



**USAID**  
FROM THE AMERICAN PEOPLE

# MEASURING FRAGILITY

INDICATORS AND METHODS FOR RATING STATE  
PERFORMANCE

**JUNE 2005**

This publication was produced for review by the United States Agency for International Development. It was prepared by ARD, Inc.



Prepared by the ARD Consortium: ARD Inc., University of Maryland, and ISciences, L.L.C., for the USAID Bureau for Democracy, Conflict, and Humanitarian Assistance/Office of Conflict Management and Mitigation (DCHA/CMM)

# MEASURING FRAGILITY

INDICATORS AND METHODS FOR RATING STATE  
PERFORMANCE

## **DISCLAIMER**

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

**CONTENTS**

- ACRONYMS AND ABBREVIATIONS..... ii**
- I.0 INTRODUCTION ..... I**
  - 1.1 EXPLAINING FRAGILITY AND STATE PERFORMANCE OUTCOMES..... I
- 2.0 PROPOSED OUTCOME INDICATORS ..... 5**
  - 2.1 POLITICAL OUTCOMES..... 6
    - 2.1.1 EFFECTIVENESS..... 6
    - 2.1.2 LEGITIMACY..... 9
  - 2.2 SECURITY OUTCOMES..... 14
    - 2.2.1 EFFECTIVENESS..... 14
    - 2.2.2 LEGITIMACY..... 17
  - 2.3 ECONOMIC OUTCOMES ..... 20
    - 2.3.1 EFFECTIVENESS..... 20
    - 2.3.2 LEGITIMACY..... 24
  - 2.4 SOCIAL OUTCOMES ..... 26
    - 2.4.1 EFFECTIVENESS..... 26
    - 2.4.2 LEGITIMACY..... 30
- 3.0 RECOMMENDED METHODS FOR CONSTRUCTING COMPOSITE INDICES..... 36**

## **ACRONYMS AND ABBREVIATIONS**

C/FACTS	Conflict and Fragility Alert, Consultation, and Tracking System
CMM	Office of Conflict Management and Mitigation
DPT	Diphtheria, Polio, and Tetanus
GDP	Gross National Product
GDP/capita	Gross Domestic Product per Capita
IMF	International Monetary Fund
PPP	Purchasing Power Parity
SIPRI	Stockholm International Peace Research Institute
UNAIDS	Joint United Nations Program on HIV/AIDS
UNHCR	United Nations High Commissioner for Refugees
USAID	United States Agency for International Development
USCR	United States Committee for Refugees
WHO	World Health Organization

## I.0 INTRODUCTION

The United States Agency for International Development (USAID) is currently developing a Conflict and Fragility Alert, Consultation, and Tracking System (C/FACTS). Whether the focus is on conflict or fragility, there is strong evidence that early response is more effective and less costly than post-conflict reconstruction and recovery. One requirement for creating the basis for this early response is timely access to relevant information. When fully implemented, C/FACTS will combine quantitative and qualitative data to:

- Document current conditions within countries, especially those conditions that are associated with fragility; and
- Identify long-term and short-term trends within countries that indicate an improving or deteriorating situation with respect to fragility, political instability, or violence.

C/FACTS will help USAID identify those countries at greatest risk for violent conflict and with the greatest need for early intervention in order to reduce the potential for conflict. C/FACTS will also supply program planners in fragile states with data on the particular sources of fragility, and generally help measure country-level impact of USAID programs in conflict-prone and fragile states.

Accordingly, C/FACTS will obtain a variety of quantitative and qualitative information and assemble it in ways to support timely decisions by senior managers and their staffs. More precisely, C/FACTS will utilize *structural data* (data that describes aggregate conditions, such as GDP/capita, infant mortality rate, or regime type), *behavioral data* (data collected through “events monitoring” that records the actions of individuals or organizations, in this case, a count of events that are conflictual/destabilizing and cooperative/stabilizing), and *opinion data* (data collected through surveys and focus groups on individuals’ attitudes and opinions). Such data will help to both *describe* a country’s level of fragility and instability and, in formal models, *predict* which countries are at greatest risk for violent conflict or other forms of political instability.

The purpose of this document is to present initial recommendations on the design and use of state performance outcome indicators and a fragility index specifically developed to correspond to USAID’s Fragile States Strategy. Additional information on C/FACTS is available from the Office of Conflict Management and Mitigation. Those with access to the USAID intranet can find the latest information at: [inside.usaid.gov/DCHA/CMM/cfacts](http://inside.usaid.gov/DCHA/CMM/cfacts).

## I.1 EXPLAINING FRAGILITY AND STATE PERFORMANCE OUTCOMES

The key difference between fragile states and states that are more capable is the nature of the relationship between the government and the governed. In fragile states, this relationship is poor. The government acts in ways that create conditions (or outcomes) that are broadly seen as ineffective, illegitimate, or both. The perceived lack of effectiveness and/or legitimacy undermines the ability and willingness of citizens to engage with the government—and often, with one another—in constructive or productive ways. Such reserve makes it all the more difficult for the government to produce effective and/or legitimate outcomes, even if the will to do so is present.

To assist in the analysis of fragility, USAID identifies four categories of outcomes, or domains, that are particularly salient: political, security, economic, and social. Considering each of these domains in terms of effectiveness and legitimacy produces the State Performance Outcomes Matrix (below). The overarching objective of this exercise is to identify indicators that capture the essence of each of the

eight cells.

	<b>EFFECTIVENESS</b>	<b>LEGITIMACY</b>
<b>POLITICAL</b>	Well-functioning political institutions and processes that ensure accountability and timely allocation of resources to address citizen needs	Political institutions and processes that are transparent, respect societal values, and do not favor particular groups
<b>SECURITY</b>	Provision of military and police services that secures borders and limits crime	Military and police services that are provided equitably and without violation of civil rights
<b>ECONOMIC</b>	Economic institutions that provide for economic growth (including jobs), shield the economy from external shocks, and ensure adaptability to economic change	Equitable distribution of the benefits and costs of economic growth and change
<b>SOCIAL</b>	Provision of legal protections and social services, in particular to meet the special needs of vulnerable and minority groups	Tolerance for diversity, including opportunities for groups to practice customs, cultures, and beliefs

Both effectiveness and legitimacy contain a mix of objective facts (inflation rate, access to education) and subjective judgments about those facts (the inflation rate is too high, access to education is biased). As a result, measuring effectiveness and legitimacy is best undertaken through interviews, public opinion polls, and surveys. Following this logic, the ARD Consortium<sup>1</sup> reviewed a number of efforts designed to produce data that are comparable across countries (e.g., World Values Survey, AfroBarometer, Eurobarometer, and LatinBarometer). Unfortunately, country coverage is still spotty, and there are often considerable lag periods before comparative results are published. Thus, we were forced to consider indirect measures. Several categories of indirect measures were evaluated:

1. Observable outcomes that are directly associated with either poor effectiveness or low legitimacy (e.g., coups d'état, armed conflict, displaced populations, and political terror);
2. Aggregated expert and public opinion surveys (e.g., Kaufmann Corruption, Government Effectiveness, and Voice and Accountability Scales);
3. Widely accepted pre-requisites (e.g., rule of law and property rights); and
4. Comparisons to countries with similar economies (e.g., deviance from GDP or predicted infant mortality).

As a result of the review, the ARD Consortium identified thirty-three appropriate outcome indicators that nonetheless remain proxies. The sixteen Effectiveness indicators and the seventeen Legitimacy indicators are summarized in Table I.2. This report also provides a detailed description of and steps required to acquire the recommended indicators for each of the eight cells of the State Performance Outcomes Matrix and suggests methods that can be employed to compute composite indices for each cell in the matrix.

<sup>1</sup> ARD Inc., University of Maryland, and ISciences, L.L.C.

**TABLE I.2. PROPOSED OUTCOME INDICATORS**

	<b>EFFECTIVENESS</b>	<b>LEGITIMACY</b>
<b>POLITICAL</b>	1. Quality of public service/ government effectiveness 2. Number of coups d'état in last five years 3. Government revenues, as percentage of GDP	4. Nature of political participation (absence or presence of factionalism) 5. % of population experiencing political discrimination 6. Extent of citizen participation in selecting government 7. Asylum requests, as % of population
<b>SECURITY</b>	8. Intensity of most severe ongoing armed conflict 9. Size of displaced population 10. Proportion of area affected by ethnic or revolutionary war	11. State use of political terror 12. Extent of state repression of citizens 13. Presence/change in support for militant groups
<b>ECONOMIC</b>	14. Three-year change in real GDP (PPP) per capita 15. Change in foreign investment 16. Poverty rate (% of population living on <\$2 [PPP]/day) 17. Primary commodity exports/total exports 18. Three-year inflation rate	19. % of population experiencing economic discrimination 20. Corruption 21. Extent of rule of law/protection of property rights 22. Number of days to start a business
<b>SOCIAL</b>	23. Infant mortality rate 24. Youth literacy rate 25. Change in % of population living with HIV/AIDS 26. DPT and measles immunization rates 27. % of population with access to improved water supplies/ sanitation	28. Male/female literacy ratio 29. Male/female life expectancy ratio 30. % of GDP spent on military 31. Deviance from GDP-predicted infant mortality 32. Deviance from GDP-predicted primary school completion rate 33. Cultural and religious freedoms

The selection of these indicators emerged from a review that weighed four criteria:

1. **Data Coverage.** Indicator choices were eventually constricted to data that had broad coverage and that have been collected regularly for several years, thereby making comparison possible across regions and time.
2. **Data Quality.** Indicators were assessed across several dimensions of data quality. More information on the quality assessment process is provided in the next section.
3. **Citizen Relevance.** Of greatest interest were those indicators that could be expected to form the basis of citizens' judgment on the effectiveness and legitimacy of governance in their country. Infant mortality and inflation, for instance, are tangible to everyone in society. Thus, it would be expected that an increase in infant deaths or inflation would affect opinions on the effectiveness of a country's health service delivery or economic management.
4. **Programming Relevance.** Indicators were also chosen according to how well they related to the programming USAID would typically support in weak and low performing states.

Sustainable management of natural resources and vulnerability to natural disasters are not explicitly called out in the fragility framework described above, though both are clearly components of fragility. Both concepts are extremely difficult to measure quantitatively. The Millennium Challenge Corporation has initiated a concerted effort to identify and operationalize indicators that measure the effectiveness of natural resource management policies. Rather than duplicate that effort, we expect to incorporate its results into the outcome indicators when available. A similar problem exists with respect to natural



disaster vulnerability. This vulnerability is a complex function of exposure (frequency and magnitude), sensitivity of the population and infrastructure, and the ability of the society to effectively respond to and recover from disasters. Indicators of response and recovery capacity would be the most useful as an outcome indicator for fragility. Unfortunately, we are not aware of good comparative datasets that assess this concept.

In assembling the proposed outcome indicators matrix, we came to recognize that many variables could apply to more than one sector, or cell. In general, we found that it is easier to decide whether a recommended indicator belongs in the Effectiveness column or in the Legitimacy column. As a case in point, we had considerable internal discussion about the placement of “displaced populations” in the Security Effectiveness cell when “asylum requests” is placed in the Political Legitimacy cell. It is often difficult to distinguish between those who become displaced due to security concerns (e.g., fleeing warfare or persecution) and those that have become displaced due to livelihood concerns (e.g., drought or unemployment). Thus, it is a matter of opinion as to whether this indicator belongs in Security Effectiveness, Economic Effectiveness, or Social Effectiveness. However, these are Effectiveness measures, and we placed the indicator in Security Effectiveness because the estimates produced by the US Committee on Refugees emphasize security issues. In contrast, political asylum is more a measure of direct oppression of dissidents by the government. While this oppression can take on political, security, economic, and social dimensions, it is generally motivated by political considerations. Thus, we have placed asylum requests in the Political Legitimacy cell.

