Indicators were: ambiguously worded Indicators; double-barreled Indicators; inconsistent terminology from Indicator to Indicator; Indicators covering similar subject matter yet differently worded and therefore inconsistent; disjunctures between Indicators and international standards; Indicators that intrinsically called for 2 by 3 rather than 3 by 3 assessment; and Indicators that lack normative guidance about what constitutes a compliance problem. (Root and Verloren, 2009, p. 32).

The following is an example of an ambiguous NAS Indicator:

Whether there are defects in the government’s complaint process, such as excessive delays or expenses, light penalties, or nonpunishment of offenders

An analyst applying this Indicator may be unsure about several questions: (1) To which complaint process does the Indicator refer? Administrative complaints? Judicial complaints?

⁸ Note that the latter problem is also a question of clarity in the definition of Indicators.

Appellate complaints? And how is the analyst to score the Indicator if all of these complaint mechanisms are considered? (2) In assessing “delays,” which time periods should the analyst consider? Delays between filing a complaint and the start of preliminary proceedings? Delays between preliminary proceedings and trials? Delays between trials and final disposition? Delays between final disposition and enforcement of the remedial order? And, for each time period, what constitutes “excessive” delay? (3) In assessing “expenses,” which expenses are relevant? The cost of hiring an attorney? Filing fees, if any, exacted by the relevant tribunal? Litigation costs, such as the cost of gathering evidence or conducting depositions? The cost of enforcing a remedy? And what constitutes “excessive” expenses? (4) In assessing “penalties,” which remedies are relevant? The text of the Indicator asks about punitive remedies (“penalties” or “nonpunishment”) but many if not most remedies for labor violations are non-punitive awards, such as compensatory damages, reinstatement, and other injunctive relief. Very likely, the intention behind the Indicator was to measure these non-punitive remedies as well as punitive remedies – since the Indicator begins by asking about defects in complaint processes and does not refer to criminal indictments that would initiate most punitive proceedings. But the text of the Indicator does not make that intention clear.

The following is an example of a double-barreled NAS Indicator:

the adequacy of personnel and budgets of labor regulation departments compared to number of workplaces, the frequency and adequacy of labor inspections, the caseloads of labor administrative bodies and labor court, and whether bribes are paid to labor inspectors by employers without effective prosecution

This Indicator asks for between six and eight variables, depending on the analyst's interpretation of ambiguities in the Indicator. Several of the variables are incommensurable (number of people, monetary budget, time periods between inspections, adequacy of inspections, number of cases handled by two or more tribunals, existence or degree of bribery), making it difficult for the analyst to arrive at a single gradated score for the Indicator.

Note also that the just-mentioned Indicator fails to give normative guidance to the analyst about what constitutes a compliance problem. The Indicator asks for caseloads of administrative and judicial bodies. Is a high caseload a normatively positive metric, on the ground that it shows high government capacity and easy access to tribunals by aggrieved workers? Or is it a negative metric, on the ground that it shows many violations and slow case-processing?

The following two NAS Indicators, applicable to rights against employment discrimination, are non-uniform and perhaps inconsistent with the preceding Indicator applicable to freedom of association, the right to organize, and the right to bargain collectively:

the breadth of labor inspections in the country, in terms of number of visits, frequency of visits, number of workers covered, etc.

the level of resources devoted to the labor inspectorate in terms of personnel and budget, absolute or relative to number of workers or spending

While, technically, there may be no logical contradiction among these two Indicators and the previous Indicator, the analyst may be puzzled why the collective bargaining Indicator folds the personnel and budget of the labor inspectorate into the personnel and

budget of the overall “labor regulation departments,” while the nondiscrimination Indicators disaggregate the personnel and budget of the labor inspectorate. The non-uniformity in wording creates a risk of double-counting of labor inspectorate personnel and budget. Also potentially confusing is the fact that the collective bargaining Indicator asks for a comparison between personnel and budgets, on the one hand, and the number of workplaces, on the other; while the nondiscrimination Indicators ask for a comparison of personnel and budget with the “number of workers” or with “spending.” Note that the definition of “budget” in the latter indicator may be circular, since it measures “budget” relative to unspecified “spending.” The analyst may also be puzzled why the collective bargaining Indicator asks about the “frequency and adequacy” of inspections, while the nondiscrimination indicator asks about the “breadth” of inspections, including frequency and other variables. In addition, the following Indicators pertaining to acceptable conditions of work also ask about the “breadth” of labor inspection, but define that term inconsistently with the Indicator pertaining to the “breadth” of labor inspection for nondiscrimination:

the breadth of labor inspections in the country (number and frequency of visits, geographic regions, or industry sectors inspected)

This inconsistency in terminology is compounded by the ambiguity of this Indicator, which gives the analyst discretion to define “breadth” along one or more of three alternative dimensions, and by analogous ambiguities in the two preceding indicators.

The following Indicator is an example of a binary NAS indicator:

Whether the country has ratified ILO Convention No. 87

Since this indicator calls for a yes/no answer, analysts are unable to place it in the triadic

gradation of the NAS matrix.


In addition to these types of problems with many individual Indicators, the panelists also expressed concern about the limited information available through the WebMILS database. First, searches using WebMILS typically lead to links to the home page of other websites, not to webpages with information referring to the specific country, labor standard, and Indicator type. For example, a search for Brazil/Freedom of Association/Government Performance leads to a list of links to the home pages of “ILO Global Reports,” “International Trade Union Confederation,” “Government of Brazil,” “International Labor Rights Forum,” “Catholic Relief Services,” “Human Rights Watch,” “Social Accountability International,” “International Industrial Relations Association,” and other websites.

Second, the pilot analysts were frustrated by the assumption that research would be conducted exclusively through online papers, many of which are secondary sources. “[T]here was [also] a recurring theme that what one needed was a local expert – someone ‘on the ground’ – who could provide up-to-date and contextual information.” (Root and Verloren, 2009, p. 42).

Finally, the *Michigan Report* found that “[t]he effort in the pilot test to make an overall assessment of the legal framework, government performance, and outcomes continually came up against the question of how one pulls together the individual assessments of Indicators. This was often articulated in terms of establishing priorities reflecting the relative importance of Indicators.... [D]eveloping some guidance for how to move from the specifics to an overall assessment would be a significant advance in the use of the Indicators.” (Root

and Verloren, 2009, p. 44).

2. The Traditional Compliance Model



As noted in the Introduction, the critical first step in formulating a body of Indicators is to precisely analyze and understand the underlying concepts that the Indicators will measure. (OECD, 2008). In our case, the underlying concepts are “denial of internationally recognized worker rights,” “taking steps to afford internationally recognized worker rights,” and “effective enforcement of labor law.”⁹

This Part makes three straightforward conceptual points. First, U.S. trade legislation and treaties are legal instruments. The rights and obligations contained in the labor provisions of those instruments are legal concepts. Second, the NAS committees saw their conceptual task as a different one than the stated purpose of this project and therefore did not systematically apply legal-interpretive methods and regulatory analysis to the formulation of its Indicators. The NAS Indicators were based on a variety of desiderata: labor rights; labor market indicia; public policy criteria; and social science measures. The stated purpose of this research is to develop a methodology for measuring compliance with the legal concepts contained in trade legislation and trade agreements.

Third, if the refined methodology is to serve the purpose of enabling the Executive Branch and other actors to determine whether countries are in compliance with legislation and treaties, it is useful to begin with a thought experiment: What would the Indicator

⁹ See footnote 6 above.

methodology look like, if it were refined through systematic application of the traditional method of regulatory compliance? Part 3 then discusses the weaknesses in the traditional method, and the features of Indicators that respond to those weaknesses.

2.1. International Labor Rights are Legal Concepts

Several federal trade statutes impose obligations on the USTR and the President to ensure that countries that trade with the U.S. comply with internationally recognized worker rights or make progress toward such compliance, subject to certain caveats. Several treaties impose obligations on trading partners and the U.S. to “effectively enforce” core labor rights, internationally recognized worker rights, or domestic labor law, to ensure that labor courts satisfy certain procedures, and to meet other collateral standards, again subject to various qualifications. When Congress ratifies a treaty and enacts legislation implementing the treaty, as it has done for the treaties just mentioned, the terms of the treaty constitute binding U.S. law.

The President of the United States is obligated by Article II of the Constitution to faithfully execute federal legislation and treaties. The President is therefore required, both by his general Executive powers and by the specific terms of trade legislation and implementing legislation, to enforce the labor provisions of the statutes and treaties just mentioned.¹⁰

In short, the labor provisions that bind our trading partners and the United States are

¹⁰ As noted in the Introduction, ILAB’s role is to inform the President’s decision by formulating a rigorous methodology and, within the constraints of the resources appropriated to ILAB by Congress, gather data and apply that methodology to the data.

quintessentially *legal* rights and obligations, codified in quintessentially legal instruments (legislation and treaties).

In some instances, those legal instruments incorporate by reference other quintessentially legal rights and obligations, such as rules promulgated in Declarations and Conventions of the International Labor Organization, and labor laws enacted by national governments. The ILO Declaration and Conventions – as well as other relevant Declarations and Conventions of the United Nations, the European Communities, the Inter-American System, and other regional bodies – are *understood* as legal instruments by the ILO itself, by the UN, by the regional bodies, by domestic agencies and courts, by human rights organizations, by labor federations, by business federations, and by scholars. And the rights and obligations contained in those instruments and in national systems of labor law are also universally understood as legal concepts. The juridical nature of labor rights is exemplified by the fact that every member of the ILO Committee of Experts on the Application of Conventions and Recommendations – the ILO body that authoritatively determines whether a country is complying with ratified ILO Conventions – is a professor of labor or international law, a labor or high court judge, or other distinguished labor or international lawyer.

Likewise, the primary national *institutions* whose performance must be measured are legal institutions – labor inspectorates, labor prosecutiates, labor boards, and labor courts (or courts of general jurisdiction, in countries without specialized labor courts, such as the United States). These are “legal” institutions in a fourfold sense. First, their primary function is to elaborate and enforce substantive legal norms. Second, their formal structure and

function are established by law. Third, their formal performance is supervised by higher level legal institutions, such as national Ministries of Labor or appellate courts and, at least in an attenuated way, international legal bodies such as ILO supervisory organs, UN committees, regional commissions, and the like. Fourth, their actual performance is strongly influenced by their relationship with collateral laws and legal institutions – such as corporate law and regulatory bodies, property law enforced by courts of general jurisdiction, and antitrust law enforced by administrative and judicial bodies – as well as nongovernmental norm-generating institutions such as collective bargaining agreements, labor-relations regimes, corporate managerial systems, modes of production, supply chains, local worker communities and cultures, and so on. That is, the formal enforcement institutions are embedded in a multi-level and polycentric network of relationships with other institutions and actors occupying a spectrum of legal to quasi-legal to non-legal.

2.2. Legal Concepts Call for Sustained Interpretive and Regulatory Analysis

Although the *NAS Report* refers to ILO and other legal sources alongside a variety of other secondary and non-legal materials, it does not use an explicit interpretive or regulatory methodology. As a result, the NAS Indicators do not fully capture the underlying concepts that concern us – that is, “denial of internationally recognized worker rights,” “taking steps to afford internationally recognized worker rights,” and “effective enforcement of labor law.”¹¹

¹¹ As already noted, this is not intended as criticism of the NAS committees, since their project and the current project pursue somewhat different objectives. The point rather is to carefully take those different objectives into

Rather than interpreting primary legal sources, the *NAS Report* (1) directed analysts to secondary reports – such as ILO committee reports, State Department Country Reports on Human Rights Practices, and ITUC Papers – which do not systematically focus their findings on the provisions of trade legislation and treaties that are the focus of this paper, and (2) formulated Indicators based for the most part on novel definitions of labor standards, untethered to Congressional intent and the settled meanings of rights and standards in international and domestic labor law, including the legislation and treaties.

The *NAS Report's* discussion of employment discrimination is characteristic. The discussion begins with Webster's Dictionary definition of "discrimination," which tells us that discriminating means distinguishing. The *Report* then posits that whether a distinction is unacceptable depends on the grounds for the distinction (e.g., race, gender, and so on) and the range of activities in which the distinction is impermissible (e.g., wages, hiring, and so on). (National Research Council, 2004, p. 197). From the standpoint of the international and domestic law of employment discrimination, however, those questions are merely the presumptive starting points of legal-regulatory analysis. That is, we start with the axiomatic rule, for example, that employers are barred from discriminating against a racial group in setting the level of wages. We must then specify more complex legal rules that have more traction in evaluating compliance in particular factual contexts. Such legal rules specify the circumstances under which differential wages constitute wrongful discrimination based on race. Those legal rules must be articulated in the Indicators, if the Indicators are to serve our

account as we refine the methodology to serve our particular objectives, in light of the problems found in the pilot testing of the NAS methodology.

policy purposes.

To continue with the example of employment discrimination, there are three principal settled tests – in international law and across leading domestic legal systems – of whether differential treatment constitutes wrongful discrimination: (1) “disparate treatment,” defined as the disadvantaging of a member of a protected group based on animus toward that group, such as a racist or sexist motive; (2) “harassment,” defined as the employer’s failure to ensure that co-workers and supervisors do not create a hostile work environment for members of a protected group;¹² and (3) “disparate impact,” defined as an employment practice that is not motivated by animus toward members of a protected group, that nonetheless disproportionately burdens them, and that is not justified by the objective qualifications for the job or by other business necessity.

In assessing a government’s substantive law, we need to ask whether it proscribes all three forms of discrimination.¹³ Since the existing Indicators ask only the general question whether the government prohibits “discrimination,” the Indicators should be refined to direct the analyst to these three rules (and to other rules measuring key dimensions of discrimination). In the absence of such refined Indicators, an analyst is left to define “discrimination” anew each time he or she applies the Indicators. And different analysts are left with discretion to make up for themselves the core meaning of employment

¹² In the case of gender discrimination, “harassment” also denotes an employer’s so-called *quid pro quo* demands for sexual favors from an employee under threat of retaliation or promise of benefit.

¹³ The body of revised Assessment Indicators in Appendix A of this paper therefore includes those three tests as three separate Indicators, for each protected group, as well as the expanded ILO and UN definitions discussed presently.

discrimination.

Further, since the existing Indicators do not measure the *settled* legal meanings of employment discrimination, they also do not direct the analyst to the more controversial questions at the *boundaries* of international law. For example, ILO law and authoritative interpretations of the various UN Covenants against discrimination expand anti-discrimination principles to include government action that causes differential workplace outcomes (i.e., disparate impact) even when the government actions are not targeted at the employment relation or the labor market. (ILO, 2007b; UN CESCR, 2009). ILO and UN law state, for example, that if the government allocates educational resources or locates housing and transportation in a way that makes it disproportionately difficult for members of one racial group to qualify for, find, and commute to work, then the government has committed an act of *employment* discrimination, even though the challenged government action is a matter of *educational, housing or transportation* policy.¹⁴ This rule goes a significant step beyond the proscriptions contained in Title VII of the U.S. Civil Rights Act and in the Equal Protection clause of the U.S. Constitution.¹⁵

If we focus on the particular objective of applying legislative and treaty standards to domestic law, then the central task of analysts is to use the Indicators to evaluate domestic labor law norms. In order to formulate metrics to assess the variegated phenomena of

¹⁴ The *NAS Report* does note an ILO Director General Report stating that ILO law prohibits both “direct” and “indirect” discrimination. (National Research Council, 2004, p 199). But the *Report* does not define those terms and does not include any analogous version of those terms in its Indicators.

¹⁵ These matters are discussed in this project’s *Literature Review and Bibliography*. (Barenberg, 2010, pp. 232-234, 250-261).

alternative domestic legal systems, those phenomena should be subject concurrently or preliminarily to systematic analysis.

The NAS Indicators on “Legal Frameworks” appropriately require analysts to determine the prevailing domestic labor law of the trading partners in question. From the standpoint of legal analysis, determining the content of domestic labor law is an exercise in comparative labor law research. As to matters of domestic labor law, the *NAS Report* asks analysts to rely principally on reports of international organizations and nongovernmental bodies. These are useful sources; but to determine the legal framework of a country, an analyst is best advised to turn in the first instance to the up-to-date Constitution, labor code, and regulations of the country in question – in other words, the primary texts at the national level. For many countries, there is a well-known, authoritative digest on domestic labor law (an official publication) and one or more equally well-known, pre-eminent treatise(s) on domestic labor law (by leading legal scholars or labor judges) which comprehensively collect or summarize the Constitution, labor code, regulations, judicial and administrative interpretations, and practice.¹⁶ These sources are discussed in the *Literature Review and Bibliography* for this project. (Barenberg, 2010, pp. 270-285, 307-328). In addition, there are analogous digests and treatises on civil procedure, criminal procedure, and administrative

¹⁶ There are analogous materials for international labor law. For the ILO, for example, there is the Committee on Freedom of Association’s Digest (a quasi-authoritative primary source) and the three-volume Code of International Labor Law (a secondary treatise that compiles the primary texts and authoritative decisions of the ILO supervisory organs). Other ILO sources, and analogous materials for the relevant UN and regional bodies and domestic systems are discussed in this project’s *Literature Review and Bibliography*. (Barenberg, 2010, pp. 250-270, 272-285, 307-328).

law, which are relevant to identifying the authority and procedures of courts of general jurisdiction (which, in many countries, hear labor cases of various kinds) and administrative agencies (such as labor boards, inspectorates, and prosecutorates). An analyst who instead starts with international reports will often find partial and out-of-date national laws and regulations, and therefore can have little confidence that Indicators will assess actual prevailing legal frameworks. Even if an analyst finds in an international report a seemingly solid compendium of national labor laws for a given country, the analyst will still have to turn to the primary texts (Constitution, labor code, regulations) to be sure that the compendium is in fact comprehensive and up-to-date.

Academic scholarship in the fields of comparative labor law, comparative legal sociology, comparative labor relations, and comparative regulation is also the primary source of information and conceptualization about well-performing and poorly performing legal institutions (matters that underlie the Indicators on Government Performance). That literature is discussed at length in the *Literature Review and Bibliography* for this project. (Barenberg, 2010, pp. 105-155, 195-197, 199-202, 270-317) The use of indicators as a method of public or private governance is itself one recent focus of these fields. Since the publication of the *NAS Report*, the academic study of “new modes of regulation” that often, and sometimes centrally, make use of indicator methodologies has grown exponentially.

In order to formulate refined Indicators that measure legal or regulatory phenomena, we must reflect on the various *types* of legal and regulatory indicators that might be best suited to measure different elements of the rights in question. It is useful to begin that

reflection by setting out the conceptual underpinnings of traditional legal-compliance and regulatory-compliance methods.

2.3. Traditional Legal Interpretation and Sources of Law

This and the next sub-part describe how the original Indicators would be refined using traditional methods of legal interpretation and regulatory compliance. Part 3 then examines the conceptual weaknesses of the traditional compliance model, and proposes alternative types of Indicators that are best suited to measure different aspects of compliance with labor rights.

There are (to put it mildly) large bodies of international, regional, and domestic jurisprudence giving specific content to the highly abstract concepts of “workers’ freedom of association,” “rights to organize,” “rights to bargain collectively,” “rights against employment discrimination,” “minimum wages,” “hours of work,” and “occupational safety and health.”

At the international, regional, and domestic levels, there are recognized primary sources of legal norms (rules and standards¹⁷), recognized hierarchies among those norms, and recognized principles and methodologies for interpreting the norms.¹⁸ The primary sources range from the Constitutions of international and regional organizations;

¹⁷ The distinction between these two major types of legal norms – rules and standards – is discussed in sub-part 3.2.

¹⁸ Here, the term “sources of law” refers to the legal instruments and rules that constitute authoritative, binding, substantive domestic law. The previous sub-part, by contrast, refers to the legal materials that are textual “sources” for research about these *substantive* “sources of law.”

Conventions, Covenants, and Declarations promulgated by those organizations; multilateral, regional, and bilateral treaties; decisions rendered by the competent bodies established by international organizations and treaties; and domestic Constitutions, legislation, regulations, judicial decisions, administrative decisions, and in some instances private standards (such as collectively bargained norms). Traditional legal methodology requires that interpretation begin with the identification of these primary sources of substantive legal rules, standards, and principles. Then, the analyst must establish the hierarchies among them. That is, domestic Constitutions are in most instances (but not always) supreme over legislation, legislation is preemptive of administrative regulations, and so on. Finally, the analyst must apply settled principles and methodologies of interpretation to the applicable norms (i.e., the applicable rules and standards).

The sources of law, hierarchies among different types of norms, and principles for interpreting each type of norm vary from jurisdiction to jurisdiction. That is, the ILO has its own hierarchy of primary texts, its own organs that provide authoritative or quasi-authoritative interpretation of those texts, and its own principles that guide those organs' authoritative interpretation.¹⁹ Each country has its own, usually well-specified, primary sources of norms, hierarchy among the norms, and principles for interpreting domestic labor norms – although, of course, national legal systems may cluster along those dimensions, either for historical

¹⁹ For a detailed treatment, see this project's *Literature Review and Bibliography* (Barenberg, 2010, pp. 211-270).

reasons or as a matter of deliberate harmonization.²⁰

Each national legal system also has its own norms and principles for incorporating international standards into domestic law and interpreting those standards. Of course, for our purposes, the United States' norms and principles for incorporating international standards are of particular importance. What count as binding treaties, and how to interpret such treaties, are the subject of special rules and principles developed by the U.S. Supreme Court.

These are elementary and highly simplified recitations of basic legal methodology. I begin with such a recitation merely to set the stage for the next sub-part.

2.4. Traditional Compliance Methods

This sub-part engages in a thought experiment: If we were to apply traditional methods of regulatory compliance, what Indicators and methodology would we formulate?

As noted in the previous sub-part, the labor rights and obligations set forth in trade legislation and treaties are framed in highly abstract terms – about as abstract as one can find in any legal document. “Freedom of association,” “rights to organize,” “rights to bargain collectively,” “rights against employment discrimination,” “acceptable conditions of work with

²⁰ The fact that the rules for identifying the sources of law vary from jurisdiction to jurisdiction does not in itself pose a conceptual problem for applying a uniform set of Indicators as a yardstick for measuring the adequacy of domestic labor law. While the hierarchy of *sources* of authoritative substantive norms varies, the substantive norms themselves can be identified and placed along the yardstick of internationally recognized worker rights or core labor rights as interpreted by the U.S. Executive Branch. Where treaties require (in addition or alternatively) that countries effectively enforce their substantive labor norms, there is again no conceptual problem in the fact that the relevant substantive norms are identified by varying hierarchical sources of law.

respect to minimum wages, hours of work, and occupational safety and health” – these concepts are cast at a level of generality typical of the most abstract Constitutional texts or of the mere heading or title of a statute.

The entire mass of United States “collective bargaining law” – with all its rules, sub-rules, sub-sub-rules, and so on – fills thousands of pages of treatises and hundreds of volumes of labor board and judicial decisions (or, these days, scores of gigabytes of electronic data). This mass of rules, sub-rules, etc., constitutes an elaboration or specification of the abstract concepts of “workers’ freedom of association,” “rights to organize,” and “rights to bargain collectively.” Likewise, the entire mass of United States “employment discrimination law” constitutes an elaboration of the general concept of “freedom from employment discrimination.” There are analogous masses of legal rules for occupational safety and health, and somewhat less copious rules for minimum wages and hours of work. There are comparable masses of rules and sub-rules in other advanced legal systems.

The ILO, UN, and regional bodies have developed analogous bodies of rules and sub-rules fleshing out freedom of association and rights of collective bargaining, rights of nondiscrimination, and the three conditions of work – although the mass of international rules is not as fully articulated as the mass of national rules, at least in advanced legal systems like the United States’.

Further, each national system of *labor administration* (to use the ILO’s terminology) is a congeries of institutions, procedures, material and human resources, and the actual

enforcement practices of those institutions. From an ideal standpoint, these institutions, resources, and practices constitute specifications of the concepts of “rule of law,” “effective enforcement,” “taking steps,” and other abstract desiderata proclaimed by international and national legal systems.

How are we to understand the relationship between the abstract concepts and the highly specific rules, sub-rules, sub-sub-rules, institutions, procedures, resources, and practices (hereafter, denoted by the shorthand “specific rules and practices”) that purport to give practical meaning to the abstractions? To put the question more precisely, for our purposes: In formulating Indicators of compliance with the abstract concepts found in trade legislation and treaties, what use should we make of the specific rules and practices, and what use should we make of more abstract principles and more general institutional features?

The traditional answer in legal theory and public administration (as well as in popular thinking) is the strategy of “regulatory compliance.” Conceptually, this is at first blush a simple and seemingly obvious strategy. The indicators of compliance are simply the mass of specific legal rules and practices themselves. These rules and practices are identified and interpreted using the traditional interpretive methodology summarized in the previous sub-part.

Applying the strategy of traditional compliance to our project, it is important to begin by emphasizing that none of the U.S. trade legislation and agreements specify that authoritative rulings of the ILO, the UN, or other international or regional bodies provide the *exclusive*

content of the “internationally recognized worker rights” contained in those statutes and treaties, however useful and sensible it may be to start with those internationally authoritative sources. The concept of “internationally recognized worker rights” is a legal artifact of U.S. law (when we speak of trade legislation) and the mutual understanding of the parties (when we speak of trade agreements).²¹

Even as to the treaties, as a matter of Supreme Court jurisprudence, the law that binds U.S. actors, including the Executive Branch, is found in the implementing legislation enacted by Congress, not in the treaty *per se*.²² The President’s formal task, then, is to interpret Congressional intent at the time the implementing legislation was enacted, not the intent of the parties at the time they negotiated and signed the treaty. Of course, the intent of the parties at the time of signing may be critical *evidence* of Congress’s intent at the time of enacting legislation that implements the treaty; and the negotiating history of the treaty, in turn, is evidence of the parties’ intent.²³ (Bederman, 2008). Indeed, the Supreme Court has recently affirmed that even the parties’ practice *subsequent* to signing may also be evidence of their mutual intent about the meaning of the treaty.²⁴

When Congress delegates authority to the President by statutes containing highly generalized and discretionary language – as is true of the statutes that concern us – the

²¹ See note 6 above. Again, as explained in note 6 above, this is true even of the U.S.-Peru agreement, notwithstanding its reference to ILO law.

²² See 1 Restatement (Third) of the Foreign Relations of the United States § 111 cmt. H (1987) (“[S]trictly speaking, it is the implementing legislation, rather than the agreement itself, that is given effect as law in the United States. That is true even when a non-self-executing agreement is ‘enacted’ by, or incorporated in, implementing legislation.”)

²³ E.g., *Zicherman v. Korean Air Lines Co.*, 516 U.S. 217, 226 (1996).

²⁴ *Medellin v. Texas*, 552 U.S. 491, 507 (2008).

President is constitutionally permitted to implement any rules that constitute “reasonable” interpretations of the statute. That is, the President may implement the statute in any way that is not arbitrary, capricious, or manifestly contrary to the statute.²⁵

Hence, in interpreting “internationally recognized worker rights,” “taking steps to afford internationally recognized worker rights,” and “effective enforcement of labor law,” we can turn the vice of abstraction into the virtue of drawing on multiple legal sources. As a general strategy, the President would reasonably implement Congressional intent by fashioning Indicators of internationally recognized worker rights (and, *a fortiori*, Indicators of ILO core rights where treaties refer to such rights) that draw on the specific legal rules and practices elaborated by the ILO and UN.²⁶ Where those rules and practices are ambiguous or incomplete, the President may reasonably base the Indicators on the consensus about rules and practices among regional bodies such as the Council of Europe’s Commission on Social and Economic Rights and the Inter-American Commission on Human Rights. Where those regional rules and practices are ambiguous or incomplete, or where there is insufficient consensus among regional bodies, the President may reasonably base the Indicators on the

²⁵ *Chevron, U.S.A., Inc. v. NRDC, Inc.*, 467 U.S. 837 (1984).

²⁶ It is generally accepted that Congress did not intend “internationally recognized worker rights” to be homologous with “core labor rights” as interpreted by the International Labor Organization. While this point adds a dimension of uncertainty to the former concept, it also adds a degree of freedom to the U.S. Executive Branch’s interpretation of U.S. legislation and treaties that use the concept of internationally recognized worker rights. Hence, the pronouncements of various UN bodies that authoritatively or quasi-authoritatively interpret UN Covenants other than ILO instruments are also pertinent sources of “internationally recognized worker rights.” Indeed, the legislative history of the relevant text of the Generalized System of Preferences refers explicitly to the UN Declaration proclaiming that labor rights are universal human rights: “The United States has embraced labor rights, in principle, as well as political rights for all the people of the world upon adoption of the Universal Declaration of Human Rights in 1948. The Declaration specifically affirms for each person the right to a job, the right to form and join unions, and the right to an adequate standard of living.” H. Rep. No. 98-1090 (1984) (Ways and Means Committee) at p. 12, reprinted in 1984 U.S.C.C.A.N. 5101, 5112.

consensus in interpreting international labor norms by national legal regimes, where such regimes have had occasion to authoritatively interpret international norms. Where such interpretations are absent or at loggerheads, it is reasonable to base the Indicators on the consensus among leading national legal regimes in interpreting domestic labor rights that correspond to international labor rights (for example, the consensus among national legal systems' interpretations of "rights of collective bargaining" as a matter of domestic law). Where such consensus is absent or manifestly falls short of the general principles and goals of international and regional labor law, the labor Indicators may reasonably be specified either by the average performance of similar countries or by the "best practice" in rules and practices exhibited by vanguard national legal regimes – perhaps by identifying the best-practice rules and the vanguard regime(s) within particular regions, within particular regime types, or within clusters of countries at similar levels of economic development or similarly situated in other relevant respects.²⁷

While these various sources provide crucial guidance, we must at all times bear in mind that the legal concepts that concern us are artifacts of U.S. law, subject to any reasonable interpretation the Executive Branch chooses to give those concepts.

At what level of specificity and discreteness should Executive Branch officials frame the rules and practices that constitute the Indicators? "Specificity" here denotes both the linguistic detail and precision with which the Indicator is framed, and (in legal terms) the particularity of the facts to which the Indicator applies or (in social science terms) the level of

²⁷ The appropriate use of comparative indicators is discussed below in Part 6. The use of control variables that identify similarly situated clusters of countries is discussed in Part 8.

detail of the data necessary for scoring the Indicator. “Discreteness” denotes that the Indicator is distinct from other Indicators, in the sense that it does not overlap or subsume other Indicators, and other Indicators do not subsume it.

In the purest conceptualization of traditional compliance, there are at least three reasons for applying the most specific and discrete Indicators identified by the Executive Branch’s reasonable interpretation of the pertinent statutory and treaty language, drawing on the most relevant legal systems and other indicators of Congressional intent (ILO, UN, regional bodies, consensus national rules, vanguard national systems, average national systems).

First, the traditional compliance method contends that analysts will face less ambiguity in applying each Indicator, if the Indicators are more specific. As just mentioned, the Indicator is specific in the sense that it precisely identifies a highly concrete factual context to which it applies. It guides the analyst more clearly and surgically to the facts necessary to accurately apply the Indicator – facts that show whether the Indicator is satisfied or not satisfied. For example, consider the following two Indicators:

- (A) Do tribunals provide effective remedies for workers who are unlawfully discharged?**
- (B) In the previous two years, have the country’s labor courts ordered reinstatement of workers in all but a trivial percentage of cases in which the courts have found the worker to have been discharged for engaging in union activity?**

Indicator B is much more specific and precisely stated than Indicator A. We can therefore

expect a significantly higher rate of variance when different analysts apply Indicator A to a particular country than when different analysts apply Indicator B. By reason of its generality, Indicator A is ambiguous on at least five key points: (1) What is the range of possible “tribunals” the analyst should examine? (Specialized labor courts? Courts of general jurisdiction? Labor boards? Other administrative agencies? Private arbitrators?) (2) What are the various categories of cases in which a worker might have been unlawfully discharged? (Anti-union discrimination? Employment discrimination based on gender? Employment discrimination based on race? Discharge of a disabled worker, for whom the employer refuses to provide reasonable accommodation? Retaliation against a whistleblower? Cases of constructive discharge or only cases of actual discharge?) (3) What are the different types of remedies the analyst should look for? (Reinstatement? Back pay? Severance pay? Other compensatory damages, such as pain and suffering or medical costs? Punitive civil damages? Injunctive remedies, such as requirements that the employer permit union organizers to meet with workers on employer property during non-working-time? Criminal fines? Incarceration?) (4) Which of these remedies, alone or in combination, should the analyst consider “effective” in individual cases? (Should the analyst look to effective *ex post* compensation or effective *ex ante* deterrence? What frequencies of different remedies across the docket of discharge cases should count as “effective” for all “workers who have been unlawfully discharged”?) (5) To which time period does the Indicator apply?

An analyst charged with applying Indicator A may well feel bewildered if not paralyzed

in moving ahead. We can expect similar perplexity on the part of the government officials and employers who are supposed to comply with the Indicator, and the legal advocates who monitor such compliance, even if they are well-trained in the domestic legal system under scrutiny – precisely because Indicator A is a broad, imprecisely drafted legal rule, and would be recognized as such by lawyers. For these reasons, many ILAB analysts strongly urged this author to formulate Indicators taking the form of specific bright-line rules. (Barenberg, 2009, p. 19).

Second, traditional compliance methodology maintains that analysts will more easily apply Indicators that are not only specific but are also discrete. If Indicators overlap, there is a high likelihood that the Indicators will contain inconsistent or blurred concepts, since the same factual evidence will be described by the differing language of multiple Indicators. If one Indicator subsumes another, it may be difficult for analysts to distinguish the evidence that is sufficient to satisfy the more particular Indicator from the evidence necessary to satisfy both the more particular Indicator and the more general Indicator that subsumes the more particular Indicator. For example, consider the following two Indicators:

(A) Do remedies punish the employer for violating rights of collective bargaining?

(B) Is there effective remediation for violations of rights of collective bargaining?

The concepts of punishment and remediation may overlap and create inconsistencies. For example, analysts may take the concept of “punishment” in Indicator A to include only imprisonment; or to include both imprisonment and criminal fines; or to include imprisonment,

criminal fines, and “punitive damages” awarded in civil cases; or to include imprisonment, criminal fines, punitive civil damages, and compensatory civil damages. On the other hand, analysts might think that the phrase “effective remediation” in Indicator B is intended to measure only whether the victim of a rights violation is compensated or made whole. In that case, the concept of “remedy” might be used inconsistently in Indicators A and B – using that term in the sense of some or all forms of judicial or administrative sanctions and awards in Indicator A, but using the very same concept with more narrow scope in Indicator B. Alternatively, the analyst may include imprisonment and punitive monetary sanctions in both Indicators (that is, she may subsume the first Indicator in the second); but different analysts may diverge in determining which remedies, in addition to punitive remedies, are necessary for a finding of effective compliance. Apart from the problem that different analysts may apply these Indicators inconsistently, the conceptual ambiguity or inconsistency generates potential deviation between the Indicators and international standards. For example, in international law as well as in fact, punitive remedies (however defined) may not be required to achieve full compliance.

Equally important, if Indicators are not discrete, then the weighting of individual Indicators and the aggregation of such weighted Indicators will be conceptually difficult to achieve – for the simple reason that double-counting will result from Indicators that overlap or subsume one another.

Third, for similar reasons, the traditional compliance strategy contends that it will be more difficult for governments to game a particular Indicator to the extent it is specific and

discrete. This may be especially true if the Indicator calls for a binary rather than a graded response. A labor court has either ordered reinstatement in the previous two years in all but a trivial percentage of cases in which it has found a worker to have been discharged for anti-union reasons, or it has not.

True, the traditional compliance model recognizes that a government may attempt to game even this specific, discrete rule by, for example, limiting aggrieved workers' access to the labor court, or raising evidentiary standards for proving the wrongful discharge. According to the compliance model, however, this is not a problem in the specificity and discreteness of the Indicator *per se* but rather a problem in the relation of this Indicator to other well-specified, discrete Indicators in the overall body of Indicators

That is, the traditional compliance strategy has a response to the problem of gaming specific, discrete Indicators. The answer is to ensure that each Indicator is not only specific and discrete but also is lodged in a body of Indicators that are as comprehensive as possible. Hence, sufficiently precise and numerous Indicators regarding access to the courts, evidentiary standards, and non-derogation of labor standards will block, or at least adequately impede, the gambits just proffered. The comprehensive Indicators fit together like a jigsaw puzzle, leaving no gaps and therefore neutralizing opportunities for gaming.

If this logic were to inform our conceptual framework, it might call for formulating Indicators that are even more specific and comprehensive than those previously promulgated or elaborated by the relevant legal authority. For example, we might start with the specific rules contained in ILO Conventions and authoritative elaborations of the Conventions by

relevant ILO committees; but we may find that many of those rules are stated at a level of generality more akin to Indicator A above or at a higher level of generality still.

How is it possible to formulate more specific Indicators than those announced by authoritative international bodies? First, and simplest, we would (hypothetically) put ourselves in the position of the authoritative international bodies, such as the ILO Committee of Experts, and ask: What specific rules would those bodies articulate, if faced with the particular factual context that is the subject of the Indicator we are constructing? This may sound like casuistry, especially to non-lawyers, but it is the routine stuff of legal interpretation. Indeed, it is mandated by the interpretive methodology of the United States Supreme Court, when interpreting U.S. legislation and treaties. That is, if a legal system has not yet decided a particular case, the lawyer's job is to "predict" the rule that would be announced by the highest court if the latter were presented with the case. The most accurate possible prediction *is* the law, as Oliver Wendall Holmes, Jr., famously argued. For example, the concept of "rights to organize and to bargain collectively" comprehensively covers the domain of all possible factual contexts that might arise in the field of union organizing, collective bargaining, collective contract enforcement, and striking and other group protesting. Even if the ILO Committee of Experts or the ILO Committee on Freedom of Association has not yet *pronounced* the rule that applies to some specific factual situation, such a specific rule *exists* in the sense that outside observers can interpret the law using the interpretive methods that those bodies themselves mandate and follow.

Second, as already described, in determining which rules of labor law are

“internationally recognized,” we can look not only to the ILO’s Conventions and committee reports, but also to UN Covenants and committee reports, to regional conventions and decisions, to consensus rules among leading national legal systems, and to the best practice of national systems that are internationally recognized as vanguards in protecting workers rights. This hierarchy of sources constitutes the most natural reading of the phrase “internationally recognized worker rights.” Where there is an overlapping consensus among several of these interpretative sources, there is a strong foundation for a specific rule or set of rules.

In any event, as explained earlier in this sub-part, the structure of U.S. Constitutional obligations and the multiplicity of relevant international and domestic legal sources permit the U.S. Executive Branch to “reasonably” stipulate specific rules and practices that most effectively protect worker rights, which is the deepest underlying purpose of the legislation and treaties that the Executive Branch is constitutionally bound to faithfully interpret and enforce.

2.5. Conclusion

This Part explained that the labor rights provisions of U.S. trade legislation and agreements are quintessentially legal rules. When the U.S. Executive Branch determines whether a country is complying with those provisions, it must apply those legal rules to the evidence in the manner of any judicial or administrative determination of an actor’s compliance or non-compliance with legal rules. Under traditional compliance methods, that

exercise calls for the formulation of Indicators that are specific, discrete, consistent, and comprehensive. The Indicators are, in essence, well-specified sub-rules fleshing out the more general legal rules contained in the legislation and treaties.

If we were to unequivocally accept traditional compliance methods, then, the initial step would be to scrub the NAS Indicators to give them the characteristics of specificity, discreteness, and consistency. The second step would be to draft additional Indicators to fill the gaps in the NAS Indicators' coverage of important aspects of labor rights, in order to achieve comprehensiveness.²⁸

Before throwing ourselves into that exercise, however, the next Part examines significant challenges to applying the naïve compliance strategy outlined in this Part. That examination prompts us to conceptualize alternative types of legal norms and Indicators. The traditional compliance method suits some Indicators, but other Indicators optimally rest on different conceptual foundations.

²⁸ These gaps resulted from the fact that the NAS committees did not see as their mandate the systematic parsing of the elements of labor law.

3. Challenges to the Traditional Compliance Strategy



Part 2 made the banal point that the substantive rights and obligations of trade legislation and trade agreements are legal concepts. That recognition, however, is just the beginning of the conceptual inquiry. There is in fact a range of alternative types of legal norms, and a range of conceptualizations of those types. Hence, we must ask: what *kind* of legal concepts are the rights and obligations (norms) in U.S. trade legislation and treaties? Only after examining this conceptual problem can we confidently proceed to the next step of formulating the Indicators that measure compliance with those rights and obligations.

