**THIS WHITE PAPER IS PUBLISHED JOINTLY BY STRATEGIC MULTI-LAYER ASSESSMENT (SMA)
AND U.S. ARMY CORPS OF ENGINEERS RESEARCH AND DEVELOPMENT DIRECTORATE**

Perspectives on Political and Social Regional Stability Impacted by Global Crises - *A Social Science Context*

January 2010

The views expressed in this document are those of the authors and do not reflect the official policy or position of the organizations with which they are associated.

Editorial Board:
Hriar Cabayan (OSD), Barbara Sotirin (USACE-HQ, CERD), Robert Davis (ERDC-NH), Robert Popp (NSI), Andy Reynolds (U.S. Dept. of State), Anne Ralte (USAID), Herm Myers (OUSDP), Phil Kearley (JFCOM), Robert Allen (TRADOC), Lynda Jaques (PACOM), Sharon Halls (AFRICOM), Carl Dodd (STRATCOM/ GISC), Martin Drake (CENTCOM), MAJ Bradley Hilton (EUCOM), Tom Rieger (Gallup), Tom Mullen, (PA), Peter Steen (JS/J-5), Larry Kuznar (NSI)

Contributing Authors:
Clare Lockhart (State Effectiveness Institute), Philip Martin (UC-Davis), Jack A. Goldstone (GMU), Elizabeth Leahy Madsen (PAI), David L. Richards (Memphis), Ronald Gelleny (Akron), Peter Steen (JS/J-5), Jerome Delli Priscoli (USACE), J. Rolf Olsen (USACE), Kathleen D. White (USACE), Julie E. Kiang (USGS), D. Phil Turnipseed (USGS), Levi D. Brekke (Reclamation), David A. Raff (Reclamation), Roger S. Pulwarty (NOAA), Robert Webb (NOAA), Alex O.Dehgan (DOS), Laura Jean Palmer-Moloney (USACE), James Diaz (LSU), Kenneth S. Yalowitz (Dartmouth), Ross A. Virginia (Dartmouth), Richard J. Stoll (Rice), Douglas J. Arent (NREL), Jeffrey Steiner (USDA), Timothy Griffin (Tufts), Donald Suarez (USDA), Debarati Guha-Sapir(CRED), Daniel Strickman (USDA), Victor Asal (Albany), Stephen Shellman(William and Mary), Lawrence A. Kuznar (NSI), Laura Leets (NSI), J. Joseph Hewitt (CIDCM), Tom Rieger (Gallup), Frank C. Zagare (SUNY Buffalo), Robert Axelrod (Michigan), George Cybenko (Dartmouth), Douglas Madory (Dartmouth), Stephen M. Millett (Futuring Assoc), Ame Stormer (DMHA), Jessica Wambach (DMHA), and Lt Gen (Ret) John F. Goodman (DMHA), Robert Popp(NSI), Tom Mullen (PA)

Compiled by: Rosa Affleck (ERDC-NH) – Rosa.T.Affleck@usace.army.mil
Rose Rainey (IIA) – rraney@iiaweb.com
April Hartman (NSI) - ahartman@NSIteam.com
Deborah Pyle (SRC) – Deborah.pyle@js.pentagon.mil

Table of Contents

PREFACE (RDML John E. Roberti) **3**

Executive Summary (Rosa Affleck, Todd Anderson, Roy Cooper, Rose Rainey) **4**

1. Stability and State Building (Clare Lockhart) **12**

2. Independent Trend and Trajectory of the Future: Issues Impacting Social Regional Stability **23**

2.1. Demographics 23

2.1.A. Demographic and Economic Trends: Implications for International Migration (Philip Martin) 23

2.1.B. Demography and Security (Jack Goldstone) 32

2.1.C. Demographic Security (Elizabeth Leahy Madsen) 44

2.2. Economics 51

2.2.A. Banking Crises, Collective Protest and Rebellion (David Richards, Ronald Gelleny) 51

2.2.B. Economics and Political Instability within the Global Economic Crisis (Peter Steen) 73

2.3. Water 78

2.3.A. Water and Security: Cause War or Help Community Building? (Jerome Delli Priscoli) 78

2.3.B. Water Resources and Stability (J. Rolf Olsen, Kathleen White, Julie Kiang, D. Phil Turnipseed, Levi Brekke, David Raff, Roger Pulwarty, Robert Webb) 89

2.3.C. Water Security and Scarcity: Potential Destabilization in Western Afghanistan (Alex Dehgan, Laura Jean Palmer-Moloney) 98

2.4. Changing of Climate and State Weather 110

2.4.A. Maintaining Geopolitical and Social Stability throughout a Global Economy in an Era of Climate Change (James Diaz) 110

2.4.B. The Arctic Region: Prospects for a Great Game or International Cooperation (Kenneth Yalowitz, Ross Virginia) 125

2.4.C. Changing Climate Impacts to Water Resources: Implications for Stability (Kathleen White, J.Rolf Olsen, Levi Brekke, David Raff, Roger Pulwarty, Robert Webb) 131

Energy 144

2.4.D. Energy, Africa, and Civil Conflict: What Does the Future Hold? (Richard Stoll) 144

2.4.E. Energy: A National and Global Issue (Douglas Arent) 150

2.5. Food 170

2.5.A. World Food Availability and the Natural Land Resources Base (Jeffrey Steiner, Timothy Griffin) 170

2.5.B. Food Production in Arid Regions as Related to Salinity (Donald Suarez) 178

2.6. Epidemics 184

2.6.A. Epidemics: A Thumbnail Sketch of the Past, Present and Future (Debarati Guha-Sapir) 184

2.6.B. Infectious Disease and Social Instability: Prevent, Respond, Repair (Daniel Strickman) 190

3. Social Science Modeling and Analysis Approaches 196

3.1. Analyzing Political and Social Regional Stability with Statistics: Challenges and Opportunities (Victor Asal, Stephen Shellman) 196

3.2. The Social Stability of Societies: An Anthropological View (Lawrence Kuznar) 203

3.3. Quantitative Content Analysis (Laura Leets) 213

3.4. The Peace and Conflict Instability Ledger Ranking States on Future Risks (J. Joseph Hewitt) 219

3.5. Perception is Reality: Stability through the Eyes of the Populace (Tom Rieger) 236

3.6. Assessing the Stability of Interstate Relationships Using Game Theory (Frank Zagare) 242

3.7. Theoretical Foundations of Partnerships for Economic Development (Robert Axelrod) 247

3.8. Process Query System as a Framework for Modeling and Analysis of Regional Stability (George Cybenko, Douglas Madory) 258

3.9. The Use of Cross-impact Analysis for Modeling, Simulation, and Forecasting (Stephen Millett) 267

3.10. Development of a Framework for Action: Community Resiliency as a Means of Achieving Stability (Ame Stormer, Jessica Wambach, Lt Gen (Ret) John Goodman) 277

4. Case Studies 283

4.1. Sudan Strategic Assessment: A Case Study of Social Science Modeling (Robert Popp) 283

4.2. Analyzing Stability Challenges in Africa: Case Study of an Approach (Tom Mullen) 289

Appendix A. Acronyms 296

PREFACE (RDML John E. Roberti)

This timely and important compilation of papers forms an essential step in understanding the importance and interdependency of factors effecting stability. The authors, both academicians and military subject matter experts, provide invaluable insight by adding their accumulated knowledge and expertise to the project.