As a result of these placement dilemmas, we suggest computing two over-arching composite indices, one for Effectiveness and one for Legitimacy. These two overall indicators can then be complemented with composite indices computed from indicators that are applicable to each matrix cell. These complementary indices avoid the need for concern about mutual exclusivity of indicators across cells. By structuring the selection of indicators to ensure that all eight cells are covered, we can be assured that the most relevant concepts are included in the two over-arching indicators, and if desired, we could compute cell-specific composite indices.

## 2.0 PROPOSED OUTCOME INDICATORS

A wide range of proposed outcome indicators were reviewed for availability and quality for each dimension of the State Performance Outcomes Matrix in order to identify a set of indicators that USAID could use to monitor and track its programming as well as the performance of recipient governments over time. Toward this end, a small group of US Government and non-governmental experts who are well-versed in USAID's conflict and fragile states efforts worked through several iterations of the matrix. We began with an overly large sample of potential indicators and attempted to match them with a list of relevant USAID programming activities or objectives inherent to each of the eight cells. We then further discussed both the narrowed list of indicators and the conceptual and operational definitions of the respective cells with a larger group. We again narrowed the list, debating relevance and placement, before discussing data availability. At this point, the draft matrix was presented to yet another group, internal to USAID. Further revisions were made, and the candidate indicators were then subjected to a final evaluation using the following criteria:

<i>Relevance:</i>	The degree to which the indicator captures the specified dimension quantitatively;
<i>Coverage:</i>	The number of countries covered by the indicator;
<i>Update Frequency:</i>	The frequency with which the indicator is updated;
<i>Lag Period:</i>	The lag time between publication of the indicator and the observations upon which they are based. For example, indicators reflecting the condition of a country in 2001 and that are published in 2003 would have a lag period of 2 years;
<i>Authority:</i>	The degree to which the indicators are published by a respected authority in the field, such as the World Bank or the United Nations Development Program, or by independent actors with less well-established credentials; and
<i>Robustness:</i>	The degree to which the methods used to assemble the indicator have been peer-reviewed, characterize sources and magnitudes of error, and can be understood and replicated or verified by others.

Each indicator is then subjectively graded on the following scale:

*Excellent:* The indicator can be used as is, and little further improvement is likely;

*Good:* The indicator can be used as is, but further improvement is possible;

*Acceptable:* The indicator can be used initially, but improvement is desirable as part of C/FACTS; and

*Deficiencies:* Data suffers from various insufficiencies that should be evaluated before use in C/FACTS.

For each cell of the State Performance Outcomes Matrix, we identify a set of readily available indicators, characterize their applicability to C/FACTS, and describe the steps required to acquire and process the data. The results of this effort are presented below.

## 2.1 POLITICAL OUTCOMES

### 2.1.1 Effectiveness

**The degree to which political institutions ensure accountability and timely allocation of resources to address citizen needs**

Political effectiveness is a difficult concept to measure directly because outcomes related to political ineffectiveness tend to be most readily observed in the security, economic, and social dimensions of fragility. That said, we focused on three indicators that measure components of political effectiveness: the Kaufmann Government Effectiveness Scale, coups d'état in the last five years, and government revenues as a percentage of GDP. Both government effectiveness and government revenues as a percentage of GDP describe the “inputs” required for the government to be able to produce good policies and deliver public goods. Thus, they are necessary, but not sufficient, indicators of political effectiveness. Coups d'état represent the inability of the government to prevent or peacefully manage conflict over resource allocation. Instead, the conflict escalates until the government itself is replaced by an opposing faction.

In addition to the indicators described below, we also sought indicators for the tax revenues as a percentage of GDP. This indicator is considered to be a good measure of the degree to which the government has established a social contract with its population to deliver public goods. Values that are too low indicate that such a contract does not exist, and values that are too high are indicative of burdensome and often unresponsive governments. However, we were unable to find an indicator with sufficient coverage and timeliness.

#### 2.1.1.1 Quality of Public Service (Kaufmann Government Effectiveness Scale)

**Source:** *Governance Indicators: 1996-2004 Dataset* (<http://www.worldbank.org/wbi/pdf/2004kkdata.xlf>)

**Description:** The World Bank *Governance Indicators* dataset (Kaufmann, Kraay, and Mastruzzi; 2005) contains six aggregate indicators of governance for 209 countries for 1996, 1998, 2000, 2002, and 2004. Updates are expected to continue every other year with lag periods of approximately one year. One of the six indicators—government effectiveness—aggregates data from multiple surveys on questions about the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies. The main focus of this index is on inputs required for the government to be able to produce and implement good policies and deliver public goods. This is not as desirable as measuring the output of political effectiveness, but it is one of the few measures that attempts to quantify the effectiveness of the political system to make and implement policies. Figure 2.1 shows the estimated values for three selected countries.

**Rationale:** The quality of public service provision is a good, directly observable outcome of effective governments.

**Specifics:** The *Governance Indicators* dataset contains both point estimates (ranging from -2.5 to +2.5) and standard errors for government effectiveness, as aggregated through a reanalysis of microdata from 37 sources, including the *State Capacity Survey*, *Political Risk Services*, the *Economist Intelligence Unit*, and the *Global Competitiveness Survey*. We recommend using the point estimate in the construction of the overall political effectiveness index, but then using the standard errors to inform a Monte Carlo sensitivity analysis to see how changes in this point estimate would effect of the overall country rankings.

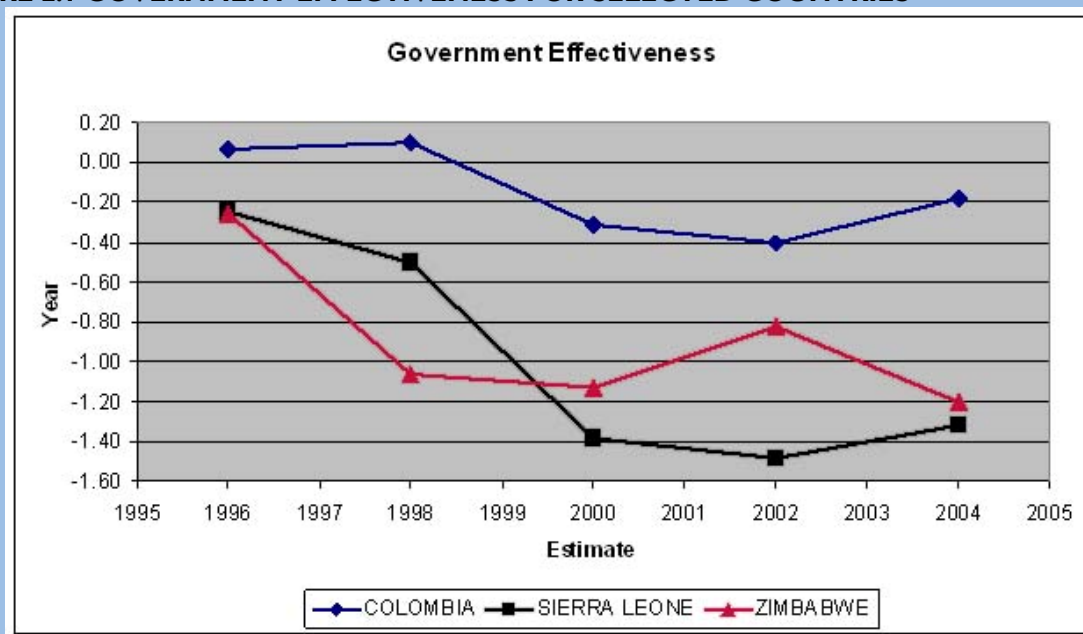
#### **Assessment Criteria:**

Relevance:                      Excellent

Coverage: Acceptable  
 Update Frequency: Acceptable  
 Lag Period: Acceptable  
 Authority: Good  
 Robustness: Good

**Next Steps:** Acquire the *Governance Indicators* dataset and merge it into a common fragile states database using common country and time period codes.

**FIGURE 2.1 GOVERNMENT EFFECTIVENESS FOR SELECTED COUNTRIES**



Source: Kaufmann, Kraay, and Mastruzzi; 2005

### 2.1.1.2 Number of Coups d'État in the Last Five Years

**Source:** Political Instability Task Force

**Description:** Number of state coups in a given year (SFTPCOUP).

**Rationale:** Coups d'état represent an inability of the government to manage conflict over resource allocation. Instead, the conflict escalates until the government itself is replaced by an opposing faction.

**Specifics:** This variable is coded as a component of the Political Instability Task Force Problem Set (see section 2.2.1.3 below) and counts the "number of state coups" that occurred in each country-year. The publicly available Problem Set code book does not describe the specifics of the coding scheme. This information should be available upon formal request from the Political Instability Task Force.

#### Assessment Criteria:

Relevance: Good  
 Coverage: Good  
 Update Frequency: Good

*Lag Period:* Good

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the SFTPCOUP variable from the Political Instability Task Force, merge it into a common fragile states database using common country and time period codes, and sum the variable for each country for the past five years.

### 2.1.1.3 Government Revenues, as Percentage of GDP

**Source:** CIA World Factbook (<http://www.cia.gov/cia/publications/factbook/>), International Monetary Fund *International Financial Statistics* (<http://ifs.apdi.net/imf/logon.aspx>), and World Bank *World Development Indicators* (<http://www.worldbank.org/data/wdi2004/index.htm>)

**Description:** All three sources listed above provide estimates of government revenues as a percentage of GDP. However, the coverage in the International Monetary Fund (IMF) and World Bank databases is spotty and must be filled in with estimates from the CIA World Factbook. It may be necessary to treat this variable non-linearly, as over-taxation may be a sign of a poor ability to manage the overall economy.

**Rationale:** The ability of a government to fund itself through taxes and other forms of revenue is a good indicator of the overall effectiveness of government institutions.

**Specifics:** Acquire the three datasets, extract government revenues as a percentage of GDP from all three datasets, and construct a combined variable that uses IMF data, if available (otherwise construct a variable that uses World Bank data, with CIA World Factbook data to be used as the final option). Evaluate how best to treat the non-linear characteristic of the variable. One option is to create a binary indicator that is “on” if revenues are either less than 10 percent or greater than 50 percent of GDP.

#### **Assessment Criteria:**

*Relevance:* Good

*Coverage:* Deficiencies

*Update Frequency:* Acceptable

*Lag Period:* Acceptable

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the data variable from the three databases described above, merge them into a common fragile states database using common country and time period codes, compute a synthetic result using IMF, then World Bank, and finally CIA World Fact Book data, and evaluate methods for characterizing the non-linear characteristics of the variable.

## 2.1.2 Legitimacy

**The degree to which political institutions and processes are transparent, respect societal values, and do not favor particular groups**

Political legitimacy describes the degree to which the population is willing to accept the outcomes of the political process, regardless of their parochial self-interest. In particular, to what degree do all of the major groups within the society believe that their voices and concerns are adequately addressed through the institutionalized negotiation and conflict resolution processes? It is important to note that this concept is distinct from a pure popularity measure. It may be possible for a government to be popular with a majority of the population by the exclusion of minority groups from meaningful political discourse. In such cases, the regime is fragile because the minority groups are unlikely to view it as a legitimate political authority.

To identify measures of political legitimacy, we first looked to attitudinal surveys such as the World Values Survey, AfroBarometer, Eurobarometer, and LatinBarometer. These surveys all contain questions designed to assess various aspects of political legitimacy. Unfortunately, country coverage is spotty, and there are often considerable lag periods before comparative results are published. Thus, we focused on measures that indirectly describe aspects of political legitimacy. Factionalism, group discrimination, and asylum requests all measure behavior that one would expect to see in countries with poor political legitimacy. Factionalism represents an implicit belief among the population that the “winners” will only represent the interests of one group or coalition at the expense of those who are not members. Group discrimination measures political exclusion. High levels of political asylum requests indicate a level of oppression that makes it difficult for large numbers of people to express dissent. In contrast, voice and accountability measures various aspects of the political process, civil liberties, and political rights that are considered to be necessary components of legitimate political systems.