This conceptual question may seem well settled and, in part for that very reason, needlessly abstract or even puzzling to readers who have not been basted in legal theory. After all, Part 2 set forth the canonical mode of identifying and interpreting relevant norms, including norms about institutional structure and function. Those norms take the form of specific, discrete, consistent, comprehensive rules.

But it is not so simple. Recognition that different types of norms serve different purposes is currently strong among legal scholars, judges, administrators, and other practitioners of legal regulation. In fact, there has always been differentiation in the conceptual underpinnings of various aspects of legal rights; but the degree of recognition of the differentiation is now particularly high. More important, the current attention to such

differentiation is a product of transformation in the reality of regulation across many legal domains, including the domains of labor regulation. This attention is especially strongly associated with a family of new regulatory frameworks that go under several names: the New Governance, the New Public Management, Reflexive Regulation, Democratic Experimentalism, and others.

The changing landscape of legal-regulatory practice and theory and the correlative conceptualization of different types of legal norms are not just scholastic matters. Quite the opposite. The use of indicators as a means of implementing legal norms is particularly associated with the new regulatory concepts and frameworks, and those concepts and frameworks are self-consciously pragmatic responses to significant weaknesses in the traditional compliance model. (See, e.g., Hood, et al., 2008; Rosga and Satterthwaite, 2009; Noonan, et al., 2009.) This project's *Literature Review and Bibliography* discusses over twenty such indicator-driven regulatory initiatives in the field of labor alone, some tied directly to regulatory programs, others generated by academic studies having indirect consequences for practical programs. (Barenberg, 2010, pp. 24-96, 105-210, 285-307). But the new regulatory concepts have also been applied to fields as diverse as environmental regulation, education, child welfare, prison reform, policing, corporate regulation, consumer protection, human rights, public health, and even constitutionalism. (E.g., Noonan, et al., 2009; Karkkainen, 2004, 2002, 2001; Garrett and Liebman, 2004; Sabel and Simon, 2004; UNHCHR, 2008; De Burca and Scott, 2006).

In this Part, the first sub-part sets out the pragmatic problems – relevant to this project

– with the traditional compliance model. Sub-parts 2 and 3 then describe some alternatives or complements to the traditional compliance model’s conceptualization of legal norms.

These alternatives or complements, elaborated in the literature on new regulatory frameworks, purport to overcome or at least mitigate the problems with the conventional compliance model. Sub-part 3 proposes that Indicators be categorized into four types: (1) fixed bright-line rules; (2) revisable bright-line rules; (3) fixed standards, and (4) revisable standards. Sub-part 4 concludes by setting forth criteria for distinguishing which of these four types is optimal for measuring different aspects of the rights that concern us.

3.1. Weaknesses in the Traditional Compliance Strategy

The traditional compliance model has some now well-recognized limitations that cannot always be overcome by the strategies of maximal specificity, discreteness, and comprehensiveness noted above in Part 2. The limitations are closely intertwined.

First, maximizing a rule’s specificity to neutralize a government’s gaming of the rule may be a never-ending chess match. The discussion above proposed provisionally confident preemptive measures against an ill-intentioned government seeking to limit access to labor courts or to raise evidentiary standards for proving wrongful discharge. But the preemptive measures were confident only because the proposed gambits were themselves overtly articulated. Consider the gambit of limiting access to the courts. There are innumerable ways that a government might seek to do so, and with sufficient furtiveness to undercut the hope that a body of indicators will block all significant avenues of evasion. This applies a

fortiori to any effort to construct a body of *ex ante* indicators that lacks a nimble mechanism for continuous revision and reticulation. (Sabel, 2006).

Second, if the body of indicators is formulated by a centralized technocracy, it cannot hope to foresee the fine texture in each of the local contexts to which it must apply. That is, a body of indicators cannot be fully specified and comprehensive not only because the evasive tactics of an ill-intentioned government are innumerable and unpredictable, but also because the cogs of a single body of even highly detailed indicators cannot mesh tightly with the variegated gears of local enforcement institutions, labor relations systems, complex workplace arrangements, and the like. (See, e.g., Sabel and Simon, 2004.)

Third, a pre-fixed set of detailed rules cannot anticipate the transformations over time in labor markets, workplace relations, and legal institutions. Just as one-size-fits-none across countries, so one size cannot fit the shifting contours across time in each country. Even when countries are transformed by common global forces, they change each in their own way; the global variables are mediated and inflected by each country's internal dynamics.

Fourth, and for similar reasons, indicators that are overly specific may miss opportunities to expand worker rights and protection over time. Production systems may evolve in ways that potentially enable workers to expand their autonomy or bargaining power. Increases in productivity have the same potential. Technological and organizational innovations may provide occasions for strengthening safety and health protocols and performance. (Piore, 2004). Rules that are pre-fixed and rigid may not be responsive to

these opportunities.

Fifth, there are some labor violations that are intrinsically resistant to sharply defined rules. Sex segregation by occupation, for example, may be sustained by complex cultural and psychic variables, including sub-conscious ones, as well as by subtle forms of harassment and coercion that may be embedded in organization-specific protocols and pathologies. Domestic legal systems – at least the more sophisticated ones – have therefore widely dealt with problems of sex segregation and sexual harassment through norms framed as “standards,” discussed presently, rather than as bright-line “rules.” (Sturm, 2001; Noonan, et al., 2009).

Sixth, there is an intrinsic tension between the two tactics discussed above in sub-part 2.4 – specificity and comprehensiveness – that seemed necessary to neutralize gaming of the rules. Recall that a precisely defined, narrow rule, taken on its own, is desirable because it clearly demarcates compliance and non-compliance. Analysts and governments know when *that* individual rule is violated. However, that same quality of individual rules makes it challenging to generate a *body* of rules that, taken together, leaves no gaps or havens that can be exploited by evasive governments. The web of narrowly threaded rules must be radically fine-meshed to create a fabric that not only stretches across the entire domain of the right in question, but does so without leaving abundant, inconspicuous interstices.

One way to respond to this dilemma – to develop a body of indicators that comprehensively covers the domain space, leaving no gaps – is to frame each indicator at a somewhat more general level.

For example, suppose prosecutors in Country X are failing to pursue co-conspirators of paramilitary officers who plead guilty to murder of trade unionists. That is, the government gives the false appearance of cracking down on such murders by trumpeting convictions based on guilty pleas, while allowing large numbers of perpetrators to go unprosecuted. To respond to this gaming, we might construct a narrowly framed Indicator such as:

In the preceding two years, in all cases in which a defendant pleads guilty to the murder of trade unionists, have prosecutors filed charges of homicide against alleged co-conspirators whenever the prosecutor had or should have had reasonable cause to believe the allegations?

Country X might game this indicator by in various ways impeding evidence from reaching prosecutors, such as denying subpoenas for collecting relevant evidence. Or Country X may obfuscate whether such evidence has in fact reached prosecutors by, for example, maintaining firewalls between police investigators' files and prosecutors' files. We might respond by formulating additional specific rules, such as:

In the preceding two years, in all cases in which a defendant pled guilty to the murder of a trade unionist, did the relevant authority grant every request for a subpoena seeking evidence against alleged co-conspirators?

In the preceding two years, in all investigations of alleged murders of trade unionists, were complete copies of police investigators' records provided to, and fully examined by, prosecutors?

Or even more bluntly:

Is there an irrebuttable presumption that prosecutors have seen all evidence contained in any government record pertaining to alleged murders of trade unionists?

But these refined rules can themselves be gamed. Prosecutorial supervisors can

discourage prosecutors from applying for subpoenas in the first place. Police supervisors can discourage investigators from recording inculpatory evidence.

So, rather than formulating more and more specific Indicators we might fashion somewhat more general Indicators that, at least conceptually, cover the entire domain in question – umbrella Indicators, so to speak. To continue the above example, we might recur to an Indicator such as:

In the preceding two years, have government officials aggressively investigated and prosecuted all alleged perpetrators of murder of trade unionists, including alleged co-conspirators and others who allegedly aid and abet such murders?

Or even more broadly:

In the preceding two years, have government officials aggressively investigated and prosecuted all alleged perpetrators of violent crimes against trade union supporters?

The key conceptual point is this: By abandoning specificity in favor of comprehensiveness, we have moved from one kind of legal norm to another. In the parlance of jurisprudence, we have moved from “rules” to “standards.”

3.2. From Rules to Standards

A “rule” dictates specific behavior. An analyst, judge, or other government official can apply the rule with a high level of clarity and precision. The rule-form of norms minimizes ambiguity. The previous sub-part contains examples of bright-line rules, such as the rule requiring the authorities to grant every request for a subpoena seeking evidence against

alleged co-conspirators of defendants who plead guilty to the murder of trade unionists. A “standard,” by contrast, is a norm that articulates a general principle or policy that is to guide action and assessment of actions.²⁹ The most prominent example of a highly general standard is found in the common law of tort, which requires actors to take “reasonable precautions” against injuring others. An analogue in statutory law is the requirement that employers take all reasonable, feasible measures to protect workers against significant workplace hazards. Those two illustrative standards are framed at a level of abstraction that is too high for our purposes. An example of a less abstract standard is the final Indicator set out in the previous sub-part, asking whether the government aggressively investigates alleged violent crimes against trade union supporters.

Rules and standards, when used as indicators, each have their characteristic vices and virtues. The virtue of a precise rule is that it makes clear what constitutes compliance or non-compliance, when our only concern is compliance with the individual norm. (The virtue does not apply when our concern is with compliance with a broader set of norms and principles, for reasons just offered.) In procedural terms, the virtue of precise rules is that they give clear guidance not only to the regulated actor – in our case, the trading partner’s officials or employers – but also to the analyst who must determine whether the rule is violated or not. The virtue of a rule – providing constraint on actors and analysts – is also its vice. A bright-line rule is rigid and non-adaptable to contextual variations. It is a one-size-fits-all norm. The rule may not achieve fairness, equality, efficiency, or other goals of legal

²⁹ For a good recent discussion, see Noonan, et al., 2009.

policy in many specific contexts to which it is applicable. What gives it its constraining quality (virtue) is precisely what makes it inflexible and non-adaptable (vice).

A standard inverts these vices and virtues. The virtue of a standard is that the actor who must implement or assess compliance with the indicator can apply it in a way that is well adapted to context. As a conceptual matter, standards have these qualities because, as just noted, they express a general principle or policy that the legal norm or indicator is intended to fulfill. The law therefore effectively delegates to the actor applying the standard – the administrative official, judge, employer, or ILAB analyst – the authority to determine how that principle or policy will be best realized in the various factual contexts to which the indicator might apply.

In each instance of applying the standard, the actor is effectively creating a new rule (a more precisely specified sub-indicator). And those more specific sub-indicators are well adapted to the context, in two senses. They are well adapted to the precise factual setting; and they are well adapted in the sense that the sub-indicator fulfills the more general principle or policy expressed by the standard. Because of the flexibility in applying the general principle to the specific facts, the standard can achieve “situated justice” or “well-adapted policy.” A standard is therefore the kind of norm that is well-suited to subject domains in which there is (a) social complexity and therefore (at least *ex ante*) indeterminacy respecting the behavior that we wish to regulate, as in the field of sex segregation and sexual harassment noted above, and (b) relatively greater fluidity in, and uncertainty about, the likely transformations in the social setting to which we wish to apply the principle or policy in the

future.

As with the rule-form of indicators, the standard-form's virtue is also its vice. The standard permits principles or policies to be well realized in context (virtue) but at the cost of delegating discretion to the regulated actor or analyst (vice). Such discretion is a vice because it carries the potential of inconsistent and unpredictable application of the indicator. This is a vice from two standpoints: first, from the standpoint of the higher-level principles embodied in the even more abstract standard called the "rule of law," which calls for norms to be predictably and uniformly applied across similarly situated actors; and, second, from the standpoint of multiple analysts who seek to apply indicators to yield consistent, predictable assessments.

The conceptual point here is this: We might prefer Indicators to take the form of precise, bright-line rules where the relevant principle or policy need not be adaptable, whether because the relevant, desired social behavior is itself clearly demarcated and relatively fixed as a descriptive matter (for example, weekly maximum hours can be clearly described in cardinal units) or because the normative (that is, values-based) principle or policy is relatively fixed (how many weekly hours *should* we tolerate?). We might also prefer the rule-form of Indicators where we have less confidence in the capacity of trading partners' officials, employers, or our own analysts to exercise discretion in ways that treat similarly situated actors uniformly and predictably. And, we might prefer Indicators to take the form of standards where the opposite conditions obtain.

But note again the point made above: Over time, repeated application of the Indicator-

framed-as-standard may generate a body of more specific sub-indicators-framed-as rules.

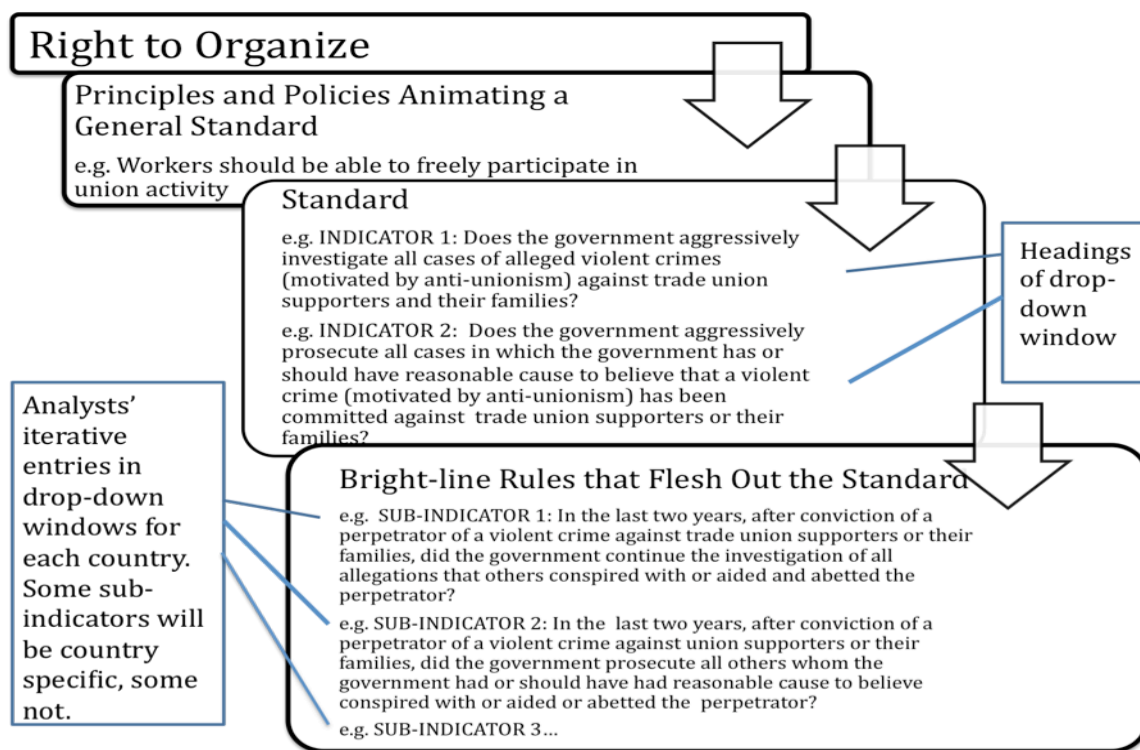
This will be the case if the Indicator methodology requires the analyst to articulate and record the specific, decisive factual basis for her successive determinations that the Indicator-framed-as standard does or does not apply to the factual context at hand.

This process is illustrated in Figure 3 below. That Figure returns to our example about prosecution of crimes against trade unionist, starting with the indicator-framed-as-standard: “Do government officials aggressively investigate and prosecute all alleged perpetrators of violent crimes against trade unionists?” After repeatedly applying this standard, the analyst will accumulate a body of sub-indicators framed as specific rules, such as: “Does the relevant authority grant every request for a subpoena seeking evidence against alleged co-conspirators of defendants who plead guilty to violent crimes against trade unionists?”

Students of law will note that the indicator methodology just described is very similar to the method of the common law. Over time, common-law judges develop bright-line rules in their case-by-case decisions. Each case presents the judge with a new factual context. Even two very similar cases will have some factual elements that differ, alongside many facts that are the same. In deciding the outcome in specific contexts, the common-law judge must determine whether one case is sufficiently analogous to another to apply the same rule to both cases, or whether instead the two cases are sufficiently disanalogous to apply another existing rule or to formulate a new rule. To do this, the judge must decide whether two cases are sufficiently similar in *relevant respects* or sufficiently dissimilar in *relevant respects*. How does she know which similarities or dissimilarities are the relevant ones? That determination

is guided, precisely, by overarching principles or policies, by the standards already embedded in common-law jurisprudence (such as the standard of reasonableness in tort law). The specific rules are made “coherent” by the principles and policies contained in the overarching standards.

FIGURE 3. Converting Standards into Rules



This combination of a standard and accumulated rules that instantiate the standard has the apparent merit that it mitigates the tension between specificity and comprehensiveness. The standard ensures that government officials or employers cannot game the system by finding the gaps in the web of accumulated specific rules, since the

umbrella standard will capture such behavior. At the same time, we can expect that the accumulated body of specific rules will, over time, provide precise specifications of the most significant behaviors that count as compliance or non-compliance.

This appealing resolution of the tension between specificity and comprehensiveness depends on the methodological mandate mentioned above: The Indicator methodology must require analysts to articulate and record their specific applications of the standard. This paper proposes one method for carrying out and visualizing this requirement. As indicated in Figure 3 above, the Indicator-framed-as-standard constitutes the heading of a drop-down window. For each country, there is a drop-down window with the identical heading (since the same Indicators apply to all countries). After clicking on the heading, the drop-down window shows the specific rules (sub-indicators) that analysts have accumulated. The analysts may learn, when applying the same standard over time to different countries, that special problems arise in the factual context of particular countries. Hence, some of the specific sub-indicator will be country-specific, and some will be common to all countries. And, bear in mind that many of the Indicators – that is, the headings for drop-down windows – will be bright-line Indicators that will not require as much specification through the creation of sub-indicators as will the Indicators framed as standards.

While this strategy is appealing, it poses some challenges. First, the methodology must take care to avoid double-counting of the standard or rule that serves as the Indicator (the heading of the drop-down window) and the more specific sub-indicators subsumed within that standard or rule. This problem figures into our evaluation of alternative strategies

of weighting and aggregation, discussed below in sub-part 11.1. Suffice it to say here that each Indicator will be a composite Indicator that aggregates each country's set of sub-indicators, thus retaining cross-country comparability, longitudinal consistency, and uniform weighting of the area of law demarcated by the standard or rule that constitutes the Indicator. Second, the effective result, over time, will be a body of fixed, specific rules, sheltered to some degree against gaming by the fixed umbrella of the relevant Indicator-framed-as-standard or Indicator-framed-as-bright-line-rule. This is not flawless because it lacks a critical dimension of adaptability, and fails to protect against a critical tactic of gaming. These points are addressed in the next sub-part.

3.3. From Fixed Norms to Revisable Norms

Although Indicators-framed-as-standards are adaptable to variegated contexts in ways that Indicators-framed-as-rules are not, the discussion above did not contemplate that the Indicator-framed-as-standard or Indicator-framed-as-rule itself would change over time. For example, an Indicator asking "whether women are widely subject to sexual harassment" is adaptable, in the sense that it can capture many types of specific behavior that violate a principle of fair treatment, types of behavior that could not be fully specified *ex ante*. And, as suggested above, in applying the Indicator, analysts over time might effectively generate many precise bright-line rules (sub-indicators) as they identify specific behaviors that violate the principle of fair treatment. But the fact that the Indicator is in *that* sense adaptable does not necessarily mean that the general Indicator itself *changes* over time. The Indicator

continues to ask, in unchanged language, “whether women are widely subject to sexual harassment.” The Indicator acts as a fixed umbrella, designating the boundaries of the principle that animates the standard, and sheltering the specific rules that are generated and accumulated over time. The drop-down window is opened time after time by clicking on the same, unchanging Indicator.

By the same token, the previous discussion presupposed that analysts will successively fashion sub-indicators taking the form of bright-line rules but that those rules, once fashioned, would not change.

To put it differently, we have thus far assumed that the Indicators are synchronically adaptable but not diachronically revisable. An Indicator, once formulated, does not adapt to changing contexts over time – changing production methods, changing productivity, changing governmental capacity, and other variables that may enable governments to provide more expansive enforcement of worker rights or even redefine the substance of rights in an increasingly worker-protective way.

Can we imagine a methodology in which Indicators are revisable over time based either on iterative application of Indicators to variable circumstances or simply on continuous learning about the best (re)formulation of Indicators applied to unchanging circumstances? The next sub-part describes the subject domains (i.e., the aspects of labor rights) for which revisable norms are well-suited, giving examples of major elements of labor rights that are already conceived as revisable within internationally recognized law. Sub-part 6.2 below makes the case that revisable norms are consistent with conceptions of labor rights as

universal human rights.

3.4. Conclusion: Four Types of Indicators and the Circumstances to Which They are Best Suited

We have, so far, set out the conceptual foundations for four different categories of Indicators or norm-types, each of which may be best suited for particular Indicators that apply to different subject domains.³⁰

FIGURE 4. Four Categories of Norm-Types for Particular Indicators

	Fixed	Revisable
Bright-Line Rule	Fixed Bright-Line Rule	Revisable Bright-Line Rule
Standard	Fixed Standard	Revisable Standard

First, Indicators taking the form of fixed bright-line rules may be optimal where normative values are fixed and factual contexts are predictable and easily specified. “Does the law require that each non-managerial, non-supervisory worker is paid at least 125

³⁰ Note that these four “types” are orthogonal to (i.e., they cross-cut) the four “categories” of Substantive Law Indicators, Enforcement Indicators, Capacity-Building Indicators, and Outcome Indicators. So, for example, a particular Enforcement Indicator might be framed as any one of the four types in Figure 4. To see how this 4-fold categorization of norm-types fits into the overall methodology for formulating individual Indicators, see the flow chart in Figure 1 in sub-part 4 of the Introduction.

percent of their regular hourly wage for hours worked in excess of 48 per week?” may be such an Indicator. The normative commitment to 48-hour weeks and overtime pay may be well-entrenched; and the passage of hours per week is a predictable, easily specified phenomenon.³¹

Second, Indicators taking the form of revisable bright-line rules may be optimal where normative values are fixed and factual contexts are presently specifiable but are expected to change. For example, an Indicator based on the ILO’s Code of Practice for Safety and Health in Underground Coalmines is this: “Does the law require that employers in the underground coalmining sector, at a minimum, provide each mineworker with a gas mask that provides at least one hour of protection against carbon monoxide inhalation, for purposes of escaping the mine?” (ILO, 2006j). The value commitment to enabling mineworkers to escape and the factual performance of masks can be well specified under current conditions; but we may anticipate that the current standard will be made obsolete either as the cost-effectiveness of self-contained breathing devices or mine ventilation technology improves.

Third, Indicators taking the form of fixed standards may be optimal where normative values are clear but the myriad factual contexts to which they may apply cannot be specified

³¹ To achieve gradation and recognize the variation across countries, the example in the text can be combined with additional binary Indicators such as: “Does the government ensure that each non-managerial, non-supervisory worker is paid 150 percent of their regular hourly wage for each hour worked in excess of 48 hours per week?;” “Does the government ensure that non-managerial, non-supervisory workers are paid 125 percent of their regularly hourly wage for each hour worked in excess of 40 hours per week?;” and “Does the government ensure that non-managerial, non-supervisory workers are paid 150 percent of their regularly hourly wage for each hour worked in excess of 40 hours per week?”

ex ante. An example, discussed above, is an Indicator measuring the government's compliance with the norm against sexual harassment: "Does the law require that employers ensure that supervisors and co-workers do not create a hostile work environment for women?" The normative value is clear: ensuring that women are not systematically disadvantaged in the workplace by reason of their gender. As a descriptive matter, however, there are innumerable ways that supervisors and co-workers might create a work environment that is hostile to women, and many alternative ways that employers might educate, monitor, and take corrective measures against supervisors and workers on the subject.

Fourth, Indicators taking the form of revisable standards may be optimal where normative values are strongly felt but are not congealed because the values are still emerging, are in flux, or are intrinsically hard to define. We may be likelier to see such a condition of normative fluidity where factual contexts are also in rapid flux, are complex, or are unpredictable. For example: "Does the law protect workers against serious psychological abuse in retaliation for the exercise of any of the rights set forth in this body of Indicators?" While this may seem uncontroversial to many, the question of whether retaliatory psychological abuse constitutes a violation of all employment rights is still not fully settled in international law or in the vanguard domestic legal systems. For example, until recently, the U.S. Supreme Court deemed that a showing of psychological coercion was insufficient to establish a violation of forced labor – precisely on the ground that the norm was excessively amorphous and might proscribe such cases as parents threatening to withdraw

their affection if a child did not work in the family business.³² However, an emerging international norm – stated most prominently in the more recent Victims of Trafficking and Violence Protection Act of 2000,³³ a subject matter in which the U.S. is currently in the vanguard – now supports the Indicator just mentioned.

Before concluding this Part, two important conceptual points should be noted about this four-fold typology of Indicators.

First, the four categories of Indicators arrayed above are not rigidly distinguished from one another. They provide heuristic guidance in thinking about the kind of Indicator that may be best suited for particular subject matters. Any particular Indicator may stand along a spectrum from relatively static to relatively revisable, and along an orthogonal spectrum from relatively specific rule-form to relatively general standard-form. The point of the four-fold typology is to provide a conceptual framework – precisely the kind of framework on which a body of Indicators must be built.

Second, even these four types of Indicators are not conceptually exhaustive. Part 7 makes the case for framing Indicators in terms of a variety of alternative normative (values-based) metrics, and Part 9 introduces the concept of Twin Indicators. The four types set out in Figure 4 cross-cut these other types of Indicators. For example, a comparative Enforcement Indicator on employment discrimination might be framed as a fixed bright-line rule, a revisable bright-line rule, a fixed standard, or a revisable standard, and it may be one of a pair of Twin Indicators. Or a zero-tolerance Substantive Law Indicator on safety and

³² See *United States v. Kozminski*, 487 U.S. 931 (1988).

³³ 18 U.S.C. § 1589.

health might be framed in any of those four ways. The flow chart in sub-part 4 of the Introduction (Figure 1) shows the relationship among the various categorizations and typologies of Indicators.

4. Capacity-Building Indicators



The conceptual discussion above addressed the general types of legal norms that might be appropriate for different subject domains: Indicators framed as specific bright-line rules vs. Indicators framed as somewhat more general standards; and static Indicators vs. revisable Indicators.

This Part turns to another major conceptual question. The NAS constructed three major categories of Indicators, applicable to each labor right: Legal Framework Indicators, Government Performance Indicators, and Overall Outcome Indicators. Are the three NAS categories conceptually optimal? This Part discusses that question by using as a foil the United Nation's recent three-fold categorization of human rights indicators. Sub-part 1 concludes that the NAS categories are conceptually sound – indeed, more so than the UN indicators. That sub-part proposes a minor re-labeling of the categories to clarify their content: Substantive Law Indicators, Enforcement Indicators, and Outcome Indicators. More important, sub-part 2 proposes adding a fourth category: Capacity-Building Indicators. As discussed in that sub-part, Capacity-Building Indicators are mandated by international and domestic law, and by compelling rationales of policy and principle.

4.1 The Three Categories of NAS Indicators vs. the Three Categories of UN Indicators

The question whether the NAS three-fold categorization is conceptually sound requires examination, in light of the fact that the United Nations recently adopted a somewhat different three-fold categorization, in its large-scale project to formulate indicators of compliance with human rights, including certain labor rights.

In 2006 and 2008, the United Nations High Commissioner for Human Rights (UNHCHR) issued *Reports* on the formulation of “indicators for promoting and monitoring the implementation of human rights.” (UNHCHR, 2006; 2008). The *UNHCHR Reports* adopt three categories of indicators: (1) structural, (2) process, and (3) outcomes. “Structural” indicators are intended to measure government “commitment.” The “process” indicators “relate State policy instruments with milestones that cumulate into outcome indicators.” (UNHCHR, 2008, p. 12). The “process” indicators are therefore conceptualized as the “cause” of the desired effect, namely the actual realization of rights. The “outcome” indicators measure that “effect.”

While the UN rubrics are superficially similar to the NAS categories, both their conceptual foundations and actual formulation are somewhat different. The UN’s “structural” indicators subsume both substantive legal norms and institutional mechanisms that promote and protect the right, while the NAS Legal Framework Indicators for the most part include only the substantive norms.

Although it initially seemed that the UN's "process" indicators would be designed, like the NAS Government Performance Indicators, to measure government effort in enforcing rights, the UN's final conceptualization of the process indicators became murkier. In the UN's conceptualization of the "cause-effect relationship," the process indicators include not only government effort but other causal variables.

The 2008 *Report* is candid about the ambiguity of the concept of a process indicator:

There is some similarity in process and outcome indicators which comes from the fact that any process can either be measured in terms of the inputs going into a process or alternately in terms of the immediate outputs or outcomes that the process generates. Thus, a process indicator on the coverage of immunization among children can be measured in terms of the public resources or expenditure going into the immunization program (which is the input variant) or in terms of the proportion of children covered under the program (which is an outcome or impact variant). In terms of the definition outlined in this note, both these indicators are process indicators. They contribute to lowering child mortality....

(UNHCHR, 2008, p. 12 n.12).

Hence, some of the indicators in the UN category of "process" cover resources, case loads, and other elements of government performance that have a causal impact on the actual realization of rights. These are similar to the NAS Government Performance Indicators. However, most of the UN process indicators would most likely count as Overall Outcome Indicators in the NAS scheme. For example, the UN process indicators on the "right to work" include such things as: "proportion of inspected enterprises that conform with labor standards;" "average time spent on unpaid domestic family care work as well as unpaid work of family business by women, men and children;" "proportion of informal sector workers shifted to formal sector employment in the reporting period;" and "proportion of labor force

undergoing some training during their employment.”

At the same time, some of the UN process indicators for the “right to work” would fit more aptly in the NAS Legal Framework category – for example, the proportion of the workforce covered by minimum wage legislation.

These are not just taxonomic quibbles. In full-scale assessments, our ultimate policy goal is to measure the effectiveness of *government* enforcement efforts. Our conceptual concern is with *government* compliance. Hence, the “input” or “causal” indicators should – as Kevin Davis has argued in his trenchant critique of “rule of law” indicators (Davis, 2004) – be limited to “legal information” or perhaps more broadly “government policy information.” The policy goal of U.S. legislation and treaties – and therefore the goal of our Indicators – is to create incentives for governments to engage in legal and policy reform to meet their legal obligations. “Input” Indicators of government performance or process should therefore specify the legal institutions and policies that are *both* causally significant *and* within the government’s control. Socio-economic indicators of the kind formulated by the UN may be relevant measures of “outcomes” or “effects” or they may be relevant control or independent variables in a multivariate model, but they are not causal indicators that directly measure the input of government effort.³⁴

It is true, of course, that the UNHCHR scheme of structural-process-outcome indicators has the imprimatur of the United Nations. And the 2008 *UNHCHR Report* itself notes that the ultimate objective of human rights indicators is to measure government

³⁴ The relation between input Indicators and output measures is discussed with greater nuance in Part 5 below.

commitment and effort to protect rights – conceptually quite similar to the goal of U.S. legislation and treaties. For the reasons just offered, however, that goal is better achieved by the NAS categories. The government’s substantive legal norms (its “Legal Framework”) and the government’s institutions and activities (its “Performance”) in advancing those norms are the relevant inputs that we wish to measure.

While the three NAS categories seem optimal, the NAS classified several Indicators in categories that did not match the content of the Indicators.³⁵ In particular, the label “Legal Framework” appeared to create confusion about whether a domestic law establishing enforcement institutions should be classified as a matter of the “Legal Framework” or of “Government Performance.”³⁶ In order to make the classifications as clear as possible, this paper proposes that the three NAS categories be relabeled as follows: (1) Substantive Law Indicators, (2) Enforcement Indicators, and (3) Outcome Indicators. This terminology is used in the remainder of the paper. In this terminology, measures of the legal design of enforcement institutions fall within Enforcement Indicators, along with measures of the resources, procedures, and actual functioning of those institutions.

³⁵ For example, one Legal Framework Indicator states: “provision in national law for labor inspection of workplaces;” while one Government Performance Indicator states “whether the country has a labor inspectorate;” and another Government Performance Indicator states simply “a labor inspectorate.” Another Legal Framework Indicator asks about an outcome: “assessment of the extent to which ‘right to work’ laws or other ‘free rider’ provisions undermine the ability of workers to organize....” (National Research Council, 2004, p. 114).

³⁶ See the preceding footnote.

4.2. A Fourth Category: Capacity-Building Indicators

This sub-part makes the case that an additional category of Indicators – Capacity-Building Indicators – is necessary to validly measure the ultimate concepts that concern us.³⁷

The Capacity-Building category includes four sub-categories of Indicators: (1) Indicators on data-collection by the government; (2) Indicators on the government’s formulation and application of its *own* indicators and targets for improved compliance, and its success in meeting those targets; (3) Indicators on the transparency of government institutions and on the participation of stakeholders in the enforcement of worker rights, in the formulation of targets, and in the evaluation of the government’s success in meeting those targets; and (4) Indicators on the government’s participation with other governments in coordinated evaluation of the comparative performance of the various governments.

The discussion in this sub-part proceeds as follows: First, it makes the case that international law *requires* governments to comply with such measures, as part and parcel of their obligation to effectively enforce or take steps to enforce substantive worker rights. Second, it discusses the conceptual reasons why such measures are essential to effective enforcement. Third, it addresses a question that is both taxonomic and conceptual: Why put these Indicators in a new category, rather than in the existing category of Enforcement Indicators?

³⁷ As emphasized throughout this paper, the three ultimate concepts that our body of Indicators must measure are “denial of internationally recognized worker rights,” “taking steps to afford internationally recognized worker rights,” and “effective enforcement of labor law.”

4.2.1. Requirements of Capacity-Building in International and Domestic Law

There are several good reasons for adding the category of Capacity-Building Indicators, but the first reason alone is decisive. International law *requires* governments to build their own capacity in the four ways just enumerated, as a matter of effectively enforcing the substantive right in question and, even more clearly, as a matter of taking adequate steps to afford workers the substantive right. These obligations pervade ILO Conventions, UN Conventions, and regional instruments.

Hence, ILO Conventions and UN Conventions require governments to collect comprehensive, accurate data on freedom of association, rights to organize, rights to bargain collectively, rights against employment discrimination, wages, hours, and occupational safety and health. Those instruments also require governments to collect data on various features of institutions devoted to enforcement of worker rights, including data on violations and penalties imposed. These obligations are set out in ILO Convention no. 160 on Labor Statistics, 1985, as well as in the core ILO Conventions and UN Covenants pertaining to labor administration, labor inspection, and the right to work. See Convention no. 81 on Labor Inspection, 1947; Convention no. 129 on Labor Inspection in Agriculture, 1969; Convention no. 150 on Labor Administration, 1978; and various authoritative Comments, discussed below, on the UN Covenant on Economic, Social, and Cultural Rights. More specific requirements are set out in the various substantive Conventions and Covenants on each of the rights codified in U.S. trade legislation and trade agreements.

Article 2 of Convention no. 160 requires governments, when collecting, compiling, and publishing labor statistics, to use “concepts, definitions and methodology” that “take into consideration the latest standards and guidelines established under the auspices of the International Labor Organization.” The most formal ILO standards are announced in Resolutions of the ILO International Conference of Labor Statisticians. ILO standards are also found in the ILO *Yearbook of Labor Statistics* (ILO, 2010c); *Key Indicators of the Labor Market* (ILO, 2010b); and other ILO documents cited below.

First, as to data-collection on wages and hours: Articles 1, 9-13 of Convention no. 160 require governments to collect comprehensive data on wages and hours. Article 9(1) requires that data on average earnings and hours actually worked “cover[] all important categories of employees and all important branches of economic activity, and in such a way as to be representative of the country as a whole.” Article 9(2) imposes the obligation to collect data on wage rates and normal hours of work in “important occupations or groups of occupations in important branches of economic activity, and in such a way as to be representative of the country as a whole.” Article 10 requires the collection of statistics on wage structure and distribution, for employees “in important branches of economic activity.” Article 11 requires the collection of statistics on labor costs “in important branches of economic activity.” Article 12 requires the computation of consumer price indices covering consumption patterns of “significant population groups or the total population.”

On protocols for data collection on work hours, the ILO International Conference on Labor Statisticians released a 2008 *Report* entitled *Measurement of Working Time* (ILO,

2008f), and adopted a *Resolution* on December 5, 2008, codifying the definitions of working hours contained in that *Report*. (ILO, 2008l). Additional explanation of the concepts in the *Report* and *Resolution* are found in the April 2008 *Report of the Meeting of Experts on Labor Statistics*. (ILO, 2008k). These two *Reports* and *Resolution* are the most authoritative internationally recognized standards on working hour concepts and definitions.

Governments' obligations to collect data on specific variables pertaining to wages and hours are found in the relevant Conventions addressing those specific subject matters, including: ILO Convention no. 1 on Hours of Work (Industry), 1919; Convention no. 14 on Weekly Rest (Industry), 1921; Convention no. 30 on Hours of Work (Commerce and Offices), 1930; Convention no. 47 on Forty-Hour Week, 1935; Convention no. 101 on Holidays with Pay (Agriculture), 1952; Convention no. 131 on Minimum Wage Fixing, 1970; and Convention no. 132 on Holidays with Pay (Revised), 1970.

As to occupational safety and health, Article 21 of Convention no. 81 and Article 14(2) of Convention no. 160 require governments to compile statistics on occupational diseases covering "all branches of economic activity." Article 14(1) requires the compilation of statistics on occupational injuries "in such a way as to be representative of the country as a whole, covering, where possible, all branches of economic activity." ILO Convention no. 187 on Promotional Framework for Occupational Safety and Health, 2006, Protocol no. 155 of 2002 to the Occupational Safety and Health Convention, 1981, Convention no. 155 on Occupational Safety and Health, 1981, and Convention no. 161 on Occupational Health Services, 1985, require governments to collect, analyze, and annually publish data on

workplace hazards, injuries, diseases, and dangerous occurrences – again, “taking into account relevant ILO instruments.” See Article 4(3)(f) of Convention No. 187 and Article 11(e) of Convention no. 155. Article 6 of Protocol no. 155 imposes essentially the same obligation, with the additional requirement that the statistics follow classification schemes established by the ILO or other competent international organizations. As to occupational accidents, the most authoritative classification scheme is found in the Annex to the *Resolution Concerning Statistics of Occupational Injuries (Resulting from Occupational Accidents)*, adopted by the 16th International Conference of Labor Statisticians. (ILO, 1998d).

More specific data-collection requirements pertaining to safety and health are found in the ILO’s various industry-by-industry *Codes of Practice*, which are referenced in the body of Assessment Indicators in Appendix A to this paper, and in the following ILO Conventions: Convention no. 184 on Safety and Health in Agriculture, 2001; Convention no. 176 on Safety and Health in the Mines, 1995; Convention no. 174 on Prevention of Major Industrial Accidents, 1993; Convention no. 170 on Chemicals, 1990; Convention no. 167 on Safety and Health in Construction, 1988; Convention no. 162 on Asbestos, 1986; Convention no. 148 on Working Environment (Air Pollution, Noise and Vibration), 1977; Convention no. 139 on Occupational Cancer, 1974; Convention no. 136 on Benzene, 1971; Convention no. 120 on Hygiene (Commerce and Offices), 1963; Convention no. 119 on Guarding of Machinery, 1963; Convention no. 115 on Radiation Protection, 1960; Convention no. 62 on Safety Provisions (Building), 1937; Convention no. 13 on White Lead (Painting), 1921.

ILO Conventions and UN Conventions also require governments to collect

comprehensive, accurate data on freedom of association, rights to bargain collectively, and rights against employment discrimination. As for nondiscrimination, Convention no. 160 on Labor Statistics, 1985, requires governments to collect data on all major groups of workers, as categories that are transversal to the wage, hour, and employment data discussed above. As for both sets of rights – freedom of association, rights to organize, and rights to bargain collectively; and rights against employment discrimination – in its first authoritative General Comment on obligations under the UN Covenant on Economic, Social, and Cultural Rights, the Committee on Economic, Social, and Cultural Rights articulated the obligation of each government to routinely collect sufficient data to monitor whether these rights are “being enjoyed by all individuals within its territory.” (UN CESCR, 1989).³⁸ It emphasizes that “this objective cannot be achieved only by the preparation of aggregate national statistics or estimates, but also requires that special attention be given to...specific groups or sub-groups which appear to be particularly vulnerable or disadvantaged.” (UN CESCR, 1989).