The ability to assess and foresee potential instabilities presents a challenging problem. Recent crises in political, social, economic, climatologic and energy issues represent just a few indicators of continuing global instability.

As we consider the complexities of the 21st century, quantitative and qualitative approaches to the study of instability will be integral to developing an appropriate understanding of the strategic context for national security. This compilation assists in building a more rigorously informed understanding of these key issues.

It is critically important, for us to understand global strategic risk, that we identify and fully comprehend the interdependence among these complex issues influencing regional and global stability. By determining the probability and consequence of this risk, we can take a prioritized approach to reducing instability by a well-informed application of all elements of national power.

John E. Roberti

Rear Admiral, USN

Deputy Director, Strategy and Policy, J-5

The Joint Staff

Executive Summary (Rosa Affleck, Todd Anderson, Roy Cooper, Rose Rainey)

Authors and Organizations: Rosa Affleck, U.S. Army Engineer Research and Development Center CRREL; Todd Anderson, Roy Cooper, Rose Rainey - Information International Associates, Inc.
Contact Information: Rosa.T.Affleck@usace.army.mil, tanderson@iiaweb.com, rcooper@iiaweb.com, rraney@iiaweb.com

A Challenging Problem

Cooperative security, stability operations and irregular warfare missions require a better understanding of the complex operational environment, notably through rich contextual understanding of the factors affecting stability. Further assessment, policy, and planning need to consider factors associated with institutional performance (community organizations, government ministries, legal structures, etc.) based on how societies emerge, develop, and function, as well as attributes that provide resiliency and flexibility.

Today, stability experts are faced with a new environment in which the world is highly interconnected, change is very rapid, and threats are multifaceted; all of these pose very different challenges to the US Government (USG). The current financial situation underscores both the rapidity and global extent of economic collapse, and it has exacerbated problems in other areas. Solutions in one area can have first-, second-, and third-order effects in other areas; these effects can be both positive and negative. Global average food prices increased by more than 80% during 2005-2008, sparking food riots in Africa, Asia, the Middle East, the Former Soviet Union, and Central and South America. Contributing factors are complicated; for example, shifts in food demands may be contributing to the food price increases, as increasing use of food crops for bio-fuels and increasing demand for protein-rich diets dramatically decreases efficient use of grain calories. Estimates of future water availability are alarming, and while earthquakes often impact manmade water management structures, reports also suggest that geophysical changes caused by large dams may have triggered earthquakes, including China's 7.9 magnitude earthquake last year along a fault near the Zipingpu Dam and Reservoir that left 80,000 people missing or dead. In addition to food and water issues, severe weather events and climate change, shifts in demographics, increasing energy demands, pandemics, and threatened usage of nuclear weapons are threats that individually, or worse, in combination, can significantly increase the fragility of world stability.

While in today's increasingly interconnected world, global crises and unstable regions pose an acute risk to world security and could provide unforeseen circumstances ripe for manipulation and exploitation, these same threats can serve as rallying instruments, catalyzing disparate groups to work in concert to develop coordinated responses and preparedness mechanisms. This coordination can also result in development of negotiation venues for other issues. First-, second- and third-order effects can have positive impacts. In addition to the factors identified in the frameworks above, other dimensions of consideration include (1) institutional performance (community organizations, government ministries, legal structures, etc.) as a function of how societies emerge, develop, and operate, and (2) attributes that provide resilience and flexibility.

So the lingering question remains, where will the next occurrence of regional instability be that requires U.S. intervention? How should we shape the structure of our future force to respond to such instability? This white paper brings some of best minds together to examine the following aspects of this challenging problem:

- Individual factors impacting regional stability
- How these individual factors can combine to create tipping points that drive significant regional instabilities
- Approaches to forecast and anticipate where future instabilities may occur to give our warfighters an unfair advantage in the future.

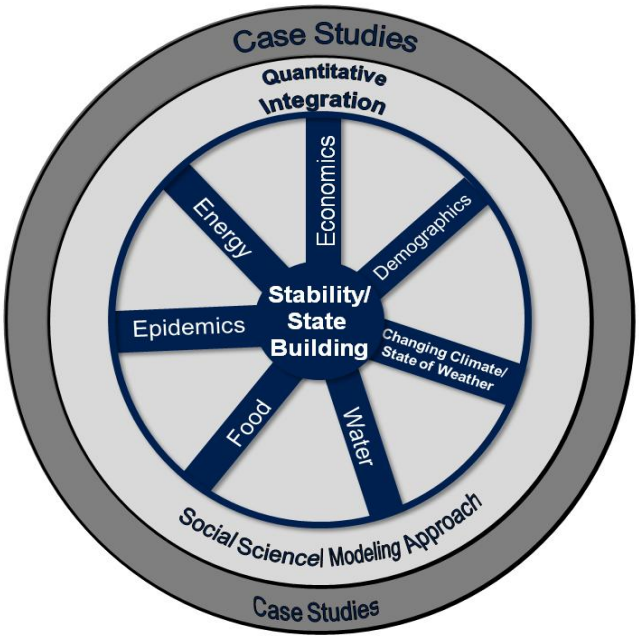
Collecting the Best Ideas

To capture the breadth of this issue, this paper examines issues ranging from political, infrastructural, demographic, economic, resource-related, climatologic, energy-related, epidemiological, sociological, and analytical issues as they impact regional stability. Decision makers are dealing with an evolving situation in which societies are increasingly interconnected; changes occur rapidly, often with unexpected consequences; and threats to stability are complex. This white paper is a comprehensive compilation and assessment of complex issues influencing regional stability by a diverse group of authors that discuss the assessment of regional stability through the lens of crises in food, water, changing climate, energy, economy, demographics and epidemics. These articles, some of which have been previously published, provide critical insights into factors and tools that can apply to this complex challenge. Contributors to this white paper have provided a rich analysis of the factors affecting stability, as well as new perspectives based on careful study and experience germane to today's issues. This compilation yields an overarching view of regional stability as it impacts global and regional crises. Through detailed citations of the scholarship and current thinking throughout their communities, the authors make a compelling case for how multiple factors can coalesce to result in radical impacts on regional stability.