### 2.1.2.1 Nature of Political Participation (Absence or Presence of Factionalism)

**Source:** Polity IV data series (<http://www.cidcm.umd.edu/inscr/polity/>), as reported by the Political Instability Task Force

**Description:** This database evaluates the regime characteristics of states as observed in the structure of their political institutions. The database is updated twice a year under present contracts with the US Government. The May coding cycle updates annual polity scores for all countries; the November research cycle examines specific regime changes that have occurred since January 1 of the coding year.

**Rationale:** The presence of factionalized political competition is an observable outcome of weak political legitimacy.

**Specifics:** The Polity IV data series describes the nature of political competition (PARCOMP) using five categories: 0 = not applicable; 1 = repressed; 2 = suppressed; 3 = factional; 4 = transitional; and 5 = competitive. Figure 2.2 shows the coding for factionalism from 1980 to 2002 (0 = not factional; 1 = factional) for Columbia, Sierra Leone, and Zimbabwe. Zimbabwe is shown as factional from 1980–1982, Colombia from 1995–2002, and Sierra Leone in 1996 and 2002.

#### Assessment Criteria:

*Relevance:* Good  
*Coverage:* Excellent  
*Update Frequency:* Good

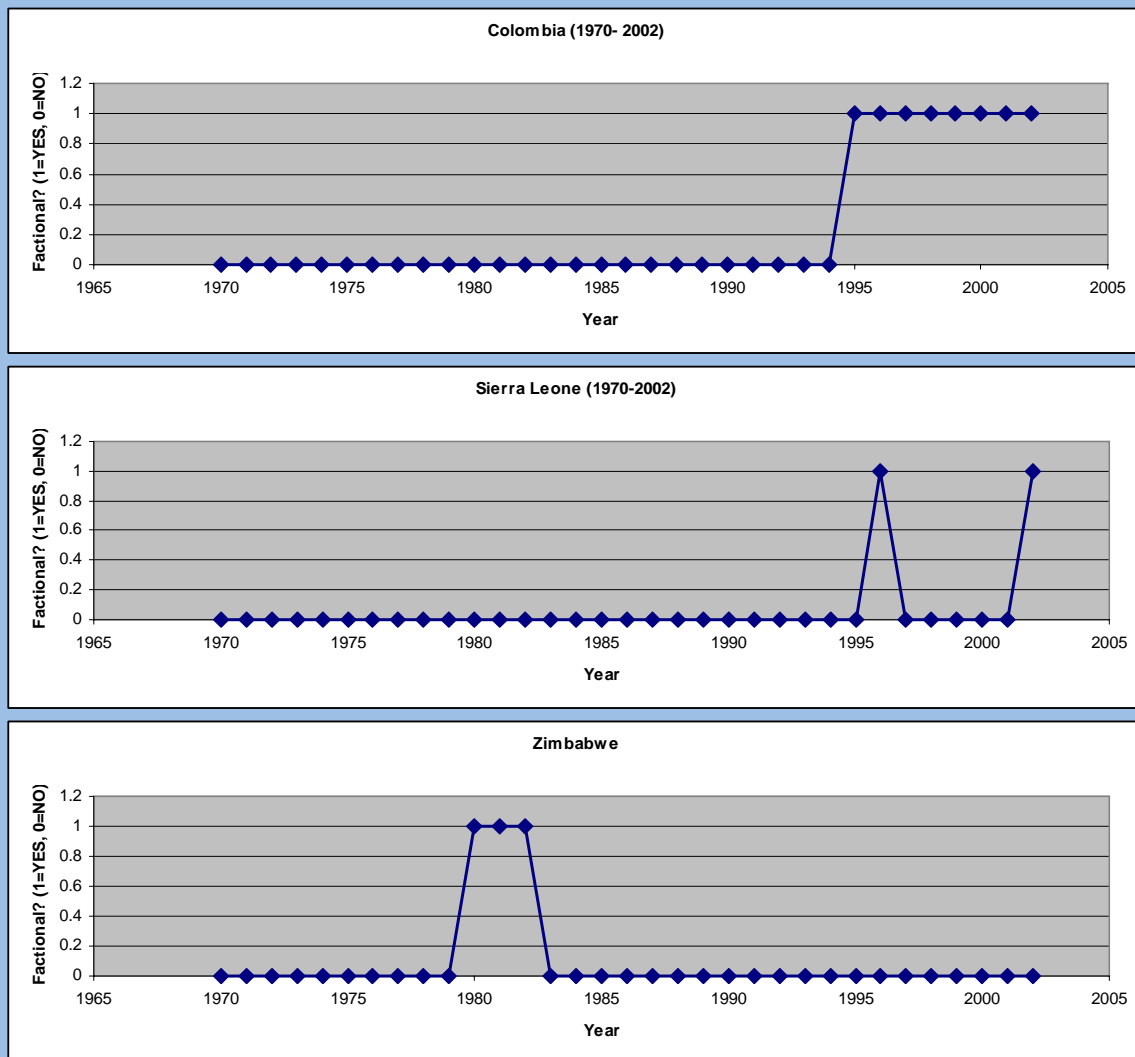
Lag Period: Good

Authority: Good

Robustness: Good

**Next Steps:** Acquire the Polity IV data series, as reported by the Political Instability Task Force, and merge it into a common fragile states database using common country and time period codes.

**FIGURE 2.2 FACTIONALISM IN SELECTED COUNTRIES (1980–2002)**



Source: Polity IV

### 2.1.2.2 Percentage of Population Experiencing Political Discrimination

**Source:** Minorities at Risk (<http://www.cidcm.umd.edu/inscr/mar/>), as summarized by the Political Instability Task Force

**Rationale:** The ability for all groups to participate effectively in the political process is at the core of the concept of political legitimacy. If groups feel excluded or marginalized in the political process, they are likely to seek non-political means to pursue their interests.

**Specifics:** The variable DISPO3P in the Political Instability Task Force dataset reports the percentage “of population experiencing levels 3–4 of political discrimination. The levels of political discrimination are defined as follows: 3 = group members are substantially under-represented due to prevailing (deliberate) social practice by dominant groups; formal public policies toward the group are neutral or, if positive, inadequate to offset widespread discriminatory practices; coded if remedial policies are enacted by governments, but subverted in practice by officials who choose not to implement them. 4 = public policies (formal exclusion or recurring repression or both) substantially restrict the group’s political participation in comparison with other groups.”

**Assessment Criteria:**

Relevance:	Excellent
Coverage:	Acceptable
Update Frequency:	Acceptable
Lag Period:	Deficiencies
Authority:	Good
Robustness:	Good

**Next Steps:** Acquire the Political Instability Task Force dataset, extract the DISPO3P variable, and merge the data into a common fragility dataset using standard codes for countries and time periods.

**2.1.2.3 Extent of Citizen Participation in Selecting Government (Kaufmann Voice and Accountability Scale)**

**Source:** *Governance Indicators: 1996-2004 Dataset* (<http://www.worldbank.org/wbi/pdf/2004kkdata.xlf>)

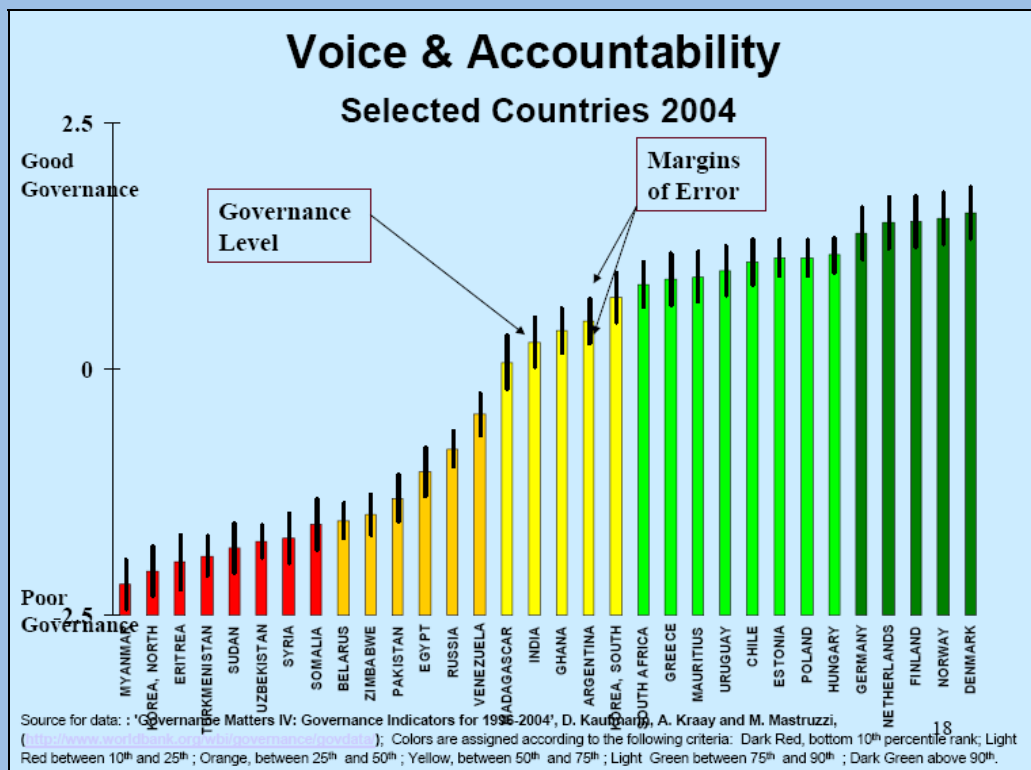
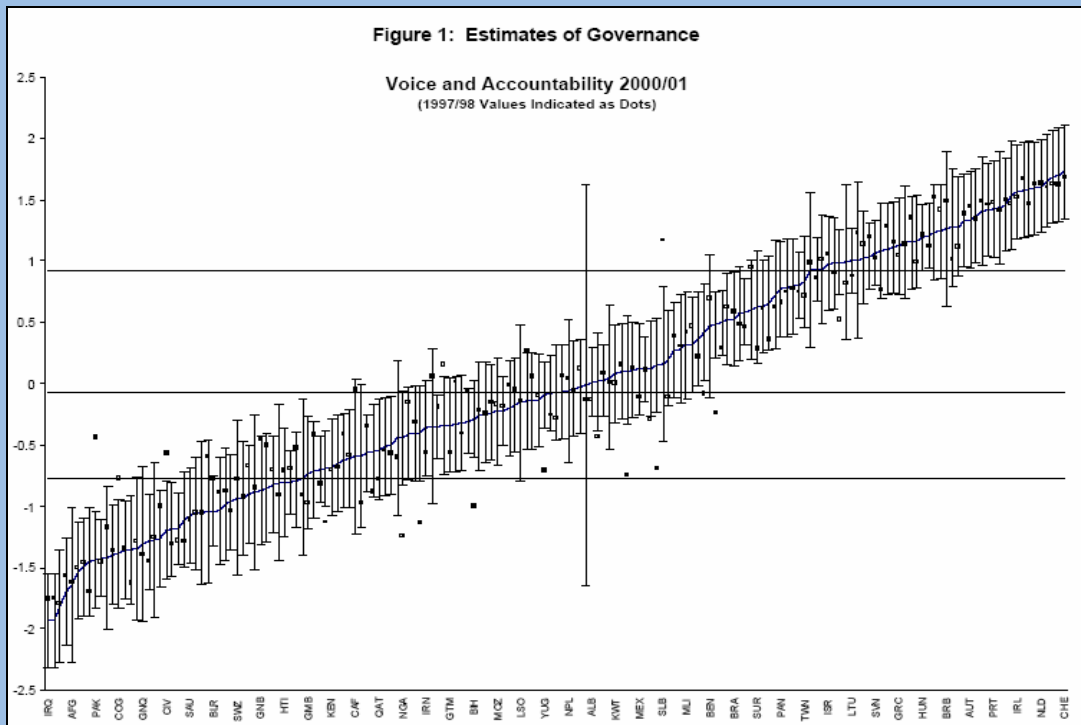
**Description:** The World Bank *Governance Indicators* dataset (Kaufmann, Kraay, and Mastruzzi; 2005) contains six aggregate indicators of governance for 199 countries for 1996, 1998, 2000, and 2002. Updates are expected to continue every other year with lag periods of approximately one year. The voice and accountability measure is aggregated from a number of indicators that measure various aspects of the political process, civil liberties, and political rights. These indicators measure the extent to which citizens of a country are able to participate in the selection of governments. They also include in this category indicators measuring the independence of the media, which serves an important role in monitoring those in authority and holding them accountable for their actions. The measure is illustrated in Figure 2.3.