These and other ILO and UN instruments also require governments to formulate their own indicators and numerical targets to ensure that they comply with, or are taking effective steps to comply with, the rights and standards that concern us. Hence, the UN Covenant on Economic, Social, and Cultural Rights requires governments to formulate “specific

³⁸ That these obligations apply to rights of association, rights to organize, rights to bargain collectively, rights against employment discrimination, and standards of minimum wages, hours of work, and occupational safety and health is clear from the face of the Convention itself, which explicitly states these rights; from General Comment no. 18 on the Right to Work, which emphasizes not only the rights to organize, to bargain collectively, and to acceptable conditions of work, but also additional international and regional instruments recognizing the universal rights against discrimination based on race, gender, and citizenship status (UN CESCR, 2005, ¶¶ 2-3, 7, 13, 36, 38); and from General Comment nos. 16 and 20 on nondiscrimination, which also explicitly include nondiscrimination based on age and disability. (UN CESCR, 2005; 2009).

benchmarks” and a “detailed plan of action” based on “carefully targeted policies, including the establishment of priorities” reflective of the international instrument. (UN CESCR, 1989, ¶¶ 4, 6). Governments must also assess their success or failure in meeting those targets. (UN CESCR, 1989, ¶¶ 6). General Comment no. 18 on the Right to Work explicitly requires governments to identify “indicators and benchmarks,” “numerical targets” and “time frame[s] for implementation” for freedom of association, the right to organize, the right to bargain collectively, the right of nondiscrimination, and acceptable conditions of work. (UN CESCR, 2005, ¶¶ 2, 7, 12,13, 36, 38.) Governments are directed to use statistical definitions developed by the ILO. (UN CESCR, 2005, ¶ 46).

ILO Conventions are pervaded with mandates that the social partners participate in all phases and functions of labor administration. The UN Covenant on economic and social rights also requires governments to take measures to ensure public scrutiny and stakeholder participation “in the formulation, implementation, and review” of the targeted policies that aim to meet the specific benchmarks and targets. (UN CESCR, 1989, ¶¶ 5; UN CESCR, 2005, ¶ 42).

Finally, the Covenant requires governments to take measures that “facilitate the exchange of information among States and to develop a better understanding of the common problems faced by States and a fuller appreciation of the types of measures which might be taken” to effectively realize the rights (UN CESCR, 1989, ¶¶ 5).

4.2.2 Conceptual Rationales for Applying Capacity-Building Indicators

Apart from the unequivocal requirements of international law, the relevant legal instruments explicitly or implicitly provide conceptual rationales why governments cannot “effectively” enforce rights or adequately “take steps” to enforce rights unless they build their capacity to do so. These conceptual rationales provide further understanding of the underlying concepts that our Indicator methodology measures.

The conceptual importance of data collection is obvious enough: Governments cannot ensure compliance or take steps to comply unless they have information about the current and recent degree of compliance. As stated in the authoritative Comment No. 1 of the UN’s CESCR, “the essential first step towards promoting the realization” of economic rights (including those rights of concern to us) is “diagnosis and knowledge of the existing situation.” (UN CESCR, 1989, ¶ 3).

The importance of the second sub-category of Capacity-Building Indicators is also straightforward. A government’s project to improve compliance is meaningless if it is aimless. It must have targets at which to take aim. The government must formulate its own specific indicators and targets, and must implement protocols to measure its success in achieving those targets.

In a real sense, ILAB’s program of applying Indicators to U.S. trading partners, and comparing countries based on their success or failure in satisfying those Indicators,³⁹

³⁹ Even if ILAB, the USTR, and the President do not wish to compare or rank countries in their compliance with the overall body of Indicators, there are good reasons for framing *individual* Indicators in comparative terms, as discussed below in Part 6.

presupposes that governments are expected, over time, to improve their compliance – at least if we start with the plausible assumption (indeed, the manifest fact) that most if not all countries fall short of full compliance with every significant dimension of every labor right.

That assumption is confirmed by the very fact that U.S. legislation uses the language of “taking steps.” In this respect, it is deeply relevant that the authoritative UN body’s determination that the government must formulate and apply indicators and targets for compliance with worker rights, and that the government must assess its success in achieving those targets, is precisely an interpretation of the “taking steps” language in international law. (UN CESCR, 1990; 2006).⁴⁰

What are the conceptual rationales for the additional measures of transparency and civil-society participation in the government’s enforcement activities, including in its formulation and application of its own indicators and targets and in its self-assessment of whether it has met those targets?⁴¹ First, of course, there is intrinsic value in democracy. Second, the stakeholders and the public are, at least in many instances, very likely to have the better knowledge (compared to centralized bureaucrats) of actual realization of specific

⁴⁰ A government’s formulation of its own indicators, targets, and self-evaluation is especially important in light of the fact that many of the Indicators proposed in this paper take a comparative form. As explained below in Part 6, many of the Indicators measure whether the government is enforcing a certain dimension of a labor right to a degree that exceeds that of the average enforcement among countries in the same quintile of real income per capita. While such Indicators serve the function of measuring current or past performance relative to the performance of other countries, they plainly serve a second, concurrent function – namely, inducing countries to continuously improve their performance in order not to become laggards in their quintile. As discussed in Part 6, many other Indicators are even more explicit about comparing improvement over time: They measure whether a country’s rate of increase in enforcing a particular dimension of a right exceeds the average rate among countries in the same quintile of real income per capita.

⁴¹ Recall that, regardless of any functional rationales, international law explicitly stipulates that these are constitutive elements of compliance with substantive rights.

dimensions of rights and of obstacles to improving compliance. They are therefore likely to have useful information about how to set and achieve targets for improvement.

Third, the stakeholders may play a vital role in articulating priorities in the allocation of scarce enforcement resources to the matters that are most urgent to workers and their communities. Fourth, and closely related, in the ongoing process of converting general standards and bright-line rules into more specific sub-indicators and potentially revising the Indicators themselves, it is the beneficiaries of the right (workers and other civil-society groups⁴²) and the parties that proximately implement the right (managers) who are best situated to understand the technological and organizational constraints and opportunities for applying and revising rules and standards in ways that provide greater protection of worker interests as workers themselves understand those interests.⁴³

Fifth, such accountability in the implementation of domestic law also promotes institutional compliance with general principles of due process and with particular procedural and substantive requirements measured by Substantive Law Indicators and Enforcement Indicators. That is, the strongest political force for ensuring that governments enforce worker rights through fair process is workers' and the public's well-informed participation in defining, implementing, and monitoring such enforcement.

There is a final reason for democratizing the process of enforcement, stemming from the nature of a project such as this, committed to applying a body of Indicators on a global

⁴² Among the civil-society groups relevant to government performance pertaining to the right against employment discrimination are organizations of women, racial groups, non-citizens, the disabled, and other protected groups.

⁴³ This proposition is one way to express the non-derogation principle.

scale. As many scholars, commentators, and advocates have cautioned, the trend toward increasing use of such measures runs the risk of diminishing democratic politics and local deliberation about the meaning of rights. (Rosga and Satterthwaite, 2009; Merry, 2009; Davis and Kruse, 2007; Davis, 2004; Jacobsson, 2000). Especially when an Indicator gains first-mover prominence (whether the UN Human Development Index or the World Bank Doing Business Reports), it carries a technocratic authority – and a “branded” supremacy – that not only makes it difficult to contest and revise, but often difficult to understand. (Davis, et. al., 2010; Hoyland, et al., 2009). A reasonably well-defended set of Indicators deployed by the U.S. government may provide a global benchmark with substantial authoritative weight. It is therefore desirable to encourage governments to develop their own indicators and targets through democratic participation. (See Rosga and Satterthwaite, 2009, p. 313-314). Those “decentralized” indicators may serve as useful interlocutors for the U.S. Indicators.

On this point, there is an inevitable tension. From the standpoint of the U.S. Executive Branch, the U.S. Indicators are decisive and are intended to create incentives to meet those very measures. It is, however, both undesirable and impractical to expect that the U.S. Indicators can simply be stamped into the enormously varying (not to mention resistant) politics and labor relations of U.S. trading partners. And yet, someone must have the last word in determining whether the terms of U.S. trade legislation and treaties are violated.⁴⁴ If the U.S. government cannot be expected to “delegate” that determination to the very

⁴⁴ Even the critics of “centralized” or “technocratic” indicators recognize that decentralized determination of human rights by local actors is problematic, in just this respect. (Rosga and Satterthwaite, 2009, p. 297, 301).

countries being assessed, the U.S. government might nonetheless construct an Indicator methodology that foments some degree of (formal or informal) iterative dialogue and mutual learning prior to the inevitable closure by U.S. government decision. Since we self-consciously conceive the body of U.S. Indicators to be revisable, those revisions may be informed by the indicators proffered by trading partners after their domestic processes of democratic deliberation – the very democratic deliberation measured by the U.S. Capacity-Building Indicators.⁴⁵

Finally, what is the conceptual rationale for the fourth sub-category of Capacity-Building Indicators – Indicators of the government’s participation in coordinated discussion with peer governments?⁴⁶ Granted, the U.S. government could well make its judgment about the performance of its trading partners without any process of dialogue among those countries, and without formal or informal hub-and-spoke dialogue between the U.S. and each of its trading partners.

Still, many of the Indicators proposed in this paper indeed rest on the conceptual view that government performance is often best measured comparatively – at least in the framing and application of *individual* Indicators, even if the U.S. government abjures rankings or comparisons of *overall* country performance. The reasons for framing individual Indicators in comparative terms are discussed at length in Part 6 below. One reason, relevant to the present discussion, is forward-looking: we want to induce countries to keep apace with and

⁴⁵ Part 12 below discusses some of the challenges and strategies for implementing such iterative dialogue.

⁴⁶ Recall once again that, regardless of the functional rationales, international law requires such participation, as a defining feature of compliance with the substantive rights.

surpass their peers. (Barenberg, 2001). Since not all countries can perform above average or above some other threshold within their peer group, the inducement to meet those benchmarks encourages continuous improvement in enforcement.

The absence of dialogue among the peers, however, increases the risk that countries will “game” the race to the top, either by collusively withholding data or by separately pressing specious arguments to justify their worse performance (that is, to ostensibly show that they are not in fact similarly situated with other countries that are doing better). If, however, countries are required to justify their differential performance as to particular Indicators in open discussion with their peers, or in hub-and-spoke communication with the U.S. government,⁴⁷ there may be a greater likelihood that their arguments will be put to the test. They will have to produce the data and the analysis to defend their arguments justifying their worse performance. They will be met with counter-evidence and counter-arguments why their circumstances in fact permitted them to have performed better. The dialogue may thereby flush out data and analysis showing what works and what does not work in efforts to effectively enforce particular rights or particular dimensions of particular rights.

The requirement of participation in disciplined discussion with peers or the U.S. government offers at least the potential to accelerate the cross-border process of learning about and improving domestic enforcement. In the best case, it also offers the possibility that

⁴⁷ Here, we refer to hub-and-spoke discussions between U.S. officials and the officials of various trading partners, for purposes of disciplined comparisons about compliance. Part 12 below discusses the possible implementation of hub-and-spoke dialogues between U.S. officials and a trading partner’s officials and stakeholders, with the aim of flushing out data and justifications after a poor provisional assessment of the trading partner by the U.S. government.

countries can improve not only enforcement but also the substantive specification of rights, including the conversion of general standards and rules into more specific sub-indicators. That is, they may learn which sub-indicators best advance the general values and principles animating the more abstract standard or rule. For example: Which specific rules about an employer's obligation to monitor and punish supervisors best protect workers against sexual harassment? Or, more expansively: Which specific strategies of organizational reform, if any, concurrently reduce sexual harassment and counter other pathologies or inefficiencies in business performance? (See, e.g., Sturm, 2001.)

4.2.3. Why Not Include Capacity-Building Indicators in the Category of Enforcement Indicators?

While this question is taxonomic, it is more than that. It is true that "capacity-building" is a feature of government performance and could therefore be included in the category of Enforcement Indicators. But there are good conceptual grounds for treating capacity-building as a separate, co-equal category.

To uncover those grounds, let us first take a step back and examine the institutional features and practices that are at the core of Enforcement Indicators.

"Enforcement" includes at least four major components: (1) the institutional architecture of enforcement agencies (the structure of major organs of labor administration, such as labor inspectorates, labor prosecutoriates, labor courts, labor boards, appellate courts); (2) the material and human resources devoted to those organs; (3) the functional

processes of those organs (labor inspections, prosecutions, filing of civil claims, case processing by labor courts and boards, remedies ordered by those courts and boards); and (4) the actual activities and practices of the enforcement institutions.

Recall that in discussing the specificity and revisability of Indicators in sub-parts 3.2 – 3.4, our examples focused largely on the kinds of measures that might be conceptually desirable as Substantive Law Indicators – such as the fixed Indicator-as-bright-line-rule for maximum hours and overtime pay; the fixed Indicator-as-general-standard for sexual harassment; the revisable Indicator-as-bright-line-rule for carbon monoxide masks in coal-mining; or the revisable Indicator-as-general-standard for psychological abuse of workers.

Much of the discussion in Part 3 applies as well to Indicators for measuring Enforcement. Hence, each Indicator of structure, resources, and processes could be framed as a fixed bright-line rule, as a fixed standard, as a revisable bright-line rule, or as a revisable standard. The various considerations mooted in sub-part 3.4 for determining which norm-form is optimal are therefore relevant to formulating Enforcement Indicators: For example, is the institutional process in question relatively fixed, subject to precise descriptive demarcation, and governed by relatively stable value-commitments? An Indicator asking “Are all court proceedings in which evidence is taken in cases alleging violation of rights against employment discrimination open to the public?” might capture such a process, calling for a fixed bright-line rule. By contrast, an Indicator asking “whether judicial proceedings are expeditious” may call for a fixed standard rather than a bright-line rule, since the underlying value is clear and stable, but it seems impracticable to pre-define the time limits for each

phase of a civil or criminal case, especially in light of varying docket loads from country to country and tribunal to tribunal, and varying types of cases heard by those tribunals.

Nonetheless, there are also conceptual issues specific to legal norms governing institutional structure, resources, processes, and practices. Enforcement measures must apply both to specific complaint-driven assessments and to systemic evaluations.

Some Indicators of effective enforcement will ask whether government enforcement institutions are operating effectively in individual cases or in an identifiable set of individual cases. For example, where a union leader was the victim of anti-union-motivated violence, did the government swiftly and rigorously investigate and, if there is reasonable cause to believe the allegation, effectively prosecute the alleged assailant?

However, institutional performance also requires measurement of variables that are more comprehensive than the government's response to specific complaints or indictments. This requirement arises from the fact that, by definition, Indicators measure the effective or ineffective functioning of the enforcement institutions as a whole, or the steps taken to achieve effective enforcement by institutions in the aggregate.

These sorts of systemic performance Indicators might be called *first-order* Enforcement Indicators, in that they directly measure the government's capacity to investigate possible violations, to engage in fact-finding and case processing, to remedy violations, and to deter future violations. First-order Enforcement Indicators may also measure the government's performance in more nuanced "promotional" activities – such as training managers and workers to better achieve compliance, or properly supervising frontline

government officials to ensure that they do not wield their institutional powers to violate worker rights (such as supervising riot police officers who might otherwise use their powers to violently suppress strikes, or frontline Ministry of Labor officials who might otherwise impede the registration of newly established labor unions). And, where institutions have in fact been charged with enforcing norms-framed-as-standards, first-order Indicators may measure the government's own performance in fleshing out standards through rule-specification – that is, its performance in converting standards into rules, an exercise that ILAB analysts themselves will undertake in applying Indicators, as discussed in sub-part 3.2 and diagramed in Figure 3 above. Likewise, first-order Indicators may measure the government's performance in revising substantive norms (whether rules or standards) to ensure that such revision does not derogate from existing protections but instead heightens rights-protection in response to economic and social changes that provide opportunities for stronger protection or in response to an evolving international consensus on value commitments.

We may also go a step further, and formulate Indicators that measure not just the institution's first-order capacity and performance in promulgating norms, in converting standards into rules, in revising substantive and procedural norms, in enforcing those norms, and in promoting public awareness about substance and procedure, but also its *second-order capacities* to self-monitor and strengthen its first-order capacities and performance.

It is these second-order Indicators that rate the new category of Capacity-Building Indicators. Conceptually, the goal of Enforcement Indicators is to measure the government's

existing capacity to enforce labor rights. The goal of Capacity-Building Indicators is conceptually something different: to measure the government's capacity for sustained improvements in enforcement – its capacity to continue to “take steps” to better afford workers their rights. Granted, this is a high-powered way of measuring whether a government is taking steps to achieve compliance. It asks not just about steps that have been taken or are being taken in first-order enforcement activities. It asks as well about the government's capacity to take first-order steps *in the future*; about its capacity to take *increasingly large* future strides; and about the *relative assurance* that the government will in fact take those future strides – since the forward motion of ongoing institutions, protocols, and constituents for building capacity may have a momentum of its own.

If this seems like asking too much, recall again that international law requires as much, and it does so in the name of measuring, precisely, whether a government is “taking steps” to realize worker rights. The government's collection and analysis of data, its formulation of indicators and targets, its self-assessment about whether it is meeting those targets, its inclusion of civil-society stakeholders in all those processes, and its participation in disciplined comparisons of its performance with that of its peers – these are all sound measures of whether the government has built an architecture that is likely to improve enforcement in the future and therefore whether it “has or is taking steps” *toward actually affording* workers their internationally recognized rights.

5. Input or Output Indicators? – Conceptual Problems



Three categories of Indicators – Substantive Law Indicators, Enforcement Indicators, and Capacity-Building Indicators – are “input” Indicators. The first category measures the adequacy of the substantive legal rules enacted by the government; the second measures the governmental institutions, procedures, resources, and practices to enforce those rules; and the third measures the government’s capacity to strengthen substantive rules and enforcement and to do so in a way that is participatory, transparent, and accountable.

The fourth category of Indicators – Outcome Indicators – measures “output,” that is, actual compliance with the substantive labor law norms. Sub-part 4.1 made the point, in passing, that in evaluating country performance our only concern is with input Indicators since we want to measure variables within the control of the government. But, to many, that may seem counter-intuitive. It might seem obvious to maintain instead that the ultimate conceptual question – the only question that “really” counts – is overall outcomes, that is, the actual terms and conditions of employment enjoyed by workers in their workplaces. Those terms and conditions are determined principally (or at least proximately) by employers. So, our only practical concern should be the degree to which employers retaliate against union supporters, fail to engage in good faith collective bargaining with legitimate worker organizations, maintain discriminatory practices, fail to protect workers against injuries and

fatalities, and so on. Hence, if we can construct and reliably apply Outcome Indicators that are good measures of employer compliance, we need not dwell on government inputs.

Taking employer compliance as our ultimate conceptual concern seems natural in large part because compliance by private actors is unusually salient in the field of labor law compared to other fields of human rights. While most human rights are dyadic, providing individuals a shield against government abuse, most labor and employment rights instead have a triangular quality. The government's obligation is to secure either a worker's negative entitlement against interference by employers – for example, the government must protect workers against employer retaliation motivated by the worker's gender or race – or a worker's positive entitlement to receive some benefit from the employer – for example, the government must ensure that the employer pays workers a certain wage or provides workers certain safety gear. (While most worker rights have this triangular quality, it bears emphasizing that not all do. The government also has obligations to refrain from direct violation of worker rights. For example, the government must not repress lawful strikes or create unjustified impediments to the establishment of worker organizations. These are classical dyadic rights, much like other human rights that focus on government harm to persons.)

However, from a formal legal standpoint our ultimate concern appears, at least at first glance, to be not compliance by employers with the rules laid down by international or domestic law, but rather compliance *by governments* with the terms of U.S. trade legislation and treaties. Under the terms of those instruments, the obligation to enforce rights or to take

steps toward affording those rights rests solely with sovereign entities, not with private employers. The USTR's and President's responsibility is, in turn, to use trade measures and other constitutional powers to create incentives for trading partners (*qua* government bodies) to meet their obligations, as well as to ensure that the U.S. government meets its own legislative and treaty obligations.

From this formalistic standpoint, the USTR's and President's ultimate concern is with input Indicators – that is, Indicators that measure legal information: Has the trading partners' government adequately defined the relevant substantive labor rights? Has that government brought adequate institutions, resources, procedures, and practices to bear on employers' behavior? Has that government refrained from committing direct violations of worker rights?

Under this formalistic analysis, the attention might shift not just from outputs to inputs, but also to labor violations that come before (rather than after) government enforcement efforts. That is, one seemingly relevant conceptualization of effective enforcement may be the degree and kind of government regulatory effort relative to the degree and kind of “pre-existing” employer and government violations.

But, in fact, the strict dichotomy between *ex post* employer compliance and government compliance poses a false choice. We need not choose between “sovereign compliance” and “employer compliance” as ultimate concerns. Both are conceptually relevant – for at least two reasons. First, the common-sense intuition that employer compliance is vital can in fact be framed in the proper terms of formal legal interpretation. Statutes and treaties are properly interpreted in light of their purposes; and one of the key

purposes behind the labor provisions of U.S. trade legislation and treaties is unquestionably to protect workers' rights against misdeeds and omissions both by governments *and* by employers. Congress made this explicit when it enacted the labor rights provisions of Section 301 of the Trade Act:

[P]romoting respect for internationally recognized rights of workers is an important means of ensuring that the broadest sectors of the population within [developing countries] benefit from [access to U.S. markets]. The capacity to form unions and to bargain collectively to achieve higher wages and better working conditions is essential for workers in developing countries to attain decent living standards and to overcome hunger and poverty. The denial of internationally recognized worker rights in developing countries tends to perpetuate poverty, to limit the benefits of economic development and growth to narrow privileged elites, and to sow the seeds of social instability and political rebellion.⁴⁸

Hence, outcomes matter. To take an extreme example, a level and type of enforcement that has no discernable effect on actual employer violations could not be called adequate, let alone "effective." We are, at least for some aspects of enforcement, equally if not more concerned with the "units" of government enforcement (and the qualitative character of those "units") relative to the "units" of *ex ante* and *ex post* violations – precisely because the Congressional purpose behind requiring "effective enforcement" and "taking steps" is to actually secure rights.

However, this raises an immediate caution. If *ex post* outcomes are taken as an Indicator of government enforcement efforts, then we run the risk of double counting if we aggregate such output indicators with the input Indicators that caused the effect. If a

⁴⁸ H. Rep. No. 98-1090 (1984) at p. 11, reprinted in 1984 U.S.C.C.A.N. 5101, 5111.

government ranks high on performance Indicators, then a high ranking on outcome Indicators may simply be a causal byproduct of the former. While this may not be problematic in the case of Probative and Diagnostic Indicators, which do not purport to be statistically valid and which must grasp whatever reliable and relevant information is available in a data-poor environment, it does raise a methodological problem for Assessment Indicators serving a high-stakes evaluative function.

For purposes of Assessment, then, we should take care not to use Outcome Indicators as *evidence* of compliance with input Indicator. We may instead, in appropriate instances, treat outcome measures as elements in the *definition* of input Indicators – via, for example, a metric that requires government inputs to be calibrated to achieve a specified average level of inputs per extant violation, or to achieve a specified average or marginal improvement in outcomes per unit of inputs. That is, setting aside for the moment questions of data availability, Indicators that compare government efforts to outcomes may serve as compelling normative measures of the “effectiveness” of government inputs and of the adequacy of “steps taken.”⁴⁹

⁴⁹ Measuring inputs relative to outcomes is just one of several possible strategies for defining absolute quantitative normative metrics for particular Indicators: (1) An Indicator might apply to all countries a metric stipulating a specific, fixed, absolute number of units of inputs per year (scaled to population, income per capita, or other pertinent variable). (2) The metric might instead stipulate a specific, absolute level of outcome (again scaled to population, income per capita, or other variable). (3) The metric might stipulate a specific percentage increase in inputs per year. (4) The metric might stipulate a specific percentage improvement in outcomes per year. (5) The metric might measure inputs relative to pre-existing violations, for example by using a linear metric that measures compliance against a specified ratio of units of inputs per units of pre-existing violations. (6) The metric might instead be geometric – that is, the ratio of inputs to pre-existing violations would increase as the absolute level of pre-existing violations increases – on the normative assumption that widespread violations call more urgently for enforcement responses. (7) Alternatively or concurrently, we might require all countries to achieve a specific, fixed marginal increase in unit of outcomes per unit of inputs. (8) To add a dimension of contextualization, we might demand higher marginal improvements in outcomes for countries with more widespread problems, based on an assumption of diminishing marginal returns to enforcement inputs.

The inverse possible solution is noted in the *OECD Handbook* on constructing composite indicators – using the example of a composite indicator for the ultimate concept of “Innovation,” which we can take as an analogue to our composite Indicator for the ultimate concept of “effective enforcement” of or “taking steps to afford” worker rights:

Too often composite indicators include both input and output measures. For example, an Innovation Index could combine R&D expenditures (inputs) and the number of new products and services (outputs) in order to measure the scope of innovative activity in a given country. However, only the output indicators should be included (or expressed in terms of outputs per unit of input) if the indicator is to measure innovation performance.

(OECD, 2009, p. 22). That is, our Indicators *could* rest primarily on outcome measures.

Input measures would then serve as elements in the definition of outcomes (for example, outcomes per unit of government enforcement) or as indirect evidence of outcomes in the absence of direct evidence.

For two reasons, however, the proposed Indicators in this paper opt for the first solution – relying principally on input Indicators, at least in the formulation of Assessment Indicators (as distinguished from Probative and Diagnostic Indicators). First, data on inputs are likely to be more readily accessible than data on outcomes. Second, ILAB officials have,

For example, we *might* expect greater increases in workplace gender equality and women’s labor market participation from each additional “unit” of enforcement inputs in Southern European or Muslim countries compared to Northern European or Anglo-American countries, since the former start at such low levels of gender equality. (9) Or just the reverse: We might apply a metric that measures the ratio of inputs to outcomes but that excuses countries in which violations remain relatively high (weak outcomes) after the government brings greater enforcement inputs to bear on widespread problems, on the assumption that a high absolute level of pre-existing problems indicates that the problems are more deeply entrenched. To continue our example, we might think it more probable (and more excusable) that even large increases in enforcement efforts devoted to increasing gender equality will not significantly improve the entrenched problems in Southern European or Muslim Countries.

for the most part, expressed a preference for input Indicators over outcome Indicators as direct measures of government performance – and for good reason. The methodology proposed in this paper rests substantially on creating high-powered incentives for data collection. If we use outcome measures as evidence of effective enforcement, on the ground that direct evidence of inputs is unavailable, then we diminish the incentives for trading partners to collect and publish data on specific enforcement efforts, and to build their systemic capacity to collect and publish data on the universe of relevant compliance matters.

True, ILAB officials have also indicated their interest in the use of outcome indicators for purposes of research on the relation between government performance and outcomes.

These multiple policy goals can be achieved in significant part by relying predominantly on input Indicators, while using outcome measures in *three ways*: first, by using direct measures of outcomes, while ensuring that such Outcome Indicators are not used for Assessment purposes when those measures would double count the specific inputs that cause the specific outcome; second, by using direct outcome measures as Probative and Diagnostic Indicators, without concern for the double counting of cause and effect, for reasons just stated; and third, by using Enforcement and Capacity-Building Indicators that, as just proposed, are in some pertinent cases *defined* by their impact on outcomes.

However, if we wish to take outcome measures as elements in the definition of input Indicators, we must address several conceptual issues, some of which are empirical and some of which are normative.

First, the number of *ex post* employer violations, taken alone, cannot tell us whether

the government has, in a normative sense, ensured adequate *ex post* compensatory and punitive remedies for the violations that have proven resistant to deterrence. That is, we need to consider whether, even if we find a large number of *ex post* violations, we should still give the government's enforcement efforts a high ranking based on the fact that the government ensured expeditious and full compensation of the victims of the violations and punitive sanctions. To put it differently, input Indicators should measure not only effective deterrence of violations relative to enforcement inputs, but also actual compensation and punishment.

A second issue stems from the relation between the strength of substantive norms and the degree of enforcement effort. For example, a country with a minimum wage that exceeds the market-clearing wage can expect greater rates of employer non-compliance even with high levels of enforcement, compared to a country with weak enforcement efforts and a minimum wage set below the market-clearing level. It seems perverse, however, to give the former country a negative assessment compared to the latter country. Hence, we must be careful to define our Indicators in a way that takes account of distinctions such as this. The ratio of outcomes relative to enforcement efforts must be calibrated to the strength of substantive norms.

Third, if we wish to define effective government efforts in terms of actual violations deterred or remedied, then we face the apparent challenge of converting pre-existing violations, enforcement efforts, and outputs into commensurable units. While units of pre-existing violations and units of *ex post* outcomes might often be apples and apples, the units

of various kinds of enforcement efforts are often oranges and apples. Several different types of enforcement inputs may be targeted concurrently at a particular type of violation. It appears that there is a normative challenge to scaling and aggregating units of those several types of enforcement efforts for purposes of defining inputs in relation to pertinent outputs. One solution is to ensure that individual Indicators measure the relation between only one input and its pertinent output(s). For example, “In the preceding five years, has the ratio between the number of workplace inspections that result in a final remedial order (outputs) and the number of labor inspectors (input) increased?” But since we recognize that there are, and should be, many variables other than the number of labor inspectors that affect whether workplace inspections lead to successful outcomes, and that there are many facets of a “successful” outcome besides a final remedial order, we must be careful to lodge the above Indicator in a set of related Indicators that are validly weighted and aggregated. Equally if not more important, we must be careful not to over-use such Indicators. Keeping input measures and output measures definitionally separated has the advantage of allowing ILAB and other researchers to engage in subsequent study of the various possible relationships between one or more particular input(s) and one or more particular output(s).

Fourth, measures of the marginal or average outcome from the government’s total or incremental units of enforcement require rich information. Problems of data availability, data accuracy, and consistent protocols for defining and collecting data across time are discussed below in Part 9. But there is a conceptual problem here as well. How do we construct and measure the counterfactual state of violations that would exist in the absence of current

enforcement efforts? Is the current level of violations *in spite of* current enforcement efforts, *because of* current enforcement efforts, or *unrelated to* enforcement efforts?

This points to a final problem: Recall that the NAS methodology asks analysts to score Indicators at one point in time. Analysts must note whether there are some problems, more extensive problems, or severe problems in compliance with the Indicator; and analysts must note the direction of change at that single point in time. This raises three conceptual hurdles, in light of the other challenges discussed above: (1) with single data points identified at one moment in time, it seems intractable to determine the number and type of pre-existing violations that would counterfactually occur in the absence of current levels and types of legal institutions and legal enforcement; (2) in some instances, as noted above, there may be no normatively sensible measure of the effectiveness of enforcement independent of pre-existing violations and actual *ex post* outcomes; and (3) as also noted above, the normative implication of outcome Indicators is often conceptually ambiguous and in any event relevant only insofar as those Indicators tell us, without double counting cause and effect, the extent to which existing enforcement efforts have reduced (indeterminate) pre-existing violations.

As to the measurement of some aspects of compliance, the explicit formulation of longitudinal and comparative Indicators may mitigate these challenges. The next Part discusses such Indicators.

6. Normative Yardsticks: Longitudinal, Comparative, and other Numerical Indicators



Every Indicator must have a normative (that is, values-based) valence. That is, is a “yes” answer to the Indicator a good or bad score for the trading partner’s compliance?

As we have seen, for all Indicators the U.S. Executive Branch has authority to stipulate normative thresholds that reasonably implement Congress’s intent in enacting trading partners’ obligations to comply with or take steps to afford internationally recognized worker rights and to effectively enforce labor law. For many Indicators, international or comparative labor law provides a clear normative yardstick that reasonably guides the Executive Branch’s stipulation. For other Indicators, there may be no such clear guidance.

The comprehensive taxonomy of normative yardsticks is this:

- (1) a **qualitative norm**, asking about the existence of some law, institution, or procedure, such as whether the law prohibits sexual harassment; or
- (2) a **quantitative norm**, including:
 - (a) an **absolute number**, such as a requirement of overtime pay for work weeks in excess of 48 hours; or
 - (b) a **zero tolerance norm** for certain quantifiable enforcement events, such as the number of failures to prosecute known murderers of trade unionists in the preceding two years, or
 - (c) a **normative threshold of a “trivial” number** of certain quantifiable enforcement events when a zero tolerance standard would be unreasonable, such as the number of judicial proceedings that were closed to the public in labor cases in the preceding two years, or

- (d) a quantitative metric based on the *country's longitudinal improvement* in enforcement, such as whether the number of labor inspectors per worker increased in the preceding two years, or
- (e) a quantitative metric based on *the country's performance compared to some quantitative measure of peer countries' performance*, such as whether the country's damage award per unlawfully discharged worker exceeded the average among countries in the same quintile of real income per capita in the preceding two years.

The logic and reasonableness of qualitative norms, absolute quantitative norms, and zero-tolerance norms are clear enough. Sub-part 1 below makes the case that, as to some matters, the U.S. Executive Branch may stipulate longitudinal and comparative Indicators as reasonable implementations of Congressional intent. Sub-part 2 then examines whether the commitment to universal human rights is undermined by Indicators defined in terms of the variable performance of a country over time and the contingent performance of peer countries, or by Indicators that are subject to revision in the manner proposed in sub-part 3.3 above. Sub-part 3 then defends the quantitative metric of “non-trivial violations” where quantitative Indicators are not reasonably based on zero tolerance or some other absolute quantitative measure.

6.1. The Case for Longitudinal and Comparative Indicators

The absence of a clear qualitative or numerical baseline in international law is not a paralyzing problem. In the absence of a clear metric, the Executive Branch (whether ILAB, USTR, or the President) has the constitutional authority to stipulate a metric, so long as the

metric “reasonably” implements Congressional intent.⁵⁰ It would be reasonable, for example, for the Executive Branch to stipulate an absolute quantitative metric – guided perhaps by ILO recommendations – for a matter such as the number of labor inspectors per 100,000 workers.

Nonetheless, there may be aspects of compliance for which an absolute metric is neither clearly given by international standards nor readily subject to reasonable stipulation by the U.S. Executive Branch. In such cases, it is highly reasonable for the Executive Branch to formulate its Indicators in longitudinal and/or comparative terms. The use of longitudinal Indicators – which ask whether the government’s compliance has improved relative to its earlier performance – is a reasonable interpretation of the progressive compliance codified by Congress in the language of “taking steps” in trade legislation and in the analogous language of trade agreements incorporated in Congressional statutes implementing those agreements.⁵¹

Comparative Indicators have several reasonable bases: First, they show the feasibility of better performance, since other similarly situated countries are able to achieve the metric defined in comparative terms. Since the countries are similarly situated in relevant respects, there is by definition no good reason for the country in question not to keep pace. As a matter of principle, then, the U.S. Executive Branch can demand that a country achieve

⁵⁰ See sub-part 2.4 above.

⁵¹ E.g., Article 18.1(2), United States-Chile Free Trade Agreement, January 1, 2004, at www.ustr.gov. (“Each Party shall strive to ensure that...the internationally recognized labor rights are...protected...”); Article 16.1(2), Dominican Republic – Central America – United States Free Trade Agreement, August 5, 2004, at www.ustr.gov. (same); Article 6(1), Agreement Between the United States of America and the Hashemite Kingdom of Jordan on the Establishment of a Free Trade Area, December 17, 2001, at www.ustr.gov. (same).

the level of compliance that its peers have shown to be possible. Second, comparative Indicators intrinsically reflect an international standard, in the sense that actual international practice sets the bar. With such an Indicator, the U.S. government is not simply positing a quantitative metric, however reasonable it may be to posit a quantitative metric on many other matters. And, third, comparative Indicators create incentives for the improved longitudinal performance explicitly required by Congress. That is, Indicators that induce trading partners to continuously improve their compliance, in order to be leaders rather than laggards in their peer group, reasonably implement Congressional intent.

As noted in the previous Part, the NAS did not engage in conceptual analysis of longitudinal or comparative indicators, since its methodology directs analysts to score the country's level of compliance problems at one point in time. For the same reason, the NAS methodology is not well suited to building outcome measures into the definition of inputs, to the extent that such definitions are longitudinal (comparing improvements in outcomes over time to increases in enforcement inputs over time).

It is true that the 3 by 3 matrix also asks for the analyst's evaluation of whether the country is showing improvement, by scoring the "problems" for each Indicator as "improving," "steady state," or "worsening." This is, in principle, a longitudinal measure. However, as the Michigan evaluation stated, "Criteria for assessing direction of change are not discussed [in the NAS *Reports*]." (Root and Verloren, 2009, p. 13). Hence, this implicit longitudinal assessment is not guided by any specific benchmarks or time frames. To the contrary, an analyst is asked only the "direction" of change at the moment of assessment. The concept of

“direction of change” at a particular moment is not the same as “actual change over a period of time” – although, it is true, an analyst could not reach a judgment about the direction of change without considering actual change over *some* time frame, however short or unspecified. The fact that in the pilot testing analysts scored the direction of change as steady state 81.6 percent of the time (contrasting with the wide variance in coding for absolute level of “problems”) is perhaps an indication that the analysts implicitly applied a short time frame, in which actual change might be indiscernible. Or, as the Michigan evaluation put it, the score of “steady state” may serve as a “default category” when analysts lack clear indications of the direction of change at the time of assessment. (Root and Verloren, 2009, p. 13).

An assessment using the NAS matrix could, in principle, compare a country’s current absolute score on a given Indicator with the country’s absolute score from a preceding round of assessment. However, in light of the intrinsic ambiguity in scoring the three gradated tiers of “problems”, as shown by the high variance across different analysts’ assessments in the Michigan pilot test, we might anticipate high levels of variance and inaccuracy in the analysts’ conclusions about the direction of change, especially if different analysts score the absolute levels from one assessment to another. More important, perhaps, if static performance is scored in only three tiers, any analysis of a country’s longitudinal performance (at least as to a single Indicator) will probably not provide sufficiently refined quantitative information to show (either in raw data or in multivariate analysis) the marginal impact of additional units of enforcement effort. Nor will the measures be sufficiently refined for convincing comparative

(cross-country) analysis.

To determine the degree to which current enforcement efforts reduce violations, we need Indicators that direct us either to rich comparative data or to rich longitudinal data. The problem of data availability is addressed in Part 9 below. But as a purely conceptual matter (i.e., bracketing problems of data availability), Indicators that are framed in comparative or longitudinal terms resolve (or at least mitigate) the problem of determining counterfactual conditions based on static, single-country assessments. Using comparative data (and perhaps multivariate models), we need not conjecture what level of violations would obtain in the absence of current enforcement efforts. We instead have the benchmark of actual violations in other countries that are, except for enforcement levels, similarly situated to the target country. Using longitudinal data (and perhaps multivariate models), we need not conjecture about the marginal gain in actual compliance that results from additional units of enforcement. We instead have the benchmark of the target country's actual level of compliance prior to the exertion of additional enforcement activities. These points will be obvious to practitioners of multivariate modeling. The key point here is that they direct us to a different conceptual foundation than that used by the NAS methodology. And they resolve some of the key conceptual questions raised in the previous Part.

The use of longitudinal and comparative indicators, however, raises a new methodological question: Should comparative and/or longitudinal dimensions be incorporated in the formulation and scoring of the Indicators themselves? Or instead should such dimensions be incorporated in the analysis of the Indicators subsequent to formulation

and application?

For example, an Indicator might be framed comparatively: “Is the real minimum wage greater than the average real minimum wage among countries in the same quintile of real income per capita?” The alternative strategy would be to apply an Indicator that asks: “What is the real minimum wage?,” and subsequently assess the indicator relative to the average real minimum wage of other countries in the same quintile of real income per capita.

The refined Indicators in Appendices A, B, and C opt for including longitudinal and comparative elements in the Indicators themselves (when, that is, we choose to use longitudinal and comparative metrics, as distinguished from the other normative metrics listed at the start of this Part). This strategy mitigates two of the key problems with the NAS Indicators identified by the Michigan evaluation: first, defining the minimum baseline of violations that reflect effective enforcement; and, second, defining the two additional thresholds for “more extensive problems” and “severe problems.”

Many of the NAS Indicators ask simply for numerical data points – such as the number of workers in each category of workers whose rights of association are restricted by law, or the level of minimum wages. (National Research Council, 2004, pp. 113, 239). Such Indicators do not provide normative guidance to analysts. That is, the relation between such absolute numbers and the three categories of “some problems,” “more extensive problems,” and “severe problems” is indeterminate. This problem is aggravated when a single Indicator asks for several numerical answers.