Bringing Wide-ranging Ideas Together

The organizational approach used to bring together the wide-ranging factors associated with stability is analogous to a wheel. At the hub, the central support, we examine the nature of stability and state building. What makes it work? This paper outlines the key aspects of functioning states, examines how they contribute to the good of all, and defines how they resist threats and provide stability.

Emanating from the hub are seven enablers of stability: demographics, economics, water, changing climate and weather, energy, food, and epidemics. Each of these articles discusses the relationship that enabler has on the stability of a society.



The rim of the wheel is a series of articles describing possible quantitative and integration methods coupled to social science modeling approaches. These approaches provide potential means to use existing and emerging information to forecast regional instabilities. The "tread" of the wheel is comprised of case studies in social science modeling. These articles provide a context and examples of unique challenges in predicting trends in social sciences.

Bringing Order to a Complex Problem

The factors in the frameworks discussed above form components of a complex system that interacts in dynamic nonlinear ways and has significant emphasis on social systems. It has become sometimes painfully obvious that evasion of independent consideration of these factors can hinder pursuing outcomes of national importance. Accordingly, the USG has considerable interest in descriptive models – with substantial attention and investment in a wide array of social science models in recent years – developed to inform a more comprehensive assessment of current security conditions and to enlighten potential future security outcomes that often ignore geo-political boundaries, that are insensitive to cultural issues, elude legal categorizations and/or expand beyond national economic conditions.

Assessment of regional stability needs to consider indicators associated with political, social and institutional performance based on how societies emerge, develop, and operate, as well as attributes that provide resilience and flexibility. Identifying geopolitical and socio-cultural indicators that target a confluence of factors, and establishing trigger thresholds may prove as important as defining the individual factors themselves. At the same time, assessment must also consider indicators associated with drivers of instability and conflict (economic decline/shock, criminalized security forces, environmental degradation, struggles for absolute power, etc.). We anticipate exploring multiple social science theories, research methods, and well-known quantitative/statistical analysis techniques (cluster analysis, nodal analysis, cross-correlation, and factor analysis) in this white paper to (1) independently select relevant indicators (climate change/change of weather, economy, water, food, pandemics or epidemics, energy), (2) operationalize and validate indicator coding schemes, (3) clarify interdependencies across indicators (impact of water policy on energy, health and economy), and (4) regroup indicators by looking at a confluence of factors, priorities and contributing factors and drivers of conflict and rallying mechanisms. Finally, assessment should function from an asset-based perspective that is focused on indigenous capabilities, perceptions, systems, interactions and activities at multiple levels, at multiple stages, and accounting for variability and rate of change. The result of the assessment should enable operators to understand what conditions exist, why they exist, and how best to transform them.

This white paper provides a comprehensive examination of the factors that influence regional stability. Points to investments and initiatives that will likely improve our ability to predict the

consequences of crisis on regional stability can help minimize the threat of instability. Through the analysis of multiple social science theories, research methods, and established quantitative and statistical analysis techniques, this white paper, *Perspectives on Political and Social Regional Stability Impacted by Global Crises - A Social Science Context*, provides a significant step along the path to resolving the complex problems critical to the future security of the U.S.

This edited volume, consisting of short contributions (5-7 pages each), will describe definitions of stability, will examine assessment approaches, and will extend to encompass strategies for tailored assessment and planning.

Clare Lockhart (Section 1) addresses the significance of "Stability and State Building." This paper describes how functioning states are a vital piece of global architecture and that as such they not only provide critical resistance to a variety of threats, but they also contribute to the collective global goals and stability in the 21st century.

Philip Martin of UC-Davis (2.1.A) reviews International Migration. The number of international migrants, defined by the UN as persons outside their country of origin a year or more, for any reason and in any legal status, more than doubled between 1990 and 2010. Current default policy is to manage what is often seen as out-of-control migration and by adjusting the rights of migrants, leading to conflicts over human rights. Martin's paper analyzes long-term factors affecting migration patterns, including aging in industrial countries, rural-urban migration that spills over national borders, and the migration infrastructure of agents and networks that moves people.

Jack Goldstone (2.1.B) continues this theme with "Demography and Security," in which he discusses the five major demographic trends likely to pose significant security challenges to the majority of developed nations in the next two decades. He notes that problems will be caused not by the overall population growth, but by the population distortions, in which populations grow too young or too fast, or become too urbanized.

In "Demographic Security," Elizabeth Leahy Madsen (2.1C) reviews findings regarding population and two security issues—outbreaks of civil conflict and level of democratic governance—at the global scale. The theme of her paper is that population influences security and development, and it is an important underlying variable in global stability because of its interactions with other factors.

David Richards and Ronald Gelleny (2.2.A) address the relationship between banking crises and domestic agitation/internal conflict in "Banking Crises, Collective Protest and Rebellion." They examine a dataset of 125 countries for the years 1981 to 2000 and find banking crises to be systematically associated with greater levels of collective protest activities such as riots, anti-government demonstrations, and strikes.

An overview of work done at the Joint Staff J-5 is outlined by Peter Steen (2.2.B) in "Economics and Political Instability within the Global Economic Crisis." This paper provides a strategic national-level understanding of the ongoing global economic crisis, as well as investigates the relationship between economics and political instability within that context.

Jerome Delli Priscoli (2.3.A) of the U.S. Army Corp of Engineers discusses the role of water and its relationship to international stability with "Water & Security: Cause War or Help Community Building." He posits that the water and security debate is driven by our notions of scarcity, and

that ultimately, the strategic aspects of water lend themselves to finding means for cooperation rather than conflict.

An interagency team - Olsen and White (USACE), Brekke and Raff (US Bureau of Reclamation), Kiang and Turnipseed (USGS) and Pulwarty and Webb (NOAA) - representing the two largest water resources operating agencies (USACE and Reclamation) and two major water resources data and science agencies (USGS and NOAA) (2.3.B) continue the water dialogue in "Water Resources." Their paper describes the multiple factors that can impact and stress water resources. Due to those interactions, solutions to water resource problems should follow a comprehensive approach that integrates multiple objectives across the proper spatial and temporal scales with all relevant stakeholders participating in the decision-making process.