**Rationale:** Governments that are open to participation in the political process and accountable to independent media outlets build political legitimacy with their constituents.

**Specifics:** The dataset reports both the mean estimate and standard error for voice and accountability. As above, we suggest that the point estimates be used for constructing composite indices and that the standard errors be used to inform a Monte Carlo sensitivity analysis of the country rankings.



**FIGURE 2.3 VOICE AND ACCOUNTABILITY**



Source: Kaufmann, et al, 2005

### Assessment Criteria:

Relevance:	Excellent
Coverage:	Acceptable
Update Frequency:	Acceptable
Lag Period:	Acceptable
Authority:	Good
Robustness:	Good

**Next Steps:** Acquire the *Governance Indicators* dataset and merge it into a common fragile states database using common country and time period codes.

#### 2.1.2.4 Asylum Requests, as Percentage of the Population

**Source:** United Nations High Commissioner for Refugees, as reported by the Migration Information Source (<http://www.migrationinformation.org/DataTools/asylum.cfm>) of the Migration Policy Institute

**Rationale:** Applications for political asylum represent people who express a loss of faith in the political legitimacy of their home government by taking diplomatic steps to leave the country.

**Specifics:** “The data were compiled by the United Nations High Commissioner for Refugees (UNHCR) Population Data Unit. The UNHCR received the data primarily from governments, based on national registration and data collection practices and nationally reported data (i.e., no adjustments were made). In some countries, UNHCR carried out, or continued to be engaged in, refugee status determination under its mandate and/or on behalf of the government.

The data are limited to asylum applications *only* and are not the number of persons who were granted asylum or resettled. The data generally refer to ‘first’ applications only and thus exclude re-opened or repeat applications or appeals. Also, the data generally refer to the *number of applicants* or individuals rather than the *number of applications* or family groups. Exceptions include: Luxembourg (for 1996 and 1997 only), Malta, the Netherlands, Portugal, the United Kingdom (from 1985 to present), and the United States.

The UNHCR data include the number of asylum applications submitted in 38 countries, including: Australia, Austria, Belgium, Bulgaria, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, the Republic of Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States.”

Table 2.1 shows a selected listing of asylum seekers, by country of origin. This raw data should be normalized by global year-to-year trends in asylum seeking and the population of the country of origin.

**TABLE 2.1 SAMPLE LISTING OF ASYLUM SEEKERS BY COUNTRY OF ORIGIN**

Country of Origin	2000	2001	2002
Afghanistan	34,827	54,430	28,524
Albania	7,869	6,083	5,675
Algeria	8,447	11,061	9,931
Andorra	—	*	*
Angola	5,587	8,556	7,417
Antigua and Barbuda	5	*	33
Argentina	1,468	1,528	1,063
Armenia	8,685	8,795	8,326
Australia	6	9	9
Austria	9	*	19
Azerbaijan	4,325	4,053	4,346
Bahamas	*	*	*
Bahrain	10	7	6
Bangladesh	6,410	6,093	5,753
Barbados	12	18	46
Belarus	2,683	2,984	3,611
Belgium	15	21	8

Source: The Migration Information Source

— Data is not available or was unreported \* Number is below 5

### Assessment Criteria:

Relevance: Acceptable

Coverage: Acceptable

Update Frequency: Acceptable

Lag Period: Acceptable

Authority: Good

Robustness: Acceptable

**Next Steps:** Acquire the asylum seekers by host country dataset, merge the data into a common fragile states database using a standard coding scheme for countries and time periods, and explore methods for normalizing by global year-to-year trends in asylum seeking behavior and the population of the country of origin.

## 2.2 SECURITY OUTCOMES

### 2.2.1 Effectiveness

#### The provision of military and police services that secures borders and limits crime

Security effectiveness measures the ability of the state to provide for the security of its population by maintaining the dominant role in enforcing law and order, enforcing justice, and protecting minorities. The measures we recommend below focus on major breakdowns in the ability of the state to provide security: armed conflict, population displacement, and areas affected by revolutionary and ethnic wars. We also sought measures of less extreme problems such as violent crimes; however, crime statistics are notorious for their lack of comparability and reporting biases.

### 2.2.1.1 Intensity of Most Severe Ongoing Armed Conflict

**Source:** Marshall, M.G. and Severn, M.D., Center for Systemic Peace, *Major Episodes of Political Violence: 1946–2004* (<http://members.aol.com/cspm/mgm/warlist.htm>) Note: these data are also available in country-year format in the Political Instability Task Force dataset.

**Description:** This dataset lists and characterizes episodes of armed conflict from 1946 to 2004. The characterization includes information on the episode type, magnitude of impact, location, and an estimate of directly related deaths. It is a synthesis of sixteen existing catalogues of armed conflict and is updated on an annual basis.

**Rationale:** States that fail to keep their population safe from armed conflict can be considered to be failing or failed, depending on the magnitude of the conflict.

**Specifics:** The dataset assigns a magnitude score of 0–7 (with 0 indicating “no episodes” and 7, “most severe episodes”) for each of the following types of conflict: civil violence (CIVVIOL); civil warfare (CIVWAR); ethnic violence (ETHVIOL); ethnic warfare (ETHWAR); warfare associated with independence (INTIND); international violence (INTVIOL); and international warfare (INTWAR). In situations where there are multiple episodes of armed conflict occurring simultaneously, we will use the maximum magnitude score.

#### Assessment Criteria:

Relevance:	Excellent
Coverage:	Excellent
Update Frequency:	Good
Lag Period:	Good
Authority:	Acceptable
Robustness:	Acceptable

**Next Steps:** Acquire the country-year form of the *Major Episodes of Political Violence* dataset from M.G. Marshall or the Political Instability Task Force, merge it by country and year to a standard set of country codes and time codes, and compute a composite magnitude indicator across all of the conflict types.

### 2.2.1.2 Size of Displaced Population

**Source:** United States Committee for Refugees (USCR), *World Refugee Survey (Annual Series)*, Washington, DC [Compiled in electronic form as: Marshall, M.G., *Forcibly Displaced Populations, 1964–2002*, Center for Systemic Peace: Severn, MD] <http://www.cidcm.umd.edu/inscr/FDPCCodebook.doc> (Codebook), and <http://www.cidcm.umd.edu/inscr/fdp2002.sav> (SPSS Saveset)

**Description:** This dataset compiles statistics reported by the United States Committee for Refugees in its annual *World Refugee Survey*. In particular, it includes estimates of: the number of refugees ( $\times 1000$ ) originating in the named country at the end of the designated year (SOURCE); the number of internally displaced persons ( $\times 1000$ ) in the named country at the end of the designated year (IDP); and the number of refugees ( $\times 1000$ ) hosted by the named country at the end of the designated year (HOST).

**Rationale:** States with large displaced populations are failing, or have failed, to provide sufficient human security to their populations. It is often difficult to distinguish between those who become displaced due to security concerns (e.g., fleeing warfare or persecution) and those that have become displaced due to livelihood concerns (e.g., drought or unemployment). Therefore, it is a matter of opinion as to where

this indicator belongs in the State Performance Outcomes Matrix (i.e., in Security Effectiveness, Economic Effectiveness, or Social Effectiveness). We have chosen to utilize it as a security indicator because, on balance, the estimates produced by the US Committee on Refugees emphasize security issues.

**Specifics:** Both internally-displaced populations and those that leave the source country represent evidence of poor security effectiveness. Thus we will sum these two variables to produce an overall measure of displaced population.

**Assessment Criteria:**

Relevance:	Excellent
Coverage:	Acceptable
Update Frequency:	Deficiencies (The compiled database was last updated in 2003, with data through 2002. However, the US Committee on Refugees has published the 2004 edition of the <i>World Refugee Survey</i> with data through 2003. The 2005 edition with data through 2004 was just released at time of publication.)
Lag Period:	Acceptable
Authority:	Acceptable
Robustness:	Acceptable

**Next Steps:** Acquire the country-year form of the *Forcibly Displaced Populations* dataset from M.G. Marshall, update the dataset with the most recent editions of the *World Refugee Survey* (2004 and 2005), merge it by country and year to a standard set of country codes and time codes, and compute a the sum of SOURCE and IPD.

**2.2.1.3 Proportion of Area Affected by Ethnic or Revolutionary War**

**Source:** Marshall, M.G., Gurr, T.R., and Harff, B., *State Failure Problem Set: Internal Wars and Failures of Governance*, Center for International Development and Conflict Management: College Park, MD, 2001. <http://www.cidcm.umd.edu/inscr/stfail/SFPScodebook.rtf> (Codebook), <http://www.cidcm.umd.edu/inscr/stfail/sfdata.htm> (Data Files)

**Description:** The *State Failure Problem Set* catalogues and characterizes four distinct types of events associated with state failure: revolutionary wars, ethnic wars, adverse regime changes, and genocides and politicides. For revolutionary and ethnic wars, the characterization includes an assessment of the area of a country affected by fighting (MAGAREA).

**Rationale:** The inability of the state to exert a monopoly on the use of force over all of its territory (i.e., areas of the state operating outside the effective control of the state) is a strong indicator of poor security effectiveness. Unfortunately, no existing dataset precisely characterizes this indicator. Rather, many conflict models use proxies such as the percentage of mountainous terrain. But such static indicators only demonstrate a propensity for lack of control, as opposed to a direct observation of lack of control. Thus we use a measure that is somewhat more blunt than ideal, the proportion of a country's area affected by fighting that is associated with ethnic or revolutionary wars. These areas are certainly outside the effective control of the state (or there would be no fighting), however, there may be other areas outside the effective control of the state where there are no ongoing ethnic or revolutionary wars.

**Specifics:** MAGAREA has six possible values as shown in the following table:

<b>MAGAREA</b>	
<b>Level</b>	<b>Interpretation</b>
0	Less than one-tenth of the country is affected and no significant cities are directly or indirectly affected.
1	One-tenth of the country (one province or state) and/or one or several provincial cities are directly or indirectly affected.
2	More than one-tenth and up to one quarter of the country (several provinces or states) and/or the capital city are directly or indirectly affected.
3	Between one-quarter and one-half of the country and/or most major urban areas are directly or indirectly affected.
4	More than one-half the country is directly or indirectly affected.
9	No basis for judging

In situations where more than one ethnic or revolutionary war is active at a given time, we will use a weighted sum of MAGAREA (after filtering out 9s; 0s will receive a weight of 0.05, 1s will receive a weight of 0.10, 2s will receive a weight of 0.175, 3s will receive as weight of 0.375, and 4s will receive a weight of 0.75). The alternative would be to use the maximum value of MAGAREA (after filtering out 9s).