True, we could define the three tiers of “problems” by specified percentage shortfalls

of the statutory minimum wage above or below the average minimum wage of countries in the same quintile of real income per capita. Analogously, we might formulate an Outcome Indicator that measures the percentage of workers who fail to receive the minimum wage; and we might define the three tiers of “problems” by specified percentages of workers who fail to receive the minimum wage. But using the NAS triadic structure in this way leads to severe complexities in drafting, weighting, and communicating Indicators. Part 7 below elaborates these problems and proposes the use of binary Indicators to mitigate the problems.

6.2. Are Longitudinal, Comparative, and Revisable Indicators Consistent with Human Rights?

When Indicators are framed longitudinally, the country’s performance is held up against the contingent standard of its own previous performance. When they are framed comparatively, the country’s performance is held up against the equally contingent standard of other countries’ performance. Does measuring a country’s performance against the yardstick of its previous performance or against the yardstick of other countries’ performance therefore contradict the idea that labor rights are universal human rights? A similar question was flagged above in sub-part 3.3, in the discussion of revisable indicators. If Indicators are subject to revision by ILAB analysts, have we abandoned the concept of absolute human rights?

An easy – perhaps too easy – way to take this contentious question off the table is by

stipulating foundational Indicators (whether standards or rules) that define the minimum protection necessary to fulfill the right in question. Comparison or revision would be permitted (indeed, encouraged) only above the baseline. This is a basic concept of domestic constitutional and international legal orders. In the United States, the federal and state governments are free to redefine rights (such as rights of speech, association, or nondiscrimination), so long as the revised right exceeds the federal constitutional baseline and does not trench on some other right. More to the point, many ILO Conventions explicitly provide that the Convention's stipulation of rights must not be taken as a ceiling and does not roll back more expansive protections that may obtain at the national or international level.

The familiar objection that a stipulated floor is likely in practice to become a ceiling does not carry force. The same objection can be leveled against any strategy for defining rights. That is, any strategy for defining a right stipulates a minimum level of permissible treatment that has the potential to become a ceiling in practice; and collateral interpretive rules or standards must make clear that the rights should be viewed as minima but not as ceilings. For example, adding a non-derogation Indicator would ensure that revisability is a one-way ratchet. Greater enforcement and more expansive interpretation of the right are permitted. Retraction or constriction, even above the minimum, is not.

There are two challenges that this solution must face. First, does the strategy of specifying a minimum baseline of each right, as a means of maintaining universality and cross-country uniformity, run into a conceptual obstacle when we recognize that U.S. trade legislation and treaties require governments (merely) to take steps to comply with worker

rights? Does that longitudinal requirement, in effect, make the received conception of absolute, universal rights a mere formal conception? That is, if rights are what governments are obligated to provide their citizens, then does the obligation that the government merely take adequate steps to protect a given interest of workers become the substantive, even if not formal, specification of the workers' entitlement?

One might respond to these questions in the following way: By taking account of whether a trading partner is “merely” taking steps toward compliance, we are not necessarily watering down the right-defined-as-minimal-absolute-baseline itself. Those trading partners that have ratified relevant international Conventions may still be obligated to fully satisfy those human rights even if the U.S. government, pursuant to trade legislation and treaties, does not impose sanctions or withdraw special trade benefits when the country is merely making progress toward satisfying those rights. Moreover, the U.S. government might – acting outside the labor rights provisions of the trade legislation and agreements discussed in this paper – deny *additional, discretionary* economic benefits (carrots rather than sticks) to a country that has not achieved full compliance with the substantive right. The U.S. government might then incrementally increase the economic benefits as the trading partner achieves incremental benchmarks (by “taking steps”) toward full compliance, and provide full benefits only when the trading partner achieves full compliance.

This counter-argument has substantial force – to the extent that (a) Congress authorizes such discretionary economic benefits and (b) a given right can in fact be defined in absolute, universal terms, so that “taking steps” is defined by a fixed endpoint. However,

the previous sub-part made the case for framing many Indicators in comparative or longitudinal terms, where international or domestic labor law does not provide a compelling normative “end point.” To the extent that adequate “steps” are determined by a trading partners’ performance relative to other countries in a relevant cluster, then a country’s obligation is not to take steps until achieving an absolute level, but to *continually* take steps to achieve or exceed an ever-improving standard – to continuously compete in the proverbial race to the top.

Note that this comparative strategy does not mean that a standard such as “acceptable conditions with respect to minimum wages” requires that minimum wages will or should continually rise at a maximum pace that is inconsistent with absolute levels of productivity, increases in productivity, tradeoffs between wages and employment that optimize welfare, and other empirical and normative considerations. The question of what constitutes the “best” comparative performance need not be mechanically determined exclusively or at all by the simplest metrics – such as “highest real minimum wage” or “greatest increase in real minimum wage.” Put somewhat differently, what constitutes “improvement” from a given starting point is not uncontroversial. Whether the *direction* and *magnitude* of a particular change in government performance is an optimal “improvement” (compared to the performance of other similarly situated countries) will in some instances require complex analysis. Such analysis will therefore include both identification of the relevant variables that define whether the given country is in fact similarly situated to other countries that appear to be better performing *and* the normative grounds for concluding that

the direction and magnitude of change are in fact improvements.⁵²

But, second, this very rationale for conceptualizing indicators as revisable takes us back to the problem with which we started: It seems to place great stress on the received notion that “internationally recognized labor rights” are universal. If a worker is entitled to the most capacious (or optimal) definition and enforcement of a given right that can be achieved by countries similarly situated with the worker’s own, then there are as many different specifications of that right as there are clusters of similarly situated countries, and those specifications are themselves provisional (contingent on individual countries’ and clusters of countries’ constricted or enlarged capacity to provide improved definitions and enforcement). That is, it appears we are no longer holding the country to a universal standard or an absolute outcome, but instead demanding that the country “do its best” to improve conditions within the constraints of its initial baseline, its level of development, its government capacity, or other contextual constraints.⁵³

Put slightly differently, each worker in the country is no longer entitled to fully secure her personal interests as a rights-bearing individual. Instead, her “entitlement” is to work in an economy regulated by a political system that is continuously improving labor-market and workplace outcomes – or improving those outcomes at a faster pace than other, similarly situated countries – even if her personal interests are not always secured by such statistical

⁵² Whether it is desirable to undertake this analysis when first formulating the body of Indicators or instead in the process of applying the Indicators is a matter discussed below in Part 8.

⁵³ Whether these enumerated constraints should in fact be taken as hard or soft constraints is discussed below in Part 8.

improvement. In the traditional concept of rights each worker who is, for example, fired in retaliation for union activities is entitled to reinstatement, back pay, and perhaps other remedies. In the concept of comparative systemic improvement, even if such remedies are not received by every wrongfully discharged worker, their “rights” might nonetheless be deemed vindicated because the system is maximally or optimally improving its capacity to provide system-wide deterrence and remediation of such retaliatory discharges.

Perhaps we can have it both ways: Individuals who file complaints under trade statutes or trade agreements are entitled to absolute vindication of their personal rights; at the same time as the government is obligated to maximize or optimize its continuous improvement of *ex ante* deterrence and *ex post* remediation on an economy-wide scale.

This solution is not perfect, however, even if it is the best we can do. Only those individuals fortunate enough to file complaints under the trade legislation or treaty gain actual personal vindication of the absolute right. All others whose rights are violated are left to the mercies of systemically flawed but “maximally improving” or “optimally improving” enforcement institutions. And there is no guarantee that workers whose rights are violated will in fact be able to vindicate their rights even by filing complaints under U.S. legislation or treaties, since the domestic enforcement mechanism by definition has imperfect, albeit improving, administrative and judicial procedures and remedies. The government, after all, is “taking steps” to afford worker rights.

In any event, it is once again critical to remember that, while the Executive Branch may be well advised to conform its definition of “internationally recognized worker rights” to

some appealing conception of “universal human rights,” it need not do so. Again, under U.S. law, the Executive Branch can give that term any specification that reasonably implements Congressional intent, which is not tied exclusively to international human rights.

These considerations show three things. First, at a minimum, it is important for the U.S. government to demand systemic remedies when individuals prove violations of their rights. That is, the U.S. should require the trading partner not only to correct the violations of the petitioners’ rights but also to take measures that will correct or prevent similar violations against other workers. Second, it is critical to give capacious meaning to the term “taking steps” – in the sense that the U.S. government should not be too quick to take limitations on a foreign government’s capacity to “take steps” as hard constraints that thwart improvement rather than soft constraints that the government can and should be expected to overcome. Parts 8 and 12 below elaborate this point.

Third, in light of the conceptual challenges to revisable, longitudinal, and comparative norms just mooted, when we construct individual Indicators we must take care to identify the subject domains for which such norms are well-suited and the subject domains for which such norms are not well-suited. Sub-part 3.4 carried out that exercise with respect to revisable norms. This Part has done the same for comparative and longitudinal norms. And the following sub-part does the same for Indicators that use the quantitative metric of “a non-trivial number of violations.”

6.3. The Case for the Normative Metric of “Non-Trivial Number of Violations”

Many of the revised *non-comparative, non-longitudinal* Indicators for Enforcement set out in Appendices A, B, and C take the following form: “In the preceding two years, has the government ensured that in all cases in which workers and worker organizations alleged violation of their rights of association, *all non-trivial proceedings* of the judicial or administrative tribunal hearing such cases were open to the public?” or “In the preceding two years, has the government ensured that when a case (in which workers or worker organizations alleged violation of their rights of association) reached final decision on the merits, the labor tribunal issued public, written decisions, in *all but a trivial number* of cases?”

The logic of these examples is that while in some subject areas, such as investigation of the murder of trade unionists, we can define “effective enforcement” or “taking steps” based on an absolute baseline of zero violations, in other subject areas that stringent baseline would not be a reasonable. Even the most worker-protective governments cannot ensure, for example, that every judge publicly issues all final decisions in writing. Hence, in many binary Indicators that rest on putatively absolute baselines (rather than on comparative or longitudinal baselines), some normative threshold other than zero violations must be specified.

It is true, of course, that the concept of a “trivial” numbers of cases is not crystal clear. But it has significant advantages over indicator methodologies proposed by certain ILO staff, which explicitly “treat[] one [anti-union] dismissal the same as a thousand.” (Kucera, 2007, p. 163, 165.) More relevant for our purposes, it has two advantages over the NAS concepts of

“some problems,” “more extensive problems,” and “severe problems.” First, it is clearly numerical. Recall that, in contrast, each of those three NAS concepts explicitly mixes three different dimensions: numerical frequency, breadth in the workforce or country, and qualitative severity – and these three dimensions are to be judged in light of a fourth concept, namely, the causal impact of the first three dimensions on the degree of difficulty posed to government compliance. (See WebMILS at <http://webapps.dol.gov/webmils/matrix-assessment-tool.aspx>, visited 10-26-10.) The interaction of these four criteria does not provide a clear numerical metric.

Second, within jurisprudence “non-trivial” is a familiar term, however squishy. In legal discourse, there is an array of terms that are familiar to the craft sense of judges, administrators, legislators, and other legal professionals – terms such as “beyond a reasonable doubt,” “clear and convincing evidence,” “substantial harm,” and so on. The concepts of “some problems,” “more extensive problems,” and “severe problems” do not have common usage. The concept of “non-trivial instances” does.

More important, determinations made by analysts applying the concept of “trivial instances” are likelier to fall within a more constrained range than determinations about the concepts of “more extensive” and “severe,” since “trivial” spans the narrower range expressed by a concept that connotes “extremely small” – as opposed to the concepts of “some,” “extensive,” and “severe,” which connote a broad numerical range.

Finally, and perhaps most important, the relevant statutory and treaty language itself guides us to the concept of “non-trivial” violations. The pertinent language of the U.S.-Peru

free trade agreement – the current template for labor rights provisions – prohibits the signatory states from failing to effectively enforce labor rights by a “recurring course of action or inaction.”⁵⁴ Earlier bilateral and regional trade agreements used the same language.⁵⁵ The germinal agreement using such language is the North American Agreement on Labor Cooperation (NAALC), which explicitly defined a “pattern of practice” as a course of action that “does not include a single instance or case.”⁵⁶ Indeed, the central obligation of the NAALC is that governments not fail to effectively enforce labor laws, unqualified by the requirement that they avoid a pattern of practice of violations.⁵⁷ The NAALC introduces the concept of a “pattern of practice” of non-enforcement of labor standards only as the standard of evaluation in reports of the Evaluation Committee of Experts.⁵⁸

The relevant treaty language therefore clearly indicates that “ineffective enforcement” is established by a non-trivial number of violations: “recurring,” in the sense of “[more than] a single instance or case.” In their iterative application of Indicators, ILAB analysts might confer and simply agree on a specific, small numeric threshold for “non-trivial,” which they may choose to vary from one Indicator to another, as appropriate to the substance of each

⁵⁴ Article 17.3(1)(a), United States-Peru Trade Promotion Agreement, April 12, 2006, at <http://www.ustr.gov>.

⁵⁵ See, e.g., Article 16.2(1)(a), Dominican Republic-Central America-United States Free Trade Agreement, August 5, 2004 (CAFTA-DR), at <http://www.ustr.gov>; Article 6(4)(a), United States-Jordan Free Trade Agreement, December 17, 2001, at <http://www.ustr.gov>; Article 16.2(1)(a), United States-Morocco Free Trade Agreement, January 1, 2006, at <http://www.ustr.gov>; Article 18.2(1)(a), United States-Chile Free Trade Agreement, January 1, 2004, at <http://www.ustr.gov>; Article 16.2(1)(a), United States-Oman Free Trade Agreement, January 1, 2009, at <http://www.ustr.gov>.

⁵⁶ Article 49(1), North American Agreement on Labor Cooperation, United States-Canada-Mexico, September 14, 1993, at <http://www.ustr.gov>.

⁵⁷ Article 3(1), North American Agreement on Labor Cooperation, United States-Canada-Mexico, September 14, 1993, at <http://www.ustr.gov>.

⁵⁸ Article 23(2), North American Agreement on Labor Cooperation, United States-Canada-Mexico, September 14, 1993, at <http://www.ustr.gov>.

Indicator.

7. Binary Indicators



This Part proposes that all of ILAB’s Probative, Diagnostic, and Assessment Indicators take binary form. That is, each Indicator should be scored yes/no, rather than gradated. Why are binary Indicators preferable to the 9-cell matrix of the NAS methodology? The primary answer is simplicity – simplicity in formulating, applying, and communicating Indicators.

At its simplest, a comparative, bright-line minimum wage Indicator might ask for comparison of the country’s real minimum wage with the average among countries in the same decile or quintile of real income per capita. Going a step further and formulating the comparative minimum wage Indicator as a “percentage of average wages,” we might ask: “Does the ratio of the minimum wage to average wages exceed the average ratio among countries in the same quintile of income per capita?”⁵⁹

Note that gradations can still be captured in a body of binary Indicators. For example, the body of Indicators could include both the binary Indicator just mentioned (“does the

⁵⁹ Note that the two Indicators just set out are Substantive Law Indicators taking the form of bright-line rules. The rules might be fixed; or they might instead be conceived as revisable. For example, we might expect that ILAB analysts may place a higher burden on governments over time, such as measuring their performance against 120 percent of the average performance of their peers, rather than 100 percent of the average performance. This higher measure is discussed in the next paragraph of the text.

country's ratio exceed the average ratio among countries in the same quintile?") and a second binary Indicator asking "does the country's ratio exceed 120 percent of the average ratio among countries in the same quintile?" The latter Indicator might seem to inject a relatively arbitrary normative judgment: Why 120 percent and not some other percentage? But this merely uncovers the fact that comparing one country with the *average* among other countries in the same quintile itself rests on an "arbitrary" normative baseline – namely, the baseline of 100 percent of the average, rather than 120 percent of the average or some other percentage.

This reveals that the real difference between the binary Indicators formulated in this paper and the triadic Indicators formulated by NAS is twofold: (1) the binary Indicator is simpler for the analyst to apply, and (2) the Indicators formulated in this paper rest, wherever possible, on bright-line, numerical distinctions rather than on undefined thresholds of "more extensive" and "serious."

As the earlier examples show, binary Indicators can rest concurrently on longitudinal and comparative yardsticks. To carry forward the current example of minimum wages, an Indicator might ask: "In the five preceding years, has the rate of increase in the ratio of the minimum wage to average wages of non-managerial, non-supervisory workers exceeded the average rate of increase among countries in the same quintile of real income per capita (with decreases taken as negative increases)?"

This Indicator might appear complex, but notice that it requires the availability of data

only on the statutory minimum wage level and average wages in each of the last five years.⁶⁰ (In fact, ILO Conventions require governments to collect data on average wages.⁶¹) Yet because the Indicator is framed as a binary, longitudinal, comparative Indicator, it yields important information for comparing the effectiveness of country law and for evaluating each country's progress ("taking steps"). True, identifying a single country-wide minimum wage may not be as easy as just intimated. In some countries, there is a readily identifiable minimum wage for the entire country. In others, the analyst may need to pro-rate multiple minimum wages fixed in different sectors or regions to arrive at a country-wide minimum wage; and similar problems may arise with the average wage. In the latter countries (with multiple minimum wages), analysts must in any event do this calculation in order to undertake *any* meaningful assessment of minimum and average wages. Also true, as discussed above in sub-part 6.2, we may not consider a continuous increase in minimum wages relative to average wages a valid measure for optimal enforcement of worker rights, in light of the potential tradeoff between minimum wages and employment. The Indicator above is merely an illustration – for purposes of the present conceptual discussion – of a binary Indicator that is both comparative and longitudinal.

Granted, if we stick with the 3 by 3 matrix, a single Indicator might ask: "Does the statutory minimum wage constitute 'some problems' because it is more than 120 percent of the average minimum wage among countries in the same quintile, or 'more extensive

⁶⁰ Note that these features follow in part from the fact that the indicator is a static bright-line indicator, in the four-fold categorization of Part 3.4 above. It is, orthogonally, a binary, longitudinal, comparative indicator.

⁶¹ Article 9, ILO Convention no. 160 on Labor Statistics, 1985.

problems’ because it is between 80 and 120 percent of the average among countries in the same quintile, or ‘severe problems’ because it less than 80 percent of the average among countries in the same quintile?”

Nonetheless, the binary form has the advantage of simplicity in three senses. First, formulations like the preceding example will often become linguistically convoluted to an unacceptable degree, when we explicitly incorporate longitudinal and comparative elements together with the three orthogonal gradations of “problems” – and even more convoluted when we incorporate other desirable elements discussed in the next sub-part. Second, when the triadic structure is incorporated in the Indicator, the weights attached to various laws and government actions will often become more occluded to the audience than are the weights attached to binary Indicators. To continue our example, the decision to create thresholds at 120 percent of the average and 80 percent of the average is a mode of relative weighting of compliance problems. But such weighting is less visible when those thresholds are embedded in the text of a single Indicator. Moreover, that Indicator may itself be weighted differentially from other Indicators, requiring the audience to comprehend the resulting compounding of normative thresholds.

And third, most matters of substantive law and many matters of enforcement lend themselves “naturally” to binary questions, as the analysts in the Michigan pilot-testing discovered. A typical Substantive Law Indicator requires the analyst to determine, for example, whether a government has a law prohibiting retaliation against union supporters or does not. Similarly, many Enforcement Indicators require the analyst to determine, for

example, whether the government has a mechanism to fix minimum wages or does not, or whether the labor inspectorate covers employment discrimination issues or not. Even Outcome Indicators can be relatively easily framed in binary form. For example: “Is the percentage of employees in the informal sector who are female no greater than the percentage of employees in the formal sector who are female?”

This is not to say that by converting Indicators to the binary form we remove normative choice in the formulation of the Indicators, or that we squeeze all discretion out of the analyst’s application of Indicators. Clearly not. As just noted, regardless whether we compare a country’s performance to the “average” of other countries or to “X percent of the average,” we are making a normative judgment. And, as discussed in sub-part 6.3, the baselines of “zero violations” and “non-trivial violations” reflect normative judgments, even though they mitigate the discretion of analysts, compared to the broader numerical ranges of “more extensive problems” and “severe problems.” The bottom-line conceptual point is this: The bright-line, numerical definition of Indicators reduces normative subjectivity in the sense that the analyst’s subjective discretion is diminished, compared to the NAS triadic structure.

8. Control Variables: Hard and Soft Constraints



The labor rights compliance of countries may depend on many country-specific contextual variables. A critical question, then, is whether the evaluation of each country should be “adjusted” – that is, whether a country’s score on each Indicator or on the composite index – should be upgraded or downgraded based on such contextual variables. Candidates for such control variables include: level of economic development, rate of economic growth, type of political regime, type of labor relations system, government capacity, colonial history, religion, legal origin (e.g., civil vs. common law), and many others.

The explicitly comparative and longitudinal Indicators proposed in this paper almost always ask for a comparison between the specified legal norm or data point of the target country and the average norm or data point among countries in the same quintile of real income per capita.

Hence, “quintile of real income per capita” is the one control variable built into the Indicators. (Note that that control variable is built into only the comparative Indicators. There is no control variable in any of the other Indicators, which rest on non-comparative normative metrics.) Why use that control variable? The level of income per capita is relevant, in light of the fact that United States trade statutes mandate either that the President shall not take action against a trading partner if acts that deny internationally recognized worker rights are

“not inconsistent with the level of economic development of the foreign country” or with the country’s “living standards” or that he consider that factor in granting eligibility for special trade benefits.⁶² True, the quintile of real income per capita is a crude proxy for level of development or standard of living. But that variable is just a starting point in the first iteration of the Indicators, and can be refined by ILAB analysts as they successively apply and revise the Indicators in the manner proposed in sub-part 3.3 and Part 12 of this paper.

Why not build one or more of the other potential control variables enumerated above into the text of particular Indicators, or into the scheme for weighting and aggregating Indicators? There are good reasons for not presuming that such variables are hard constraints.

First, what might seem *a priori* to be a hard constraint that excuses poor compliance

⁶² 19 U.S.C. § 2411(d)(3)(C)(i)(II) (Section 301 of the Trade Act of 1974, as amended). Unlike Section 301 of the Trade Act, *id.*, the Generalized System of Preferences is ambiguous on this matter. It first mandates that “the President shall not designate any country as a beneficiary developing country” if “[s]uch country has not taken or is not taking steps to afford internationally recognized worker rights to workers in the country....” 19 U.S.C. § 2462(b)(2)(G). It then, however, states that in determining whether to designate any country as a beneficiary developing country, the President “shall take into account” a list of seven factors, including “the level of economic development of such country, including its per capita gross national product, the living standards of its inhabitants, and any other economic factors which the President deems appropriate” and “whether or not such country has taken or is taking steps to afford its workers in that country...internationally recognized worker rights.” 19 U.S.C. §§ 2462(c)(2), (7). The Act also states that the President is authorized to take into account those seven factors as well as other considerations enumerated in § 2462 when designating a country as a “least developed beneficiary country;” this language seems to impose the very same preconditions as does the definition of “developing beneficiary country.” 19 U.S.C. § 2462(a)(2). In order to avoid internal inconsistency, the best reading of this language is that the President shall take account of the country’s labor-rights record and shall not designate the country as a beneficiary country if it has not taken or is not taking steps to afford internationally recognized labor rights. Under that reading, the “level of economic development” is not a factor that can override the determination of whether the country has or is taking steps to afford worker rights. Instead, if a country is complying with internationally recognized worker rights, the level of economic development may act as an independent ground for the President to decline to designate a country as a beneficiary country. The same statutory analysis applies to regional preference programs such as the Caribbean Basin Economic Recovery Act. See 19 U.S.C. §§ 2702, 2467.

performance may in fact be a variable that is susceptible to mitigation by feasible government policy instruments. Indeed, some of those variables may be within the government's control *and* may be root causes of various forms of compliance failure. For example, a government's weak "capacity" to enforce rights may be the key problem to be solved, rather than an excuse for poor compliance. Therefore, the various candidates for control variables should not be treated as hard constraints until the data gathered through repeated ILAB assessments show they are not in fact soft constraints.

Second, and closely related, the current data are too thin and existing multivariate models are too crude to provide statistically significant conclusions about the causal impact of the candidate control variables.⁶³ Again, conclusions about treating any particular variables as hard constraints should await ILAB's medium- or long-term collection and modeling of relevant data. Indeed, this is another reason for building high-powered incentives for data production into the Indicator methodology itself.

For example, the "Southern European" model is a regime type that is purportedly resistant to women's participation in the workplace; an Indicator of the regime type might be taken, *a priori*, to excuse weak performance on gender equality. In the EU-15, Greece, Italy, and Spain have the highest percentage of couples in which the man works full-time outside the home and the woman does not participate in the paid labor market. Yet we see that Portugal, which is otherwise similar in key respects to its Southern European peer countries, has an unusually high degree of female labor-force participation. And, we see that Portugal's

⁶³ This point is elaborated in sub-part 11.4 below.

better performance in this dimension is based on contingent historical events – showing that the sexual division of labor is not hard-wired into that model. In the 1960s, the Portuguese male workforce was depleted by that country’s colonial wars, bringing women into the workforce and deflecting the country from its apparent “Southern European” fate. (Karamessini, 2008).

Looking backward from the present, one might be tempted to say that Portugal has been locked into a different path-dependent course than Greece, Spain, and Italy. But precisely the opposite may be true. This is a case not of “path-dependency” (in which the identified variable is genuinely resistant to political change), but rather of “historical contingency” (in which a seemingly resistant variable may be quickly transformed by the right conjuncture of political or economic events).

Even more apropos, a Cambridge University research team has shown convincingly that countries are not locked into particular labor law rules based on their colonial origins or on the historical or categorical type of their legal system. (Deakin and Reberieux, 2009; Deakin and Sarkar, 2008; Deakin, et al., 2007b).⁶⁴

At a minimum, these examples and innumerable others like them cast doubt on the assumption that “regime types” are marked by “institutional complementarities” that excuse poor performance in certain dimensions of labor rights compliance on the ground that one institution or practice cannot be changed without concurrently changing a host of entrenched collateral institutions or practices. Social scientists and comparative labor relations scholars

⁶⁴ For further discussion, see this project’s *Literature Review and Bibliography* at pp. 106-137.

have recently launched powerful theoretical and empirical critiques of the theory of institutional complementarities, citing many cases analogous to the Portuguese case above. (See, e.g., Crouch, 2005; Hyman, 2009.)

9. Data Problems, and Two Responses: Twin Indicators and Capacity-Building Indicators



9.1. The Problems of Unavailable, Unreliable, or Inconsistent Data

Data unavailability, inaccuracy, and inconsistency pose tremendous challenges to the formulation and application of a body of Indicators that validly measures the underlying concepts of “effective enforcement,” “taking steps,” and the various elements of the substantive rights that concern us. The primary challenge is quite obvious: If analysts cannot find information pertaining to an Indicator, they cannot answer the revised yes/no binary indicators. One equally obvious solution is to provide three answer options: (1) yes, (2) no, (3) no information. But this is not a genuine solution until we have a method for weighting the “no information” option, and conceptually convincing justifications for that method.

Researchers have deployed four main strategies for dealing with missing data. One is to simply delete indicators for which data is not broadly available. This is a candid strategy of one research organization in a provisional, unpublished draft report on formulating indicators for certain aspects of labor compliance. The document states that “[t]he amount of missing

data for an item was the central consideration” in selecting labor indicators, although “relevance of the item and reliability were also considered.” (Anonymous, 2009).

The caveat reveals one of the weaknesses in the strategy. In deciding whether to erase an indicator, the unavailability of data was given more weight than the relevance of the indicator. This may be less problematic in policy fields in which composite indices aim to measure government performance or country conditions not immediately tied to compliance with legal rights – such as environmental quality, public health, corruption, or ease of doing business. But the strategy of deleting relevant items is problematic when measuring compliance with rights. If data are hard to find on the murder of trade unionists or workplace exposure to the worst toxins, it is not proper for the President or USTR to simply throw up their hands (or, to mix metaphors, to plunk their fingers on the delete button). Legislation and treaties are legal documents that the USTR and the President are constitutionally obligated to enforce.⁶⁵ Setting aside significant elements of the multidimensional rights enumerated in those instruments because good data is unavailable is an unsatisfying solution. This is particularly true in light of the fact that governments that are willing to cause or tolerate severe abuses of worker rights may well see little interest in collecting and publishing data on the point.

A second strategy is to keep the indicator but treat missing data as neither a positive nor negative score. If “yes” is scored as “1” and “no” is scored as “-1,” then “no information” is scored as “0.” This is sometimes described as treating the absence of data “neutral[ly]” or

⁶⁵ Those legal instruments *mandate* that the President decide whether the rights are or are not violated or whether steps are being taken to comply with those rights.

passing “no judgment” on the missing data. (Id., 2009). It amounts, however, to deleting the indicator for that particular country – again, as though one dimension of a legal right is irrelevant in assessing that country.

A third strategy is to treat the absence of data equivalently with a violation of the indicator. If “yes” is scored as “1,” then both “no” and “no information” are scored as “-1.” Some have suggested that this is unduly harsh, since the lack of information may reflect difficulties for researchers in finding a country’s extant legal rules rather than a failure on the part of the government. (Id., 2009). This is not entirely convincing. Finding a country’s legal rules is not difficult. That argument has somewhat more force, though, if applied not to Substantive Law but to measures of Enforcement, Capacity-Building, and Outcomes, for which existing data may be difficult to uncover. But even on that score, if data exist we may want to place the burden on the trading partner’s government to come forward with them.

A fourth strategy, or set of strategies, is to treat “no information” as a negative score, but not as strongly negative as the score for an answer of “no.” If “yes” gets “1” and no gets “-1,” then “no information” might get “-.5,” or alternatively the score for “no information” may be variable depending on indicator-by-indicator variations in the positive “yes” scores and negative “no” scores generated by principle component analysis. (Id., 2009). The latter strategy is conceptually unsatisfying, in that principal component analysis effectively weights indicators based not on the importance of the feature of the right measured by a particular indicator, but instead on the indicator’s contribution to variance in the composite indicator.⁶⁶

⁶⁶ This problem is discussed further in sub-part 11.3 below.

That strategy does not delete rights altogether, nor does it delete the fact that data pertaining to that right is absent, but it does depreciate a legal entitlement for reasons other than the intrinsic importance of the right.

All of these strategies are efforts to make the best of a bad situation. A better strategy is to formulate Indicators to make the bad situation better. The four strategies just mentioned are backward-looking, in the sense that they assume that the job of the analyst is to evaluate past performance without using the assessment itself as an occasion to change incentives for improving the government's future performance, including most crucially future performance in producing information. It is true that the third and fourth strategies just listed will create an incentive for governments to improve their data collection, even if that is not the primary intention of the negative score for missing data. But it is better to create this incentive in a conceptually deliberate way. The following sub-parts propose two such means.

9.2. Twin Indicators

Rather than modulating the scoring of a substantive Indicator based on the analyst's inability to find information about past performance, we should construct Indicators that make evident to trading partners and advocates that the assessment process is forward-looking. The guiding conceptual framework should not be based on an image of the target government as a flickering, evasive astral object whose ancient light is viewed through a foggy telescope wielded by the U.S. analyst. The conceptual design should instead rest on a predictably interactive, iterative process of applying Indicators, inciting data-collection and

improved future performance, and (potentially) revising the Indicators in light of the information revealed.⁶⁷

The first element of such a conceptual design is to formulate Twin Indicators. For each pair of twins, the first Indicator will take the familiar form (akin to the NAS Indicators) of a substantive question to be answered by the ILAB analyst's research. The analyst will use whatever sources are available to the analyst within the constraints of ILAB time and resources. In each round of assessment, ILAB analysts will successively build up their cache of sources appropriate to each country. These country-specific sources will – in the drop-down window for each Indicator – be recorded alongside the increasingly detailed country-specific sub-indicators that also appear in that drop-down window, as discussed above in sub-part 3.2 and illustrated in Figure 3 in that sub-part.

The second Indicator essentially repeats the first, except it asks whether the government “has convincingly and verifiably demonstrated” that the answer to the first Indicator is “yes.” In other words, the burden of proof in answering the first Indicator rests with ILAB analysts, using source material that is independent of the trading partner's government as well as data produced by that government. The burden of proof in answering the second Indicator rests with the trading partner's government. Ideally, that government would “convincingly and verifiably demonstrate” its compliance with the substantive Indicator

⁶⁷ In the most expansive version of the iterative process – discussed in Part 12 on iterative dialogue in the probative, diagnostic, and assessment phases – trading partners would offer not only new data in each round of appraisal but also justifications for their failure to satisfy U.S. Indicators or to match the performance of peer governments. Those proffered justifications might be rejected by U.S. analysts, or might be accepted. If accepted, then analysts might revise country-specific sub-indicators to take account of the (inevitably) contextual justifications.

by collecting data that substantially matches the definitions in the Indicator – or by directing the U.S. analyst to convincing and verifiable independently produced information. For this reason, the refined U.S. Indicators refer to definitions promulgated by the ILO and other internationally authoritative statistical agencies – to avoid imposing on trading partners the additional burden of applying a set of definitions peculiar to the U.S. Indicators.

Alternatively, the target government might satisfy the second Indicator by demonstrating that it has put into place sufficiently sound protocols for collecting data and setting targets on the subject matter covered by the Indicator, for transparently assessing its own success in achieving those targets, and for including relevant stakeholders in each step of these processes.⁶⁸

The obvious purposes of the second Twin Indicator are to (1) “flush out” data and information that ILAB analysts might not otherwise find in their independent research, and (2) create incentives for the target government to produce better data or install more rigorous protocols in the future. This template puts great pressure on the trading partner’s government to produce and come forward with reliable data. If it does, it will score positively on each of the Twin Indicators. If it does not, it risks getting a negative score for each of the Twin Indicators – although it remains possible that ILAB analysts will find independent data

⁶⁸ It is true that these features of government performance are also measured by the new Capacity-Building Indicators discussed in the next subsection. The Twin Indicators, however, focus on the government’s application of those features to specific substantive norms, enforcement activities, and outcomes, while the Capacity-Building Indicators focus predominantly on the government’s more general architecture for collecting data, setting targets, evaluating its success in meeting those targets, and including stakeholders in those processes. As a strict conceptual matter, then, the Capacity-Building Indicators do not subsume the Twin Indicators. The latter measure the independent question whether the government’s general capacity-building machinery is being *applied* to the specific norms, activities, and outcomes that define the key elements of the rights that concern us.

that award a positive score for the first twin, even though the government fails the second Indicator by not itself producing convincing and reliable evidence on the matter.

Twin Indicators are most appropriate for Enforcement Indicators, Capacity-Building Indicators, and Outcome Indicators, where the relevant existing data may not be easy for ILAB researchers to obtain. Such data are likelier to be within the possession of the government being assessed or comparatively easy for that government to obtain or produce. Twin Indicators are not vital for Substantive Law Indicators, which can be answered by the ILAB analyst reading the government's laws-on-the-books or communicating with a country legal expert.

By way of example, below are two sets of twins. The first pair is on a matter of Enforcement. The second pair is on a matter of Outcomes.

[twin 1] In the preceding two years, did the competent body⁶⁹ for filing complaints in fact file a complaint (after a finding by the labor inspectorate of any violation of workers' freedom of association, rights to organize, or rights to bargain collectively) in all but a trivial number of cases received by that body upon referral by the inspectorate?

[twin 2] Has the government convincingly and verifiably demonstrated that, in the preceding two years, the competent body for filing complaints in fact filed a complaint (after a finding by the labor inspectorate of any violation of workers' freedom of association, rights to organize, or rights to bargain collectively) in all but a trivial number of cases received by that body upon referral by the inspectorate?

* * *

[twin 1] Does the ratio of the secondary school enrollment rate of females to the

⁶⁹ Note that in the actual body of Probative, Diagnostic, and Assessment Indicators proposed in this paper, there are Annotations explaining individual Indicators. Hence, there is an Annotation explaining what is meant by "the competent body for filing complaints...upon referral by the labor inspectorate."

secondary school enrollment rate of males exceed the average ratio among countries in the same quintile of real income per capita?

- [twin 2] Has the government convincingly and verifiably demonstrated that the ratio of the secondary school enrollment rate of females to the secondary school enrollment rate of males exceeds the average ratio among countries in the same quintile of real income per capita?

9.3. Capacity-Building Indicators

The new category of Capacity-Building Indicators, proposed above in Part 4, includes Indicators on the government's capacity to collect, analyze and publish accurate data that comprehensively covers the matters addressed in the bodies of Probative, Diagnostic, and Assessment Indicators.⁷⁰

Both Twin Indicators and Capacity-Building Indicators therefore generate incentives for trading partners to collect, analyze, and publish data. The distinction between the two, however, is clear: Twin Indicators measure the trading partner's collection and publication of data as to the *specific substantive matters* contained in the pertinent Indicator. Capacity-Building Indicators, in contrast, measure the trading partner's construction and maintenance of the *systemic architecture* for collecting and publishing accurate, comprehensive data relevant to the body of Indicators and other policy purposes.

The potential interaction between Twin Indicators and Capacity-Building Indicators is discussed in the next Part.

⁷⁰ As mentioned above, the Indicators refer, wherever possible, to international protocols for collecting data and international standards for defining variables, to avoid overburdening governments with additional data-collecting obligations that might otherwise be imposed by the U.S. Indicators.

10. Probative Indicators, Diagnostic Indicators, and Assessment Indicators



10.1. Why Three Bodies of Indicators?

There is a strong case for formulating three bodies of Indicators – one set of Probative Indicators, one set of Diagnostic Indicators, and one set of Assessment Indicators. Among the ILAB staff interviewed by this author, there was near-universal consensus about the need for at least one set of Indicators that would serve the purpose of “triage” or “prioritization” or “an early warning system.” Such a set of indicators – call them Probative Indicators – can point ILAB analysts to countries that are likely to show across-the-board non-compliance and non-enforcement, or at least severe non-compliance and exceptionally weak enforcement across important elements of the rights in question. They can also alert well-intentioned government officials of U.S. trading partners to areas in which their enforcement efforts need to be strengthened. And, apart from the process of determining compliance with U.S. trade legislation and agreements, Probative Indicators might help ILAB decide where to allocate resources for technical assistance and research.

The Probative Indicators are a species of diagnostic indicators; and in fact the body of Probative Indicators set out in Appendix C are effectively a “short list” of the Diagnostic Indicators set out in Appendix B. The body of Diagnostic Indicators may therefore be

applied, in a second stage of screening, to the set of countries initially identified by the body of Probative Indicators. The Diagnostic Indicators provide more intensive testing to determine those countries that warrant deeper research, technical assistance, or comprehensive evaluation.

At the same time, the nature of the labor rights provisions in trade legislation and agreements calls for a more comprehensive set of Assessment Indicators. As discussed in sub-part 1 of the Introduction to this paper, it is the constitutional obligation of the USTR and the President to determine whether our trading partners are complying with those provisions. That determination requires a bottom-line evaluation of whether a country is or is not effectively enforcing or taking steps to afford the enumerated labor rights. By gathering data and applying Assessment Indicators, ILAB can inform the judgment of USTR and the President.

10.2. What Types of Indicators are Best Suited for Probative Indicators, for Diagnostic Indicators, and for Assessment Indicators?

The question, then, is which Indicators should serve as Probative Indicators, which should serve as Diagnostic Indicators, and which should serve as Assessment Indicators?

Plainly, the body of Assessment Indicators should encompass a comprehensive set of well-specified, consistent measures of Substantive Law, Enforcement, Capacity-Building, and Outcomes.⁷¹ The four strongest reasons for well-specified, consistent, and comprehensive

⁷¹ Recall that a “well-specified” Indicator need not take the form of a detailed, fixed bright-line rule. It might also take the form of a well-specified revisable bright-line rule; a well-specified fixed standard; or a well-specified

measures are discussed in Parts 2 and 3 above, and will be reprised here in capsule form:

First, if Indicators are well-specified and use consistent terminology, we can expect less variance in application of the Indicators both across countries and across time, especially when different analysts are charged with applying the Indicators to different countries and at different times. Second, clarity and comprehensiveness together will serve the goal of capturing the complex, contextual variables that are relevant from country to country. Those variables will fall through the gaps in a small set of ambiguous Indicators.

Third, when complex, multidimensional rights are at stake, Assessment Indicators must be applicable to the fine-grained facts of individual disputes that span myriad factual contexts. The latter rationale is especially strong in complaint-driven investigations and assessments, but is also relevant to systemic assessments that must measure enforcement institutions' capacity to provide relevant remediation and deterrence across the gamut of individual disputes.