In "Water Security and Scarcity: Potential Destabilization in Western Afghanistan," Alex Dehgan, and Laura Jean Palmer-Moloney (2.3.C) offer a water resources case study. The paper highlights the implications of plans for upgrading and developing Afghanistan's water infrastructure in the Helmand River watershed. While crucial to the social and economic development of Afghanistan, these plans will also impact transboundary water flow and as a result, Afghanistan's relations with Iran.

In "Maintaining Geopolitical and Social Stability throughout a Global Economy in an Era of Climate Change," James Diaz (2.4.A) describes the uniform agreement in the international scientific community that the earth is warming from a variety of climatic effects. Ultimately, this change will have far reaching impacts on human health and public safety. Diaz believes the challenge for the U.S. will be to assume leadership in maintaining geopolitical and social stability throughout the global economy in an era of climate change.

Kenneth S. Yalowitz and Ross A. Virginia (2.4.B) address the role of the Arctic in the changing climatic environment in "The Arctic Region: Prospects for a Great Game or International Cooperation." Their theory is that the pace of ecological, political, social, and economic change in the Arctic region is accelerating due to the warming climate. The paper evaluates the prospects for contrasting outcomes in the Arctic region: a return to international power politics as states seek to claim Arctic energy, extend continental shelves, and enforce their wills through military means versus the emergence of increased international cooperation around environmental protection and sustainable development.

"Changing Climate Impacts to Water Resources: Implications for Stability" is authored by Kathleen D. White, J. Rolf Olsen, Levi D. Brekke, David A. Raff, Roger S. Pulwarty, and Robert Webb (2.4.C), all from the U.S. Army Corp of Engineers and discusses climate change in the context of water resources, including the potential implications for stability issues. They argue that water resource managers are already accustomed to dealing with changes and therefore offer a potential resource to utilize for the larger issue if they are prepared to act quickly. The authors propose a strategy encompassing both the potential for increased conflict over water and increased cooperation by water resources managers to enhance planning for stability.

An introduction to the role of energy begins with "Energy, Africa and Civil Conflict: What Does the Future Hold?" by Richard J. Stoll (2.5.A). Stoll notes that Africa will likely become increasingly important to the U.S. as a source of resources, including oil. However, Africa is also rife with civil conflict. Ongoing civil conflict makes it very difficult to establish or continue the exportation of natural resources. Therefore, it is in the U.S.'s best interest to address the issue of civil conflict in Africa.

Douglas J. Arent (2.5.B) continues this topic of Energy with "Energy: A National and Global Issue," and states that new energy pathways are a necessity to balance the increasingly complex policy goals of accessibility, environmental concerns, geopolitical issues, and affordability. Continuing reliance on geographically concentrated oil and natural gas to feed the ongoing demand will threaten international energy security.

Jeffrey Steiner and Timothy Griffin (2.6.A) address the role of food in relation to the larger picture of Stability in "World Food Availability and Natural Land Resources Base." Steiner and Griffin note that there is a need to consider how changing population and wealth patterns will not only impact food availability and consumption patterns, but also our inter-related needs for energy and water. This is increasingly important in relation to the Earth's finite land mass.

Donald Suarez (2.6.B) follows with "Food Production in Arid Regions Due to Salinity." He discusses the issue of the water supply, the impact of salinity, the potential for water reuse—both of irrigation drainage and municipal waste water—and utilization of saline waters for crop production. Improvements in irrigation practices, investment in new technologies, and development of salt-tolerant plant varieties may enable these regions to utilize more abundant brackish and saline waters for irrigation and may minimize degradation of fresh water supplies.

The discussion of Epidemics begins with "Epidemics: A Thumbnail Sketch of the Past, Present and Future," by Debarati Guha-Spair (2.7.A). Disease outbreaks have significant impacts on factors that are critical for national and international stability. There are clear disparities between rich and poor nations and their abilities to react and control the situation. Balancing policies to address the problems will also be a challenge for global disease control.

The Epidemics discussion continues with "Infectious Disease and Social Instability: Prevent, Respond, Repair," by Daniel Strickman (2.7B). The effects of infectious diseases on social instability can be devastating and society's ability to prepare and respond to an epidemic can offer social stability. The concept of Integrated Disease Management is presented as a construct to mitigate the societal impacts of infectious disease using the functions of risk assessment, surveillance, prevention and control, and sustainable support.

The role of social science modeling and its approaches to analysis is presented in Section Four. Victor Asal and Steve Shellman (3.1) begin with "Analyzing Political & Social Regional Stability with Statistics: Challenges and Opportunities." They provide an overview of some of the research on the causes of stability and instability done using statistical analysis. A discussion of the efforts made in the area of forecasting is presented, as well as the challenges of statistical analysis related to issues of data and method.

Larry Kuznar (3.2) presents the anthropological view in "The Social Stability of Societies: An Anthropological View." He notes that instability does not appear overnight and that a longer term historical perspective is necessary to understand the latent factors that accumulate slowly and then result in dramatic social collapse. Anthropology provides this perspective, and this chapter reviews insights concerning why some societies fail while others prosper.

"Quantitative Content" is presented by Laura Leets (3.3) and focuses on the central concepts underlying content analysis and how to conduct effective research. Content analysis is a technique for gathering and analyzing the content of text. Text can be anything written, visual or spoken, which serves as a medium for communication. Content analysis can be utilized in either qualitative or quantitative format; however, this submission focuses on the quantitative uses.

Joe Hewitt (3.4) contributes "The Peace and Conflict Instability Ledger Ranking States on Future Risks." He presents country rankings from the *2010 Peace and Conflict Instability Ledger* which are based on newly calculated risk estimates. The ledger represents a synthesis of some of the leading research on explaining and forecasting state instability. Hewitt also discusses some of the key results from the analysis, including the pivotal relationship between democratization and risk of instability.

Tom Rieger (3.5) provides a discussion into the problems related to developing stability models in "Perception is Reality: Stability through the Eyes of the Populace," the largest problem being the limitations due to what sources are available. He describes how having a robust model of the level of stability in a given population based on perceptions of conditions would be a major contribution to the ability to plan for—and possibly help avoid—significant human suffering as a result of instability.

In "Assessing the stability of Interstate Relationships Using Game Theory," Frank Zagare (3.6) explains the sense in which game models can be used to establish the stability, or lack thereof, of typical deterrence relationships and to understand the context of their policy recommendations. Game-theoretic models are a natural and intuitively satisfying framework in which to assess the stability of contentious inter-state relationships.

Robert Axelrod (3.7) presents a simple theoretical framework to enhance insight about partnerships for economic development. "Theoretical Foundations of Partnerships for Economic Development" clarifies the idea of theoretical foundations of partnerships by analyzing partnerships using game theory of an iterated prisoner's dilemma. The analysis then reveals implications of selecting a partner, setting up a partnership, choosing a *modus operandi*, building trust, achieving selectivity, and performing monitoring and evaluation.