#### **Assessment Criteria:**

Relevance: Acceptable

Coverage: Excellent

Update Frequency: Good

Lag Period: Good

Authority: Good

Robustness: Good

**Next Steps:** Acquire the most recent edition of the *State Failure Problem Set: Internal Wars and Failures of Governance* from the Political Instability Task Force, merge it by country and year to a standard set of country and time codes, and compute the weighted sum and maximum of MAGAREA for each country-year.

## **2.2.2 Legitimacy**

### **Military and police services are provided equitably and without violation of civil rights**

In evaluating candidate indicators for security legitimacy, we searched for data indicating that: one or more groups are systematically subjected to violence or deliberately not provided security by the state; private militias are formed, in part to protect their own group; or there is evidence that alternate means (i.e., not the state) are sought to provide for security. We were able to find good data for political terror, repression, and the presence/change in support for minority-based militant organizations. We searched for more comprehensive measures of militias (e.g., private militias) but were unable to find satisfactory data.

#### **2.2.2.1 State Use of Political Terror**

**Source:** Gibney, M., *Political Terror Scale* (<http://www.unca.edu/politicalscience/faculty-staff/gibney.html>)

**Description:** This database codes annual country-by-country human rights reports produced by the US Department of State and Amnesty International using a five-level scale as described in the following table. Scores are reported separately for each source.

<b>POLITICAL TERROR SCALE</b>	
<b>LEVEL</b>	<b>INTERPRETATION</b>
1	Countries are under a secure rule of law, people are not imprisoned for their views, and torture is rare or exceptional. Political murders are extremely rare.
2	There is a limited amount of imprisonment for nonviolent political activity. However, few persons are affected, and torture and beatings are exceptional. Political murder is rare.
3	There is extensive political imprisonment or a recent history of such imprisonment. Execution or other political murders and brutality may be common. Unlimited detention for political views, with or without a trial, is accepted.
4	The practices of Level 3 are expanded to larger numbers. Murders, disappearances, and torture are a common part of life. In spite of its generality, this level only affects those who interest themselves in politics or political ideas.
5	The terrors of Level 4 have been expanded to the whole population. The leaders of these societies place no limits on the means or thoroughness with which they pursue personal or ideological goals.

**Rationale:** State-sponsored political terror, by definition, targets groups opposed to the state with various forms of coercion. In this portion of the fragile outcomes matrix, we are particular interested in coercion directed at personal security, as opposed to economic, social, or political forms of force.

**Specifics:** The US Department of State and Amnesty International each produce annual essays on human rights for most countries of the world. M. Gibney reads these reports and codes them as described above. He updates his database annually, with a combined lag period of approximately six months. Thus, we would expect his dataset to be updated with 2004 data in June/July of 2005. The reporting from Department of State and Amnesty International are occasionally at odds with one another. In such situations, we will use the maximum of the two codes.

**Assessment Criteria:**

Relevance:                   Excellent  
 Coverage:                    Good  
 Update Frequency:        Good  
 Lag Period:                 Acceptable  
 Authority:                  Good  
 Robustness:                Good

**Next Steps:** Acquire the most recent edition of the *Political Terror Scale* from either M Gibney or the Political Instability Task Force, merge it by country and year to a standard set of country and time codes, and compute a the maximum of the US Department of State and Amnesty International codes.

**2.2.2.2 Extent of State Repression of Citizens**

**Source:** *State Capacity Survey*, Political Instability Task Force

**Description:** The Political Instability Task Force has conducted two rounds of an expert survey designed to comparatively assess various aspects of state capacity across a country. The results of the

survey have been used by the World Bank *Governance Indicators* project, Transparency International, and others.

**Rationale:** One of the survey questions asks, “To what extent does the state and/or its allied groups engage in repression of its citizens?” The answers are presented as mean and standard deviations of scores across respondents for a given country.

**Specifics:** Country coverage and number of respondents per country are limited. Thus, results are not as robust as may be desired. However, where there is overlap with the Political Terror Scale data above, these data may serve as a useful check to reinforce the validity of the Political Terror Scale data.

**Assessment Criteria:**

<i>Relevance:</i>	Excellent
<i>Coverage</i>	Deficiencies
<i>Update Frequency:</i>	Deficiencies
<i>Lag Period:</i>	Good
<i>Authority:</i>	Acceptable
<i>Robustness:</i>	Good

**Next Steps:** Acquire the *State Capacity Survey* dataset from the Political Instability Task Force (Note that the “merge” datasets do not contain question-by-question results from the survey; a separate request must be made for the complete survey dataset.), merge these data into a common fragility database using a standard set of country and time period codes, and compare these data with those provided by Political Terror Scale.

### 2.2.2.3 Presence/Change in Support for Militant Groups

**Source:** Minorities at Risk (<http://www.cidcm.umd.edu/inscr/mar/>)

**Description:** The Minorities at Risk project monitors and analyzes the status and conflicts of politically-active communal groups in all countries with a current population of at least 500,000.

**Rationale:** Minorities at Risk codes for the presence and change in support for militant organizations—a clear sign that groups within the state are seeking alternate means for providing for security.

**Specifics:** For each group-country pairing, the Minorities at Risk project codes a variable (MILOR\_) as defined in Figure 2.4. Two overall indicators can be constructed from this coding to characterize the presence and change in support for militant organizations for the country as a whole. The first is a sum for all groups within a country of the number of times MILOR\_ is valued 0, -1, or 1 (indicating the presence of a militant organization acting on behalf of the group), essentially a count of groups within the country with militant organizations. The second is a sum of changes in the level of support for militant organizations (i.e., excluding codings of -99, 9, and 99, and then summing the value of MILOR\_ for each group in the country). This second variable will give a picture of whether overall support for militant organization is increasing or decreasing within the country.

**Assessment Criteria:**

<i>Relevance:</i>	Excellent
<i>Coverage:</i>	Acceptable



*Update Frequency:* Acceptable

*Lag Period:* Deficiencies (Minorities at Risk took over two years to release an update of data through 2003.)

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the Minorities at Risk dataset, compute the indicators suggested above, and merge them into a fragile states dataset using a common set of country and time period codes.

### FIGURE 2.4 MINORITIES AT RISK CODING FOR PRESENCE/CHANGE IN SUPPORT FOR MILITANT ORGANIZATIONS

MILOR_	Change in support for militant organizations, 1996-2000	261
	Missing Values: -99	
	Value	Label
	0	no change in support
	1	significant increase in support
	9	no militant organizations known
	99	no basis for judging
	-1	significant decline in support

Source: Minorities At Risk Dataset Users Manual 030703

## 2.3 ECONOMIC OUTCOMES

### 2.3.1 Effectiveness

**Economic institutions that provide for economic growth (including jobs), shield the economy from external shocks, and ensure adaptability to economic change**

While there are many ways to assess economic performance, we focused on three outcome indicators: change in real GDP per capita over a three-year period, poverty rates, and the three-year inflation rate. In addition, we include changes in foreign investment as an indicator of international market expectations for future growth and primary commodity exports, as a percentage of total exports, as a measure of economic vulnerability. We considered additional macro-economic measures such as foreign indebtedness per capita. However, the interpretation of these indicators with respect to fragility is often difficult. For example, foreign debt becomes a dominant issue when economic growth does not meet the rates required to satisfy repayment schedules.

We are aware that the US Treasury and the IMF have ongoing programs to assess likelihood of currency collapses. However, the results of these efforts are understandably kept confidential.

#### 2.3.1.1 Three-year Change in Real GDP (PPP) per Capita

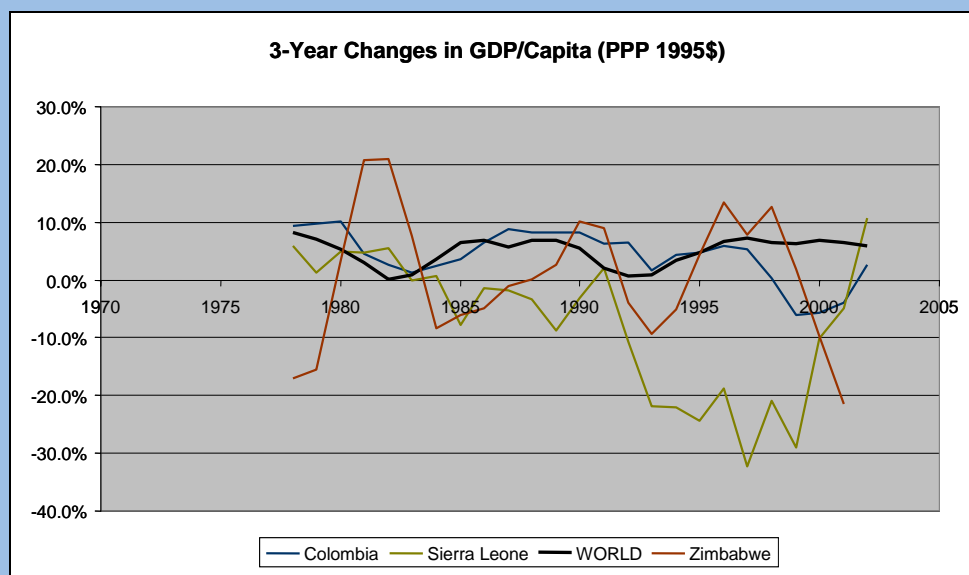
**Source:** World Bank World Development Indicators

**Rationale:** Gross domestic product per capita (GDP/capita) is the most widely accepted measure of economic development. The most suitable measure of GDP/capita for this purpose is in terms of purchasing power parity (PPP) measured in constant 1995 international dollars.

**Specifics:** The World Bank World Development Indicators dataset contains a variable defined as “GDP per capita based on PPP. PPP GDP is gross domestic product converted to international dollars using

purchasing power parity rates. An international dollar has the same purchasing power over GDP as the US dollar has in the U.S.” This variable is reported using constant 1995 international dollars to allow comparison over time. This allows one to measure percentage change over time. Changes in GDP/capita vary considerably from year to year. To remove the effect of a single year jump or decline that may be due to measurement problems (particularly in less-developed countries), we suggest calculating changes over a three-year period and normalizing by changes in world GDP/capita over the same three-year period.

**FIGURE 2.5 THREE-YEAR CHANGES IN GDP/CAPITA (PPP, 1995 INTERNATIONAL DOLLARS, WORLD BANK)**



**Assessment Criteria:**

- Relevance: Good
- Coverage: Good
- Update Frequency: Good
- Lag Period: Good
- Authority: Good
- Robustness: Good

**Next Steps:** Acquire the World Development Indicators dataset, merge the GDP/capita (PPP, 1995 constant international dollars) variable into a common fragile states database, compute three-year changes normalized by global three-year changes.

**2.3.1.2 Change in Foreign Investment**

**Source:** World Bank World Development Indicators, 2004

**Description:** Foreign direct investment is net inflows of investment to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor.

**Rationale:** Increasing levels of foreign direct investment are evidence of strong investor confidence in the ability of the economy to produce repatriatable profits. Decreasing levels of foreign direct investment are evidence of declining confidence in the ability of the economy to produce repatriatable profits.

**Specifics:** The precise variable name is “foreign direct investment, net inflows (% of GDP)”

**Assessment Criteria:**

<i>Relevance:</i>	Good
<i>Coverage:</i>	Acceptable
<i>Update Frequency:</i>	Good
<i>Lag Period:</i>	Good
<i>Authority:</i>	Good
<i>Robustness:</i>	Good

**Next Steps:** Acquire the World Development Indicators dataset, merge the foreign direct investment, net inflows (percent of GDP) variable into a common fragile states database, and compute three-year changes normalized by global three-year changes.