Fourth, the validity of a composite index for evaluating country performance requires, among other things, that the body of Indicators be well-specified and well-balanced and consistently capture the variables that measure all significant elements of the underlying concept – in our case, all significant elements of the effective enforcement of, or taking steps

revisable standard. Recall also that a “comprehensive” set of measures does not necessarily entail that contextual sub-indicators be specified *ex ante*. In the iterative process of assessing countries, ILAB analysts can flesh out more general standards and more general rules, by entering sub-indicators and information sources in the drop-down window for each such general standard and rule – that is, for each Indicator. Some of these sub-indicators and information sources will be country-specific; others will not be. In that process, ILAB analysts can also revise and strengthen Indicators and sub-indicators to account for changes over time in both country-specific contexts and cross-country variables (such as increased productivity, transformations in managerial systems, and internationally recognized redefinitions of rights).

to afford, each of the multi-dimensional rights. Comprehensiveness is especially important when the Indicators take the binary form.

Which Indicators are best suited to serve as Probative and Diagnostic Indicators? Since a Probative Indicator is a species of diagnostic indicator, the following discussion will treat them together, under the genus of Diagnostic Indicators. Some of the proposed Diagnostic and Probative Indicators in Appendices B and C are drawn from the list of Assessment Indicators in Appendix A. But others are newly formulated to suit the special diagnostic purpose of probing for pathologies and severe problems of non-compliance. Since the diagnostic process is neither high-stakes nor evaluative, in the sense that it does not trigger economic sticks or carrots, it is not critical that the bodies of Probative and Diagnostic Indicators meet the more rigorous standards of comprehensiveness, balance, weighting, sensitivity testing, and statistical validity for aggregation that the body of Assessment Indicators must satisfy. And, compared to the Assessment Indicators, the Probative and Diagnostic Indicators need not be as tightly bound to the precise obligations imposed by the terms of U.S. trade legislation and treaties and by the internationally recognized legal sources that undergird the Assessment Indicators. As befitting their purpose, the Probative and Diagnostic Indicators are more idiosyncratic and probing.

The construction of Probative and Diagnostic measures rests on two desiderata: First, the measures should exploit whatever reasonably reliable data is available in an otherwise data-poor environment. Since application of the Probative and Diagnostic Indicators precedes the more intensive exercise of comprehensive assessment and is directed to a

universe of countries that includes many that are seriously non-compliant, the Indicators will likely be applied in settings of limited accurate data. For that reason, and simply to expedite the screening process, it seems especially desirable to use Indicators that guide analysts to the most readily available, reasonably reliable information. The Probative and Diagnostic Indicators may therefore include not only direct measures of inputs and outputs, but indirect measures as well.⁷²

Second, such Indicators should be formulated in response to the following question: “What phenomena do we expect, *a priori*, to find in workplaces, labor markets, and government institutions where there is pervasive lack of enforcement of the labor right in question?” The answer might in some instances be the same as Assessment Indicators constructed largely by reflecting on, and collecting data about, the question of what *positive* outcomes we expect to find where governments are compliant. For example, a high rate of government collection of accurate data about wage payment may indicate strong enforcement and a low rate may indicate weak enforcement.

But binary Indicators may be asymmetric; that is, we may need to frame Indicators to measure bad performance that are different from Indicators to measure good performance. For example, a substantial rate of failures to prosecute known murderers of trade unionists might be a good Indicator of across-the-board non-compliance with workers’ right of association. Yet even a zero rate of failures to prosecute such cases might be a very weak measure of across-the-board effective enforcement of freedom of association. Think of the

⁷² Part 12 proposes one iterative method for applying and refining such direct and indirect measures over time, in environments where accurate data is especially thin.

many countries where there are no such assassinations, yet workers' right to organize is extremely curtailed by weak deterrence or remediation of large numbers of anti-union discharges. In other words, a seemingly strong Probative or Diagnostic Indicator may produce few false positives but many false negatives. Probative and Diagnostic Indicators should therefore combine obvious, salient measures – such as impunity for perpetrators of violence against trade unionists – with more subtle probative measures, some examples of which are set out below.

The Probative and Diagnostic Indicators would be high stakes if, for example, they led immediately to the imposition of economic penalties or rewards, followed by application of Assessment Indicators to determine the benchmarks for incrementally reducing the penalties or increasing the rewards. But Probative and Diagnostic Indicators are plainly not intended for that purpose. Rather, they are intended for low-stakes or intermediate-stakes purposes. Probative or Diagnostic Indicators are low stakes, for example, if they are designed to aid the U.S. government in allocating resources for technical assistance or for research. Probative or Diagnostic Indicators would be “intermediate stakes” if their application cannot trigger immediate carrots or sticks but instead merely poses some risk of a full assessment that might or might not be followed by economic incentives.

Should Probative and Diagnostic Indicators focus on substantive law, enforcement, capacity-building, or outcomes? There is a case to be made for each category. Input Indicators measure the first-order activities (substantive law and enforcement) at the heart of the concept of compliance, and the second-order concepts (capacity-building) that measure

the sustained strengthening of the first-order capacities. As discussed in Part 5, in formulating Assessment Indicators, we must take care to not double-count the cause (good government performance) and its effect (good workplace outcomes). Recall that, as well, in assessments we will not use Outcome Indicators as proxies for inputs, since there are many variables that influence outcomes other than government performance and, moreover, we want to avoid diminishing the incentives for governments to collect data about enforcement efforts. For assessment purposes, then, we will take care to use outcome measures exclusively as constitutive elements in the definition of input Indicators, and as potential dependent variables in ILAB research about the impact of enforcement efforts.

But these issues are less pressing in the formulation of Probative and Diagnostic Indicators, in light of their non-evaluative purpose. If reliable data are available on outcomes that are *a priori* compelling measures of very poor enforcement (e.g., widespread assassination of trade unionists), then the possibility of double-counting is not a substantial concern. Outcome measures may yield false positives but this is not so costly, since the stakes are not as high in the screening stage as in the evaluation stage.

Can Capacity-Building Indicators serve as important probative and diagnostic tools? On the one hand, measures of second-order government processes may seem too distant from our deepest concern, namely, the actual application of government efforts to workplaces and the actual experience of workers. That is, our ultimate concern is the product of government self-monitoring systems, not the systems themselves. To put it tendentiously, should we forego a deeper assessment of a country where trade union leaders are

assassinated, on the ground that the government is good at collecting data, setting targets, and making action plans?

But that question is indeed tendentious. There is more to capacity-building systems than collecting data, setting targets, and making plans. The government must also implement its action plans, and must ensure that stakeholders participate in formulating, implementing, and evaluating the success of those plans. That is, Capacity-Building Indicators require that the government have ongoing mechanisms that secure its first-order efforts to define and enforce worker rights – in this case, its efforts to stop trade union assassination.

Using Capacity-Building Indicators as Probative Indicators or Diagnostic Indicators has some other merits. First, if we are concerned with more than the one-time or short-term impact of applying Indicators – more, that is, than producing a snapshot of government performance – then creating incentives to build government capacity, especially in the critical area of data-collection, may be a priority. And, indeed, the logic of the labor rights provisions in statutes and treaties is not to impose sanctions as punishment for bad current or past performance. Since the point is to create incentives to achieve future compliance, trade sanctions, once imposed, will be reduced incrementally as the government meets benchmarks of increasing compliance, and lifted altogether when compliance is fully achieved. Without insisting on the capacity for data collection, the U.S. government fails to signal clearly that it takes seriously the eventuality of rigorously applying such benchmarks.

Second, even if we assume *arguendo* that first-order inputs and outputs are ILAB's

key concerns and that the diagnostic inquiry is backward-looking rather than forward-looking, we need rich, reliable data to assess those current or recent problems. Creating government capacity to collect and analyze data is the necessary platform for the second step of holding that data up against the yardsticks of our first-order input and output Indicators.

Third, using Capacity-Building Indicators at the probative and diagnostic stages may mitigate the incentive for governments to suppress, falsify, or simply not collect data. In the context of high-stakes assessments, placing demands on a government to collect rich, reliable data may perversely collide with demands for substantive improvement. If data-collection is high-stakes – if the government knows that accurate data will reveal first-order violations and therefore trigger immediate trade sanctions – then the government has strong incentives to avoid data collection or to falsify data and corrupt its self-monitoring machinery. But if the government’s data-collection and capacity-building systems are measured probatively or diagnostically, then the stakes are lowered at least to some degree. The government will have some incentive to build those capacities in order to avoid deeper assessment.

True, some countervailing incentive cannot be fully avoided. If the government knows that reliable data-collection will reveal serious violations of first-order alarm-bell Indicators (i.e., Probative Indicators), then it will be reluctant to collect that data. However, that perverse incentive will apply more strongly to the specific data applicable to those particular alarm-bell Indicators, and more weakly to the construction of a sound data-collection system overall, which is the central inquiry of Capacity-Building Indicators addressing data collection.

Third, there is a good probability that private advocates will come forward with data that documents brazen alarm-bell violations (such as murder of trade unionists or widespread mandatory pregnancy testing of workers). Hence, even if the probative use of data-collection Indicators would otherwise create an incentive for the government to pervert its collection of data respecting alarm-bell violations, the government's awareness that private advocates will unmask its skullduggery mitigates that perverse incentive, especially in light of the countervailing incentive to score positively on the data-collection Indicators that measure the comprehensiveness and overall quality of data-collection protocols. There is little sense in suppressing data and scoring poorly on data-collection Indicators if the suppressed information will be reported anyway.

Hence, the incentives cut in two directions: The target government may still suppress or fail to collect data on a Probative Indicator, to avoid "admitting" wrongdoing. But the government will also be aware that its suppression of data may simply inculcate the government twice over: once for the underlying violation that private advocates and U.S. analysts are likely to uncover, and twice for the suppression of data itself.

Based on this analysis, we cannot conclude that any single category of Indicators is unrivaled for probative or diagnostic purposes. Rather, there is reason to think that alarm-bell Indicators taken from across the several categories will be synergistic. Hence, on the input side, a combination of first-order Probative or Diagnostic Indicators (such as failure to prosecute widespread violence against trade unionists) with second-order Indicators (such as the failure to collect data on first-order alarm-bell Indicators and to set targets to improve

those egregious violations) seems optimum. In addition, simple Outcome Indicators may be highly probative. For reasons just mooted, this combination may best mitigate perverse incentives, while maximizing the positive incentives rooted in low- to intermediate-stakes probative and diagnostic screening.

What of the fact that Probative and Diagnostic Indicators can be more readily gamed since they are non-comprehensive and non-specific? In light of this problem, we can expect the value of such Indicators to be attenuated over time, as governments adapt to U.S. probative and diagnostic expectations. (Rosga, 2009, p. 285). This is a variant of Goodhart's Law or the Lucas Critique: Once a measure is defined as a policy target, it may lose its evaluative capacity by inducing changes in actual performance that seek to track or evade the target. (See, e.g., Chrystal and Mizen, 2001).

One simple solution is to change the sets of Probative or Diagnostic Indicators from time to time. This solution does not pose problems for the validity of the Indicators, since they are not intended as comprehensive measures for purposes of evaluating compliance with legislation and treaties, and in many cases are not longitudinal. The U.S. government will not want to change those Probative or Diagnostic Indicators that are in fact longitudinal, since they must consistently track performance over time. For that reason, it is best to attach especially high weight to the government's accurate data-collection on such longitudinal Indicators, to at least reduce the incentive for gaming the Indicators.

10.3. Which Individual Indicators Should Serve as Probative and Diagnostic Indicators?

Finally, the more substantive question: Based on the above considerations, what individual substantive Indicators can serve most compellingly as Probative or Diagnostic Indicators?

First, and most obviously, we can begin with a well-chosen sub-set of Assessment Indicators. Assessment Indicators comprehensively measure hundreds of features of substantive law, enforcement, capacity-building, and outcomes. For probative and diagnostic purposes, then, we can begin with a targeted sample of those Assessment Indicators (a) that are *a priori* likeliest to correlate with the largest number of other Assessment measures in core domains of each labor right, and (b) for which reliable data are most likely to be available in the most data-poor contexts or for which the absence of data is most likely to correlate with a large number of Assessment measures in core domains.

It might seem, at first glance, that radical failures to establish basic labor codes and basic enforcement institutions would meet these criteria. These measures should indeed be included in the bodies of Probative Indicators and Diagnostic Indicators. As noted above, it is easy enough for a government to formally enact a labor code and to formally establish a labor inspectorate, an occupational safety and health agency, and other labor administrative bodies – but perhaps for that very reason, the failure to do so is a sign of the government's weak commitment to enforcement. Such an Indicator will yield many false negatives but perhaps some true positives as well.

The more telling, albeit crude, measure is whether the government provides minimal resources to such agencies in the form of budgets and personnel per worker or per workplace. These measures are crude for any number of reasons. Corrupt government officials may fill the labor inspectorate with patronage appointments. Even if labor inspectors are relatively numerous, they may perform their functions very poorly⁷³ due to low salaries, limited training, limited resources for transportation, limited data-tracking, corruption, and many other factors – although, it is true, corruption may be picked up in the corruption Indicator proposed below. But what level of resources counts as “minimal”? It is precisely in this context that a comparative Indicator may be warranted. For example: “Is the labor inspectorate budget per worker less than one-third of the average budget per worker among countries in the same quintile of real income per capita?”

Assessment Indicators that measure first-order and second-order failures of data collection may be prime candidates to adapt for probative or diagnostic purposes. A government’s failure to collect data about workplace fatalities may be a strong measure of weak government commitment to enforcing workplace safety and health. Similarly, failures to disaggregate wage and employment data by gender or ethnicity may be strong alarm-bell Indicators showing weak government commitment to enforcing rights against employment discrimination. Parallel failures of data collection are probative of weak commitment to enforce the other rights that concern us. Such first-order failures of data collection are compounded if the government is making no second-order effort to build its data-collection

⁷³See the discussion in this project’s *Literature Review and Bibliography* at pp. 317-327.

capacities on those matters.

For measures of actual enforcement *practices* included in Assessment Indicators (as distinguished from the enforcement institutions' structure, resources, procedures, and data collection), official data is likely to be so thin and unreliable that an alternative strategy is required. The most reliable probes are the investigations conducted by the most rigorous nongovernmental monitors, such as the Worker Rights Consortium and Human Rights Watch, by public international organizations such as the ILO, by international union federations such as the ITUC, and by U.S. government reports such as the State Department's Country Reports on Human Rights Practices. Such investigations are "probes" in the figurative medical sense: slender instruments that test the depth and direction of a wound or pathological cavity. The nongovernmental investigative reports are "slender," in the sense that they are not large-n surveys from which we can draw statistically significant conclusions. But they often penetrate deeply and surgically into the details of factories that supply large Western corporations. Notwithstanding the limitations of small-n studies, if nongovernment organizations reliably document more than one case of mass discharge or blacklisting of union supporters, more than one case of systematic pregnancy-testing, more than one case of exposure to banned toxins, or a drastic increase in the use of temporary employment contracts, there is reason to suspect a deeper pathology – in part because such supplier factories are likely to have better compliance records than the typical local firm not linked to global supply chains. (Mosley, 2008).

Second, we can formulate new Indicators that are insufficiently precise to serve as

valid Assessment Indicators, or are too causally remote from labor rights enforcement to serve as Assessment Indicators, but that otherwise have the qualities mentioned above – namely, likelihood of correlating with a large number of Assessment measures in core domains, and likelihood that reliable data are available or that absence of data correlates with other core violations.

There are, *a priori*, broad minimal preconditions for enforcement and capacity-building. If a country is autocratic, it is very unlikely to enforce freedom of association and the right to organize, since independent unionization (like other civil-society organizing) may produce a center of opposition to the ruling clique. Moreover, a non-democratic government may generally be less likely than democratic governments to enforce the rule of law, including labor law. Further, an autocratic government is less likely to be committed to securing the basic rights and needs of its citizens, including rights to a minimum wage and to safe and healthful workplaces. There are exceptions, such as autocratic corporatist systems (for example, classic Peronist regimes) that mobilize and protect workers; or idiosyncratic regimes like Apartheid-era South Africa that maintain a functioning judiciary imbued, within limits, with a rule-of-law culture. But the general rules seem *a priori* well-grounded.

This seems to call for one or more Indicators of basic democratic features of the government. These features may not be sufficiently proximate causes of labor compliance to serve as valid Assessment Indicators, but may be compelling indirect measures for probative and diagnostic purposes; yet such features may be more difficult to precisely specify and confidently apply to the facts than more direct measures of labor enforcement. The U.S.

government may be politically reluctant to bluntly label a group of governments as non-democratic. A promising candidate for an alarm-bell Indicator of lack of enforcement or capacity-building, which does not entail such labeling, may be this: “In the preceding ten years, has no more than one political party won more than 10 percent of the votes in an election for the national legislature?”⁷⁴

As just implied, it may be useful to identify a proxy for the government’s general commitment not just to democracy but also to the “rule of law” since enforcement of any right, labor or non-labor, flows from such commitment. Are there promising candidate Indicators that capture the concept of “minimal commitment to the rule of law”? One proxy for lack of such commitment may be pervasive corruption in the judiciary or in labor-related enforcement bodies such as the labor inspectorate, labor boards, or specialized labor courts. But how to precisely frame such an Indicator? In-country labor lawyers, unions, managers, academics, and journalists may have general knowledge of such pervasive corruption, but ILAB analysts or U.S. labor attaches may not have access to relevant surveys. One candidate based on available information is this: “Does the country rank in the lowest 20 percent of countries in Transparency International’s most recent application of its Corruption Perception Index?”

Analogously, even though this research is not focused on rights against labor trafficking and forced labor, Indicators measuring violations of such fundamental labor rights may serve as alarm bells for general non-compliance with labor rights. One such Indicator is: “Is the country ranked in Tier 3 of the most recent United States government’s Trafficking in

⁷⁴ The score for this Indicator would be “yes” if no legislative elections were held during that period.

Persons Report?” Another is: “In the preceding two years, has a nongovernmental or governmental organization reliably documented more than one instance of the use of forced labor?”

As to Indicators dedicated to the separate rights that concern us: For freedom of association, rights to organize, and rights to bargain collectively, the most important measure in principle is the annual number and percentage of union supporters who are discharged or blacklisted without effective remediation. However, reliable country-wide data on those measures are very likely unavailable. Still, if a nongovernmental or governmental source has reliably documented significant instances of governmental failure to deter, penalize, or compensate anti-union discharges or blacklisting, this can be taken as a strong Indicator of ineffective enforcement – even if such an Indicator also yields many false negatives. Another candidate that in principle is a strong proxy and would likely not have such a strong asymmetry between true positives and false negatives is the number of union elections conducted as a percentage of the workforce and the number of union victories as a percentage of the those elections. One advantage of this measure is that governments generally keep public records of elections held and election outcomes. However, that otherwise appealing measure has at least two problems. First, many countries do not require an election as a precondition to a union’s gaining rights of collective bargaining. Second, in the pool of countries that most interest us – that is, the weakest-complying governments – many “victorious” unions may in fact be protection unions.⁷⁵ Hence, even if the overall rate of

⁷⁵ “Protection union” denotes a labor organization installed by managers or government officials to block workers from organizing genuine, independent organizations.

false negatives is lower for this measure than for reported anti-union retaliation, the false negatives may cluster in the countries the Indicators aim to identify.

For just this reason, a more promising Probative or Diagnostic Indicator may be the number of collective agreements that stipulate wages and benefits that do not exceed, or barely exceed, those already required by statute or regulation – since the agreements “negotiated” by protection unions rarely provide wages and benefits above the legal minima. And, in most countries, collective agreements must be filed with a public body and are publicly available. Similar Indicators that in principle probe for protection unionism – whether the union fails to process grievances on behalf of workers, or whether workers are simply unaware that they are “represented” by the union and are covered by a collective agreement – would require field surveys of workers and are therefore not suitable for ILAB’s purposes.⁷⁶

A second promising Indicator for labor union rights may be: “Has the government failed to convincingly and reliably show that, in more than two cases in the export sector⁷⁷ in the preceding two years, it has imposed and actually collected – either on its own behalf or on behalf of the aggrieved worker(s) – fines or monetary awards equal to or exceeding the back pay lost by worker(s) who were discharged for anti-union motives, calculated from the time of discharge to the time of the final order by the court or other tribunal?” This information should be within the control of the trading partner’s government. The

⁷⁶ Again, the widespread existence of protection unionism may be a fact that is generally known to the trading partner’s labor lawyers, officials of genuine unions, managers, academics, and journalists – but, again, ILAB analysts may not have access to reliable survey data.

⁷⁷ As explained below, it is desirable to limit the question to the export sector for probative and diagnostic purposes, notwithstanding that ILAB analysts are concerned with labor rights and conditions in the broader workforce for other purposes, including for comprehensive assessment.

government's failure to come forward with more than two such instances of meaningful legal sanctions is symptomatic of broader non-compliance.

With respect to rights against discrimination, promising Indicators may be those that focus on employment and wages in the public sector or in certain occupations in the public sector – for at least three reasons.⁷⁸ First, governments are more likely to have data on public sector employment than private sector employment. Second, it may be possible to identify occupational categories that are present in the public sector of every country. And third, discrimination is very likely no worse in the public sector than in the private sector; so poor compliance in the public sector can serve as a good alarm bell for compliance in the private sector as well. Hence, one promising set of Indicators may measure the percentage share of public sector jobs held by women and racial or ethnic minorities relative to their percentage share in the working-age population or in the overall paid workforce. Another set may compare wages of women and racial or ethnic minorities in the public sector with wages of majority-race or majority-ethnic men. Specific public-sector occupations that might be suitable for cross-sector, longitudinal, or cross-country comparisons are postal work and mass transit (bus and train) station attendants. Other occupations – such as teachers, nurses, and public sector construction workers – might be probative of widespread occupational segregation, but are not likely to differentiate the worst performing countries from better performers. In addition, Indicators of instances of actual imposition of meaningful legal sanctions, analogous to that mentioned above for rights to organize, may be good

⁷⁸ Professor Andrew Schrank of the University of New Mexico broached the ideas in this paragraph.

Probative or Diagnostic Indicators of employment discrimination.

An open-ended Indicator of acceptable conditions with respect to wages, hours, and safety and health such as “Is the percentage of workers in the informal sector larger than the average percentage among countries in the same quintile of real income per capita?” may be too imprecise since it provides no easily applied definition of the “informal sector.” A proxy such as percentage of workers with temporary contracts (as distinguished from indefinite, long-term, or permanent employment contracts) may serve as a more precise probe for informal employment. Apart from serving as a good proxy for informal employment and therefore weak enforcement of wages, hours, and safety and health, the use of temporary contracts also serves as a pointed measure of enforcement of rights to organize, since employers may simply decline to renew the contracts of temporary workers who support unions. Another proxy for formal employment may be the percentage of workers who receive social security payments.

Also with respect to acceptable conditions of work, more pointed Indicators of enforcement institutions may be *a priori* more convincing than the general measures of labor inspection proposed above. For example: “Does the occupational safety and health body employ a non-trivial number of industrial hygienist, occupational physicians, and occupational nurses as a percentage of the total workforce?” Over time, in successive iterations of such an Indicator, ILAB analysts can refine a more precise percentage (or a more precise absolute number) than “non-trivial.” And again, Indicators measuring imposition and actual collection of fines and damage awards, analogous to the Indicators for rights to organize and

nondiscrimination set out above, may be good probative and diagnostic measures of enforcement of wages, hours, and safety and health.

While ILAB's interest extends beyond the export labor market and the collateral labor markets that influence terms and conditions in the export labor market, for probative and diagnostic purposes it may be desirable to limit some Indicators to the export sector or manufacturing sector. From a cross-country perspective, reliable documentation by governmental, nongovernmental, and international organizations is likelier to focus on those sectors, as explained above.

11. Alternative Strategies to Weight and Aggregate Indicators



In the general field of composite indicator-formulation, there are many different strategies for weighting and aggregating indicators. In the more specific field of labor regulation, the *Literature Review* for this project describes and analyzes more than twenty efforts to create composite indicators or indices. Many of those efforts, however, serve policy goals other than ours – such as the measurement of “quality of working life,” “decent work,” “efficient labor markets,” “the legal origins of labor regulation,” “the political sources of labor regulation,” and many others. This Part canvasses the alternative strategies that seem to be the most promising candidates for use in creating composite Indicators for compliance with U.S. trade legislation and treaties.

The OECD and the European Commission collaborated in authoring the *Handbook on Constructing Composite Indicators: Methodology and User Guide* (OECD, 2008) (“*Handbook*”). The *Handbook* is the state-of-the-art on the general subject of constructing composite indicators. In addition, the Joint Research Centre of the European Commission sponsors a website on the same subject. (See <http://composite-Indicators.jrc.ec.europa.eu>) (“*Website*”). There are, of course, innumerable other papers and reports setting forth desiderata for the construction of composite indicators, many discussed in the *Literature*

Review for this project and some published subsequently. (See, e.g., Hoyland, et al., 2010; Hood, et al., 2008; Rosga and Satterthwaite, 2009.)

The *Handbook* and *Website* suggest a ten-step methodology for constructing a composite indicator: (1) developing a theoretical framework, (2) identifying indicators (“variables”) and applying the indicators to relevant data, (3) imputing missing data, (4) applying multivariate analysis to analyze the underlying structure of the data, (5) normalizing the data, (6) attaching weights to, and aggregating, indicators, (7) applying uncertainty and sensitivity analysis to gauge the robustness of the composite index, (8) de-constructing composite indicators to identify and analyze the contribution of individual indicators and sub-categories of indicators, (9) testing the explanatory power of the composite indicator by linking it to well-known and measurable phenomena, and (10) effectively communicating the composite indicator and the underlying methodology to end-users. (OECD, 2008).

Steps (1) and (2) have been discussed fully above, in Parts 1-10. Potential responses to data unavailability – step (3) – have also been discussed above in Part 9. However, completing the remaining seven steps for each of the alternative strategies discussed below is beyond the scope of this paper. Pursuant to the terms of the research proposal, this Part discusses only step (6) – alternative strategies for weighting and aggregating Indicators. As indicated in the research proposal, the point of this discussion is not to reach a settled conclusion about the optimal strategy, but instead to canvass the strengths and weaknesses of the alternative strategies.

The most fundamental point about the weighting of indicators is this: “Regardless of

which method is used, weights are essentially value judgments.” (OECD, 2009, p. 31).

11.1. Equal Weighting and Simple Summation of Indicators

The most widely used – and simple – strategy for weighting indicators is to give each indicator equal weight and to sum the indicator scores. (OECD, 2009, p. 31). For example, this is the strategy used by the European Commission in its first, recent effort to construct a composite indicator for the European Employment Strategy (EES) of the European Union, also known as the Lisbon Strategy. The European Commission adopted this strategy from the work of researchers at the Paris School of Economics and the Sorbonne. (Davoine, et al., 2008b). The Commission candidly states: “The choice of equal weights is largely arbitrary, although being [sic] transparent, simple and in line with the literature which does not establish any clear ‘hierarchy’ between the different components of job quality.” (European Commission, 2008, p. 164).

This strategy is also used by Cambridge University researchers who constructed an index of the overall strength of labor regulation, for purposes of assessing whether civil-law countries have stronger and more rigid regulation than common-law countries. (Deakin and Reberieux, 2009; Deakin and Sarkar, 2008; Deakin, et al., 2007b). The same strategy is used by Botero, *et al.*, whose labor regulation indicators were incorporated in the now-abandoned labor indices of the World Bank’s *Doing Business Reports*. (Botero, *et al.*, 2004). Although the Cambridge researchers are otherwise critical of the Botero, *et al.*, methodology, they do not reject the equal-weighting strategy.

At first blush the equal-weighting strategy seems to conflict with our conceptual

framework. After all, it fails to differentiate Indicators that are more relevant to the ultimate concept of compliance from those that are less relevant. However, that “failure” may in fact be a conceptual advantage for our particular policy purpose. Unlike other composite indicators that use this strategy (for example, the EU’s Indicator of “job quality” or the ILO’s various indicators of “decent work”), our composite Indicator is a measure of compliance with legal *rights*. In this context, since each Indicator states a rule or standard that defines an element of a universal legal right, a case can be made that *full* compliance requires the government to comply fully *and equally* with each Indicator. Stated in operational terms, the government’s failure to comply with each Indicator or adequately take steps to comply with the Indicator is of equal conceptual significance with compliance with any other Indicator, since the Indicators are of co-equal, universal importance.

This is especially true, when the Indicators are used in response to complaints or petitions. In such case-driven assessments, we can expect the complainant to allege violations of quite specific Indicators that might not, *a priori*, seem to be critical elements of the body of Indicators viewed as a whole. For example, a group of complainants might allege that, after a labor court’s finding that they were discharged for attempting to organize a union, the court ordered severance pay and back pay but did not order reinstatement on the ground that the workers had taken jobs elsewhere at the same pay as the jobs they lost. This might at first seem a relatively minor violation of an Indicator; the workers, after all, seem to be “made whole” for all their monetary damages and are now employed. And this might seem to be a “mere” procedural question of the type of remediation provided to plaintiffs in a case where the court has in fact gone a long way in protecting the workers’ right

to organize, in the sense that the court has found a violation and given a make-whole remedy.

However, on closer inspection, such an order may not in fact provide the critical remedy for damage to the union organizing drive at that location and elsewhere. Those drives may be utterly stymied by the discharges, since other workers at both the workplace in question and at other workplaces will see the risk they run in attempting to unionize. Conversely, if union supporters were immediately reinstated, workers would see that the legal system gives them powerful protection if they wish to exercise their right of association. What seems like a small component of the overall body of rules on freedom of association and rights to organize may in fact be critical to the complaining workers' genuine exercise of those rights.

Stated in less consequentialist and more formal legal terms: We have constructed the Indicators based precisely on the explicit or implicit universal, *non-derogable* importance that the specific rule or standard is accorded either in international labor law or in some instances, in the consensus of well-developed, democratically formulated regional or domestic systems of labor law. Stated alternatively, the Indicators are the Executive Branch's reasonable interpretation of Congress's intent in enacting the obligation to comply with "internationally recognized worker rights." In short, we have deliberately constructed the Indicators on the conceptual assumption that they are justifiably co-equal in weight.

Among the potential problems with the equal-weighting strategy identified by researchers is this: Suppose sub-indices are formulated for each major right (rights to organize, rights against employment discrimination, and so on), and the sub-indices are then

aggregated through simple summation or averaging. A problem arises when different numbers of Indicators are summed or averaged to construct each sub-index. Each of the Indicators comprising a sub-index containing relatively few Indicators implicitly receives greater weight in the overall composite index than each of the Indicators comprising a sub-index containing relatively more Indicators.

This problem, however, may not be significant in our case, since U.S. trade legislation and agreements require trading partners to comply with each of the major rights and standards. The “sub-indices” (one for each of the major rights) are not aggregated. Nonetheless, if the equal-weighting strategy is employed, then we must take care not to create sub-sub-indices within each right and standard that have imbalanced numbers of Indicators, to avoid the same problem at the level of composite sub-indices. In fact, in constructing the body of candidate Assessment Indicators in Appendix A below, we have sought to provide balance in the number of Indicators contained in each of the equally significant sub-areas within each right – and as we prune the candidate list of Assessment Indicators we can take care to maintain such balance. Achieving this balance in formulating the Indicators themselves strengthens the validity of equal weighting of the Indicators. Each equally significant sub-area of the right will be given equal weight in the composite “sub-index” for the respective right in question.

Alternatively, we can simply set aside the strategy of aggregating Indicators within each sub-area (“sub-sub-indices”) and instead stick to the strategy of aggregating all equally weighted Indicators within each right or standard (“sub-indices”). We then need not be overly concerned even about the problem of balance among sub-areas *within* each right – that is,

the balance among “sub-sub-indices” within each “sub-index”. If each specific legal rule or standard measured by an Indicator is of co-equal significance with the specific legal rule or standard measured by other Indicators, then it may be normatively less problematic if the various sub-areas of the overall right contain different numbers of Indicators and therefore have different weights. Each sub-area *should* have different weights, precisely because each Indicator has co-equal status.

Note also that the nature of our underlying concept – compliance with all elements of universal rights – may mitigate the problem of correlation between different measures of the same aggregate, a problem sometimes raised when determining weights for indicators. (Sometimes for this reason, researchers give lower weight to each of two correlated indicators.) First, again, we might maintain that our “aggregate” concepts – such as “freedom of association” – are each comprised of a group of sub-rules (Indicators), and each of those sub-rules is essential to the aggregate concept. Those sub-rules will indeed be highly correlated since they each are *constitutive* of the same larger concept. In that sense, they are correlated *by design*. Second, in formulating the Indicators, we have deliberately sought to avoid overlapping measures of the same sub-rule.

And recall that even if, as proposed above in sub-part 3.2, analysts begin with a significant sub-rule (i.e., the Indicator, which constitutes the heading of a drop-down window) and then flesh out that sub-rule with a set of more specific, contextual rules (i.e., the sub-indicators shown in that window), the latter set of sub-indicators will then be aggregated to determine the score for the Indicator. The weights of the various Indicators will remain unchanged and equal to one another. This proposition, then, presumes that the sub-sub-

rules (sub-indicators) will have variable weights, while the sub-rules (Indicators) are equally weighted. It is the equally weighted Indicators that remain uniformly weighted across countries.⁷⁹

One great virtue of the equal-weighting strategy is that it may be the method that is most comprehensible to the public, to our trading partners, and to advocates using the methodology for filing petitions or for independent monitoring. This virtue should not be underestimated – especially in light of the fact that the net analytic benefit of other weighting schemes is ambiguous. The equal-weighting strategy may be most comprehensible not only in terms of the public understanding of a simple arithmetic algorithm, but in terms of the normative idea conveyed above: Many workers, nongovernmental organizations, and advocates adhere to the view that each significant, co-equal element of a universal right deserves equal protection. Indeed, some ILAB officials emphasized just this point in interviews with this author.

⁷⁹ To return to one of our key examples: An Indicator (sub-rule) might state: “Does the government aggressively prosecute all cases of violence against trade unionists and their families, when the government has or should have reasonable cause to believe that such crimes have been committed?” This Indicator, recall, is a “fixed standard,” because the basic norm is well-settled and unchanging (“fixed”), but the content of the norm cannot be fully spelled out in advance and must be fleshed out from country to country (“standard”). Over time, ILAB analysts will add sub-indicators to that window, some of which are country-specific and some of which are not. For example, in the window for Colombia, the analyst might enter the following sub-indicator: “When members of the paramilitary or other perpetrators plead guilty to violence against trade unionists or their family, do prosecutors continue aggressively to pursue investigations and prosecutions against co-conspirators or other perpetrators whom the government has or should have reasonable cause to believe committed the same or similar crimes.” Over time, ILAB analysts will have to construct explicit or implicit weights for the various sub-indicators that are entered in the window for the given Indicator, while the Indicator itself maintains an equal weight within the body of Indicators and across countries.

11.2. Weighting of Indicators Based on Hierarchies in Authoritative Legal Sources

The previous sub-part argued that equal weighting of Indicators could be justified by the claim that all Indicators were constructed to exceed a relatively high threshold of legal significance as a universal, non-derogable component of the right in question.

An alternative position is to acknowledge that the relevant legal sources on which the Executive Branch reasonably relies – international labor law and, in some instances, the consensus among well-developed, democratically promulgated labor laws at the regional or national levels – contain explicit or implicit hierarchies of legal norms. Traditional legal analysis enables those with expertise in general jurisprudential methods and, more specifically, in international and comparative labor law to tease out and map that hierarchy. The weighting of Indicators can thus be grounded, with some degree of confidence, in the existing, authoritative hierarchy of norms.

This strategy has the merit, at least conceptually, of relying on objective,⁸⁰ existing prioritizations of labor norms. More important, the strategy by definition relies on the particular prioritization scheme of authoritative, internationally recognized labor law. In the absence of such an objective, existing prioritization, the researcher or analyst would unacceptably impose her own subjective view of the relative importance of each Indicator. True, the researcher or analyst might apply some objective theory other than traditional legal

⁸⁰ By “objective,” I here mean only that the weighting methodology is not based on the personal, subjective preferences of the researcher who identifies the weights for each Indicator. Of course, the hierarchy in existing legal materials is value-laden, just as Rawlsian analysis and welfare economics (to which the text presently refers) are value-laden. That is in the nature of normative analysis.

analysis (for example, welfare economics or Rawlsian concepts of equality) to prioritize the Indicators. But such theories, like the researchers' subjective personal preferences, are external to the normative prioritization embodied in existing law and are therefore inappropriate for our project, which is to measure compliance with legal rights.

This strategy, however, poses several challenges. First, prioritization is one thing. It is another matter to attach numerical weights to the different tiers of the hierarchy. One organization, in the unofficial, unpublished pilot-testing of a recent effort to create a composite index for freedom of association, uses equal weighting, but notes that the method can be readily adapted to differential weighting of more and less important rights or aspects of rights. (Anonymous Organization, 2010). The organization gives no numerical examples. The organization does tell us that the number of violations for most indicators will be coded in several cardinally-numbered bands (with 1 representing the most numerous violations). Still, the bands indicate only the severity of the number of violations of an indicator; they do not prioritize or weight the importance of the norms embodied in the indicator. The organization does indicate that some matters are so important that a single violation of the indicator will be given the default score of "1." This, of course, tells us that such indicators are weightier than others, in the sense that a single violation of some indicators is deemed as weighty as numerous violations of other indicators. But that differential weighting will have no consequence for the composite indicator where there are so many violations of such "important" indicators that the latter would be coded "1" even without the default rule. In such instances, the "important" indicator will implicitly have equal weight with "ordinary" indicators.

Hence, whether certain indicators are actually afforded differential weight will depend on the contingent empirical reality of the relative numerosity of violations.

Second, it is one thing to say that a good lawyer can use his or her craft skill to extract the prioritization of different elements of labor rights that is “objectively” embedded in the legal materials. It is another to carry out the process of extraction and articulation in a way that is convincingly “objective” rather than being or appearing to be the product of the lawyer’s subjective judgment. While some notable legal philosophers defend the view that such an exercise can be fully objective (see Dworkin, 1986), others may not be so trusting of the ILAB’s legal advisors and analysts.

Still, as one research consultant writes, many if not most composite indicator methodologies, if they do not use equal weighting, choose instead a “somewhat arbitrary weighting based loosely on theory.” (Anonymous, 2009, p. 17.) In our case, the prioritization of Indicators would not be loosely based on theory, but instead based relatively tightly on the underlying concept – namely, the prioritization of different elements of rights found in international law or in a broad consensus of regional and national systems with well-developed labor law, as reasonably interpreted by the U.S. Executive Branch.

But, third, the very fact that U.S. jurisprudence authorizes the U.S. Executive Branch to stipulate any Indicators that reasonably implement Congressional intent means that the Indicators need not in fact be seamlessly bound to international or comparative sources of labor law. This adds an element of indeterminacy to the otherwise “objective” mapping of established legal hierarchies.

Fourth, formulating a hierarchy of Indicators creates new opportunities for governments to game the methodology. Governments determined to deny certain rights (such as suppressing union organizing) might adopt a new strategy of avoiding violation of the weightiest Indicators for the right to organize and instead engaging in sufficiently numerous violations of the less weighty Indicators. Such gaming opportunities might be diminished if each Indicator is scored for severity of the number of violations of that Indicator and then adjusted for weightiness of the norm embodied in the Indicator. The most prominent weighting scheme proposed by ILO staff uses neither the first nor the second of these adjustments. (See, e.g., Kucera, 2004, 2007.) The NAS Indicators, the Cambridge researchers, and the anonymous organizational proposal mentioned above use the former scoring (i.e., based on number of violations) but do not combine it with the latter weighting (i.e., the normative importance of the norm) – although, as just mentioned, the anonymous organization noted that such weighting could be used in the future by it or by others.

Another shield against gaming, as noted above in the discussion of Probative and Diagnostic Indicators in sub-part 10.2, is to simply revise the body of Indicators at regular intervals.⁸¹

⁸¹ Here we are speaking of “revision” in the sense of more or less random substitutions of some Indicators with others that are of relatively equal, *prima facie* importance or that measure slightly different elements of similar subject matters. Governments cannot game the Indicators if they do not know which Indicators will be used in the next round of appraisal.

As discussed in sub-part 10.2, this strategy is problematic in the case of Assessment Indicators. Since Assessment Indicators are intended to comprehensively protect all the key elements of worker rights, and since our goal is to ensure that governments deliberately measure their behavior against those comprehensive yardsticks, it is undesirable to randomly interchange Indicators. As noted above, this is especially true of Indicators that are explicitly longitudinal.