"Process Query System as a Framework for Modeling and Analysis of Regional Stability" is authored by George Cybenko and Douglas Madory (3.8). In response to similar problems across a variety of application domains, suggesting an underlying common analytic foundation, they propose two technologies—Process Query Systems (PQS) and Human Behavioral Modeling Language (HBML)—which could form the foundation for a standard, common computational capability. This capability could then be used to represent economic, health, political and environmental models related to regional stability, and reasoning about those models in the context of data, observations and other evidence.

The focus of Stephen M. Millett's contribution (3.9), "The Use of Cross-Impact Analysis for Modeling, Simulation and Forecasting," is to assert that cross-impact analysis may be just as effective, and arguably quicker, less expensive, and more robust, than systems dynamics. Millett asserts that it is a complementary approach that provides further foresight for the benefit of both forward-looking analysts and decision-makers, and he recommends further exploration as a supplementary approach to system dynamics, as well as other modeling, simulation and forecasting methods.

The Center for Excellence in Disaster Management and Humanitarian Assistance (3.10) submits "Development of a Framework for Action: Community Resiliency as a Means of Achieving Stability," which outlines their process of creating a defining framework. The framework would determine what makes an individual household, community, and society resilient before, during, and after disasters.

Finally, Section Five looks at two case studies applicable to the Stability issue. First, Robert Popp (4.1) reports on the Sudan in "Sudan Strategic Assessment: A Case Study of Social Science Modeling." This assessment was a strategic level proof-of-concept study in which a combination of quantitative and computational social science modeling and analysis approaches were developed and applied to better understand a complex "state" lacking true borders and encompassing many competing interests and complexities. Results demonstrated how multiple quantitative and computational social science models in conjunction with SMEs and other analyses are an effective, evolutionary step in the analyst toolkit, especially when the need is to provide additional lenses to look at highly complex and ambiguous stability problems (like the Sudan) to inform the decision-making process.

Second, Tom Mullen (4.2) presents "Analyzing Stability Challenges in Africa: A Case Study." He notes that, with resources spread thin, we need better ways to rapidly understand what matters in each new situation, and to better understand why particular actions worked (and others did not), to aid in determining where and how those lessons might be more likely to work well. Assessment of highly complex situations quickly improves the ability to take rapid, effective action. Mullen's case study attempts to provide lessons drawn from analyzing a number of stability situations on the continent of Africa over the past three years, with a focus on lessons in analyzing complex stability situations rather than specific actions to apply in a wide range of situations.

↑ Back to the top

1. Stability and State Building (Clare Lockhart)

Author and Organization: Clare Lockhart, The Institute for State Effectiveness

Contact Information: clareelockhart@gmail.com

Introduction

The international order created sixty years ago in the wake of World War II assumed a world premised on a fabric of states, each delivering core services to citizens and fulfilling its obligations to the global community of states. The visionary architects of that order succeeded in creating institutions that underpinned democratic stability and market-based economies in the United States, and in Europe and Asia. While those states are not perfect, they nonetheless provide their citizens with a degree of resilience against looming global threats, and they remain the best instrument through which the collective aspirations of 21st century humanity may be realized.

Meanwhile, nearly two billion people remain locked out of stability and prosperity because they live in states that have been unwilling or unable to deliver even core services. These states prevent their citizens from accessing the opportunities of globalization, but they have also been unable to shield their citizens from the darker side of globalization. Few citizens in these countries can participate legitimately in the unprecedented prosperity and opportunity unleashed in the last 60 years, although, through globalized media, citizens see another world full of opportunity, innovation and prosperity which defines their expectations. At the same time, their states are powerless over—if not complicit with—global networks of criminality and violence that thrive upon disenfranchised populations, uncontrolled territory, and weak state institutions. Such states cannot hope to meet the array of challenges presented by climate change, resource shortages, pandemics, and food and fuel price shocks. Nor can they help position their citizens to take advantage of emerging global opportunities. Refugee flows, humanitarian action, and trafficking of people, arms and drugs are all symptoms of this arc of state dysfunction.

The international community affords these states the same legal authority that it affords those states actually performing core functions for their people. At the same time, recognizing the constellation of problems emanating from weak or dysfunctional states, the international community draws on a pool of international institutions designed to reduce poverty, address serious humanitarian crises, and, in some cases, resort to force. These institutions, while staffed by competent and dedicated people, have relied upon instruments devised for a different world, and institutional silos and cultures have hampered collective and systemic understanding of the problem.

This paper sets out how functioning states are vital to stable global architecture, as critical providers of resilience against a range of threats, and as key instruments for the collective realization of aspirations in the 21st century. It considers (1) the functions that a state should perform for its citizens in the current context, (2) a framework through which state functionality could be catalyzed, (3) and how the approach described was applied in Afghanistan between 2002 and 2004. Finally, this paper suggests emerging lessons and considers aspects of an approach to future state-building efforts in Afghanistan.

Stability and the State

For states to be stable, they must establish and maintain relations of trust with their citizens. Citizens' loss of trust in the ability of their state to create an inclusive political, social and economic order made predictable by rule of law will erode the state's legitimacy. This erosion has been characterized by "an increase in illegality, informality, and criminality in the economy; ineffective delivery of basic services; failure to maintain or expand essential infrastructure; increase in corruption; and appropriation of public assets for private gain. As a result, administrative control weakens and the bureaucracy is seen as an instrument for abuse of power, in turn leading to a crisis in public finances—where both revenue and expenditure are unpredictable and budgeting becomes an exercise in emergency management."[1] Ultimately the state resorts to using violence internally, and armed groups begin to appear, defying state authority and taking control of sections of territory.

State Dysfunction and Regional Instability

The kind of instability described above poses profound regional and international challenges. For their neighbors, such states will be a constant cause for concern. Criminal networks will take advantage of such poorly controlled environments, and they will also operate across state borders through the trafficking of illicit goods such as drugs and arms, and other criminal activities. Such activities subvert rule of law and threaten stability in neighbors, particularly neighbors whose own legitimacy and capabilities are fragile. Ultimately, dysfunctional states may become bases for insurgent groups operating against neighboring states. Insurgents will interact in complex ways with criminal networks and will engage in criminal activities themselves in order to finance and equip their activities. Unless states can be rehabilitated, these dark features of globalization are extremely difficult and expensive to stifle.