**2.3.1.3 Poverty Rate (Percent of Population Living on <\$2 [PPP]/day)**

**Source:** World Bank World Development Indicators, 2004  
(<http://www.worldbank.org/data/wdi2004/pdfs/table2-5.pdf>)

**Description:** The World Bank World Development Indicators contains country-by-country estimates of the percentages of people living on income below \$1 and \$2 per day. However, these estimates are based on survey dates ranging from 1990 to 2001.

**Rationale:** Economic effectiveness should also be measured by the ability of the national government to keep people out of poverty.

**Assessment Criteria:**

<i>Relevance:</i>	Excellent
<i>Coverage:</i>	Acceptable
<i>Update Frequency:</i>	Deficiencies
<i>Lag Period:</i>	Deficiencies
<i>Authority:</i>	Good
<i>Robustness:</i>	Deficiencies

**Next Steps:** Acquire the World Development Indicators dataset and merge the percentage of population living on <\$2 (PPP)/day variable into a common fragile states database.

**2.3.1.4 Primary Commodity Exports/Total Exports**

**Source:** World Bank World Development Indicators, 2004

**Description:** The World Bank World Development Indicators contain variables for agricultural raw materials (percent of merchandise exports), ores and metals (percent of merchandise exports), and fuel exports (percent of merchandise exports).

**Rationale:** Countries highly dependent on primary commodity exports have less well developed economies than those that exports value-added goods and services.

**Specifics:** The sum of three variables described above is a good, but imperfect, approximation of primary commodity exports as a percentage of total exports.

**Assessment Criteria:**

*Relevance:* Acceptable

*Coverage:* Acceptable

*Update Frequency:* Good

*Lag Period:* Acceptable

*Authority:* Good

*Robustness:* Acceptable

**Next Steps:** Acquire the World Development Indicators dataset, merge the three export variables into a common fragile states database, and compute the sum.

### 2.3.1.5 Three-year Inflation Rate

**Source:** World Bank World Development Indicators, 2004

**Description:** The World Bank World Development Indicators contains data on consumer price indices (1995 = 100).

**Rationale:** Persistent hyper-inflation or deflation is a sign of macro-economic mismanagement.

**Specifics:** The consumer price index reflects changes in the cost to the average consumer of acquiring a fixed basket of goods and services that may be fixed or changed at specified intervals, such as by year. The Laspeyres formula is generally used. We will use the compounded three-year inflation rate for the most recent three years of data and threshold the result to flag countries with absolute values in excess of 100 percent.

**Assessment Criteria:**

*Relevance:* Acceptable

*Coverage:* Acceptable

*Update Frequency:* Good

*Lag Period:* Acceptable

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the World Development Indicators dataset, merge the consumer price index (1995 = 100) common fragile states database, compute the three-year inflation rate, and flag it if absolute value is above 100 percent.

## 2.3.2 Legitimacy

### Equitable distribution of the benefits and costs of economic growth and change

Economic legitimacy is driven by the degree to which all segments of the population benefit from the economic growth and opportunities that are available. Thus, we focus on measures of economic discrimination, corruption, and the rule of law/property rights. In addition, we considered measures such as the size of the informal economy, as a percentage of GDP, to capture the degree to which people choose to not participate in the state economy. However, we are not aware of a good comparative dataset with timely updates. In lieu of this measure, we use the number of days to start a business as a proxy for the informal economy. This measure is also used by the Millennium Challenge Corporation to assess economic freedom.

#### 2.3.2.1. Percentage of Population Experiencing Economic Discrimination

**Source:** Minorities at Risk (<http://www.cidcm.umd.edu/inscr/mar/>), as summarized by the Political Instability Task Force

**Rationale:** The ability for all groups to participate effectively in the economy is at the core of the concept of economic legitimacy. If groups feel economically excluded or marginalized, they are likely to seek non-economic means to pursue their interests.

**Specifics:** The variable DISPO3E in the Political Instability Task Force dataset reports the percentage “of population experiencing levels 3–4 of economic discrimination. The levels of economic discrimination are defined as follows: 3 = group members experience substantial poverty and under-representation due to prevailing (deliberate) practice by dominant groups; formal public policies toward the group are neutral or, if positive, inadequate to offset active and widespread practices of discrimination; 4 = public policies (formal exclusion or recurring repression or both) substantially restrict the group's economic opportunities in contrast with other groups.”

#### Assessment Criteria:

<i>Relevance:</i>	Excellent
<i>Coverage:</i>	Acceptable
<i>Update Frequency:</i>	Acceptable
<i>Lag Period:</i>	Deficiencies
<i>Authority:</i>	Good
<i>Robustness:</i>	Good

**Next Steps:** Acquire the Political Instability Task Force Dataset, extract the DISPO3E variable, merge the data into a common fragility dataset using standard codes for countries and time periods.

#### 2.3.2.2 Corruption (Kaufmann Control of Corruption Scale)

**Source:** *Governance Indicators: 1996-2004 Dataset* (<http://www.worldbank.org/wbi/pdf/2004kkdata.xlf>)

**Description:** The World Bank *Governance Indicators* dataset (Kaufmann, Kraay, and Mastruzzi; 2005) contains six aggregate indicators of governance for 199 countries for 1996, 1998, 2000, and 2002. Updates are expected to continue every other year with lag times of approximately one year. The control of corruption measure is aggregated from a number of indicators perceptions of corruption, conventionally defined as the exercise of public power for private gain. Despite this straightforward

focus, the particular aspect of corruption measured by the various sources differs somewhat, ranging from the frequency of “additional payments to get things done,” to the effects of corruption on the business environment, to measuring “grand corruption” in the political arena or in the tendency of elite forms to engage in “state capture.”

**Rationale:** “The presence of corruption is often a manifestation of a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules which govern their interactions, and hence represents a failure of governance according to our definition.”

**Specifics:** The dataset reports both the mean estimate and standard error for corruption. As above, we suggest that the point estimates be used for constructing composite indices and that the standard errors be used to inform a Monte Carlo sensitivity analysis of the country rankings.

**Assessment Criteria:**

Relevance:	Excellent
Coverage:	Acceptable
Update Frequency:	Acceptable
Lag Period:	Acceptable
Authority:	Good
Robustness:	Good

**Next Steps:** Acquire the *Governance Indicators* dataset and merge it into a common fragile states database using common country and time period codes.

**2.3.2.3 Extent of Rule of Law/Protection of Property Rights**

**Source:** Miles, M.A., Feulner, E.J., and O’Grady, M.A., *2005 Index of Economic Freedom*. The Heritage Foundation and the Wall Street Journal: Washington DC  
(<http://www.heritage.org/research/features/index/downloads.cfm>)

**Description:** This factor scores the degree to which a country’s laws protect private property rights and the degree to which its government enforces those laws. It also accounts for the possibility that private property will be expropriated. In addition, it analyzes the independence of the judiciary, the existence of corruption within the judiciary, and the ability of individuals and businesses to enforce contracts. The less legal protection of property exists, the higher a country’s score; similarly, the greater the chances of a government expropriating property, the higher a country’s score.

**Rationale:** “The ability to accumulate private property is the main motivating force in a market economy, and the rule of law is vital to a fully-functioning, free market economy. Secure property rights give citizens the confidence to undertake commercial activities, save their income, and make long-term plans because they know that their income and savings are safe from expropriation.”

**Specifics:** “Unless otherwise noted, the authors used the following sources for information on property rights, in order of priority: Economist Intelligence Unit, Country Commerce, 2003 and 2004; US Department of Commerce, Country Commercial Guide; and US Department of State, Country Reports on Human Rights Practices, 2003 and 2004.”

**Assessment Criteria:**

Relevance:	Excellent
------------	-----------

<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Good
<i>Lag Period:</i>	Good
<i>Authority:</i>	Acceptable
<i>Robustness:</i>	Acceptable

**Next Steps:** Acquire the *Index of Economic Freedom* spreadsheet and merge the results into a common fragile states database using common country and time period codes.

#### 2.3.2.4 Number of Days to Start a Business

**Source:** World Bank World Development Indicators, 2005

**Description:** This indicator measures the number of days required to start a business in each country. It is available for 140 countries as of 2004.

**Rationale:** The more difficult it is to start a business, the more likely people will opt out of the formal economy and operate within the informal economy.

**Specifics:** The data are assembled by the World Bank Private Sector Advisory Service.

#### Assessment Criteria:

<i>Relevance:</i>	Acceptable
<i>Coverage:</i>	Acceptable
<i>Update Frequency:</i>	Unknown
<i>Lag Period:</i>	Good
<i>Authority:</i>	Acceptable
<i>Robustness:</i>	Acceptable

**Next Steps:** Extract the data from the World Development Indicators and merge the results into a common fragile states database using common country and time period codes.

## 2.4 SOCIAL OUTCOMES

### 2.4.1 Effectiveness

#### Provision of legal protections and services, in particular to meet the special needs of vulnerable and minority groups

Effectiveness of social services is measured by the ability of the state to deliver core education and health care services and access to vital infrastructure. We focus on infant mortality, youth literacy, the percentage of the population living with HIV/AIDS, childhood immunization rates, and access to improved water and sanitation. Together, these indicators provide a good picture of the effectiveness of key social services. We considered additional indicators, such as the percentage of urban population in slums or the adolescent fertility rate. However, we opted against including these measures in the outcome matrix due to concerns over data quality or cultural biases.

### 2.4.1.1 Infant Mortality Rate

**Source:** World Bank World Development Indicators, 2005

**Description:** Infant mortality rate is the number of infants who die before reaching one year of age, per 1,000 live births in a given year.

**Rationale:** Infant mortality is a broad indicator that assesses the ability of the state to provide a broad range of social services, including health care, environmental quality, food, housing, and education.

**Specifics:** The World Bank makes estimates using data from the United Nations and UNICEF's *State of the World's Children*.

#### **Assessment Criteria:**

*Relevance:* Good

*Coverage:* Good

*Update Frequency:* Good

*Lag Period:* Good

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the most recent edition World Development Indicators and merge the results into a common fragile states database using common country and time period codes.

### 2.4.1.2 Youth Literacy Rate

**Source:** World Bank Development Indicators, 2005

**Description:** The youth literacy rate is the percentage of people ages 15–24 who can, with understanding, read and write a short, simple statement on their everyday life.

**Rationale:** Literacy is the principal output of the educational system, thus it is a good measure of the effectiveness of the state in providing education. Youth literacy changes more quickly than adult literacy and is, therefore, a better measure of the current educational infrastructure.

**Specifics:** The World Development Indicators database reports a variable titled “Literacy rate, youth total (% of people ages 15–24).”

#### **Assessment Criteria:**

*Relevance:* Good

*Coverage:* Good

*Update Frequency:* Good

*Lag Period:* Good

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the most recent edition World Development Indicators and merge the results into a common fragile states database using common country and time period codes.



### 2.4.1.3 Change in Percentage of Population Living with HIV/AIDS

**Source:** 2004 Report on the Global HIV/AIDS Epidemic: Fourth Global Report, Joint United Nations Program on HIV/AIDS (UNAIDS), July 2004

[http://www.unaids.org/bangkok2004/GAR2004\\_pdf/UNAIDSGlobalReport2004\\_en.pdf](http://www.unaids.org/bangkok2004/GAR2004_pdf/UNAIDSGlobalReport2004_en.pdf)

[http://www.unaids.org/html/pub/Global-Reports/Bangkok/Table\\_countryestimates\\_2004\\_en\\_xls.xls](http://www.unaids.org/html/pub/Global-Reports/Bangkok/Table_countryestimates_2004_en_xls.xls)

**Description:** “In 2004, nearly 40 million people globally were estimated to be living with HIV. The AIDS epidemic claimed more than 3 million lives and close to 5 million people acquired the human immunodeficiency virus (HIV) in 2004 ... The latest country-specific HIV and AIDS estimates were published in the UNAIDS biannual Report on the global AIDS epidemic (July 2004).”