The strategy of routinely revising Indicators is normatively more acceptable in the case of Probative Indicators and Diagnostic Indicators, for two reasons: First, the cost of the strategy is lower, since such

11.3. Principal Component Analysis

Another weighting strategy is to apply the statistical method called principal component analysis to the body of Indicators. Principal component analysis is a form of factor analysis that converts a large number of variables (which may be correlated) into a smaller number of “principal components.” The principal components are uncorrelated variables constructed to maximally express the variance in the data set. Hence, the first principal component accounts for as much of the variability as possible in the entire data set; the second principal component accounts for the greatest amount of variance in the remainder of the data set; and so on. In short, the idea of principal component analysis “is to account for the highest possible variation in the indicator set using the smallest possible number of factors.” (OECD, 2008, p. 89). For a factor analysis, only the principal components that account for the most variance are retained or assigned high weights; the rest are discarded or down-weighted.

The appeal of principle component analysis in the formulation of many types of composite indicators is clear. If one wishes to rank the relative performance of countries, it may be useful to reduce the number of variables and attach weights to those variables based on their contribution to the variance among countries.

Indicators do not purport to be valid, comprehensive, evaluative measures of universal rights. There are likely to be alternative sets of reasonably sound Indicators that point to egregious violations calling for a second round of more comprehensive assessment. Second, the benefit of the strategy is higher, since Probative Indicators and Diagnostic Indicators are designed to identify the worst-performing governments that may be more likely to try to game the Indicators.

However, principal component analysis seems ill-suited to the construction of a composite Indicator that aggregates Indicators measuring compliance with the various elements of non-derogable legal rights. Indicators may receive low weight even if they are critical to compliance and measure very important aspects of rights. This may be the case if most countries do not comply with the Indicator, so that the Indicator accounts for little variance among countries. For example, principal component analysis may attach low weights to Indicators for occupational segregation by gender, since most countries have substantial levels of such discrimination. Principle component analysis may also assign low or negative weight to positive government actions that happen to co-vary with negative government actions. (See Anonymous, 2009).

11.4. *Ex Ante* Multivariate Modeling

To make meaningful use of cross-country and time-series data, ILAB analysts might begin with a sophisticated multivariate model, including well-specified control, independent, and dependent variables. The model might generate the weights to be attached to the independent and dependent variables (Indicators) that concern us. Independent and control variables might include average real income per capita, per capita tax revenue, legal regime type (e.g., civil vs. common law), political regime type, labor relations regime type, degree of

civil stability or instability, foreign direct investment, colonial origins, and others.⁸² The Indicators that measure Substantive Law, Enforcement, Capacity-Building, and Outcomes might be either independent or dependent variables, as discussed below.

However, the policy relevance of constructing and applying even a rigorously designed multivariate model, using rich longitudinal and comparative data, is not entirely clear.

Consider a sophisticated recent attempt, in the work of Murillo and Schrank, to use multivariate modeling to uncover the variables that account for positive labor law reform.⁸³

Murillo and Schrank's regressions show, surprisingly, that most Latin American countries maintained or increased their collective labor protections in the 1980s and 1990s, even during the period of market liberalization, privatization, and (in non-labor matters) deregulation. (Murillo and Schrank, 2005; Murillo and Schrank, 2010). Their control and independent variables include "structural variables" (level of economic development and size of economy), "political factors" (alliances between political parties and labor unions, and alliances between organized labor and nongovernmental international labor rights advocates), "institutional variables" (labor-mobilizing versus labor-repressive regime types), and "conjunctural factors" (economic recovery, inflationary pressure, democratization, and trade openness).

Although their model does not disaggregate the dependent variable of "collective labor

⁸² For examples of multivariate models that attempt to show correlation or causation between some of these variables and the dependent variable of labor regulation, see this project's *Literature Review and Bibliography* at pp. 138-146, 285-307.

⁸³ The following analysis is not intended as a critique of their excellent social-science modeling. Rather, it is intended merely to demonstrate the extraordinary challenge of adapting such a model, or constructing a new multivariate model, suited to our special policy purposes.

protections” into more specific aspects of legal reform,⁸⁴ they note that in one category of regime type (“labor-repressive”) “most of the reforms we have examined negatively sanctioned antiunion activities, improved job protections for union activists, and eased the procedures for union registration.” As to the second regime type (“labor-mobilizing”), they note that “union-friendly reforms were more likely to increase the organizational resources of trade unions to help them cope with the broader impacts of liberal market reforms.” (Murillo and Schrank, 2005, p. 994). They attribute the reforms in the labor-repressive regimes to domestic labor unions’ success in enlisting international advocates and foreign governments, which preconditioned market access on compliance with core worker rights (making use, for example, of the U.S. GSP). They attribute the reforms in the labor-mobilizing regimes to the domestic labor unions’ capacity to extract concessions from their historical legislative allies in exchange for yielding to deregulation in fields other than labor law.

What are we to make of such modeling, in light of our policy purposes? If their regressions are credited, then Murillo and Schrank have shown that historically labor-repressive regimes are likelier (in certain periods of history and in one region) to improve some elements of the core rights to associate and to organize – namely, greater protection against anti-union retaliation and improved procedures for union registration – while historically labor-mobilizing regimes are likelier to increase material resources for labor

⁸⁴ Murillo and Schrank also limit their concept of labor law reform to changes in legislation, in contrast with the much broader and more detailed array of rules, institutions, and resources that concern us – that is, constitutional, legislative, administrative, and judicial rules across four discrete worker rights and three workplace standards, and the institutional machinery and material and human resources devoted to enforcing those rules and to increasing the capacity to enforce those rules.

unions. Should our Indicators therefore downgrade particular labor-repressive or labor-mobilizing regimes if they have not shown the respective improvements in labor protections that were (according to Murillo and Schrank's statistical model) characteristic of each regime type in the 1980s and 1990s? Alternatively or concurrently, should we downgrade particular labor-mobilizing regimes for not putting more resources into making the kinds of reforms that labor-repressive regimes have been able to carry out, and vice versa? Or do we make neither "adjustment," since the labor-repressive governments deserve no "credit" for the capacity of their labor unions to form alliances with foreign activists and foreign powers, and the labor-mobilizing governments deserve little credit for the capacity of their unions to extract concessions as a *quid pro quo* for market de-regulation in the non-labor sphere?

These questions show that models such as Murillo and Schrank's are ambiguous as to two of the key conceptual points discussed above in Parts 5 and 8. First, should our dependent variables be measures of substantive law, enforcement effort, and capacity-building, or measures of the outcomes of such inputs, or both? Legal rules, enforcement institutions and resources, and capacity-building efforts might be treated as independent variables (inputs) and employer compliance treated as dependent variables (outcomes). Alternatively, legal rules and enforcement efforts might be modeled as dependent variables, if rules and efforts are taken as the ultimate indicators of government compliance. That is, we may wish to know which (independent) variables *account* directly or indirectly for good government performance. Yet in Murillo and Schrank's model, variables pertaining to legal regulation and government performance appear as *both* independent variables (e.g., "labor-

repressive” regimes) and dependent variables (e.g., “union-averse” legislative reforms); and, likewise, socio-economic variables that might measure the “outcome” of labor regulation appear as both independent variables (e.g., linkages with international labor rights groups) and dependent variables (e.g., more resources for union activity).

Second, and closely related, such models do not differentiate between independent variables that are within the immediate control of current policy-makers – namely, Substantive Law Indicators, Enforcement Indicators, and Capacity-Building Indicators – and those that are not – namely, regime type, level of economic development, size of economy, foreign direct investment, historical legal origins, and so on.

From the standpoint of our policy concerns, what would be the conceptual significance of teasing out from these models the non-legal variables that correlate with or cause the independent variables of interest? Do we then treat such causal, non-legal variables (for example, partisan alignments or international vulnerabilities) as hard constraints that “excuse” poor performance in labor law reform, since the variables are outside the control of the government? Or, do we instead search for Indicators of government policy that might positively influence those causal variables (such as policy designed to change the relevant party alignments or to increase certain types of international linkages), even if such policy is relatively remote (in both institutional mechanisms and time) from the proximate enforcement of labor rights?

These difficult questions can be skirted, in light of the fact that even a state-of-the-art model such as Murillo and Schrank’s has not disaggregated the concept of “labor law reform”

into variables that are more fine-grained than the broad dichotomy between legislative changes that are “union-friendly reforms” and legislative changes that are “union-averse reforms.” The somewhat more specific but still broadly defined reforms mentioned by Murillo and Schrank – negative sanctions for anti-union activity, easier union registration, greater resources for unions – are not captured in their formal model but, rather, are anecdotal assessments. Even if these anecdotal factors were modeled as independent variables, neither the model’s dependent nor its independent variables would capture the hundreds of precise Indicators that define the four multi-dimensional rights and the three complex workplace standards comprising “internationally recognized worker rights;” that define the multiple institutions, resources, procedures, and practices that comprise “effective enforcement” of those rights and standards; and that define the four elements of building capacity to improve such enforcement.

To be sure, social science models like Murrillo and Schrank’s were not intended to illuminate the causes or consequences of the comprehensive rules, institutions, resources, and practices that constitute compliance with internationally recognized worker rights. They instead aimed to shed light on more broadly defined historical phenomena, such as highly generalized shifts in labor legislation (“union-friendly” vs. “union-averse”) implemented by government institutions defined dichotomously (“labor-repressive vs. “labor-mobilizing”) over a relatively long period of time (the 1980s and 1990s). The discussion above is therefore not intended as a critique of such exemplary social-science modeling. But the discussion above shows just how great a methodological leap would be necessary to move from such

generalized models of long-term interaction among a relatively small number of broadly defined variables to multivariate models that capture the short- or medium-term interaction among a much larger number of highly precise legal, institutional, budgetary, personnel, and actual performance variables.

In sum, we are nowhere near having a sufficiently sophisticated model and sufficiently rich data set to demonstrate which independent variables (within the control of the current government in power) account for success or failure in effectively enforcing the rights and standards that concern us, and which independent or control variables (not within the control of the current government) either neutralize the success of good governmental efforts or are instead genuinely responsible for what is ostensibly the successful fruit of government efforts.

The strategy of attaching weights to Indicators based on multivariate regressions therefore seems unpromising.

11.5. *Ex Post* Multivariate Modeling: Learning by Monitoring

Can we find methodological virtue in necessity? Even if it were imaginably feasible, would it be pragmatically desirable to try to start with a complex multivariate model, including well-specified dependent variables, well-specified independent and control variables, and clear differentiation between variables within the government's immediate control and variables that are not?

An alternative, chastened strategy is to start with a comprehensive set of Indicators, some of which build in longitudinal and comparative concepts, while building in a minimum of control variables.⁸⁵ As discussed above in Part 8, the proposed bodies of Indicators build in an admittedly crude control variable (quintile of real income per capita);⁸⁶ but ILAB analysts might refine that control variable and add new control variables based on data and explanations gathered and tested in successive rounds of probing, diagnosing, and assessing. Since the methodology of applying Indicators is self-consciously iterative, a multivariate model could be gradually constructed through a process of “learning by monitoring.” (Sabel, 1994).

This alternative strategy is suggested by recent initiatives to measure the performance of public agencies – under the banner of “New Public Management,” “New Governance,” and the other new regulatory methods mentioned above in sub-parts 3.1, 3.2, and 3.3. (See, e.g., Hood, et. al., 2008.) The strategy rests on the experimentalist nature of public interventions into a social problem. The iterative efforts to improve social outcomes, and the participation of official auditors or analysts in those efforts, may provide opportunities either for constructing and refining a rigorous multivariate model – defining and revising the model’s control, independent, and dependent variables “as we go” – or for constructing improved comparative and longitudinal Indicators short of such a multivariate model.

To take one of countless case studies that highlights strategies that might be

⁸⁵ This is the starting point defended throughout this paper and put into practice in the bodies of Assessment, Diagnostic, and Probative Indicators formulated in Appendices A, B, and C.

⁸⁶ Part 8 above gives reasons for choosing this control variable.

applicable to our project: In one large U.S. city, the Head of Schools initiated a system of accountability based on the construction and application of indicators.⁸⁷ The indicators sought to measure, school by school, both the absolute level of student and school performance and the pace of improvement in outcomes. To make the exercise meaningful, the School Board set out to compare similarly situated schools. But what variables would define whether two schools are similarly situated? That is, how should the independent or control variables be specified? And how should educational outcomes – the dependent variables – be specified?

For at least three reasons, the School Board deliberately decided not to begin by attempting to identify and model numerous independent, control, and dependent variables. First, the environment was data-poor. A complex model would not get off the ground, in terms of yielding strong empirical findings about the variables that accounted for the similar or dissimilar outcomes among schools.

Second, there was a deeper conceptual problem with trying to stipulate the relevant respects in which two schools could be deemed “similar.” Some features of schools might seem to be obvious candidates for independent or control variables, such as the demographics of each school’s student body. But some features that *a priori* might appear “relevant” grounds for excusing poor school outcomes might in fact be readily susceptible to transformation and improvement by feasible policy interventions. That is, a variable taken as an exogenous “cause” of poor outcomes (or a hard constraint) might be better viewed as

⁸⁷ The following account is based on confidential interviews, conducted in August, 2010, with the official who formulated the methodology. The City and its officials are therefore anonymous.

endogenous (or a soft constraint). The systemic intervention could probe the potential to transform the various root causes of poor outcomes; and the iterative nature of the school assessments made it possible to do so.

Third, there was no fixed understanding about what should count as good outcomes. Instead there were only crude, provisional starting points for measuring performance, such as existing math and literacy tests. And the understanding of what counts as a good education might well change as the accountability system revealed what was practically possible and what the range of desired outcomes in the eyes of the relevant stakeholders was. Indeed, the stakeholders' own understanding of their educational goals might change in the course of dialogue among themselves and the officials conducting the accountability program.

In this context – marked by (a) data scarcity, (b) uncertainty about what accounts for poor outcomes, (c) uncertainty about the degree to which root causes could and should be subject to improvement via policy interventions, and (d) uncertainty about what ultimately counts as good outcomes – the School Board's Accountability Office began pragmatically by identifying provisional indicators of strong performance or poor performance, based in part on whatever data was available and reasonably reliable. In the search for such data, provisional performance indicators were fashioned by asking the common-sense question: What might we expect to see in a school environment where children are learning well? The answer might, provisionally, be simple measures such as scores on standardized math and reading tests, or might include less constricted measures.

Teams of analysts from the Accountability Office then confidentially discussed with a team of principals and teachers at each school its provisional performance score. This simple process had the immediate effect of “flushing out” candidates for control and independent variables, new outcome and performance variables, and new data, as schools attempted to explain and justify their seemingly poor performance compared to other schools. Only at this stage – after gathering, through such hub-and-spoke discussions, the new data and candidates for new independent and dependent variables – did the econometricians in the Accountability Office begin the process of modeling exogenous and endogenous variables, while keeping vigilant about the possibility that ostensibly hard constraints were in fact not so.


Call this a strategy of “chastened multivariate modeling” – chastened, because it starts with a limited set of performance measures based on the contingency of thin available data, and iteratively works its way toward a more sophisticated model bit by bit. This strategy, in many practical variations, is now being widely implemented in indicator-driven public programs.

Could the strategy be used to gradually build a multivariate model that would yield weights for the measures of government performance that contribute most to achieving positive compliance with internationally recognized worker rights? As the preceding sub-part concluded, problems of data availability and the complexity of the social world make it imprudent to expect that the strategy would yield even serviceable multivariate models in the foreseeable future. For the same reasons, successfully communicating any such

(provisional) model to trading partners, nongovernmental organizations, and the public would be difficult.

Nonetheless, even if ILAB could not expect to realize this chastened strategy for statistically valid weighting of Indicators, it still points to some features of an alternative “twice-chastened” strategy that might guide ILAB’s overall process of applying and revising Indicators – twice-chastened, because the alternative strategy would not seek, even bit by bit, to create a high-powered multivariate model. Some speculative thoughts about this alternative strategy are offered in the following Part.

12. A Process Based on Iterative Dialogue? – Some Conjectures



The preceding Part concluded that ILAB should not set its sights on weighting Indicators based on a sophisticated multivariate model of labor regulation, even if the construction of the model were conceived as an iterative process of “learning by monitoring” (referred to above as the “once-chastened” strategy). The social world is too complex, and the data will remain too thin for the foreseeable future.

Nonetheless, the discussion of the once-chastened strategy may still be illuminating, since it points to a potential *process* for implementing the methodology proposed in this paper. Call this the “twice-chastened” process, or the process of “iterative dialogue.”

The preceding Part’s discussion of the once-chastened strategy highlights that ILAB is not engaged in a traditional social science exercise. In traditional social science modeling, the analyst typically examines a body of existing data (comparative and/or time series) and seeks to find correlation or causation among independent and dependent variables. Initiatives like ours and the school reform project summarized in sub-part 11.5 are different from traditional social science in at least two ways. First, because they are iterative policy initiatives based on the application of indicators, they entail intervention in the social world and, if the intervention is successful, transformation in the object of study.⁸⁸ That is, the

⁸⁸ This, of course, is a longstanding insight about policy interventions, especially interventions that make use of targets. Pessimistic versions of the insight go under the names of the Lucas Critique, Goodhart’s Law, and

analyst may have the capacity to influence independent or control variables, at least over a sufficiently long period of time – and therefore the capacity to test whether those variables have the desired effect (assuming they are within the proximate control of the government) or whether they in fact impede or propel government efforts (assuming they are not within the government’s control).⁸⁹

Second, the analyst has the capacity to communicate with the entities whose performance is being measured.⁹⁰ In iterative communication, the analyst has an opportunity to cajole the entities to both expand the data set and proffer explanations for poor performance. The entities have an incentive to provide exculpatory data and explanations.

In our case, U.S. officials could inform the trading partner’s officials of the U.S. government’s provisional diagnostic appraisal or provisional comprehensive evaluation, triggering a dialogue over data and variables in the local context that might explain, justify, or excuse the appraisal. The new data and explanations may either change the outcome of the appraisal or induce revisions of the Indicators for the next round of appraisal or both.

While the process of iterative dialogue may apply to both diagnosis and evaluation, the process may have different advantages in the two stages. There are two apparent advantages in using the process for probative or diagnostic purposes. First, diagnostic

Campbell’s Law, as mentioned above. The latter concepts, however, focus more on the perverse or unexpected effects of policy targets than on the iterative feedback on learning by policy-makers.

⁸⁹ The process of iterative dialogue accords with the conclusion of Part 8 – namely, that it is mistaken to begin with a surfeit of *a priori* control variables, or “adjustment factors.” What we assume to be a hard constraint that renders the government helpless to improve its compliance might not be at all.

⁹⁰ Contemporary social scientists are quite aware that data is “co-generated” by the analyst and the analysand. The argument in the text is that this process can self-consciously be made central to labor diagnosis and assessment that seek to incentivize the production of useful data, the production of useful deliberation about the significance of the data, and the production of useful improvements in the institutions of legal enforcement.

appraisal is less statistically rigorous than evaluation. Indicators and explanations can be more fluid and informal, and are therefore compatible with the give and take of interchange between U.S. officials and trading partner's officials and stakeholders. Second, iterative dialogue is useful not only for identifying priority countries that warrant full assessment, but also for gauging the utility of technical assistance and future research. Technical assistance calls for commitment and ownership by the host country, which can only be determined and achieved by dialogue. In-depth research into country conditions may also call for the trading partner's commitment to cooperate with surveys, sampling, and other data collection.

On the other hand, in the assessment stage, the incentive for trading partners to engage in disciplined presentation of data and of justifications for weak performance may be particularly high-powered. Since the stakes are higher when the trading partner faces comprehensive evaluation, its officials may feel greater urgency to produce and disclose pertinent data and to craft precise arguments that join issue with the U.S. government's provisional assessment.

How might the iterative dialogue be structured? Although Part 4 argued that transparency is critical to capacity-building, the first stage of iterative dialogue is best conducted in a non-transparent manner, for at least two reasons. First, if the "dialogue" is conducted in a public forum, there is a good chance it will devolve into grandstanding and other rhetorical posturing not related to reasoned presentation and analysis of the data. Second, if local officials expect that their justificatory arguments and data will be made public, they may be less forthcoming. Government officials may fear that, should the analysts reject

their data and analysis, they will be publicly repudiated or, worse, penalized through a comprehensive evaluation and trade sanctions triggered by their own flawed arguments. If conducted in a non-transparent way, government officials will have stronger incentives to provide accurate, relevant data and press rational arguments, since they know that the ILAB analysts will, in the next iteration, produce new Indicators and appraisals based on the analysts' evaluation of the best data and justificatory arguments. That is, trading partners' officials have nothing to gain by posturing or withholding, since it is the ILAB analysts who control the final diagnosis or assessment in the current round, the subsequent revision (if any) of the Indicators, and the subsequent rounds of diagnosis and assessment.

This strategy may seem over-optimistic — even putting aside the political obstacles to such iterative dialogue. If trading partners' officials know that in the probative or diagnostic phase weak or negative data will trigger high-stakes comprehensive assessments, they may yield to the more powerful incentive to not collect or to suppress potentially damaging data. There are two responses: First, while government officials may indeed face that incentive, they will still be motivated to produce whatever accurate, verifiable data and analysis weighs in their favor. The “dialogue” will thus have flushed out some relevant data, even if the variables so defended have a systematic bias in favor of the trading partner's interests. The ILAB analysts are unlikely to be flummoxed by such predictable bias. To the contrary, the non-transparent dialogue will be an occasion in which the ILAB analysts will press the trading partner for more complete data. ILAB analysts can make the case that relevant missing data may be presumed to cut against the trading partner's purported justifications for poor

performance. And, indeed, many of the Indicators directly impose a negative score for the trading partner's failure to convincingly and verifiably demonstrate compliance with specific aspects of its enforcement obligations.

Second, the Indicators are strongly weighted toward requirements that trading partners collect and publish accurate data on a *systemic* level – that is, on the level of entire labor markets and enforcement institutions, as well as on the level of specific legal rules and enforcement efforts – and that they continuously improve such systemic collection and publication of data. Further, the Indicators require transparency and stakeholder participation in *that* process of systemic collection and systemic improvement. In other words, while the trading partner's data and analysis, offered in initial dialogue with ILAB analysts in response to *particular* performance failures, should be non-transparent, just the opposite is true of the trading partner's *general* systems of data collection.⁹¹

Hence, over time, the trading partners' failure in the non-transparent dialogue to produce well-targeted data and to make compelling arguments against poor provisional appraisals will be more and more damning, in tandem with its ongoing failure to develop a transparent system for collecting accurate, verifiable data that goes beyond self-serving arguments. Or, to put it conversely, if the trading partner fails over time to meet Indicators for the development of a sound, public data-collection system, then its private arguments –

⁹¹ In the non-transparent dialogue, the trading partner's officials may attempt to justify their ongoing failure to collect transparent systemic data. If other trading partners are able transparently to produce such systemic data, then ILAB analysts are unlikely to excuse the country in question, let alone produce refined country-specific Indicators in the next iteration that respond positively to the country's excuses – and the trading partner knows as much.

based on fragmentary, non-transparent data proffered in defense of particular problems in its substantive enforcement – will have predictably weaker purchase on ILAB analysts' probes, diagnoses, assessments, and potential Indicator revisions.

Having said this, it is true that trading partners will be readier to produce complete data, even potentially or actually negative data, when the dialogue is low-stakes rather than high-stakes. That is, if a negative diagnosis triggers positive technical assistance rather than a comprehensive assessment that risks punitive sanctions, then trading partners are less likely to deliberately suppress or fail to collect negative data. Even then, however, the disincentive to produce negative data remains, so long as U.S. trade legislation and agreements authorize the imposition of trade penalties, leading trading partners to fear that data produced for diagnostic purposes could *potentially* be used for punitive purposes as well. Nonetheless, as discussed in sub-part 10.2, there is a countervailing incentive if trading partners know that nongovernmental or international organizations may produce damaging information; the trading partner then risks a negative appraisal both for its substantive violation and for its failure to produce data.

The iterative methodology just described seems unabashedly technocratic, in two senses: First, the dialogues between ILAB analysts and trading partners' officials are explicitly non-transparent. Second, the ultimate use of data and argumentation provided to ILAB analysts in those dialogues remains fully in ILAB's control. ILAB decides whether and how the trading partner's data and arguments will be used in a final appraisal in the current round and in subsequent appraisals in later rounds. ILAB also decides whether such data

and arguments will be incorporated in revised country-specific Indicators or revised weighting of Indicators – whether, that is, the data and arguments will alter the endogenous or exogenous variables in the overall model of enforcement (even if that “model” is not articulated and tested as a rigorous, freestanding multivariate model, but instead is merely an implicit model embodied in the body of Indicators and the weights attached to them).

Nonetheless, the methodology can still be substantially democratic or participatory. ILAB can build alternative channels for consultation with and participation by relevant stakeholders, particularly worker representatives. One possible model is a series of hub-and-spoke dialogues. ILAB is the hub; and various stakeholders in the country being assessed could engage with ILAB along separate spokes.

These separate hub-and-spoke dialogues might then be followed by a general public hearing or forum, in which all stakeholders participate. Such a second-stage public hearing would be most warranted in the case of evaluative assessments, when the stakes are high. Transparent public hearings are less significant in the case of diagnostic appraisals aimed at identifying further research or technical assistance.

Why this two-stage model, rather than simply beginning with a general private or public forum? Again, there is a risk that the usefulness of a premature multi-stakeholder forum will be undermined by grandstanding and posturing among government officials, union representatives, employer federations, and other actors. These actors may have interests that conflict across a range of issues beyond those at stake in the application and revision of a rigorous, precise Indicator methodology. In a general forum they may for political reasons

feel compelled to assert those interests in ways that distract from rational presentation of data and argumentation about the data's relevance.

The potential effectiveness of hub-and-spoke dialogue is suggested by the results of the Harvard Global Labor Survey. (Chor and Freeman, 2005). The survey asked for expert opinions about the strength and effectiveness of labor regulation. The researchers concluded that there was very low variance in responses between anonymous experts aligned with labor interests and those aligned with employer interests. One might suppose that if those opinions were initially offered in a public group forum, the variance would be higher and the survey results less reliable.

Nonetheless, wide-open participation and interchange among stakeholders is essential for purposes of (1) developing *general*, transparent systems of data collection and analysis not immediately tethered to any iteration of diagnostic or evaluative application of Indicators; (2) the trading partner's formulation of its own domestic indicators and targets for measuring the effective implementation of such systems – as required by the new Capacity-Building Indicators; and (3) as just described, multistakeholder discussion of ILAB's evaluative appraisal, at a second stage that follows the first stage of closed-door hub-and-spoke dialogues between the U.S. government and the various stakeholders of the target country.⁹²

⁹² The value of such public interchange is demonstrated by the ILO's statistical conferences, which are informed not only by the secretariat's "technocratic" papers but also by well-focused and frequently compelling "democratic" input from tripartite meetings of representatives selected by governments, worker organizations, and employer federations.

13. Next Steps



There are several possibilities for carrying forward this paper's research program:

- (1) Exploring strategies to pare down the list of candidate Assessment Indicators, without losing the comprehensive coverage of, and balance among, all significant facets of each labor right.
- (2) Exploring strategies to pare down and revise the list of candidate Probative Indicators and Diagnostic Indicators to ensure their effective application in data-poor environments.
- (3) Exploring methods to facilitate ILAB analysts' application of the large bodies of Indicators, including:
 - (a) Further elaboration of the Annotations that explain each Indicator.
 - (b) Adding various symbols or keys after each Indicator, to alert the analyst to the type of norm contained in that Indicator (e.g., fixed bright-line rule, revisable standard, etc.), for purposes of the analysts' further specification or ongoing revision of the Indicator, as proposed above in sub-part 3.3.
 - (c) Drafting a Handbook on Indicator Application, to aid analysts in consistently and expeditiously applying the Indicators.

- (d) Identifying sources targeted at each Indicator (in contrast to WebMILS's use of links to the home pages of websites), including country-specific primary legal material, treatises, legal research guides, and overviews of the structure, function, and resources of legal institutions; country-specific sources on labor market variables and trends, and on recent political and economic developments; and other pertinent country-specific data sources.
 - (e) Designing software for the entry, display, storage, and updating of the country-specific sub-indicators and pertinent information sources.
 - (f) Exploring the possibility of partnerships between research institutions and ILAB, for purposes of applying Indicators and accumulating country data and information sources. For example, law students taking seminars on international labor rights might conduct such research under their professor's supervision.
- (4) Operationalizing the preferred strategies for weighting and aggregating Indicators, canvassed in Part 11.
- (5) Sensitivity-testing the body of Assessment Indicators in tandem with operationalizing the alternative strategies for weighting and aggregation.
- (6) Pilot-testing analysts' application of the bodies of Indicators and the alternative weighting and aggregation strategies, following the methodology of the Michigan evaluation of the original NAS matrix methodology.

(7) Carrying out intensive case studies of selected countries at different levels of development, for several possible purposes:

- (a) To examine available data sources, explore the relationship between input and output Indicators, and probe for control variables other than level of economic development.
- (b) To examine the practicality of trading partners' satisfying the data-production requirements of the Twin Indicators and the Capacity-Building Indicators.
- (c) To examine the feasibility of the process of iterative dialogue between the United States and trading partners discussed above in Part 12.

ACRONYMS



CMILS	Committee on Monitoring International Labor Standards
GSP	Generalized System of Preferences
DOL	Department of Labor
ILAB	Bureau of International Labor Affairs
ILO	International Labor Organization
ITUC	International Trade Union Confederation
NAS	National Academy of Sciences
OECD	Organization for Economic Cooperation and Development
UN	United Nations
UNCESCR	United Nations Committee on Economic Social and Cultural Rights
UNHCHR	United Nations High Commissioner for Human Rights
USTR	United States Trade Representative
WebMILS	Website for Monitoring International Labor Standards

BIBLIOGRAPHY



Abramo L. and Valenzuela, M.E. (2005). *Women's Labor Force Participation Rates in Latin America*, *International Labor Review*, Vol. 144, No. 4, pp. 369-400.

Addison, J., Teixeira, P., and Zwick, T. (2010). *German Works Councils and the Anatomy of Wages*, *Industrial and Labor Relations Review*, Vol. 63, No. 2, pp. 247-270.

Aghion, P., Burgess, R., Redding, S., and Zilibotti, F. (2008). *The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India*, *American Economic Review*, Vol. 98, No. 4, pp. 1397–1412.

Ahlering, B. and Deakin, S. (2007). *Labor Regulation, Corporate Governance and Legal Origin: A Case of Institutional Complementarity?* *Law and Society Review*, Vol. 41, pp. 865-908.

Ahmad, S. (2002). *Law and Labor Market Regulation in Malaysia*, in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Ahmed, I. (2003). *Decent Work and Human Development*, *International Labor Review*, Vol. 142, No. 2, pp. 263-271.

Akandji-Kombe, J.-F. (2005). *The Material Impact of the Jurisprudence of the European Committee of Social Rights*, in De Burca, G. and De Witte, B., eds., *Social Rights in Europe*. Oxford: Oxford University Press.

Albracht, G. (2005). *Integrated Labor Inspection Systems: The Strategy of the ILO*, *Labor Education*, No. 140-141, pp. 67-73.

Algan Y., Dustmann, C., Gritz, A., and Manning, A. (2009) *The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom*, *The Economic Journal*, Vol. 120, No. 542, pp. F4-F30.

Alimenti, J. (2008). *Captive Audience Speech: Argentina*, *Comparative Labor Law and Policy Journal*, Vol. 29, No. 2 (Winter 2008).

Allen, L. (2004). *Open Season on Brazil's Labor Inspectors*, *International Journal of Occupational and Environmental Health*, Vol. 10, No. 2, p. 240.

Almeida, R. and Carneiro, P. (2007). *Inequality and Employment in a Dual Economy: Enforcement of Labor Regulation in Brazil*, IZA Discussion Paper No. 3094. Bonn: Institute for the Study of Labor.

Alston, P. (2005). *The Contribution of the EU's Fundamental Rights Agency to the Realization of Economic and Social Rights*, in Alston, P. and de Schutter, O., eds., *Monitoring Fundamental Rights in the EU: The Contribution of the Fundamental Rights Agency*. Oxford: Hart.

_____. (2005b). *Assessing the Strengths and Weaknesses of the European Social Charter's Supervisory System*. NYU School of Law, Center for Human Rights and Global Justice Working Paper: Economic, Social and Cultural Rights Series, No. 6/2005.

Amable, B. (2003). *The Diversity of Modern Capitalism*. Oxford: Oxford University Press.

Anees, M. (2009). *A Decomposition Analysis of Immigrants-Natives Wages in the UK: Evidence from the Quarterly Labor Force Survey, 2009 (October 8, 2009)*. SSRN.

Anker, R. (2005). *A New Methodology for Estimating Internationally Comparable Poverty Lines and Living Wage Rates*, ILO Policy Integration Department, Statistical Development and Analysis Group, Working Paper No. 2. Geneva: ILO.

Anker, R., Chernyshev, I., Egger, P., Mehran, F., Ritter, J. (2002). *Measuring Decent Work With Statistical Indicators*, ILO Policy Integration Department, Statistical Development and Analysis Group, Working Paper No. 2. Geneva: ILO.

Anner, M. 2008. *Meeting the Challenge of Industrial Restructuring: Labor Reform and Enforcement in Latin America*, *Latin American Politics and Society*, Vol. 50, No. 2, pp. 33–65.

Anonymous. (2009). *Provisional Draft Report on Strategies to Compile and Test an Index to Measure Government Efforts to Combat Child Labor*. (Unpublished).

Anonymous. (2010). *Indicators of Freedom of Association*. (Unpublished).

Armour, J., Deakin, S., Lele, P., and Siems, M. (2009). *How Do Legal Rules Evolve? Evidence from a Cross-National Comparison of Shareholder, Creditor and Worker Protection*, *American Journal of Comparative Law*, Vol. 57, pp. 579-630.

Bacugnan, P. and Ofreneo, R. (2002). *Development of Labor Law and Labor Market Policy in the Philippines*, in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Baliamoune-Lutz, M. and McGillivray, M. (2009). *Does Gender Inequality Reduce Growth in Sub-Saharan African and Arab Countries?*, *African Development Review*, Vol. 21, No. 2, pp. 224-242.

Barenberg, M. (2001). *Labor Federalism in the United States: Lessons for Coordinated Decentralization in Supranational Regimes*, in D. Esty and D. Geradin, eds., *Regulatory Competition and Economic Integration: Comparative Perspectives*, Oxford: Oxford University Press.

_____. (2009) *Summary and Analysis of Interviews with ILAB Staff (Submitted to the U.S. Department of Labor, Bureau of International Labor Affairs)*.

_____. (2010). *Literature Review and Bibliography for Research Project: Refining the NAS-ILAB Matrix (Submitted to the U.S. Department of Labor, Bureau of International Labor Affairs)*.

Barnard, C. (2008). *New Developments in Employment Discrimination Law: The United Kingdom Report*, in Blanpain, R., ed., *New Developments in Employment Discrimination Law, Bulletin of Comparative Labor Relations No. 68*. The Hague: Kluwer.

Batt, R., Holman, D., and Holtgrewe, U. (2009). *The Globalization of Service Work: Comparative Institutional Perspectives on Call Centers, Industrial and Labor Relations Review*, Vol. 62, No. 4, pp. 453-488.

Battu, H. and Zenou, Y. (2009). *Oppositional Identities and Employment for Ethnic Minorities: Evidence from England*, *The Economic Journal*, Vol. 120, No. 542, pp. F52-F71.

Beaudonnet, X. (2005). *L'utilisation des sources universelles du droit international du travail par les juridictions internes*, in *Bulletin de droit compare du travail et de la securite sociale*, pp. 43-48. Bordeaux.

Bederman, D. (2008). *Medellin's New Paradigm for Treaty Interpretation*, 102 *American Journal of International Law* 529 (July, 2008).

Belot, M. and Van Ours, J. (2000). *Does the Recent Success of Some OECD Countries in Lowering Their Unemployment Rates Lie in the Clever Design of Their Labor Market Reforms?* *IZA Discussion Paper No. 147, Bonn (April 2000)*.

Berg, J. and Kucera, D., eds. (2008). *In Defense of Labor Market Institutions: Cultivating Justice in the Developing World*. Houndmills: Palgrave MacMillan.

Bernhardt, A., Milkman, R., et al. (2009) *Broken Laws, Unprotected Workers: Violations of Employment and Labor Laws in America's Cities*, National Employment Law Project.

Bertozzi, F. (2007). *Measuring Flexicurity: Conceptual and Data Availability Challenges at the Macro Level*, RECWOWE Paper (May 2007).

Bescond, D., Chataignier, A, and Mehran, F. (2003). *Seven Indicators to Measure Decent Work: An International Comparison*, *International Labor Review*, Vol. 142, No. 2, pp. 79-212.

Besley, T. and Burgess, R. (2004). *Can Labor Regulation Hinder Economic Performance? Evidence from India*, *Quarterly Journal of Economics*, Vol. 119, No. 1, pp. 91-134.

Bhattacharjea, D. (2006). *Labor Market Regulation and Industrial Performance in India: A Critical Review of the Empirical Evidence*, Working Paper No. 141, Delhi Center for Development Economics.

_____. (2008). *Wage Developments, Minimum Wages and Wage Flexibility in India*, Presentation at the European Foundation for the Improvement of Living and Working Conditions Workshop, OECD Center, Paris, November 27-28, 2008.

Blanpain, R., ed. (1977-2010). *International Encyclopaedia for Labor Law and Industrial Relations*. Deventer: Kluwer.

_____, ed. (2003). *Collective Bargaining, Employment Discrimination, Social Security and European Integration*, *Bulletin of Comparative Labor Relations*, No. 48. The Hague: Kluwer.

_____, ed. (2004). *The Actors of Collective Bargaining: A World Report*, *Bulletin of Comparative Labor Relations*, No. 51. The Hague: Kluwer.

_____, ed. (2004b). *Labor Relations in the Asia-Pacific Countries*. *Bulletin of Comparative Labor Relations*, No. 49. The Hague: Kluwer.

_____, ed. (2005). *Collective Bargaining and Wages in Comparative Perspective*, *Bulletin of Comparative Labor Relations*, No. 56. The Hague: Kluwer.

_____, ed. (2007). *Decentralizing Industrial Relations and the Role of Labor Unions and Employee Representatives*, *Bulletin of Comparative Labor Relations*, No. 61. The Hague: Kluwer.

_____, ed. (2008). *New Developments in Employment Discrimination Law, Bulletin of Comparative Labor Relations, No. 68*. The Hague: Kluwer.

_____, ed. (2008b). *Globalization and Employment Relations in the Auto Assembly Industry, Bulletin of Comparative Labor Relations, No. 64*. The Hague: Kluwer.

_____, ed. (2009). *The Modernization of Labor Law and Industrial Relations in Comparative Perspective, Bulletin of Comparative Labor Relations, No. 70*. The Hague: Kluwer.

_____, ed. (2009b). *Employment Policies and Multilevel Governance, Bulletin of Comparative Labor Relations, No. 71*. The Hague: Kluwer.

_____, ed. (2009c). *European Works Councils, Bulletin of Comparative Labor Relations, No. 72*. The Hague: Kluwer.

_____, ed. (2009d). *The Laval and Viking Cases, Freedom of Services and Establishment v. Industrial Conflict in the European Economic Area and Russia, Bulletin of Comparative Labor Relations, No. 69*. The Hague: Kluwer.

Blanpain, R., Bissom-Rapp, S., Corbett, W., Josephs, H., and Zimmer, M., eds. (2007). *The Global Workplace: International and Comparative Employment Law: Cases and Materials*. Cambridge: Cambridge University Press.

Blanpain, R., Pennings, F., and Sural, N., eds. (2006). *Flexibilization and Modernization of the Turkish Labor Market, Bulletin of Comparative Labor Relations, No. 59*. The Hague: Kluwer.

Block, R. (2005). *Indicators of Labor Standards: An Overview and Comparison, ILO Policy Integration Department, Statistical Development and Analysis Group, Working Paper No. 54*. Geneva: ILO.

_____. (2007). *Indicators of Labor Standards: An Overview and Comparison, in Kucera, D., ed., Qualitative Indicators of Labor Standards: Comparative Methods and Applications, pp. 27-56*. Dordrecht: Springer.

Block, R. and Berg, P. (2009). *Joint Responsibility Unionism: A Multi-Plant Model of Collective Bargaining under Employment Security, Industrial and Labor Relations Review, Vol. 63, No. 1, pp. 60-81*.

Block, R., Berg, P., and Roberts, K. (2003). *Comparing and Quantifying Labor Standards in*

the U.S. and the European Union, The International Journal of Comparative Labor Law and Industrial Relations, Vol. 9, No. 4, pp. 441-467.