The State and Global Challenges

As globalization has accelerated, new threats have arisen. Lessons from the first wave of globalization would indicate that the spontaneous character of globalization demands that the process be led rather than treated reactively. In this context, the task of building effective states is both more formidable and more urgent. In the 21st century, global challenges ranging from climate change to sustainable energy and food security, to appropriate management of international financial markets, all demand marshaling collective resources. Not the least of these demands is the need for political resources, as no single state or group of states has the capacity to confront such challenges single-handedly. The task of catalyzing the emergence of effective states is essential for heading off global challenges that threaten stability.

Finally in this context, one of the key challenges to global stability is the arc of state dysfunction itself, given the constellation of regional and global criminal and military risks, and the likely humanitarian implications of ongoing state dysfunction.

The State and National Resilience

Citizens across the world are exposed daily to an array of powerful global forces over which, as individuals, they exert little influence. An effective state, however, can plan for and respond to the impact of such forces, whether through implementing early warning systems and evacuation procedures for hurricanes, through providing reserves of food and water to insulate populations from shortages, maintaining cash reserves, or implementing public health interventions. The ability to effectively prepare for and prevent or mitigate the effects of a crisis is a key

responsibility for states, and their success or failure in building national resilience and partnering with other states to tackle international challenges will actively generate or undermine stability, because citizens judge effectiveness by performance.

The State and Collective Aspirations

On the other hand, while globalization has extended new opportunities and released unprecedented wealth, dysfunctional states continue to prevent their citizens from accessing either. Instability stifles foreign investment, and failure to invest in human capital shuts doors of education and economic opportunity to millions of people. Furthermore, instability stifles the economy's capacity to develop in the longer term, and therefore future state revenue. Disenfranchisement naturally breeds disillusionment, and in contexts where lost generations of youth have been denied an education, steady employment, or a stake in the future, predictably volatile consequences follow. Failure to meet collective aspirations is a symptom of state dysfunction and a driver of ongoing instability.

1.1 The Functions of the State

While agreeing that ineffective states are at the heart of enormous and overlapping global problems, international responses to state failure have frequently fallen short of their goals.[2] A key reason for this has been lack of understanding or consensus on what tasks a state needs to perform in order to serve its citizens in the 21st century. Without this agreement, the international community has not been able to devise instruments for helping a government acquire the capacity to serve its citizens, nor to understand the necessary timeframes, sequencing or interdependencies between state functions. While in recent years there has been an emerging consensus on the need to build functioning states—to mitigate threats, to realize aspirations for citizens, and to act as essential pillars of a functional and stable international order—consensus on a coherent approach remains vital but pending work.

Because agreement on core functions is a prerequisite for concerted and effective international action, the Institute for State Effectiveness (ISE) has proposed that the state should perform ten core functions for its citizens in the 21st century.[3] It is hoped that these will be subjected to discussion and debate, but based on reading of history, engagement with international development, and experience with the challenge of state-building in some of its most difficult contexts, ISE has proposed the following ten functions as critical.

Rule of Law

The rule of law is perhaps the most important function in that it defines the governance arrangements for all other functions. It sets parameters for behavior within society, opening opportunities while setting constraints. It is the glue that binds all aspects of the state together, defining limits of power, rights and responsibilities for citizens and for the state itself.

Monopoly on the Legitimate Means of Violence

Weber's celebrated definition of the state emphasizes this function above all others, but nonetheless stresses the need for legitimate monopolization of the means of violence. Beyond achieving actual control over the means of violence, the state needs to subordinate decisions about the use of violence to rules recognized as legitimate.

Administrative Control

When a state is managed by a competitively recruited professional cadre, administrative control is achieved by organizing a state's functions and spatial territory via a unified body of rules and practices and a hierarchy of administrative units performing specialized functions.

Sound Management of Public Finances

Public goals can only be achieved through sound management of public finances. The budget acts as a lens through which hard decisions are made about competing priorities so that resources are allocated to translate aspirations into outcomes. Without effective systems for collecting and allocating revenues, the state will be unable to deliver services adequately.

Investments in Human Capital

Stability depends on forging paths of upward social mobility and creating stakeholders in the future. Such investments are the key to building domestic capacity across the full spectrum of state and market activity, and as human capital becomes more important than other forms of capital, public investment in building human capital becomes increasingly important.

Creation of Citizenship Rights through Social Policy

Creating stability within the state means overcoming fragmented identities and a politics defined by zero-sum competition for control over resources by competing groups. Social policy geared to equality of opportunity, rights and responsibilities—cutting across gender, ethnicity, race, class, spatial location, and religion—can overcome this divisiveness and weave a social fabric of national unity and a sense of shared destiny.

Provision of Infrastructure Services

In order to provide security, administration, investment in human capital, and support to the market economy is the need for adequate transportation, power, water, communications and pipelines. Exploring and harnessing regional potential in infrastructure can not only create efficiencies, but can also create virtuous webs of cooperation and closer ties between states.

Formation of a Market

The state has critical roles to fulfill in creation and expansion of the market. It sets and enforces the rules for commercial activity, supports the operation and development of private enterprise, and intervenes at times of market failure.

Management of Public Assets

The state wields an immense variety of public assets, including fixed assets like land and buildings, to natural capital, to intangible assets such as the licensing of businesses. These assets can be put to work to achieve tremendous collective benefits to citizens. Managing these processes by defining legal frameworks, conferring rights, regulating and arbitrating, are key tasks for the state to fulfill in this regard.

Effective Public Borrowing

Effective public borrowing enables investments in human, physical, institutional, and social capital. If made wisely, such investments can more than pay back the debts incurred. On the other hand, poor decisions can lead to debt crisis. The ability to steward borrowing effectively is therefore a key element of state effectiveness and stability.

1.2 Catalyzing State Functionality: Afghanistan 2002–2004

The global aid system was designed for a different era. It assumed a world of functioning states that lacked financial capital and needed to build infrastructure. As states themselves weakened, the aid system gradually began to assume a variety of functions on behalf of states. It evolved a series of specialized agencies that compete for resources and that exist in organizational stovepipes, with the result that they often operate in splendid isolation from one another. The main instruments of the aid system, the developmental project and use of technical assistance, are widely recognized to undermine state functionality by creating fragmentation, waste, and parallel structures, and by drawing government staff away from state posts to act as support staff for the international community.

In 2001, following the fall of the Taliban, these traditional aid approaches were mobilized by aid agencies to be applied in Afghanistan. However, cognizant of the problems that such a model would cause for Afghanistan, a small team of Afghan and international staff, in partnership with global leaders, believing conventional modes of assistance to be unsuited to Afghanistan's needs, designed and began to apply an alternative model intended to overcome and avoid the problems of the global aid system.