**Rationale:** Uncontrolled HIV/AIDS epidemics indicate an observable outcome of poor government capacity to address fundamental public health issues.

**Specifics:** UNAIDS publishes indicators on the number of adults and children living with HIV every two years. The estimates are published both as a point estimate and a high/low estimate range. We suggest using the point estimate and then performing a sensitivity analysis using the high/low estimate ranges. The most recent publication presents estimates for the end of 2003. We also suggest considering the two-year rate of change in the percentage of population living with HIV/AIDS. Since HIV/AIDS has a cumulative effect, the rate of change may present a better picture of current government effectiveness in addressing the epidemic.

#### **Assessment Criteria:**

<i>Relevance:</i>	Good
<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Acceptable
<i>Lag Period:</i>	Good
<i>Authority:</i>	Good
<i>Robustness:</i>	Acceptable

**Next Steps:** Acquire the data from UNAIDS, merge it into a common fragile states database using common country and time period codes, compute the percentage of the population living with HIV/AIDS using current population estimates data from World Development Indicators, and compute the rate of change from 2001 to 2003.

### 2.4.1.4 Diphtheria, Polio, and Tetanus (DPT) and Measles Immunization Rates

**Source:** World Bank World Development Indicators, 2005

**Description:** “Child immunization measures the percentage of children ages 12–23 months who received vaccinations before one year of age.”

**Rationale:** DPT and measles immunization rates measure the proportion of the population immunized against common childhood diseases. Immunization programs represent a complex form of socio-technical infrastructure that requires coordination among many elements of society. Immunization programs must be recreated each year and, therefore, changes in the ability of society to deliver such services tend to be reflected in year-to-year changes.

**Specifics:** Both DPT and measles immunizations are considered fundamental preventative public health measures. We will, therefore, use the minimum of the two values.

**Assessment Criteria:**

<i>Relevance:</i>	Good
<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Acceptable
<i>Lag Period:</i>	Good
<i>Authority:</i>	Good
<i>Robustness:</i>	Acceptable

**Next Steps:** Acquire the World Development Indicators, merge the DPT and measles immunization indicators into a common fragile states database using a common coding scheme for countries and time periods, and calculate the minimum of the two values.

**2.4.1.5 Percentage of Population with Access to Improved Water Supplies/Sanitation**

**Source:** *Global Water Supply and Sanitation Assessment 2000 Report*, as republished by World Development Indicators, World Health Organization (WHO)/UNICEF Joint Monitoring Program for Water Supply and Sanitation, 2000

**Description:** “Access to an improved water source refers to the percentage of the population with reasonable access to an adequate amount of water from an improved source, such as a household connection, public standpipe, borehole, protected well or spring, and rainwater harvesting ... Access to improved sanitation facilities refers to the percentage of the population with at least adequate excreta disposal facilities (private or shared, but not public) that can effectively prevent human, animal, and insect contact with excreta.”

**Rationale:** States that do not provide their populations with access to improved water and sanitation are failing to meet one of their basic obligations.

**Specifics:** WHO and UNICEF perform a global assessment of water supply and sanitation every ten years. These figures are commonly republished by secondary providers (e.g., the World Bank and World Resources). The World Development Indicators titles these variables, “Improved water source (% of population with access)” and “Improved sanitation facilities (% of population with access).” Both indicators are essential. Therefore, we will use the minimum of the two.

**Assessment Criteria:**

<i>Relevance:</i>	Good
<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Deficiencies
<i>Lag Period:</i>	Good
<i>Authority:</i>	Good
<i>Robustness:</i>	Acceptable

**Next Steps:** Acquire the World Development Indicators, merge the access to improved water and sanitation indicators into a common fragile states database using a common coding scheme for countries and time periods, and calculate the minimum of the two values.

## 2.4.2 Legitimacy

### Tolerance for diversity, including opportunities for groups to practice customs, cultures, and beliefs

For this category we examined indicators that measured differentials in the provision of key social services to particular groups (male/female literacy ratio and male/female life expectancy ratio), patterns of central government spending likely to be viewed as illegitimate (the percent of GDP spent on the military), discrepancies between expected versus achieved social performance (deviance from GDP-predicted infant mortality and primary school completion rate), and cultural and religious freedoms.

#### 2.4.2.1 Male/Female Literacy Ratio

**Source:** World Bank World Development Indicators, 2005

**Description:** “Literacy rate, youth male (% of males ages 15-24)” and “Literacy rate, youth female (% of males ages 15-24)” measure the percentage of people ages 15-24 who can, with understanding, read and write a short, simple statement on their everyday life.

**Rationale:** The ratio of male to female literacy rates provides an indication of whether education is provided equally across gender. Gross imbalances in either direction are a cause for concern.

**Specifics:** We use the ratio between youth literacy rates because these best measure current provision of educational services.

#### Assessment Criteria:

*Relevance:* Good

*Coverage:* Good

*Update Frequency:* Good

*Lag Period:* Good

*Authority:* Good

*Robustness:* Good

**Next Steps:** Acquire the World Development Indicators, merge the youth literacy indicators into a common fragile states database using a common coding scheme for countries and time periods, and calculate the maximum of the two ratios (male-to-female and female-to-male).

#### 2.4.2.2 Male/Female Life Expectancy Ratio

**Source:** World Bank World Development Indicators, 2005

**Description:** “Life expectancy at birth, male (years)” and “Life expectancy at birth, female (years)” indicate the number of years a newborn infant of a given gender would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

**Rationale:** The ratio of male-to-female life expectancy rates provides an indication of whether public health, housing, and food are provided equally across gender. Gross imbalances in either direction are a cause for concern.

**Specifics:** These measures need to be normalized to the overall world rates for each gender (e.g., divide the country male life expectancy by the world male life expectancy, divide the country female life

expectancy by the world female life expectancy, and then ratio the two resulting values). As with literacy rates, gross imbalances in either direction are cause for concern. Thus, we would then use the maximum of the male-to-female ratio and the female-to-male ratio.

**Assessment Criteria:**

<i>Relevance:</i>	Good
<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Good
<i>Lag Period:</i>	Good
<i>Authority:</i>	Good
<i>Robustness:</i>	Good

**Next Steps:** Acquire the World Development Indicators, merge the life expectancy and male/female population indicators into a common fragile states database using a common coding scheme for countries and time periods, calculate world population weighted life expectancies for each gender, calculate the country life expectancies relative to the world life expectancies for each gender, and calculate the maximum of the two ratios (male-to-female and female-to-male).

**2.4.2.3 Percentage of the GDP Spent on the Military**

**Source:** Stockholm International Peace Research Institute Military Expenditure Database, Stockholm International Peace Research Institute (SIPRI): Stockholm, Sweden  
[http://web.sipri.org/contents/milap/milex/mex\\_database1.html](http://web.sipri.org/contents/milap/milex/mex_database1.html)

**Description:** “The database on military expenditure covers more than 160 countries and contains consistent time series for the period since 1988. SIPRI military expenditure data are based on open sources only, including a SIPRI questionnaire, which is sent out annually to all countries included in the database. Collected data are processed to achieve consistent time series and as far as possible in accordance with the SIPRI definition of military expenditure.

The online military expenditure database provides military expenditure data by country in the following three formats:

- Military expenditure in local currency, at current prices, 1988–2003;
- Military expenditure in US dollars, at constant (2000) prices and exchange rates, 1988–2003; and
- Military expenditure as a share (percent) of GDP, 1988–2002”

**Rationale:** Countries that use an excessive proportion of their GDP to support their military effectively “starve” the broader population of social services in favor of the military.

**Specifics:** “You may freely download or cite text and data presented by SIPRI on the Internet. Any reproduction—in any medium, electronic or printed—of text or data presented by SIPRI, or the creation of any links to SIPRI sites, is authorized only on request. Permission may be granted provided that:

- A request for permission is submitted to SIPRI.

- Attribution to SIPRI in a source credit-line is given, including the phrase ‘Information from the Stockholm International Peace Research Institute (SIPRI), URL ADDRESS.’ This applies to both citations and reproduction.

When text or data are to be used for commercial purposes, and once permission has been obtained, the parties will negotiate a reasonable royalty for such commercial use.

Individual documents may have different copyright conditions, which are noted on those documents.”

**Assessment Criteria:**

<i>Relevance:</i>	Acceptable
<i>Coverage:</i>	Good
<i>Update Frequency:</i>	Good
<i>Lag Period:</i>	Acceptable
<i>Authority:</i>	Acceptable
<i>Robustness:</i>	Acceptable

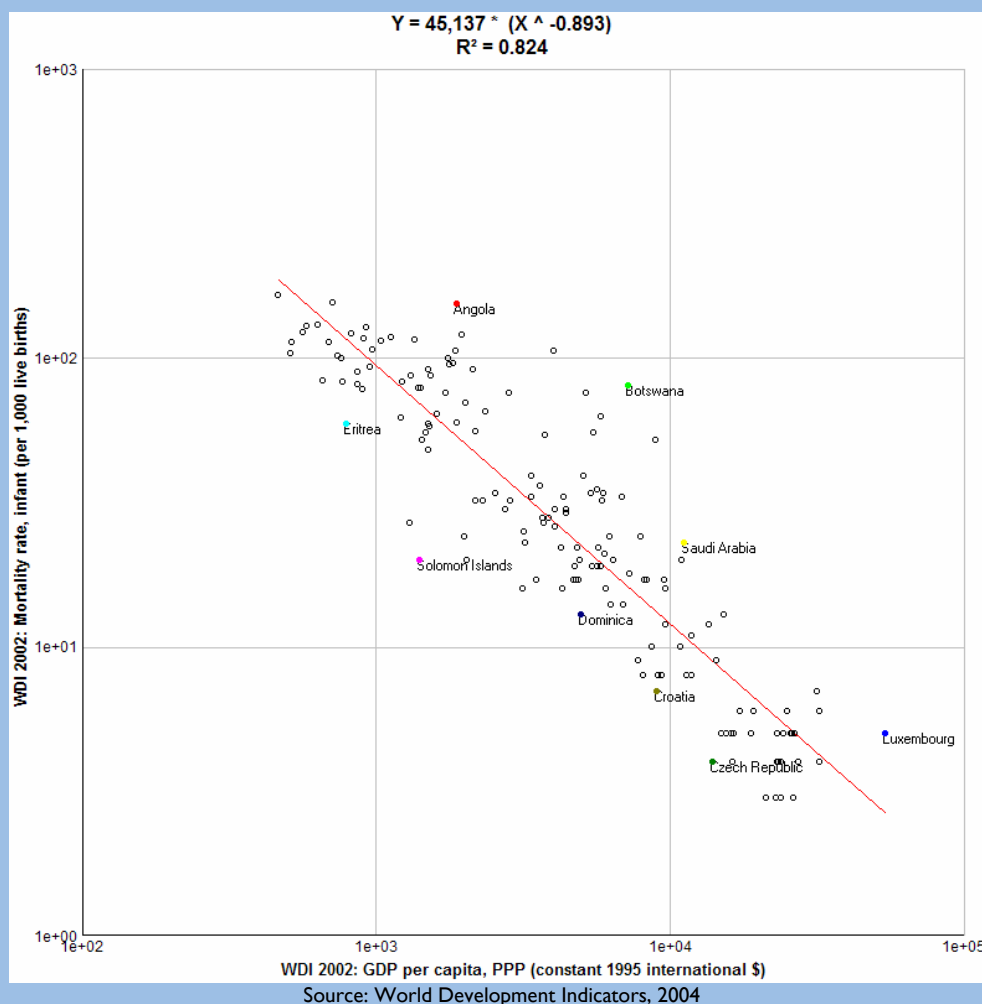
**Next Steps:** Acquire the data from SIPRI and merge the data into a common fragile states database using a common coding scheme for countries and time periods.