Bonnet, F., Figueiredo, J., Standing, G. (2003). *A Family of Decent Work Indexes, International Labor Review*, Vol. 142, No. 2, pp. 213-238.

Booth, A. and Leigh, A. (2010). *Do Employers Discriminate by Gender? A Field Experiment in Female-Dominated Occupations. CPER Discussion Paper No. DP7638.*

Botero, J., Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2004). *The Regulation of Labor, Quarterly Journal of Economics*, Vol. 119, pp. 1340-1382.

Boyer, R. (2004). *New Growth Regimes, But Still Institutional Diversity, Socio-Economic Review*, Vol. 2, No. 1, pp. 1-32.

Bronstein, A. (2005). *The New Labor Law of the Russian Federation, International Labor Review*, Vol. 144, No. 3, pp. 291-318.

_____. (2005b). *The Role of the International Labor Office in the Framing of National Labor Law, Comparative Labor Law and Policy Journal*, Vol. 26, No. 3, pp. 339-369.

_____. (2009). *International and Comparative Labor Law: Current Challenges.* Houndmills/Geneva: Palgrave/ILO.

Bronstein, A. and Thomas, C., eds. (1995). *International and European Labor Standards in Labor Court Decisions, and Jurisprudence on Sex Discrimination, Labor Management Relations Series No. 82.* Geneva: ILO.

Brown, G. and O'Rourke, D. (2003). *Occupational Safety and Health in China, International Journal of Occupational and Environmental Health*, Vol. 9, No. 4.

Bhuiyan, A. and Haq, M. (2008). *Improving Occupational Safety and Health in Bangladesh, International Journal of Occupational and Environmental Health*, Vol. 14, No. 3, pp. 231-234.

Buchanan, S., Krantz, A., Klempner, S., and Alvarado, R. (2006). *Use of a National Reporting System for Occupational Injuries in Costa Rica, International Journal of Occupational and Environmental Health*, Vol. 12, No. 2, pp. 142-147.

Caballero, R., Cowan, K., Engel, E., and Micco, A. (2004). *Effective Labor Regulation and Microeconomic Flexibility.* Department of Economics, MIT. Cambridge, MA. (unpublished).

Caraway, T. (2004). *Protective Repression, International Pressure, and Institutional Design:*

Explaining Labor Reform in Indonesia, Studies in Comparative International Development, Vol. 39, No. 1, pp. 28-49.

_____. (2006). *Freedom of Association: Battering Ram or Trojan Horse?*, *Review of International Political Economy*, Vol. 13, No. 2, pp. 210-232.

_____. (2008). *Explaining the Dominance of Legacy Unions in New Democracies: Comparative Insights from Indonesia*, *Comparative Political Studies*, Vol. 41, No. 10, pp. 1371-1379.

_____. (2009). *Labor Rights in Asia: Progress or Regress?*, *Journal of East Asian Studies*, Vol. 9., pp. 153-186.

Carnes, M. (2008). *The Politics of Labor Regulation in Latin America*, PhD Dissertation, Department of Political Science, Stanford University.

_____. (2009). *Labor Markets, Worker Organization, and Variation in Labor Codes in Latin America*, Prepared for Presentation at the 2009 Annual Meeting of the American Political Science Association, Toronto, Canada, September 3-6, 2009.

_____. (2009b). *Institutionalizing Inequality: The Political Origins of Labor Codes in Latin America*, Kellogg Institute for International Studies Working Paper No. 363 (December, 2009).

Carnes, M. and Mares, I. (2007). *The Welfare State in Global Perspective*, in Boix, C. and Stokes, S., eds., *Handbook of Comparative Politics*. New York: Oxford University Press.

Carr, M. and Chen, M. (2004). *Globalization, Social Exclusion and Gender*, *International Labor Review*, Vol. 143, No. 1-2, pp. 129-160.

Carrion, V. (2009). *Comentarios a consolidacao das leis do trabalho*. 34th ed. Sao Paulo: Revistas dos Tribunals.

Casale, G. (2003). *Collective Bargaining and the Law in Central and Eastern Europe: Recent Trends and Issues*, in Blanpain, R., ed., *Collective Bargaining, Discrimination, Social Security, and European Integration*, *Bulletin of Comparative Labor Relations No. 48*. The Hague: Kluwer.

Celin, M. (2003). *European Employment Strategy: The Right Answer for Candidate Countries?* *Transfer: European Review of Labor and Research*, Vol. 9, No. 1, pp. 88-100.

Chan, A. (2008). *China's Trade Unions in Corporatist Transition*, in Unger, J., ed., *Associations and the Chinese State: Contested Spaces*. Armonk: M.E. Sharpe.

Chan, A. and Wang, H. (2004/2005). *The Impact of the State on Workers' Conditions – Comparing Taiwanese Factories in China and Vietnam*, *Pacific Affairs*, Vol. 77, No. 4, pp. 629-646.

Chataignier, A. (2005). *Statistical Indicators Relating to Social Dialogue: A Compilation of Multiple Country Databases*, *Statistical and Development Analysis Group, Policy Integration Department, International Labor Office*. Geneva: ILO.

Chatton, G.T. (2006). *The Collective Complaints Mechanism within the European Social Charter: Making Economic and Social Rights Really Matter*, in Besson, S., Hottelier, M., Werro, F., eds., *Human Rights at the Center*. Geneva: Edition Schulthess.

Chen, M.-S. and Chan, A. (2010). *Occupational Safety and Health in China: The Case of State-Managed Enterprises*, *International Journal of Health Services*, Vol. 40, No. 1, pp. 43-60.

Chen, S.-J., Ko, J.-J., and Lawler, J. (2003). *Changing Patterns of Industrial Relations in Taiwan*, *Industrial Relations*, Vol. 42, No. 3, pp. 315-340.

Chiao, C.-K. (2008). *Employment Discrimination in Taiwan*, in Blanpain, ed., *New Developments in Employment Discrimination Law, Bulletin of Comparative Labor Relations No. 68*. The Hague: Kluwer.

Chrystal, K. and Mizen, P. (2001). *Goodhart's Law: Its Origins, Meaning and Implications for Monetary Policy*, *Paper for Festschrift in Honor of Charles Goodhart*, Bank of England, November 15-16, 2001.

Churchill, R. and Khaliq, U. (2004). *The Collective Complaints System of the European Social Charter – An Effective Mechanism for Ensuring Compliance with Economic and Social Rights?*, *European Journal of International Law*, Vol. 15, p. 417.

Chor, D. and Freeman, R.B. (2005). *The 2004 Global Labor Survey: Workplace Institutions and Practices Around the World*, *NBER Working Paper No. 11598*.

Collier, R. and Handlin, S., ed. (2009). *Reorganizing Popular Politics: Participation and the New Interest Regime in Latin America*. University Park: Penn State University Press.

Cook, L. (2001). *Trade Unions, Management, and the State in Contemporary Russia*, in Rutland, P., ed., *Business and the State in Contemporary Russia*. Jackson: Westview Press.

Cook, M.L. (2007). *The Politics of Labor Reform in Latin America: Between Flexibility and*

Rights. University Park, Pennsylvania: Pennsylvania State University Press.

Cooney, S., Lindsey, T., Mitchell, R. and Zhu, Y. (2002). *Labor Law and Labor Market Regulation in East Asian States* in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Cooney, S. and Mitchell, R. (2002). *What is Labor Law Doing in East Asia?* in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Council of Europe. (1953). *Convention for the Protection of Human Rights and Fundamental Freedoms, Adopted 4 November 1950, Entered into Force, 3 September 1953. CETS No.: 005*.

_____. (1996). *European Social Charter (revised). Strasbourg, 3.V.1996. CETS No. 163*.

_____. (1998). *Additional Protocol of 1995 to the European Social Charter Providing for a System of Collective Complaints, Entry Into Force 1/7/1998, CETS No. 158*.

_____. (2005). *Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms, Adopted on November 4, 2000 and Entered into Force on April 1, 2005*.

_____. (2008). *Digest of the Case Law of the European Committee of Social Rights*.

Crouch, C. (2005). *Capitalist Diversity and Change: Recombinant Governance and Institutional Entrepreneurs*. Oxford: Oxford University Press.

Cuyvers, L. and Van Den Bulcke, D. (2007). *The Quantification of Respect for Selected Core Labor Standards: Towards a Social Development Index?* in Kucera, D., ed., *Qualitative Indicators of Labor Standards: Comparative Methods and Applications*, pp.115-144. Geneva: ILO.

Da Silva, H. (2009). *Curso de direito do trabalho aplicado. 10 Volumes*. Rio de Janeiro: Elsevier-Campus Juridico.

Däubler, W. and Wang, Q. (2009). *The New Chinese Employment Law*, *Comparative Labor Law and Policy Journal*, Vol. 30, No. 2 (Winter 2009).

Davis, K. (2004). *What Can the Rule of Law Variable Tell Us About Rule of Law Reforms?*, *Michigan Journal of International Law*, Col. 26, No. 1, pp. 141-161.

Davis, K. and Kruse, M. (2007). *Taking the Measure of Law: The Case of the Doing Business Report, Law and Social Inquiry*, vol. 32, pp. 1095-1119.

Davis, K., Merry, S., and Kingsbury, B. (2010). *Indicators as a Technology of Global Governance*. (unpublished draft).

Davoine, L., Erhel, C., and Guergoat-Lariviere, M. (2008). *Monitoring Quality in Work: European Employment Strategy Indicators and Beyond, International Labor Review*, Vol. 147, No. 2-3, pp. 163-198.

_____, _____, and _____. (2008b). *A Taxonomy of European Labor Markets Using Quality Indicators, Centre d'Etudes de L'Emploi, Rapport de Recherche No. 45 (Mai 2008)*.

Deakin, S. (2009). *Legal Origin, Juridical Form and Industrialization in Historical Perspective: The Case of the Employment Relationship and the Joint Stock Company, Socio-Economic Review*, Vol. 7, pp. 35-65.

Deakin, S., Lele, P., and Siems, M. (2007). *The Evolution of Labor Law: Calibrating and Comparing Regulatory Regimes, Cambridge University, Center for Business Research, Working Paper No. 352*.

Deakin, S., Lele, P., and Siems, M. (2007b). *CBR Labor Regulation Index for the UK, the US, Germany, France, and India. Cambridge University, Center for Business Research (December 2007)*.

Deakin, S. and Reberioux, A. (2009). *Corporate Governance, Labor Relations and Human Resource Management in Britain and France: Convergence or Divergence*, in Touffut, J.-P., ed., *Does Company Ownership Matter?* Cheltenham: Edward Elgar.

Deakin, S. and Sarkar, P. (2008). *Assessing the Long-Run Economic Impact of Labor Law Systems: A Theoretical Reappraisal and Analysis of New Time Series Data, Industrial Relations Journal*, Vol. 39, pp. 453-487.

De Burca, G. (2007). *Reconceiving Law and New Governance, Columbia Journal of European Law*, Vol. 13, pp. 519-537.

De Burca, G. and De Witte, B. eds. (2005). *Social Rights in Europe*. Oxford: Oxford University Press.

De Burca, G. and Scott, J., eds. (2006). *Law and the New Governance in the European Union and the United States*. Oxford: Hart Publishing.

Deeg, R. (2007b). *Complementarity and Institutional Change in Capitalist Systems*, *Journal of European Public Policy*, Vol. 14, No. 4, pp. 612-631.

Delclos, G., Bright, K., Carson, A., and Felkner, S. (2005). *A Global Survey of Occupational Health Competencies and Curriculum*, *International Journal of Occupational and Environmental Health*, Vol. 11, No. 2, pp. 185-99.

De Ruyter, A., Singh, A., Warnecke, A., and Zammit, A. (2009). *Core vs. Non-Core Standards, Gender and Developing Countries: A Review With Recommendations for Policy and Practice*, Paper Presented to ILO Conference on Decent Work, Geneva, 8-10 July 2009.

Djankov, S. and Ramalho, R. (2009). *Employment Laws in Developing Countries*, *Journal of Comparative Economics*, Vol. 37, No. 1, pp. 3-13.

Doellgast, V. (2008). *National Industrial Relations and Local Bargaining Power in the US and German Telecommunications Industries*, *European Journal of Industrial Relations*, Vol. 14, No. 3, pp. 265-287.

_____. (2009). *Still a Coordinated Model? Market Liberalization and the Transformation of Employment Relations in the German Telecommunications Industry*, *Industrial and Labor Relations Review*, Vol. 63, No. 1, pp. 3-23.

Doumbia, S. and Meurs, D. (2003). *Gender Equality at Work in Sub-Saharan Africa: A Case Study of Mali's Modern Sector*, *International Labor Review*, Vol. 142, No. 3, pp. 295-316.

Dworkin, R. (1986). *Law's Empire*. Cambridge: Harvard University Press.

Ebbinghaus, B. (1998). *European Labor Relations and Welfare-State Regimes: A Comparative Analysis of their 'Elective Affinities'*, Background Paper for Conference on Varieties of Welfare Capitalism in Europe, North America and Japan, Max Planck Institute for the Study of Societies, Cologne, June 11-13, 1998.

Egger, P. (2003). *Decent Work and Competitiveness: Labor Dimension of Accession to the European Union*, *International Labor Review*, Vol. 142, No. 1, pp. 5-28.

Elson, D. (2006). *Budgeting for Women's Rights: Monitoring Government Budgets for Compliance with CEDAW*. New York: United Nations Development Fund for Women (UNIFEM).

Ernst, C. (2008). *Recent Dynamics in Brazil's Labor Market*. Economic and Labor Market Paper No. 2007/10, Employment Analysis and Research Unit, Economic and Labor Market

Analysis Department. Geneva: ILO.

Erickson, C., Kuruvilla, S., and Ofreneo, R. (2003). *From Core to Periphery? Recent Developments in Employment Relations in the Philippines*, *Industrial Relations*, Vol. 42, No. 3., pp. 368-395

Esping-Anderson, G. (1990). *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.

Estrin, S. and Mickiewicz, T. (2009). *Do Institutions Have a Greater Effect on Female Entrepreneurs?* *IZA Discussion Paper No. 4577 (November 2009)*.

Etchemendy, Sebastian. (2004). *Repression, Exclusion, and Inclusion: Government-Union Relations and Patterns of Labor Reform in Liberalizing Economies*, *Comparative Politics*, Vol. 36, No. 3, pp. 273-290.

Etherington, B. (2005) *The B.C. Health Services and Support Decision—The Constitutionalization of a Right to Bargain Collectively in Canada: Where Did it Come From and Where Will it Lead?*, *Comparative Labor Law and Policy Journal*, Vol. 30, p. 715 (Summer 2009).

European Commission. (2002). *Decision of 18 February 2002 on guidelines for member states' employment policies for the year 2002*, OJ L 060, 1.3.2002.

_____. (2006). *Indicators for Monitoring the Employment Guidelines: 2006 Compendium*. Brussels, Directorate for Employment, Social Affairs and Equal Opportunity.

_____. (2007). *Strategic Report on the Renewed Lisbon Strategy for Growth and Jobs: Launching the New Cycle: 2008-2010*, 11 December 2007, COM(2007)803.

_____. (2008). *Employment in Europe 2008*, Directorate-General for Employment, Social Affairs and Equal Opportunities (October 2008). Luxembourg: Office for Official Publications of the European Communities.

_____. (2009). *Employment Guidelines (2009) – Indicators for Monitoring and Analysis – endorsed by EMCO [Employment Committee] 24/06/09*.

_____. (2009b). *Indicators for Monitoring the Employment Guidelines Including Indicators for Additional Employment Analysis – 2009 Compendium*, 03/11/09.

European Council. (2008). *Decision of 15 July 2008 on Guidelines for the Employment Policies of the Member States*. 2008/618/EC.

_____. (2009). *Decision of 23 June 2009 on Guidelines for the Employment Policies of the Member States. 2009/10421/EC.*

European Court of Human Rights. (2008). *Demir and Baykara v. Turkey, Application no. 34503/97*, 12 November 2008.

_____. (2009). *Enerji Yapi-Yol Sen v. Turkey, Application no. 68959/01*, 21 April 2009.

European Court of Justice. (2007). *Case C-438/05, International Transport Workers' Federation, Finnish Seamen's Union v. Viking Line ABP, OÜ Viking Line Eesti, Opinion of Advocate General Miguel Poiares Maduro, 23 May 2007, ECJ decision*, 11 December 2007.

_____. (2000b). *Case C-341/05, Laval un Partneri Ltd v. Svenska Byggnadsarbetareförbundet, Svenska Byggnadsarbetareförbundet, avd. 1, Svenska Elektrikerförbundet, Opinion of Advocate General Paolo Mengozzi, 23 May 2007, ECJ decision*, 18 December 2007.

European Foundation for the Improvement of Living and Working Conditions. (2007). *Fourth European Working Conditions Survey 2005*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2007b). *Working Conditions in the European Union: The Gender Perspective*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2008). *Working Conditions of an Ageing Workforce*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2008b). *Employment Security and Employability: A Contribution to the Flexicurity Debate*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2008c). *Use of Technology and Working Conditions in the European Union*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2008d). *Working Conditions and Social Dialogue*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2008e). *Final Questionnaire for CAR on Working Conditions and Social Dialogue*. Luxembourg: Office for the Official Publications of the European Communities

_____. (2009). *Convergence and Divergence of Working Conditions in Europe: 1990-2005*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2009b). *A Sector Perspective on Working Conditions*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2009c). *Working Conditions in the European Union: Work Organization*. Luxembourg: Office for the Official Publications of the European Communities.

_____. (2009d). *Annual Review of Working Conditions in the EU 2008/2009*. Luxembourg: Office for the Official Publications of the European Communities.

European Trade Union Institute. (2009). *Benchmarking Working Europe 2009*. Brussels: ETUI.

_____. (2009b). *Minimum Wages in Europe: New Debates Against the Background of Economic Crisis, ETUI Policy Brief, European Economic and Employment Policy, No. 2/2009*.

Ewing, K. and Hendy, J. (2010), *The Dramatic Implications of Demir and Baykara, Industrial Law Journal*, Vol. 39, No. 1.

Eyraud, F. and Saget, C. (2005). *The Fundamentals of Minimum Wage Fixing*. Geneva: ILO.

Fagernäs, S. (2007). *Labor Law, Judicial Efficiency and Informal Employment in India. CBR Working Paper No. 353*. Cambridge, University of Cambridge, Centre for Business Research.

Fauth, R. and McVerry, A. (2008). *Can 'Good Work' Keep Employees Healthy? Evidence from Across the EU*. London: The Work Foundation.

Federal Statistical Office Germany. (2009). *Quality of Employment in Germany: Country Report for the Task Force on the Measurement of Quality of Employment*.

Feldman, H. (2008). *Business Regulation and Labor Market Performance Around the World, Journal of Regulatory Economics*, Vol. 33, No. 2, pp. 201-235.

Fields, G. (2003). *Decent Work and Development Policies, International Labor Review*, Vol. 142, No. 2, pp. 239-262.

Figueroa Gutarra, E. (2009). *Despido arbitrario. Estudio constitucional, comparado y jurisprudencial*. Lima: San Marcos.

Filho, R. and Lobao, R. (2008). *Captive Audience Speech in the Brazilian Labor Law, Comparative Labor Law and Policy Journal*, Vol. 29, No. 4, pp. 341-347.

Flanagan, R. (2006). *Globalization and Labor Conditions: Working Conditions and Worker Rights in a Global Economy*. Oxford: Oxford University Press.

Fontagné, L., Laborde, D. and Mitaritonna, C. (2008). *An Impact Study of the EU-ACP Economic Partnership Agreements (EPAs) in the Six ACP Regions*. CEPII Working Paper No. 2008-04 (April 2008).

Forsyth, A. (2007). *The 'Transplantability' Debate Revisited: Can European Social Partnership be Exported to Australia?*, *Comparative Labor Law and Policy Journal*, Vol. 27, p. 305 (Fall 2007).

France. (2009). *Quality of Employment in France: Country Report for the Task Force on the Measurement of Quality of Employment, 14-16 October 2009, Geneva*.

Frenkel, S. (1993). *Theoretical Frameworks and Empirical Contexts of Trade Unionism*, in Frenkel, S., ed., *Organized Labor in the Asia-Pacific Region: A Comparative Study of Trade Unionism in Nine Countries*. Ithaca: ILR Press.

Frenkel, S. and Kuruvilla, S. (2002). *Logics of Action, Globalization, and Employment Relations Change in China, India, Malaysia, and the Philippines*, *Industrial and Labor Relations Review*, Vol. 55, No. 3, pp. 387-412.

Frers, R. (2004). *Labor in Latin American and European Constitutions*. Peru: Dialogo.

Freund, C. and Bolaky, B. (2008). *Trade, Regulations, and Income*, *Journal of Development Economics*, Vol. 87, No. 2, pp. 309-321.

Frey, D. (2009). *A Diagnostic Methodology for Regulating Decent Work*, Paper Presented to Conference, International Labor Office (July 8-20, 2009).

Fudge, J. (2010). *Brave New Words: Labor, the Courts, and the Canadian Charter of Rights and Freedoms*, *Windsor Yearbook Access to Justice*, Vol. 48, No. 1.

_____. (2010b). *Constitutionalizing Labor Rights in Europe* (unpublished).

Fugazza, M. (2003). *Racial Discrimination: Theories, Facts and Policy*, *International Labor Review*, Vol. 142, No. 4, pp. 507-541.

Gacek, S. (1994). *Revisiting the Corporatist and Contractualist Models of Labor Law Regimes: A Review of the Brazilian and American Systems*, *Cardozo Law Review*, Vol. 16, p. 21 (1994).

Galli, R. and Kucera, D. (2004). *Labor Standards and Informal Employment in Latin America*, *World Development*, Vol. 32, No. 5, pp. 809-828.

Garret, B. and Liebman, J. (2004). *Experimentalist Equal Protection*, *22 Yale Law and Policy Review* 261 (2004).

Gartner, H. and Stephan, G. (2004). *How Collective Contracts and Works Councils Reduce the Gender Wage Gap*. *IAB Discussion Paper 7/2004*. Nuremberg: Institut für Arbeitsmarkt.

Gasparini, L. and Tornarolli, L. (2007). *Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Microdata*. *CEDLAS Working Paper No. 46*. La Plata: Centro de Estudios Distributivos, Laborales y Sociales.

Gatto, A. (2005). *The Integration of Social Rights Concerns in the External Relations of the European Union*, in De Burca, G. and De Witte, B., eds., *Social Rights in Europe*. Oxford: Oxford University Press.

Gaynor, J., ed. (2007). *Employment and Labor Law (2nd Edition)*. Suffolk: Patrick Wilkins.

Ghai, D. (2003). *Decent Work: Concept and Indicators*, *International Labor Review*, Vol. 142, No. 2, pp. 113-145.

Golbert, A. and Nun, Y. (1982). *Latin American Laws and Institutions*. New York: Praeger.

Gori, C. (2005). *Domestic Enforcement of the European Social Charter: The Way Forward*, in de Burca and de Witte, eds., *Social Rights in Europe*. Oxford: Oxford University Press.

Gorman, R. and Finkin, M. (2004). *Basic Text on Labor Law: Unionization and Collective Bargaining (2nd Edition)*. St. Paul: West.

Gravel, E. (2008) *Is the Appropriation of International Labor Standards by New Actors Replacing or Complementing the ILO's Traditional Standards-Related Work? Governance, International Law and Corporate Social Responsibility Research Series No. 116*, pp. 79-103. Geneva: ILO/IILS.

Greenhill, B., Mosley, L., and Prakash, A. (2009). *Trade-Based Diffusion of Labor Rights: A Panel Study, 1986-2002*, *American Political Science Review*, Vol. 103, No. 4, pp. 669-690.

Gravel, E. and Delpich, Q. (2008). *International Labor Standards: Recent Developments in Complementarity Between the International and National Supervisory System*, *International Labor Review*, Vol. 147, No. 4, pp. 403-415.

Grisolía, J. (2009). *Manual de derecho laboral. 5ª Edición*. Buenos Aires: Lexis Nexis.

Guillén A.M. and Dahl S.-A., eds. (2009). *Quality of Work in the European Union -Concept, Data and Debates from a Transnational Perspective*. Brussels: Peter Lang.

Gumedze, S. (2003). *Bringing Communications Before the African Commission on Human and Peoples' Rights*, *African Human Rights Law Journal*, Vol. 3, pp. 118-148.

Gupta, N., Oaxaca, R., and Smith, N. (2006). *Swimming Upstream, Floating Downstream: Comparing Women's Relative Wage Progress in the United States and Denmark*, *Industrial and Labor Relations Review*, Vol. 59, No. 2, pp. 243-266.

Guvenen, F., Koruscu, B., and Ozkan, S. (2009). *Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis*, *NBER Working Paper No. w15526*.

Hall, P.A. and Soskice, D. (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

Hay, D. and Craven, P., eds. (2004). *Masters, Servants and Magistrates in Britain and the Empire, 1562-1955*. Chapel Hill, N.C.: University of North Carolina Press.

Heckman, J. and Pages, C., eds. (2004). *Law and Employment: Lessons from Latin America and the Caribbean*. Chicago: University of Chicago Press.

Heintz, J., Wicks-Lim, J., and Pollin, R. (2006). *Decent Work in America: The State-by-State Work Environment Index 2005*. *University of Massachusetts, Political Economy Research Institute*.

Heyns, C. and Killandar, M. (2006). *The African Regional Human Rights System*, in Isa, F. and de Feyrer, K., eds., *International Protection of Human Rights: Achievements and Challenges*. Bilbao: University of Deusto.

Heyns, C., Strasser, W., and Padilla, D. (2003) *A Schematic Comparison of Regional Human Rights Systems*, *Africa Human Rights Law Journal*, Vol. 3, pp. 76-87.

Higgins, J., ed. (2006-2009). *The Developing Labor Law: The Board, the Courts, and the National Labor Relations Act (5th Edition)*. Washington, D.C.: BNA.

Hood, C., Dixon, R., and Beeston, C. (2008). *Rating the Rankings: Assessing International Rankings of Public Service Performance*, *International Public Management Journal*, Vol. 11, n. 3, pp. 298-328.

Hopkins, K. (2002). *Effect of an African Court on the Domestic Legal Orders of African States*, *Africa Human Rights Law Journal*, Vol. 2, pp. 234-251.

Hoyland, B., Moene, K., and Willumsen, F. (2009). *The Tyranny of International Index Rankings*, *ESOP Working Paper* (July, 2009).

Hyman, R. (2009). *How Can We Study Industrial Relations Comparatively?* in Blanpain, R., ed., *The Modernization of Labor Law and Industrial Relations in Comparative Perspective*. The Hague: Kluwer.

Inter-American Commission of Human Rights. (n.d.). *Office of the Special Rapporteur for Freedom of Expression*, "Access to Public Information in the Americas. Contributions of the Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights."

Inter-American Court of Human Rights. (2003). *Legal Conditions and Rights of Undocumented Migrant Workers*, *Consultative Opinion OC-18/03*, 17 September 2003.

International Association of Labor Inspection. (2001-2008). *Forum*. Luxembourg: IALI.

_____. (2008). *Global Code of Integrity for Labor Inspection*. Luxembourg: IALI.

International Institute for Labor Studies. (2009). *The Financial and Economic Crisis: A Decent Work Response. Preprint Edition*. Geneva: ILO.

International Labor Organization. (1972). *Code of Practice on Safety and Health in Building and Civil Engineering Work*. Geneva: ILO.

_____. (1974). *Code of Practice on Safety and Health in Shipbuilding and Ship Repairing*. Geneva: ILO.

_____. (1977). *Code of Practice on Protection of Workers Against Noise and Vibration in the Working Environment*. Geneva: ILO.

_____. (1979). *Code of Practice Guide to Health and Hygiene in Agricultural Work*. Geneva: ILO.

_____. (1980). *Code of Practice on Occupational Exposure to Airborne Substances Harmful to Health*. Geneva: ILO.

_____. (1984). *Code of Practice on Safety in the Use of Asbestos*. Geneva: ILO.

- _____. (1986). *Code of Practice on Safety and Health in Coal Mines*. Geneva: ILO.
- _____. (1987). *Code of Practice on Radiation Protection of Workers (Ionizing Radiations)*. Geneva: ILO.
- _____. (1991). *Code of Practice on Prevention of Major Industrial Accidents*. Geneva: ILO.
- _____. (1991b). *Code of Practice on Safety and Health in Opencast Mines*. Geneva: ILO.
- _____. (1992). *Code of Practice on Safety and Health in Construction*. Geneva: ILO.
- _____. (1992). *Minimum Wages: Wage-Fixing Machinery, Application and Supervision, Report III (Part 4B), General Survey, International Labor Conference, 79th Session*. Geneva: ILO.
- _____. (1993). *Code of Practice on Safety in the Use of Chemicals at Work*. Geneva: ILO.
- _____. (1996). *Code of Practice on Recording and Notification of Occupational Accidents and Diseases*. Geneva: ILO.
- _____. (1998). *Encyclopaedia of Occupational Health and Safety (fourth edition)*. Geneva: ILO.
- _____. (1998b). *Declaration of Fundamental Rights and Principles at Work, 86th Session, Geneva, June 1998*.
- _____. (1998c). *Code of Practice on Safety and Health in Forestry Work*. Geneva: ILO.
- _____. (1998d). *Resolution Concerning Statistics of Occupational Injuries (Resulting from Occupational Accidents)*, adopted by the 16th International Conference of Labor Statisticians (October, 1998). Geneva: ILO.
- _____. (1999). *Decent Work, Report of the Director-General, International Labor Conference, 87th Session*. Geneva: ILO.
- _____. (2000). *Safety and Health at the Workplace – Trade Union Experiences in Central and Eastern Europe*. Geneva: ILO.

- _____. (2001). *Code of Practice on Ambient Factors in the Workplace*. Geneva: ILO.
- _____. (2001b). *Guidelines on Occupational Safety and Health Management Systems*. Geneva: ILO.
- _____. (2001c). *Code of Practice on Safety in the Use of Synthetic Vitreous Fibre Insulation Wools*. Geneva: ILO.
- _____. (2001d). *Code of Practice on HIV/AIDs and the World of Work*. Geneva: ILO.
- _____. (2002). *Labor Inspection: A Guide to the Profession*. Geneva: ILO.
- _____. (2002b). *Combating Child Labor: A Handbook for Labor Inspectors*. Geneva: ILO.
- _____. (2003). *General Report, Report I, 17th International Conference of Labor Statisticians (November 24 – December 3, 2003), doc. ICLS/17/2003/1*. Geneva: ILO.
- _____. (2003b). *Protection of Wages, Standards and Safeguards Relating to the Payment of Labor Remuneration*. France: ILO Publications.
- _____. (2003b). *Code of Practice on Safety and Health in the Non-Ferrous Metals Industries*. Geneva: ILO.
- _____. (2004). *Report of the Conference, 17th International Conference of Labor Statisticians (November 24 – December 3, 2003), Chapter 2: Measuring and Monitoring Decent Work, Annex: Report of the Working Group on Decent Work Indicators, doc. ICLS/17/2003/4*. Geneva: ILO.
- _____. (2004b). *Statistics on the Employment Situation of People with Disabilities: A Compendium of National Methodologies, ILO Policy Integration Department, Bureau of Statistics, Working Paper No. 40*. Geneva: ILO.
- _____. (2005, reprinted 2008). *Guide to the Cambodian Labor Law for the Garment Industry, Better Factories Cambodia*. Geneva: ILO.
- _____. (2005b). *General Survey of the Reports Concerning Hours of Work, Report III (Part 1B), International Labor Conference, 93rd Sess. 2005*.
- _____. (2005c). *A Handbook on HIV/AIDS for Labor and Factory Inspectors*. Geneva: ILO.

_____. (2005d). *Code of Practice and Training Manual: Guidelines for Labor Judges and Magistrates*. Geneva, ILO.

_____. (2005e). *Code of Practice on Safety and Health in the Iron and Steel Industry*. Geneva: ILO.

_____. (2006). *From Pilot to Decent Work Country Program: Lessons from the Decent Work Pilot Program*. Geneva: ILO.

_____. (2006b). *Digest of Decisions and Principles of the Freedom of Association Committee of the Governing Body of the ILO (Fifth edition, 2006)*. Geneva: ILO.

_____. (2006c). *Special Procedures for the Examination in the International Labor Organization of Complaints Alleging Violations of Freedom of Association, Annex I, Committee on Freedom of Association, Digest of Decisions (Fifth edition, 2006)*. Geneva: ILO.

_____. (2006d). *Strategies and Practice for Labor Inspection. Governing Body, Committee on Employment and Social Policy, 297th Session. Document GB.297/ESP/3*. Geneva: ILO

_____. (2006e). *Labor inspection, General Survey of the reports concerning the Labor Inspection Convention, 1947 (No. 81), and the Protocol of 1995 to the Labor Inspection Convention, 1947, and the Labor Inspection Recommendation, 1947 (No. 81), the Labor Inspection (Mining and Transport) Recommendation, 1947 (No. 82), the Labor Inspection (Agriculture) Convention, 1969 (No. 129), and the Labor Inspection (Agriculture) Recommendation, 1969 (No. 133), Report III (Part 1B), ILC, 95th Session, Geneva, 2006*.

_____. (2006f). *Integrated Labor Inspection Training System*. Geneva: ILO.

_____. (2006g). *A Code of Ethical Behavior for Labor Inspectors*. Geneva: ILO.

_____. (2006h). *Strategies and Practice for Labor Inspection, Governing Body, 297th Session, Geneva, November 2006*.

_____. (2006j). *Code of Practice on Safety and Health in Underground Coalmines*. Geneva: ILO,

_____. (2007). *The Employment of People with Disabilities: Towards Improved Statistical Information*. Geneva: ILO.

_____. (2007b). *Equality at Work: Tackling the Challenges: Global Report under the*

follow-up to the ILO Declaration on Fundamental Principles and Rights at Work, Director-General Report I(B), International Labor Conference, 96th Session, 2007. Geneva: ILO.

_____. (2008). *Measurement of Decent Work: Discussion Paper for the Tripartite Meeting of Experts on the Measurement of Decent Work, doc. TMEMDW/2008.* Geneva: ILO.

_____. (2008b). *General Report, Report I: 18th International Conference of Labor Statisticians (November 24 – December 5, 2008), doc. ICLS/18/2008/I.* Geneva: ILO.

_____. (2008c). *Decent Work Indicators for Asia and the Pacific: A Guidebook for Policy-makers and Researchers.* Geneva: ILO.

_____. (2008d). *Chairperson's Report: Tripartite Meeting of Experts on the Measurement of Decent Work (September 8 to September 20, 2008).* Geneva: ILO.

_____. (2008e). *Global Wage Report 2008/9: Minimum Wages and Collective Bargaining: Towards Policy Coherence.* Geneva: ILO.

_____. (2008f). *General Report, Report II: Measurement of Working Time, 18th International Conference of Labor Statisticians (November 24 – December 5, 2008), doc. ICLS/18/2008/II.* Geneva: ILO.

_____. (2008g). *Freedom of Association: The Lessons Learned: Global Report Under the Follow-up to the ILO Declaration on Fundamental Principles and Rights at Work, Director-General Report I(B), International Labor Conference, 97th Session, 2008.* Geneva: ILO.

_____. (2008h). *Decent Work Country Program Sri Lanka 2008-2012.* Geneva: ILO.

_____. (2008i). *Report of the Independent Final Evaluation: ILO/Japan Asian Regional Program on Expansion of Employment Opportunities for Women - Cambodia and Vietnam Chapters.* Geneva: ILO.

_____. (2008j). *Working Conditions Laws 2006/2007: A Global Review.* Geneva: ILO.

_____. (2008k). *Report of the Meeting of Experts on Labor Statistics: Part 1 – Child Labor Statistics. Part II – Measurement of Working Time, Geneva, 1-10 April 2008.*

_____. (2008l). *ILO International Conference of Labor Statisticians, Resolution I – Resolution Concerning the Measurement of Working Time, 5 December 2008.*

- _____. (2009). *Report of the Director-General, Fifth Supplementary Report: Measuring Decent Work, Governing Body, 306th Session, doc. GB.306/17/5*. Geneva: ILO.
- _____. (2009b). *Workshop Report: Tripartite Consultation Workshop on Decent Work Indicators for Brazil (November 2009)*. Geneva and Brasília: ILO.
- _____. (2009c). *Trabalho Decente no Brasil*. Geneva and Brasília: ILO.
- _____. (2009d). *Decent Work Country Profile: Austria*. Geneva: ILO.
- _____. (2009e). *Delivering Decent Work in Europe and Central Asia: 8th European Regional Meeting, Report of the Director General, Vol. I, Part 1*. Geneva: ILO.
- _____. (2009f). *Delivering Decent Work in Europe and Central Asia: 8th European Regional Meeting, Report of the Director General, Vol. I, Part 2*. Geneva: ILO.
- _____. (2009g). *Towards Decent Work: A Review of ILO Work for 2005-08: 8th European Regional Meeting, Report of the Director General, Vol. II*. Geneva: ILO.
- _____. (2009h). *Global Wage Report: 2009 Update*. Geneva: ILO.
- _____. (2009i). *Report of the Committee of Experts on the Application of Conventions and Recommendations*. Geneva: ILO.
- _____. (2009j). *Report of the Conference Committee on the Application of Standards (International Labor Conference, 98th Session 2009)*. Geneva: ILO.
- _____. (2009k). *Global Employment Trends for Women*. Geneva: ILO.
- _____. (2009l). *Twenty-Third Synthesis Report on Working Conditions in Cambodia's Garment Sector, Better Factories Cambodia*. Geneva: ILO.
- _____. (2009m). *Labor and Social Trends in Sri Lanka 2009*. Geneva: ILO.
- _____. (2009n). *Vietnam Employment Trends Report 2009*. Geneva: ILO.
- _____. (2009o). *World of Work Report 2009*. Geneva: ILO.
- _____. (2009p). *Guide to the New Millennium Development Employment Indicators, Including the Full Set of Decent Work Indicators (June 2009)*. Geneva: ILO.
- _____. (2009q). *The ILO Global Wage Database: A Methodological Note*. Geneva: ILO.

_____. (2009r). *ILO Standards on Occupational Safety and Health: Promoting a Safe and Healthy Working Environment, Report of the Committee of Experts on the Application of Conventions and Recommendations (Articles 19, 22 and 35 of the Constitution), Report III (Part 1B), General Survey Concerning the Occupational Safety and Health Convention, 1981 (No. 155), the Occupational Safety and Health Recommendation, 1981 (No. 164) and the Protocol of 2002 to the Occupational Safety and Health Convention, 1981, International Labor Conference, 98th Session, 2009.*

_____. (2010). *The Fundamentals of Labor Administration.* Geneva: ILO.

_____. (2010b). *Key Indicators of the Labor Market, 6th edition.* Geneva: ILO.

_____. (2010c). *Yearbook of Labor Statistics 2009 Time Series.* Geneva: ILO.

_____. (2010d). *Decent Work for Domestic Workers, International Labor Conference, 99th Session, 2010.* Geneva: ILO.

_____. (2010e). *The Minimum Wage Revisited in the Enlarged EU.* Geneva: ILO.

_____. (n.d.). *Revised Office Proposal For the Measurement of Decent Work Based on Guidance Received at the Tripartite Meeting of Experts on the Measurement of Decent Work.* Geneva: ILO.

International Labor Office and International Finance Corporation. (2009). *The Better Work Program: Stage II: July, 2009 - June, 2012 (20 March 2009).* Geneva: ILO.

_____ and _____. (n.d.). *Better Work – Good Practices: Apparel.* Geneva: ILO.

International Labor Office, World Bank, CARE, and UNIFEM. (2006). *Cambodia: Women and Work in the Garment Industry.* Phnom Penh: ILO.

Italian National Institute of Statistics. (2009). *A Validation Study of the Quality of Employment Indicators. Project No. 40051280/0.*

_____. (2009b). *Quality of Employment Indicators: Country Profile Italy, Report for the Task Force on the Measurement of Quality of Employment (September 2009).*

Jackson, G. and Deeg, R. (2006). *How Many Varieties of Capitalism? Comparing the Comparative Institutional Analyses of Capitalist Diversity, MPIFG Discussion Paper 06/2.*

_____ and _____. (2007). *Towards a More Dynamic Theory of Capitalist Variety, Socio-*

Economic Review, Vol. 5, No. 1, pp. 149-179.

_____ and _____. (2008). *From Comparing Capitalisms to the Politics of Institutional Change*, *Review of International Political Economy*, Vol. 15, No. 4, pp. 680-709.