Emerging from a devastating series of conflicts lasting more than 20 years, Afghanistan faced the daunting task of reconstituting and rebuilding the entire social, political and economic fabric of the state. Afghanistan had consistently ranked among the least developed countries in the world and was struggling with an estimated $240 billion in destroyed infrastructure and foregone economic opportunities since 1979. The task seemed overwhelming, and for key decision-makers on both the government and donor sides, the choice of what to do and where to begin was not an easy one. In late 2001, an unrepresentative Interim Administration was established through a UN-mediated process in Bonn to begin the task of establishing a legitimate political center and directing national development efforts.

In the same period, a range of development and security actors arrived in the country, each with differing aims and objectives. The International Security Assistance Force (ISAF), established in Kabul, was primarily tasked with keeping order and preventing political unrest in the capital city. Coalition forces were present across the country to consolidate defeat of the Taliban and fight the nascent terrorist insurgency. The World Bank and UN agencies led a needs assessment to help the government establish the extent of reconstruction needs and to prioritize responses. A large number of donor agencies, UN bodies, and nongovernment organizations (NGOs) established themselves throughout the country to tackle problems ranging from infrastructure reconstruction to humanitarian assistance. At the end of 2001, without either an acknowledged leader or a process for the post-war reconstruction, poverty-reduction, and humanitarian efforts, development efforts were confused and often contradictory. Indeed, as early as January 2002, many reconstruction teams arrived from around the world to undertake identical projects.

As the political process began to establish a legitimate political center, attention turned to the problem of concerting national and international activities. From the outset, a team argued the case for putting a state-building approach at the center of the agenda and designed a distinctive series of "national programs" as vehicles for concerting the plethora of national and international actors behind a common strategic vision geared to building state functionality.

1.3 National Programs in Afghanistan

The first step was to forge a consensus between Afghan citizens, the Afghan government, and their international partners, that the overarching goal in Afghanistan was to create a legitimate and effective state capable of delivering services to the people and fulfilling its international obligations. This consensus was built around a vision, articulated by the government, and refined in a National Development Framework (NDF). This framework included three pillars: (1) Human Capital and Social Protection, (2) Physical Infrastructure, and (3) Trade and Investment, Public Administration and Rule of Law/Security, and articulated a series of cross-cutting issues, such as gender equity, security and rule of law, and administrative and financial reform. These three pillars were then addressed through a series of twelve National Programs. The NDF underpinned the National Development Budget, which costed the programs. The vision and strategy laid out during the course of 2002 was subsequently framed within a comprehensive document, Securing Afghanistan's Future, presented to an assembled group of 62 finance and foreign ministers from around the world on 31st March and April 1st 2004 in Berlin.

National Programs were designed to overcome the fragmentation and confusion of the project model. They serve as implementation mechanisms to enable a government to perform state functions throughout its territory in an effective and transparent manner. This is accomplished by mobilizing actors (including government, private sector and/or civil society) to perform critical tasks, translating the vision and mission for each state function into credible outcomes. By providing a unified set of rules, clear decision rights and obligations were established, generating accountability. The design of the programs was intended to harness Afghan assets and to recast international assistance as catalytic to the state-building process. The end result of these programs was an integrated architecture for facilitating good governance and development at various levels of government.

National Programs seek to avoid recognized problems with traditional approaches to development assistance by harnessing national policy oversight to national capacity and combining it with extensive outreach and consultation to ensure transparent, effective, accountable aid delivery. The Afghan economic team, in partnership with a group of creative international actors, focused relentlessly on implementation of these programs during 2001–2004. National Programs in Afghanistan included those discussed below.

National Solidarity Program (NSP)

The first program to be designed by the Government, in partnership with the World Bank and UN Habitat, the NSP was designed to empower communities to manage the reconstruction process.[4] The government provided block grants of approximately $20,000 to every village in the country against the requirement that each village elect a leadership council by secret ballot, hold participatory meetings to design recovery plans and projects, and post account documents in a public place. While the national government set the rules and managed financing, NGOs were contracted to manage the personnel and organizational issues, and an international firm was hired to provide management and oversight services. Thus the government provided a legal framework to empower a range of actors, including communities, while harnessing the considerable positive assets of international agencies, firms, and NGOs in a concerted, coherent fashion. Four years on, the program has seen more than 12,000 village development councils elected and more than 19,000 project plans approved.

The program has addressed the process of democratization from the ground level up, in parallel with national-level constitution-making and rule-writing at the centre, culminating in the first direct presidential election in the history of the country. The vicious cycle between government and Afghan citizens, through which successive regimes had preyed on villages and used factionalism as a means through which to entrench authoritarianism, was finally broken. The secret ballots empowered both men and women to directly elect village councils, with accountability ensured through transparent decision-making processes and close supervision of block grants. At the village level, democratic decision-making processes have taken hold, challenging the assumptions, practices and legitimacy of rule by kleptocratic oligarchy.

National Transportation Program

The National Transportation Program created an integrated transportation system to facilitate internal markets and to connect economic hubs across Afghanistan to regional and global markets. It was also designed as a mechanism to create internal linkages for access to education and healthcare, especially in remote areas, and to ensure cross-cultural linkages. The program facilitated completion of the Afghan ring-road and its connecting spurs to the borders of the country that has helped to transform Afghanistan into a central Asian land bridge.[5]

National Communications Program

At the heart of the National Communications Program was the formulation of a strategy for the transparent licensing of telecoms, which resulted in the provision of four licenses in two separate tranches in 2006. Through the licensing process, the number of subscriptions to mobile phones rose from about 100 to 1.5 million in 2006; telecommunication companies provided a significant boost to domestic revenues and over $550 million in private sector investments were channeled into Afghanistan. This indicates that the management of licensing and procurement processes in a post-conflict country can be managed in a productive and transparent fashion.

National Health and Education Programs

A National Health Program focused on preventive medicine through restructuring the Ministry of Health, essentially transforming it into a regulatory body through which NGOs were contracted to provide services. As a result, every child in the country was immunized. Infant (under five) and maternal mortality rates have declined markedly: a key goal of the program. The corollary in the education sector, the National Education Program, focused on school enrollment for both girls and boys, and has resulted in a massive increase in the number of children who regularly attend school.

The National Accountability and Transparency Program

The National Accountability and Transparency Program sought the complete overhaul of the public finance system. It began with an immediate and rapid currency conversion process, followed by reform of the budget process, the Treasury, revenue collection mechanisms, expenditure monitoring and banking system oversight. These changes were backed by improvements in accounting and auditing systems and were driven by timely and detailed reports on revenue and expenditure presented to the Cabinet and the gathering of national delegates at the Constitutional Loya Jirgas, and they were also elucidated regularly to the public through the media and through regular consultations with the public.