**2.4.2.4 Deviance from GDP-predicted Infant Mortality**

**Source:** World Bank World Development Indicators, 2005

**Description:** Infant mortality rates are highly correlated to GDP/capita (PPP, 1995 constant international dollars). However, they are not perfectly correlated, as shown in Figure 2.6. Some countries have higher than expected infant mortality rates (e.g., Angola, Botswana, Saudi Arabia, and Luxembourg), whereas others have lower than expected (e.g., Eritrea, Solomon Islands, Dominica, Croatia, and the Czech Republic).

**FIGURE 2.6 INFANT MORTALITY (LOG) VS. GDP/CAPITA (LOG) (2002)**



**Rationale:** The ability of the government to provide either better than expected or worse than expected infant mortality is an indicator of legitimacy. Countries that perform better than expected do well at minimizing social inequalities throughout the society. When their citizens compare themselves to those of peer countries (neighboring countries with similar GDP/capita) they will be satisfied with the results. The converse is true for those countries that perform worse than expected.

**Specifics:** Fit a univariate log-log model predicting infant mortality rates as a function of GDP/capita and compute the residuals.

**Assessment Criteria:**

- Relevance:                    Acceptable
- Coverage:                     Good
- Update Frequency:         Good
- Lag Period:                  Acceptable
- Authority:                    Acceptable
- Robustness:                  Acceptable

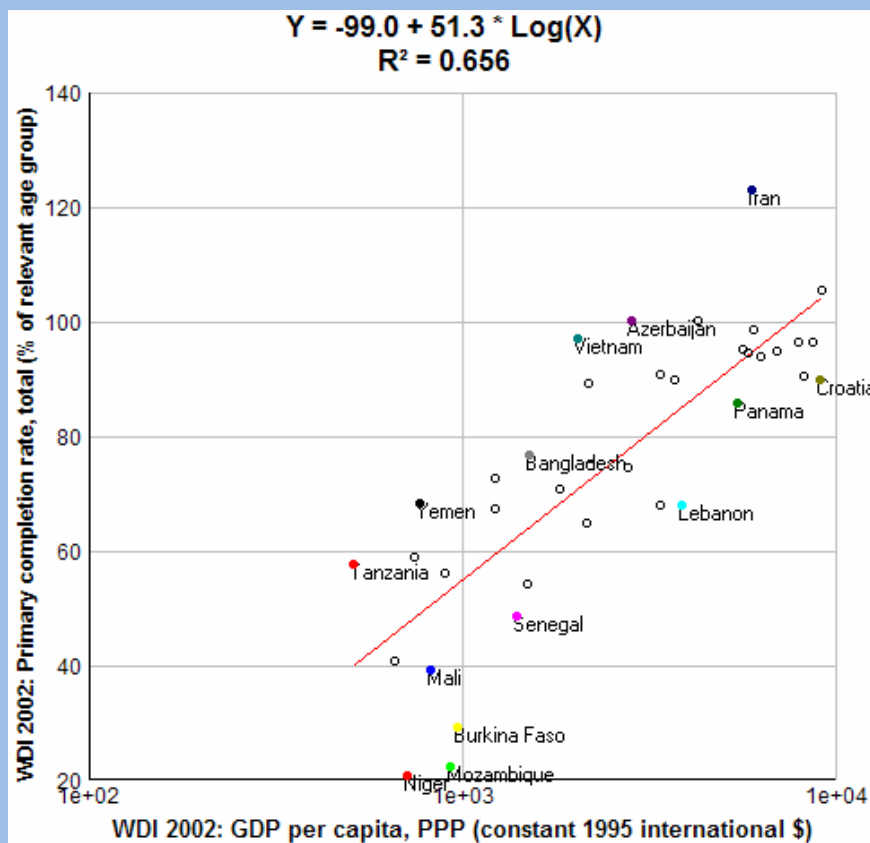
**Next Steps:** Acquire the World Development Indicators, merge the infant mortality and GDP/capita indicators into a common fragile states database using a common coding scheme for countries and time periods, and calculate the residuals from a log-log regression.

### 2.4.2.5 Deviance from GDP-predicted Primary School Completion Rate

**Source:** World Bank World Development Indicators, 2005

**Description:** The primary school completion rate is the number of students successfully completing the last year of (or graduating from) primary school in a given year, divided by the number of children of official graduation age in the population. Primary school completion rates are fairly well correlated to GDP/capita (PPP, 1995 constant international dollars), as shown in Figure 2.7. Some countries have higher than expected youth literacy rates (e.g., Tanzania, Yemen, Bangladesh, Vietnam, Azerbaijan, and Iran), whereas others have lower than expected (e.g., Niger, Mozambique, Burkina Faso, Mali, Senegal, Lebanon, Panama, and Croatia).

**FIGURE 2.7 PRIMARY SCHOOL COMPLETION RATES VS. GDP/CAPITA (LOG) (2002)**



Source: World Development Indicators, 2004

**Rationale:** The ability of the government to provide either better than expected or worse than expected primary school completion is an indicator of legitimacy. Countries that perform better than expected do well at minimizing social inequalities throughout the society. When their citizens compare themselves to those of peer countries (neighboring countries with similar GDP/capita) they will be satisfied with the results. The converse is true for those countries that perform worse than expected.

**Specifics:** Fit a univariate log-linear model predicting primary school completion rates as a function of GDP/capita and compute the residuals.

### Assessment Criteria:

<i>Relevance:</i>	Acceptable
<i>Coverage:</i>	Acceptable (Note: to achieve reasonable coverage, the most recently reported data from 2001–2003 must be used.)
<i>Update Frequency:</i>	Acceptable
<i>Lag Period:</i>	Acceptable
<i>Authority:</i>	Acceptable
<i>Robustness:</i>	Deficiencies

**Next Steps:** Acquire the World Development Indicators, merge the primary school completion rate and GDP/capita indicators into a common fragile states database using a common coding scheme for countries and time periods, and calculate the residuals from a log-log regression.

#### 2.4.2.6 Cultural and Religious Freedoms

**Source:** Minorities at Risk (<http://www.cidcm.umd.edu/inscr/mar/>)

**Rationale:** Social legitimacy is undermined when groups are denied the opportunity to practice their cultures, speak their languages, or practice their religions. Minorities at Risk codes variables describing restrictions on: religion, use of language, use of language instruction, ceremonies, appearance, family life, cultural organizations, and other cultural activities. The presence of such restrictions is a good indication there are distinct group biases in the provision of social services.

**Specifics:** For each group-country pairing, the Minorities at Risk project codes a series of variables (CULP01–CULP08) that indicate the presence of restrictions on religion, use of language, language instruction, ceremonies, appearance, family life, cultural organizations, and other cultural activities. Each variable is coded with a score of 0 (no restrictions), 1 (activity informally restricted), 2 (activity somewhat restricted), 3 (activity sharply restricted), or 99 (no basis for judgment). Thus, an overall score ranging from 0 to 8 can be constructed for each group by converting 99 values to 0 and adding CULP01 through CULP08 together. A country-wide score can be computed by adding this sum for all groups in the country.

**Next Steps:** Acquire the Minorities at Risk dataset, compute the indicators suggested above, and merge them into a fragile states dataset using a common set of country and time period codes.



### 3.0 RECOMMENDED METHODS FOR CONSTRUCTING COMPOSITE

#### INDICES

Computing composite indices, whether for effectiveness and legitimacy overall or for each cell of the State Performance Outcomes Matrix, requires a method for combining measures with different scales and weights into a believable result—one that does not seem fixed to deliver the results desired by policy makers or other interested parties. The problem is complicated by the fact that, by necessity, we are mixing measures of various types. Some are continuous (e.g., infant mortality rates), whereas others are categorical (e.g., factionalized political competition). Furthermore, if we start by computing overall indices for effectiveness and legitimacy (as suggested in the introduction), we have to somehow provide explicit weights for each component indicator since we can no longer allocate weights by sector. We present a set of suggestions for how this should initially be done.

One approach to assessing the consistency of variables in each cell of the matrix and identifying weights of variables in each composite index is to perform factor analysis on the set of its constituent indicators. Factor analysis is designed to take a collection of indirect measures of an underlying, but not directly measurable, quantity and develop an estimate of the underlying quantity by explaining correlations in the indirect measures.

Thus, given a basket of effectiveness or legitimacy indicators, factor analysis can be used to assign weights to each indicator and estimate scores of effectiveness and legitimacy. Indeed, this is the technique used to construct effectiveness and legitimacy scores from the Political Instability Task Force's State Capacity Survey.

Factor analysis also provides a diagnostic measure indicating how much of the variance among indicators can be explained with the estimate of the underlying quantity. We should question the placement of variables within baskets if we cannot explain a reasonable amount of the variance among the indicators (e.g., 75 percent). Thus, in addition to assigning weights, factor analysis can help us design baskets of reasonably consistent indicators.

The difficulty with factor analysis is that it is something of a “black box” that may produce indicator weightings that do not align with *a priori* expectations or assumptions. However, this can be an advantage if we need the ability to deny constructing the indices to achieve a pre-determined result.

An alternative to factor analysis is to establish a point system for each component indicator that can be added to compute the composite index score. The advantage of such an approach is that it is fairly simple to explain to a broad audience. If all variables in a given basket are scored to a common scale (-3 to +3, in the example above), then it is relatively easy to add up the scores and assign explicit *a priori* weights to compute an overall score. The disadvantage is that the approach is entirely heuristic. The rescaling methods and the assigned weights are both subject to judgment, and critics could assert that the judgment was biased to arrive at pre-determined outcomes. Thus, we do not recommend a point system to help define composite indicators.

Z-scores or percentile rankings provide more defensible ways to aggregate indicators into composite measures. Z-scores measure the departure of a given observation from the mean, measured in standard deviations. The advantage of z-scores is that they provide an easy way to convert any continuous measure with Gaussian distribution to a common scale. The effect of outliers can be minimized by “winsorizing,” or clipping, the results to a predetermined minimum and maximum threshold (e.g., -2

standard deviations and +2 standard deviations). Thus, for continuous variables, z-scores can be used instead of designing explicit point systems. Z-scores implicitly rank countries relative to the mean. It should be noted that this is a relativistic approach and can be criticized for ignoring the importance of critical thresholds. Furthermore, z-scores alone do not work very well for categorical data as they still need to be combined using some sort of weighting system. However, z-scores can be effectively combined with factor analysis to estimate implicit weights instead of assigning explicit weights. Z-scores are used in efforts such as the Environmental Sustainability Index and selected components of the Human Development Index.

Another data-driven approach, similar to z-scores is the conversion to percentile rankings, or percentile groupings (e.g., deciles, quintiles), and then compute averages across percentile ranks.

In the end, as any single indicator aggregation method will be subject to criticism, we recommend an approach designed to achieve three objectives: 1) to ensure that all the variables in the basket are reasonably consistent with one another using factor analysis; 2) to compute and publish aggregate indicators using a z-score based system; and 3) to augment the published rankings with an assessment of how dependent a given country's ranking is on the aggregation method (factor analysis, z-score, or percentile). Most countries will receive similar ranks with all three methods. In these cases, the results will be mutually reinforcing, and the resulting ranking needs no further explanation. However, a small number of countries will likely move substantially between ranking systems. In the final report, we would publish the results of a single ranking system (e.g., z-score), but flag countries whose rankings are highly dependent on the aggregation algorithm.

**U.S. Agency for International Development**

1300 Pennsylvania Avenue, NW

Washington, DC 20523

Tel: (202) 712-0000

Fax: (202) 216-3524

**[www.usaid.gov](http://www.usaid.gov)**