Jacobsson, B. (2000). *Standardization and Expert Knowledge*, in N. Brunson and B. Jacobsson, *A World of Standards*. Oxford: Oxford University Press.

Josephs, H. (2009). *Measuring Progress under China's Labor Law: Goals, Processes, Outcomes*, *Comparative Labor Law and Policy Journal*, Vol. 30, No. 2 (Winter 2009).

Kaplan, D., Sadka, J., Silva-Mendez, J. (2007). *Litigation and Settlement: New Evidence from Labor Courts in Mexico*, World Bank, Financial Private Sector Development Department, Policy Research Working Paper No. 4434.

Kamuzora, P. (2006). *Non-Decision Making in Occupational Health Policies in Developing Countries*, *International Journal of Occupational and Environmental Health*, Vol. 12, No. 1, pp. 65-72.

Karamessini, M. (2008). *Continuity and Change in the Southern European Social Model*, *International Labor Review*, Vol. 147, No. 1, pp. 43-70.

Karkkainen, B. (2001). *Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?*, 89 *Georgetown Law Journal* 257 (2001).

_____. (2002). *Toward a Smarter NEPA: Monitoring and Managing Government's Environmental Performance*, 102 *Columbia Law Review* 903 (2002).

_____. (2004). *'New Governance' in Legal Thought and the World*, 89 *Minnesota Law Review* 471 (2004).

Kearns, E. (1999-2009). *The Fair Labor Standards Act*. Washington, D.C.: BNA.

Keune, M. (2006). *Collective Bargaining Systems in Europe: A Schematic Overview*. Brussels: ETUI-REHS.

Khoza, S. (2004). *Promoting Economic, Social and Cultural Rights in Africa: The African Commission Holds a Seminar in Pretoria*, *Africa Human Rights Law Journal*, Vol. 4, pp. 334-343.

Koukiadis, I. (2009). *General Characteristics of Greek Labor Law*, *Comparative Labor Law and Policy Journal*, Vol. 30, No. 2 (Winter 2009).

Kristensen, H. (2009). *The European Committee of Social Rights and Its Case Law on the Negative Freedom of Association*, in *Studies in Labor Law and Social Policy – Liber Amicorum Prof. Andrzej Marian Swiatkowski*. Krakow: Rozcznik, pp. 275-284.

Krillo, K. and Masso, J. (2010). *The Part-Time/Full-Time Wage Gap in Central and Eastern Europe: The Case of Estonia*. University of Tartu Faculty of Economics and Business Administration Working Paper No. 65-2010.

Kubicek, P. J. (2004). *Organized Labor in Post-Communist States: From Solidarity to Infirmary*. Pittsburgh: University of Pittsburgh Press.

Kucera, D. (2004). *Measuring Trade Union Rights: A Country-Level Indicator Constructed from Coding Violations Recorded in Textual Sources*, ILO Policy Integration Department, Statistical Development and Analysis Group, Working Paper No. 50. Geneva: ILO.

Kucera, D. (2007). *Measuring Trade Union Rights by Violations of These Rights*, in Kucera, D., ed., *Qualitative Indicators of Labor Standards: Comparative Methods and Applications*. Geneva: ILO.

Kumamoto-Healey, J. (2005). *Women in the Japanese Labor Market, 1947-2003: A Brief Survey*, *International Labor Review*, Vol. 144, No. 4, pp. 451-472.

Kuruvilla, S. and Erickson, C. (2002). *Change and Transformation in Asian Industrial Relations*, *Industrial Relations*. Vol. 41, No. 2, pp. 171-228.

Landau, E. and Beigbeder, Y. (2008). *From ILO Standards to EU Law: The Case of Equality Between Men and Women at Work*. Leiden: Martinus Nijhoff Publishers.

Lane, D. (2005). *Emerging Varieties of Capitalism in Former State Socialist Societies*, *Competition and Change*, Vol. 9, pp. 227-247.

Lasak, K. (2009). *Health and Safety in the European Charter*, *International Journal of Comparative Labor Law and Industrial Relations*, Vol. 25, No. 2, pp. 109-122.

Lawrence, S. and Ishikawa, J. (2005). *Social Dialogue Indicators: Trade Union Membership and Collective Bargaining Coverage: Statistical Concepts, Methods, and Findings*, ILO Policy Integration Department, Bureau of Statistics and Social Dialogues, Labor Law and Labor Administration Department, Working Paper No. 59. Geneva: ILO.

Lee, C.-S. (2002). *Law and Labor-Management Relations in Korea*, in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London:

Routledge.

Lee, S. and Eyraud, F., eds. (2008). *Globalization, Flexibilization and Working Conditions in Asia and the Pacific*. Geneva: ILO.

Lee, S. and McCann, D. (2008). *Measuring Labor Market Institutions: Conceptual and Methodological Questions on 'Working Hour Rigidity,'* in Berg, J. and Kucera, D., eds., *In Defense of Labor Market Institutions: Cultivating Justice in the Developing World*. London: Palgrave MacMillan.

Lee, S., McCann, D., and Torm, N. (2008). *The World Bank's 'Employing Workers' Index: Findings and Critiques – A Review of Recent Evidence*, *International Labor Review*, Vol. 147, No. 4, pp. 416-432.

Lee, S.-W. (2008). *New Developments in Employment Discrimination in Korea*, in Blanpain, R., ed., *Bulletin of Comparative Labor Relations: New Developments in Employment Discrimination Law No. 68*. The Hague: Kluwer.

Leggett, C. (1993). *Corporatist Trade Unionism in Singapore*, in Frenkel, S., ed., *Organized Labor in the Asia-Pacific Region: A Comparative Study of Trade Unionism in Nine Countries*. Ithaca: ILR Press.

Lele, P. and Siems, M. (2007). *Shareholder Protection: A Leximetric Approach*, *Journal of Corporate Law Studies*, Vol. 7, pp. 17-50.

Lemos, S. (2007). *Minimum Wage Effects Across the Private and Public Sectors in Brazil*, *Journal of Development Studies*, Vol. 43, No. 4, pp. 700-720.

Leschke, J., Watt, A., and Finn, M. (2008). *Putting a Number on Job Quality? Constructing a European Job Quality Index*. ETUI-REHS Working Paper 2008.03.

_____, _____, and _____. (2008b). *Job Quality in Europe*. ETUI-REHS Working Paper 2008.07.

Lieberwitz, R. (2008). *Employment Discrimination Law in the United States: On the Road to Equality?* in Blainpain, ed., *New Developments in Employment Discrimination Law, Bulletin of Comparative Labor Relations No. 68*. The Hague: Kluwer.

Lindsey, T. and Masduki, T. (2002). *Labor Law in Indonesia after Soeharto*, in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Livanos, I. and Pouliakas, K. (2009). *The Gender Wage Gap as a Function of Educational Degree Choices in an Occupationally Segregated EU Country*, IZA Discussion Paper No. 4636 (December 2009).

Lobel, O. (2004). *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 *Minnesota Law Review* 342 (2004).

Locke, R., Kochan, T., Romis, M. and Qin, F. (2006). *Beyond Corporate Codes of Conduct: Work Organization and Labor Standards at Nike's Suppliers*, *International Labor Review*, Vol. 146, No. 1-2, pp. 21-40.

Lockwood, G. and Schmall, L. (2009). *UK-US Comparative Disability Law. Working Paper Series. SSRN*.

Lokiec, P. (2008). *Discrimination Law in France*, in Blanpain, R., ed., *New Developments in Employment Discrimination Law*, *Bulletin of Comparative Labor Relations* No. 68. The Hague: Kluwer.

Losa, F.B. and Origoni, P. (2005). *The Socio-Cultural Dimension of Women's Labor Force Participation Choices in Switzerland*, *International Labor Review*, Vol. 144, No. 4, pp. 473-494.

Lothian, T. (1986). *The Political Consequences of Labor Law Regimes: The Contractualist and Corporatist Models Compared*, *Cardozo Law Review*, Vol. 7. p. 1001.

_____. (1995). *Reinventing Labor Law: A Rejoinder*, *Cardozo Law Review*, Vol. 16, p. 1749.

Lozano, E. (2005). *Quality in Work: Dimensions and Indicators in the Framework of the European Employment Strategy*, *UNECE/ILO/Eurostat Seminar on the Quality of Work, May 11-13, 2005, Geneva*.

Lyon, B. and Paoletti, S. (2005). *Inter-American Developments on Globalization's Refugees: New Rights for Migrant Workers and Their Families*, *European Yearbook of Minority Issues*, Vol. 3, 2003/4 (published 2005), pp. 63-87.

Lyutov, N. and Petrylaite, D. (2009). *Trade Unions' Law Evolution in Post-Soviet Countries: The Experience of Lithuania and Russia*, *Comparative Labor Law and Policy Journal*, Vol. 30, p. 779 (Summer 2009).

Manothum, A., Rukijkanpanich, J., Thawesaengskultai, D., and Thampitakkul, B. (2009). *A Participatory Model for Improving Occupational Health and Safety: Improving Informal Sector*

Working Conditions in Thailand, International Journal of Occupational and Environmental Health, Vol. 15, No. 3, pp. 305-315.

Merry, S. (2009). *Measuring the World: Indicators, Human Rights, and Global Governance, Paper Presented at the American Society of International Law*, May 13, 2009.

Mikkola, M. (2008). *Social Human Rights of Migrants under the European Social Charter, European Journal of Social Security*, Vol. 10, No. 1, pp. 25-59.

Mock, C., Adjei, S., Acheampong, F., DeRoo, L., and Simpson, K. (2005). *Occupational Injuries in Ghana, International Journal of Occupational and Environmental Health*, Vol. 11, No. 3, pp. 238-246.

Mosley, L. (2008). *Workers' Rights in Open Economies: Global Production and Domestic Institutions in the Developing World, Comparative Political Studies*, Vol. 41, pp. 674-714.

Mosley, L. and Uno, S. (2007). *Racing to the Bottom or Climbing to the Top? Economic Globalization and Collective Labor Rights, Comparative Political Studies*, Vol. 40, No. 8, pp. 923-948.

Mundlak, G. (2007). *Fading Corporatism: Israel's Labor Law and Industrial Relations in Transition*. Ithaca: ILR Press.

_____. (2009). *The Israeli System of Labor Law: Sources and Form, Comparative Labor Law and Policy Journal*, Vol. 30, p. 159 (Winter 2009).

Murcia, J.G. (2006). *The European Commission's Green Paper on Labor Law, International Labor Review*, Vol. 146, No. 1-2, pp. 109-114.

Muriem F. (2007). *Building Safety – An International Perspective, International Journal of Occupational and Environmental Health*, Vol. 13, No. 1, pp. 5-12.

Murillo, M. and Schrank, A. (2005). *With a Little Help from My Friends: Partisan Politics, Transnational Alliances, and Labor Rights in Latin America, Comparative Political Studies*, Vol. 38, No. 8, pp. 971-999.

_____ and _____. (forthcoming). *Labor Unions in the Policymaking Process in Latin America*, in Stein, E. and Tommasi, M., eds., *Actors in Latin American Policymaking*. Inter-American Development Bank.

Naidoo, R., Jeebhay, M., Robins, T., and Meyers, J. (2006). *Addressing the Challenges of Underdevelopment in Environmental and Occupational Health in Southern Africa*,

International Journal of Occupational and Environmental Health, Vol. 12, No. 4, pp. 392-400.

Nascimento, A. (2009). *Curso de direito do trabalho*. 24th Ed. Sao Paulo: Saraiva.

National Institute of Statistics and Geography, Mexico. (2009). *Quality of Employment in Mexico: Country Report for the Task Force on the Measurement of Quality of Employment (September 2009)*.

National Research Council. (2004). *Monitoring International Labor Standards: Techniques and Sources of Information*. Washington, D.C.: The National Academies Press.

National Research Council (Rigby, Crispin, ed.). (2004b). *Monitoring International Labor Standards: International Perspectives – Summary of Regional Forums*. Washington, D.C.: National Academies Press.

National Research Council (Ulewicz, Monica, ed.). (2003a). *Monitoring International Labor Standards: Human Capital Investment: Summary of Workshop*. Washington, D.C.: National Academies Press.

National Research Council (McElrath, Roger, ed.). (2003b). *Monitoring International Labor Standards: Summary of Domestic Forums*. Washington, D.C.: National Academies Press.

National Research Council (Hilton, Margaret, ed.). (2003c). *Monitoring International Labor Standards: Quality of Information, Summary of a Workshop*. Washington, D.C.: National Academies Press.

National Research Council (Rigby, Crispin, ed.). (2003d). *Monitoring International Labor Standards: National Legal Frameworks: Summary of a Workshop*. Washington, D.C.: National Academies Press.

Neumark, D. and Wascher, W. (2004). *Minimum Wages, Labor Market Institutions, and Youth Employment: A Cross-National Analysis*, *Industrial and Labor Relations Review*, Vol. 57, No. 2, pp. 223-249.

Neves Mujica, J. (2009). *Introducción al derecho del trabajo*. Lima: PUCP-Fondo Editorial.

Nichols, T. and Zhou, W. (2010). *Disaffection with Trade Unions in China: Some Evidence from SOEs in the Auto Industry*, *Industrial Relations Journal*, Vol. 41, No. 1, pp. 19-33.

Nicholson, P. (2002). *Vietnam's Labor Market: Transition and the Role of Law*, in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Nickell, S., Nunziata, L., and Ochel, W. (2005). *Unemployment in the OECD Since the 1960s: What Do We Know?*, *Economic Journal*, Vol. 115, No. 500, pp. 1–27.

Noguer, H. (2009). *Derecho del Trabajo y de la Seguridad Social - Tomo I-III*. 18^a edición. Santiago: Jurídica de Chile.

Noonan, K., Sabel, C., Simon, W. (2009). *Legal Accountability in the Service-Based Welfare State: Lessons from Child Welfare Reform*, 34 *Law and Social Inquiry* 523 (Summer, 2009).

Novitz, T. (2003). *International and European Protection of the Right to Strike*. Oxford: Oxford University Press.

_____. (2009). *Workers' Freedom of Association*, in Gross, J. and Compa, L., eds., *Human Rights in Labor and Employment Relations: International and Domestic Perspectives*. Champaign: LERA.

Occupational Safety and Health Administration, United States Department of Labor. (2009). *Review of the Nevada Occupational Safety and Health Program (October 20, 2009)*.

_____. (2009b). *Federal OSHA to Review All State Plans, News Release (October 29, 2009)*.

_____. (2010). *OSHA Fact Sheet: State Job Safety and Health Programs*.

O'Connell, C. (2009). *Bringing Socio-Economic Rights Back to the Mainstream of Human Rights: The Case-Law of the European Committee of Social Rights as an Example of Rigorous and Effective Rights Adjudication*, *Revista Europea de Derechos Fundamentales*, No. 13, 1er Semestre, pp. 259-299.

Ochel, W. (2000). *Collective Bargaining (Centralization and Coordination)*, Ifo Institute, Munich.

Olivier, M. (2009). *Extension of Labor Law and Social Security Protection to the Informal Sector: Developing Country Perspectives, with Specific Reference to Southern Africa*, in Blanpain, R. ed., *The Modernization of Labor Law and Industrial Relations in Comparative Perspective*, *Bulletin of Comparative Labor Relations*, No. 70. The Hague: Kluwer.

Orbie, J. (2008). *Core Labor Standards in Trade Policy: The GSP Regime of the European Union*, in Fenwick C. and T. Novitz, eds., *Legal Protection of Workers' Human Rights: Regulatory Changes and Challenges*. Oxford: Hart Publishing.

Organization of American States. (1948). *American Declaration of the Rights and Duties of Man, Adopted by the Ninth International Conference of American States, Bogota, Colombia, 1948.*

_____. (1969). *American Convention on Human Rights, Adopted at the Inter-American Specialized Conference on Human Rights, San Jose, Costa Rica, 22 November 1969.*

_____. (1988). *Additional Protocol to the American Convention on Human Rights in the Area of Economic, Social and Cultural Rights - "Protocol of San Salvador."*

_____. (2008). *Inter-American Commission on Human Rights, Guidelines for Preparation of Progress Indicators in the Area of Economic, Social and Cultural Rights, Approved by the Commission During Its 132nd Regular Period of Sessions.*

Organization for Economic Co-operation and Development. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD.

_____. (2009). *2009 OECD Employment Outlook: Tackling the Jobs Crisis.* Paris: OECD.

O'Rourke, D. and Brown, G. (2003). *Experiments in Transforming the Global Workplace: Incentives for and Impediments to Improving Workplace Conditions in China, Occupational Safety and Health in China, International Journal of Occupational and Environmental Health, Vol. 9, No. 4.*

Padilla, D. (2002). *African Human Rights Court: Reflections from the Perspective of the Inter-American System, African Human Rights Law Journal, Vol. 2, pp. 185-194.*

Pagano, M. and Volpin, P. (2001). *The Political Economy of Corporate Governance, CPER Discussion Paper No. 2682.*

Pena-Casas, R. (2007). *Quality of Work and Employment in EU Policy Arena: Conceptual Frameworks and Monitoring Indicators, Paper for the RECOWE Network, Observatoire Social Europeen.*

Pencavel, J. (1999). *The Appropriate Design of Collective Bargaining Systems: Learning from the Experience of Britain, Australia and New Zealand, Comparative Labor Law and Policy Journal, Vol. 20, p. 447 (Spring 1999).*

Perspective. (2003). *The Minimum Wage as a Tool to Combat Discrimination and Promote Equality, International Labor Review, Vol. 142, No. 4, pp. 543-556.*

Piore, M. (2004). *Rethinking International Labor Standards*, in W. Milberg, ed., *Labor and the Globalization of Production: Causes and Consequences of Industrial Upgrading*. New York: Palgrave Macmillan.

Piore, M. and Schrank, A. (2008). *Toward Managed Flexibility: The Revival of Labor Inspection in the Latin World*, *International Labor Review*, Vol. 147, No. 1, pp. 1-23.

_____ and _____. (2007). *Norms, Regulations, and Labor Standards in Central America*. *Estudios y Perspectivas Series No. 77*. Mexico: ECLAC.

Pires, R. (2006). *The Forging of Regulatory Capacity: Coproduction Arrangements and Enforcement in the Monitoring of Labor Regulation*. *Doctoral Thesis. Department of Urban Studies and Planning, Massachusetts Institute of Technology*. Cambridge, Massachusetts.

_____. (2008). *Promoting Sustainable Compliance: Styles of Labor Inspection and Compliance Outcomes in Brazil*, *International Labor Review*, Vol. 147, No. 2-3, pp. 199-229.

_____. (2009). *Governing Regulatory Discretion: Innovation, Performance and Accountability in Two Models of Labor Inspection Work*, *Network on Regulating for Decent Work Conference (ILO), Geneva, July 8-10, 2009*.

Pityana, N. (2004) *Reflections on the African Court on Human Rights and Peoples' Rights*, *African Human Rights Law Journal*, Vol. 4, No. 1, pp.121-129.

Polaski, Sandra. (2004). *Cambodia Blazes a New Path to Economic Growth and Job Creation*, *Carnegie Paper No. 51* (October 2004).

_____. (2009). *Harnessing Global Forces to Create Decent Work in Cambodia*. Geneva: International Institute for Labor Studies.

Politakis, G. (2004). *Deconstructing Flexibility in International Labor Conventions*, in G. Politakis, ed., *Les Normes Internationales du Travail: Un Patrimoine Pour L'Avenir. Melanges en l'Honneur de Nicolas Valticos*. Geneva: ILO.

Pringle, T. and Frost, S. (2003). *'The Absence of Rigor and the Failure of Implementation': Occupational Health and Safety in China*, *Occupational Safety and Health in China*, *International Journal of Occupational and Environmental Health*, Vol. 9, No. 4.

Quinn, G. (2005). *The European Social Charter and EU Anti-discrimination Law in the Field of Disability: Two Gravitational Fields with One Common Purpose*, in De Burca, G. and De Witte, B., eds., *Social Rights in Europe*. Oxford: Oxford University Press.

Rachel, M. (2002). *Comparison Between the African and European Courts of Human Rights*,

African Human Rights Law Journal, Vol. 2, pp 195-223.

Republic of Moldova, National Bureau of Statistics. (2009). *Quality of Employment in the Republic of Moldova – Country Report, Presented at the Meeting on the Measurement of Quality of Employment, 14-16 October 2009, Geneva.*

Reding, V. (2010). *Towards a European Area of Fundamental Rights: The EU's Charter of Fundamental Rights and Accession to the European Convention of Human Rights, High Level Conference on the Future of the European Court of Human Rights, Interlaken, 18 February 2010.*

Rodgers, J. and Rubery, J. (2003). *The Minimum Wage as a Tool to Combat Discrimination and Promote Gender Equality, International Labor Review*, Vol. 142, No. 4., pp. 543–556.

Roe, M. (2000). *Political Preconditions to Separating Ownership from Corporate Control, Stanford Law Review*, Vol. 53, pp. 539-606.

Root, L. and Verloren, A. (2009). *Evaluation of the NAS Matrix for Monitoring International Labor Standards: Project Report, Submitted to Bureau of International Labor Affairs, Department of Labor (February 23, 2009).*

Rosga, A. and Satterthwaite, M. (2009). *The Trust in Indicators: Measuring Human Rights, Berkeley Journal of International Law*, Vol. 27, pp. 253-315.

Ross, P. and Bamber, G. (2009). *Strategic Choices in Pluralist and Unitarist Employment Relations Regimes: A Study of Australian Telecommunications, Industrial and Labor Relations Review*, Vol. 63, No. 1, pp. 24-41.

Rothstein, M., Craver, C., Schroeder, E., and Shoben, E. (2010). *Employment Law (4th Ed.)*. St. Paul: West.

Rubin, N., ed. (2005). *Code of International Labor Law: Law, Practice and Jurisprudence. 2 Volumes*. Cambridge: Cambridge University Press.

Sabel, C. (1994). *Learning By Monitoring: The Institutions of Economic Development*, in N. Smelser and R. Swedberg, eds., *The Handbook of Economic Sociology*. Princeton: Princeton University Press.

_____. (2006). *Rolling Rule Labor Standards: Why Their Time Has Come, and Why We Should Be Glad of It*, in G., Politakis, ed., *Protecting Labor Rights as Human Rights*. Geneva: ILO.

Sabel, C. and Simon, W. (2004). *Destablization Rights: How Public Law Litigation Succeeds*, 117 *Harvard Law Review* 1015 (2004).

Sabel, C. and Zeitlin, J. (2008). *Learning from Difference: The New Architecture of Experimentalist Governance in the European Union*, *European Law Journal*, Vol. 14, No. 3, pp. 271-327.

Saget, C. (2006). *Wage Fixing in the Informal Sector: Evidence from Brazil, India, Indonesia and South Africa. Conditions of Work and Employment Series No. 16*. Geneva: ILO.

_____. (2008). *Fixing Minimum Wage Levels in Developing Countries: Common Failures and Remedies*, *International Labor Review*, Vol. 147, No. 1, pp. 25–42.

Sakuraba, R. (2008). *Employment Discrimination Law in Japan: Human Rights or Employment Policy*, in Blanpain, ed., *Bulletin of Comparative Labor Relations: New Developments in Employment Discrimination Law No 68*. The Hague: Kluwer.

Salais, R. (2007). *On the Correct (and Incorrect) Use of Indicators in Public Action*, *Comparative Labor Law and Policy Journal*, Vol. 27, No. 2, pp. 237-256.

Schmitter, P. (1974). *Still the Century of Corporatism?*, in Pike, F. and Stritch, T., eds., *The New Corporatism: Social-Political Structures in the Iberian World*. Notre Dame: University of Notre Dame Press.

Schrank, A. (2007). *Asian Industrialization in Latin American Perspective: The Limits to Institutional Analysis*, *Latin American Politics and Society*, Vol. 49, No. 2, pp. 183-200.

_____. (2009). *Professionalization and Probity in the Patrimonial State: Labor Law Enforcement in the Dominican Republic*, *Latin American Politics and Society*, Vol. 51, No. 2., p 51.

Sciarra, S. (2007). *The Evolution of Collective Bargaining: Observations on a Comparison of the Countries of the European Union*, *Comparative Labor Law and Policy Journal*, Vol. 29, p. 1 (Fall 2007).

Seifert, H. and Tangian, A. (2008). *Flexicurity: Reconciling Social Security with Flexibility: Empirical Findings for Europe*, *WSI-Diskussionspapier Nr. 154 (August 2008)*.

Siems, M. (2005). *Numerical Comparative Law – Do We Need Statistical Evidence in Order to Reduce Complexity?*, *Cardozo Journal of International and Comparative Law*, Vol. 13, pp. 521-540.

_____. (2005b). *What Does Not Work in Comparing Securities Laws: A Critique on La Porta et al.'s Methodology*, *International Company and Commercial Law Review*, pp. 300-305.

Sewerynski, M. (2004). *Toward a New Codification of Polish Labor Law*, *Comparative Labor Law and Policy Journal*, Vol. 26, No. 1, pp. 55-95.

Simitis, S. (2000). *The Case of the Employment Relationship: Elements of a Comparison*, in Steinmetz, W., ed., *Private Law and Social Inequality in the Industrial Age: Comparing Legal Cultures in Britain, France, Germany and the United States*. Oxford: Oxford University Press, pp. 181–202.

Smith, B. (2008). *Australian Anti-Discrimination Laws: Framework, Developments*, in Blanpain, R., ed., *New Developments in Employment Discrimination Law*, *Bulletin of Comparative Labor Relations* No. 68. The Hague: Kluwer.

_____. (2010). *Fair and Equal in the World of Work: Two Significant Federal Developments in Discrimination Law*, *Sydney Law School Research Paper* No. 10/4.

Spurgeon, A. (2003). *Working Time: Its Impact on Safety and Health*. Geneva: ILO.

State of Israel, Central Bureau of Statistics. (2009). *Country Report on the Quality of Employment – Israel, Presented at the Meeting on the Measurement of Quality of Employment, 14-16 October 2009, Geneva*.

State Statistics Committee of Ukraine. (2009). *Quality of Employment: Country Profile for Ukraine, Prepared for the Task Force on the Measurement of Quality of Employment, 14-16 October 2009, Geneva*.

Statistics Canada. (2009). *Job Quality Country Report – Canada, Presented at the Meeting on the Measurement of Quality of Employment, 14-16 October 2009, Geneva*.

Statistics Finland. (2009). *Quality of Employment Report: Case of Finland, Prepared for the Task Force on the Measurement of the Quality of Employment*.

Storm, S. and Naastepad, C. (2009). *Labor Market Regulation and Productivity Growth: Evidence for Twenty OECD Countries (1984-2004)*, *Industrial Relations*, Vol. 48, No. 4, pp. 629-654.

Sturm, S. (2001). *Second Generation Employment Discrimination: A Structural Approach*, 101 *Columbia Law Review* 458 (April 2001).

Su, Z. (2003). *Occupational Health and Safety Legislation and Implementation in China, Occupational Safety and Health in China, International Journal of Occupational and Environmental Health*, Vol. 9, No. 4.

Sugeno, K. (2002). *Japanese Employment and Labor Law. Trans. by Kanowitz, L.* Durham: Carolina Academic Press.

Summers, C. (1998). *Exclusive Representation: A Comparative Inquiry into a "Unique" American Principle, Comparative Labor Law and Policy Journal*, Vol. 20, p. 47 (Fall 1998).

Sur, M. (2009). *General Framework and Historical Development of Labor Law in Turkey, Comparative Labor Law and Policy Journal*, Vol. 30, No. 2 (Winter 2009).

Sural, N. (2008). *Captive Audience Speeches in the Labor-Management Setting in Turkey, Comparative Labor Law and Policy Journal*, Vol. 29, No. 2 (Winter 2008).

_____. (2009). *Anti-Discrimination Rules and Policies in Turkey, Comparative Labor Law and Policy Journal*, Vol. 30, No. 2 (Winter 2009).

Swiatkowski, A. (2007). *Charter of Social Rights of the Council of Europe*. Leiden: Kluwer.

Tangian, A. (2005). *A Composite Indicator of Working Conditions in the EU-15 for Policy in the EU-15 for Policy Monitoring and Analytical Purposes, Hans Bockler Institute, Diskussionspapier Nr. 135* (March 2005).

_____. (2007). *Analysis of the Third European Survey on Working Conditions with Composite Indicators, European Journal of Operational Research*, Vol. 181, pp. 468-499.

_____. (2008). *Is Europe Ready for Flexicurity? Empirical Evidence, Critical Remarks and a Reform Proposal, Intereconomics – Review of European Economic Policy*, Vol. 43, No. 2, pp. 99-111.

_____. (2009). *Decent Work: Indexing European Working Conditions and Imposing Workplace Tax, Transfer: European Review of Labor and Research*, Vol. 15, No. 3-4, pp. 527-556.

Teitelbaum, E. (2009). *Measuring Trade Union Rights Through Observed Violations: An Assessment, Committee on Concepts and Methods Working Paper Series No. 32* (May 2009).

Thomas, C. (2003). *Information Sources and Measures of International Labor Standards on Employment Discrimination, Comparative Labor Law and Policy Journal*, Vol. 24, No. 2, pp.

365-400.

Thomas, C., Oelz, M., and Beaudonnet, X. (2004). *The Use of International Labor Law in Domestic Courts: Theory, Recent Jurisprudence, and Practical Implications*, in G. Politakis, ed., *Les Normes Internationales du Travail: Un Patrimoine Pour L'Avenir. Melanges en l'Honneur de Nicolas Valticos*. Geneva: ILO.

Thompson, C. and Benjamin, P. (1965-2010). *South African Labor Law*. Capetown: JUTA.

Tomei, Manuela. (2005). *Indigenous and Tribal Peoples: An Ethnic Audit of Selected Poverty Reduction Strategies Papers*. Geneva, ILO.

_____. (2003). *Discrimination and Equality at Work: A Review of the Concepts*, *International Labor Review*, Vol. 142, No. 4, pp. 401-418.

Toth, M. (2008). *The Right to Dignity at Work: Reflections on Article 26 of the Revised European Social Charter*, *Comparative Labor Law and Policy Journal*, Vol. 29, No. 3, pp. 275-316.

Trebilcock, A. (2004). *International Labor Standards and the Informal Economy*, in G. Politakis, ed., *Les Normes Internationales du Travail: Un Patrimoine Pour L'Avenir. Melanges en l'Honneur de Nicolas Valticos*. Geneva: ILO.

Tronti, L. (1999). *Benchmarking Employment Performance and Labor Market Policies: The Results of the Research Project, Transfer: European Review of Labor and Research*, Vol. 5, pp. 542-562.

United Nations Committee on Economic, Social and Cultural Rights. (1989). *General Comment 1: Reporting by State Parties, 24/02/89*.

_____. (1990). *General Comment 3: The Nature of State Parties Obligations, 14/12/90*.

_____. (1994). *General Comment 5: Persons with Disabilities, 09/12/94*.

_____. (1995). *General Comment 6: The Economic, Social and Cultural Rights of Older Persons, 08/12/95*.

_____. (1997). *General Comment 8: The Relationship Between Economic Sanctions and Respect for Economic, Social and Cultural Rights, 12/12/97*.

_____. (1998). *General Comment 9: The Domestic Application of the Covenant, 03/12/98*.

_____. (1998). *General Comment 10: The Role of National Human Rights Institutions in the Protection of Economic, Social and Cultural Rights*, 14/12/98.

_____. (2000). *General Comment 14: The Right to the Highest Attainable Standard of Health*, 11/08/2000.

_____. (2005). *General Comment 16: The Equal Right of Men and Women to the Enjoyment of All Economic, Social and Cultural Rights*, 11/08/2005.

_____. (2006). *General Comment 18: The Right to Work*, 06/02/2006.

_____. (2009). *General Comment 20: Nondiscrimination in Economic, Social and Cultural Rights*, 02/07/2009.

United Nations Committee on the Elimination of All Forms of Discrimination Against Women. (1989). *General Recommendation No. 9: Statistical Data (Eighth Session, 1989)*.

_____. (1989b). *General Recommendation No. 13: Equal Remuneration for Work of Equal Value (Eighth Session, 1989)*.

_____. (1991). *General Recommendation No. 16: Unpaid Women Workers in Rural and Urban Family Enterprises (Tenth Session, 1991)*.

_____. (1999). *General Recommendation No. 24: Women and Health (Twentieth Session, 1999)*.

_____. (2008). *General Recommendation No. 26: Women Migrant Workers (Forty-Second Session, 2008)*.

United Nations Economic Commission for Europe (UNECE), European Commission Statistical Office of the European Communities (Eurostat), and International Labor Organization (ILO). (2007). *Towards an International Quality of Employment Framework: Conceptual Paper of the Task Force on the Measurement of Quality of Work, Joint UNECE/Eurostat/ILO Seminar on the Quality of Work, 18-20 April 2007, Geneva, Working Paper No. 1*.

_____, _____, and _____. (2009). *Note by the UNECE Task Force on the Measurement of Quality of Employment, Introduction of the Conceptual Framework for Measuring the Quality of Employment, Meeting on the Measurement of Quality of Employment, 14-16 October 2009, Geneva*.

United Nations General Assembly. (1948). *Universal Declaration of Human Rights, General Assembly Resolution 217A(III), 8 December 1948.*

_____. (1965). *International Convention on the Elimination of All Forms of Race Discrimination, Adopted and Opened for Signature and Ratification by General Assembly Resolution 2106 (___) of 21 December 1965, Entry Into Force 4 January 1969.*

_____. (1966). *International Covenant on Economic, Social and Cultural Rights, Adopted and Opened for Signature, Ratification and Accession by General Assembly Resolution 2200A (___I) of 16 December 1966, Entry Into Force 3 January 1976.*

_____. (1966b). *International Covenant on Civil and Political Rights, Adopted and Open for Signature, Ratification and Accession by the General Assembly Resolution 2200 (___I) of 16 December 1966, Entry into Force 23 March 1976.*

_____. (1966c). *Optional Protocol of the Covenant on Civil and Political Rights, Resolution Adopted on 16 December 1966.*

_____. (1979). *Convention on the Elimination of All Forms of Discrimination Against Women, Adopted and Opened for Signature, Ratification and Accession by General Assembly, 18 December 1979, Entry Into Force 3 September 1981.*

_____. (1990). *International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, Adopted by General Assembly Resolution 45/158 of 18 December 1990.*

_____. (1999). *Optional Protocol to the Convention on the Elimination of Discrimination Against Women, Resolution Adopted 10 December 1999.*

_____. (2006). *Optional Protocol to the Convention on the Rights of Persons with Disabilities, Resolution Adopted 12 December 2006.*

_____. (2006b). *Convention on the Rights of Persons with Disabilities, Resolution Adopted on 13 December 2006*

_____. (2008). *Optional Protocol to the International Covenant on Economic, Social and Cultural Rights, Resolution Adopted on December 10, 2008, A/RES/63/117.*

United Nations High Commissioner for Human Rights. (2006). *Report on Indicators for Monitoring Compliance with International Human Rights Instruments. Fifth Inter-Committee Meeting of the Human Rights Treaty Bodies, Geneva, 19-21-June 2006, HRI/MC/2006/7.*

_____. (2008). *Report on Indicators for Promoting and Monitoring the Implementation of Human Rights, Seventh Inter-Committee Meeting of the Human Rights Treaty Bodies, Geneva, 23-25 June 2008, HRI/MC/2008/3.*

United States Agency for International Development. (2006). *Cambodia Garment Industry Workforce Assessment: Identifying Skill Needs and Sources of Supply.* Washington, D.C.: USAID.

United States Department of Labor, Bureau of International Labor Affairs, Office of Child Labor, Forced Labor, and Human Trafficking. (2009). *Proceedings Report: Child Labor Index Experts' Workshop, Washington, D.C. (August 21-21, 2009) (Draft – Not for Distribution).*

Valticos, N. and von Potobsky, G. (1995). *International Labor Law (2nd Edition).* Deventer: Kluwer.

van der Mei, A. (2005). *Advisory Jurisdiction of the African Court on Human Rights and People's Rights, African Human Rights Law Journal, Vol. 5, pp. 25-46.*

van Stel, A., Storey, D., and Thurik, R. (2007). *The Effect of Business Regulations on Nascent and Young Business Entrepreneurship, Small Business Economics, Vol. 28, No. 2-3, pp. 171-186.*

Vazquez Vialard, A. (2008). *Derecho del trabajo y de la seguridad social. 2 tomos. 10^o Edición.* Buenos Aires: Astrea.

Vega Ruiz, M. (2005). *La Reforma Laboral en America Latina: 15 Anos Despues – Un Analisis Comparado.* 1st Ed. Lima: ILO.

Venn, D. (2009). *Legislation, Collective Bargaining and Enforcement: Updating the OECD Employment Protection Indicators, OECD Social, Employment and Migration Working Papers No. 89.* Paris: OECD.

Viederman, D. and Klett, E. (2007). *Country-Level Assessments of Labor Conditions in Emerging Markets: An Approach for Institutional Investors,* in Kucera, D., ed., *Qualitative Indicators of Labor Standards: Comparative Methods and Applications, pp. 57-82.* Geneva: ILO.

Von Richtofen, Wolfgang. (2002). *Labor Inspection: A Guide to the Profession.* Geneva: ILO.

Waas, B. (2008). *New Developments in Employment Discrimination Law: Germany,* in Blanpain, R., ed., *New Developments in Employment Discrimination Law, Bulletin of Comparative Labor Relations No. 68.* The Hague: Kluwer.

Wang, H.-L. and Cooney, S. (2002) *Taiwan's Labor Law: The End of State Corporatism?* in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Weil, D. (2005). *Public Enforcement/Private Monitoring: Evaluating a New Approach to Regulating the Minimum Wage*, *Industrial and Labor Relations Review*, Vol. 58, No. 2, pp. 238-257.

_____. (2007). *Crafting a Progressive Workplace Regulatory Policy: Why Enforcement Matters*, *Comparative Labor Law & Policy Journal*, Vol. 28, No. 2, pp. 125-154.

_____. (2008). *A Strategic Approach to Labor Inspection*, *International Labor Review*, Vol. 147, No. 4, pp. 349-375.

_____. (2009). *Regulating Vulnerable Work: A Sector-Based Approach*, *Unpublished Paper (May 26, 2009)*.

Weil, D. and Mallo, C. (2007). *Regulating Labor Standards via Supply Chains: Combining Public/Private Interventions to Improve Workplace Compliance*, *British Journal of Industrial Relations*, Vol. 45, No. 4, pp. 805-828.

Weiss, M. (2007). *Convergence And/Or Divergence in Labor Law Systems?: A European Perspective*, *Comparative Labor Law and Policy Journal*, Vol. 28, No. 3, pp. 469-486.

Wisskerchen, A. (2005). *The Standard-Setting and Monitoring Activity of the ILO: Legal Questions and Practical Experience*, *International Labor Review*, Vol. 144, No. 3, pp. 253-289.

Wisskerchen, A. and Hess, C. (2001). *Employers' Handbook on ILO Standards-Related Activities*. ILO Bureau for Employers' Activities (ACT/EMP). Geneva: ILO (2001).

Woodiwiss, A. (1998). *Globalization, Human Rights and Labor Law in Pacific Asia*. Cambridge: Cambridge University Press.

Yun, J.-W. (2008). *Regulatory Contradictions: The Political Determinants of Labor Market Inequality in Korea and Japan*, *Governance*, Vol. 22. No. 1.

Zarka-Martres, M., Guichard-Kelly, M. (2005). *Decent Work, Standards and Indicators*, ILO Policy Integration Department, Statistical Development and Analysis Group, Working Paper No. 58. Geneva: ILO.

_____ and _____. (2007). *Decent Work, Standards and Indicators*, in Kucera, D., ed., *Qualitative Indicators of Labor Standards: Comparative Methods and Applications*, pp. 83-113. Dordrecht: Springer.

Zeytinoglu, A., ed. (2002). *Flexible Work Arrangements: Conceptualizations and International Experiences*. The Hague: Kluwer.

Zhao, Y. (2009). *China's New Labor Dispute Resolution Law*, *Comparative Labor Law and Policy Journal*, Vol. 30, No. 2.

Zhu, Y. (2002). *Economic Reform and Labor Regulation in China* in Cooney, S., Lindsey, T., Mitchell, R., and Zhu, Y., eds., *Law and Labor Market Regulation in East Asia*. London: Routledge.

Zubieta, I., Brown, G., Cohen, R., and Medina, E. (2009). *Cananea Copper Mine: An International Effort to Improve Hazardous Working Conditions in Mexico*, *International Journal of Occupational and Environmental Health*, Vol. 9, No. 4, pp. 14-21.