Neither the Afghan government and people nor the international community alone could have managed the design and implementation process for the National Programs, and the Afghan

experience broke significant new ground in terms of the collaborative efforts made by all stakeholders to improve aid effectiveness. The programs built stable alliances and coalitions between international program managers, Afghan officials, NGOs, private sector companies and the Afghan people. The partnership between the Afghan government and international partners created a common agenda. The National Programs provided a vehicle through which disparate assets and capabilities from across the spectrum of actors could be put to work in a concerted fashion.

The government marshaled its rule-making powers to define spheres of operations in which national and international actors could constructively participate. In order to ensure that international donors could contribute funds directly to the government, the World Bank agreed to run a Multi-Donor Trust Fund that only released funds to the Government's budget after satisfactory audit reports and required reform steps—designed by the Ministry of Finance—were completed. Funding for the Trust Fund from donors was premised on the Afghan Government's adherence to an agenda of building what ISE calls "National Accountability Systems" that create and reinforce institutional checks and balances for collecting, allocating and spending resources.

The National Programs approach faced some challenges. The UN-led needs assessment process created rival claims for resources and threatened to reproduce familiar problems, such as fragmentation of efforts, and undermining of state capacity. The National Programs also had Afghan challengers since they removed sources of revenue and decision-rights from political elites across different levels of Afghan political life. After 2004, the Afghan government was increasingly affected by corruption, and new Ministers sometimes sought to make an impact by constructing their own projects and programs. For these reasons, the Government itself began to dismantle parts of the National Program architecture, although some programs, such as NSP, had won such broad grass roots (and increasing international) support that they proved very difficult to touch. Recently there has been widespread affirmation of the National Program approach by the Afghan Government and its key partners, including the U.S. government, the UN, the World Bank, and key donors.

Emerging Lessons

The task of partnering with the Afghan people and government to catalyze the emergence of a state capable of performing core functions is daunting. Nevertheless, the National Programs approach has met with some significant successes, despite political opposition both domestically and at times internationally, as well as the difficulties posed by severe government corruption, narcotics, and the insurgency. If a state-building approach is to be pursued, it will need to grow both in light of experience and in order to remain relevant in the evolving context. The state-building agenda in Afghanistan could be enhanced through attention to improving the following six particular areas:

Market Building

In Afghanistan, as in every other post-conflict context that ISE has documented, criminalization of the economy has been a critical issue that government and international actors have been unable to prevent or control. In Afghanistan, criminalization of the economy and the infiltration of narco-mafia interests into fragile state institutions are having severe consequences across the economy, polity, and civil society. This criminalization drives corruption at all levels of government and society and undermines the rule of law by normalizing and rewarding illegality

and making large numbers of citizens active or complicit in crime. Criminality also finances and equips terrorists and insurgents.

An approach to criminality in Afghanistan must increase the number and strength of stakeholders in stability by growing the licit economy. Afghanistan is blessed with impressive mineral wealth and attention to harnessing that wealth to the benefit of Afghanistan. The international community could provide expertise on licensing, and risk guarantees to kick-start economic activity could do more than aid dollars.

Civil Society

Corruption has seriously penetrated the Afghan government, whereas most ordinary Afghan people desire decent, dependable institutions. Their disillusionment fuels insurgency and insecurity. NSP demonstrates that ordinary Afghans can be mobilized behind an agenda of solidarity and national regeneration. Mechanisms are required to create the space for civil-Afghan groups working towards peaceful change. All too often, development assistance equates domestic civil society with foreign or domestic NGOs, whereas through the lens of citizens, NGOs are often perceived as parasitic organizations. Real civil society may or may not include some NGOs as a subset, but it is a much broader category. Mechanisms become critical to channel insurgents and their empathizers away from the organizational form of the Taliban through other channels, including a national agenda of reform, if necessary, as an opposition movement.

Implementation and Accountability

Some of the failures of the effort over the last eight years in Afghanistan relate to failure of design; others relate to failure of implementation. Where good designs are drawn up, there is very little follow-through in terms of implementation, as plans, leadership, and resources are not aligned to each other. Part of this stems from a fragmentation in planning between key organizations: the ISAF, the World Bank, the UN, the U.S. and its key allies, and the Afghan Government. Part stems from the fact that, even when there is agreement, decisions are not resourced, nor are organizational structures in place. What is urgently required, whether for Afghanistan or other areas, is a change in the culture of execution and implementation. This must strike a careful balance between understanding when an external agent is needed to manage or produce something and when domestic actors and processes are needed to drive and implement an agenda. It is oft overlooked and very misunderstood that the internal budget and planning processes of a country are the key drivers of effectiveness. These internal accountability systems must be in place for implementation of revenue and expenditure rules.

Collaboration and Coordination Platforms

NSP demonstrates that where programs are successfully designed, a spectrum of disparate national and international participants can contribute diverse assets to achieve common goals. NSP resulted from collaboration between villages as social organizations and the budget of the country, with a trust fund and technical support from NGOs. It will be necessary to move from an industrial model of top-down implementation to one which mobilizes across network and hierarchy through the appropriate platforms. This approach will incorporate tribal, informal, and Islamic contexts and can harness collective energies, synchronizing top down and bottom up.

1.4 Conclusion

The international community increasingly perceives the necessity of building effective states, and in Afghanistan, the international community's ability to achieve this goal using conventional methods of assistance has been severely tested.

Despite being challenged by a range of vested interests, the National Programs in Afghanistan offered a distinctive approach to building state functions by harnessing a range of actors within an integrated architecture. In a context where corruption impedes the creation of an effective state, National Programs can still contribute the building blocks of a functional state, and while rallying those actors and groups who desire such a state, they can also create new groups of stakeholders willing to struggle towards that goal. Tipping the balance so that the stakeholders in a common, prosperous, and stable order come to outweigh interests of criminality, corruption, and state dysfunction requires concrete mechanisms capable of binding actors together in ways that create virtuous circles. In doing this, the National Programs approach offers a distinctive alternative to outdated international approaches.

The National Programs approach also suggests that the international community possesses a stock of more or less latent assets that, if harnessed, could contribute very productively. Expertise in designing licensing was critical to success in the telecom sector. Aid in the form of risk guarantees could stimulate investment and entrepreneurship.

In Afghanistan, despite receiving limited support and attention, as well as active subversion by a range of national and international interests, National Programs have achieved significant concrete successes in building state functions and creating stakeholders in a stable future. Going forward, such programs could be refined and supplemented within Afghanistan. Attention should be given to building the legal market and to strengthening civil society groups seeking to tackle corruption and apply pressure on the state. Such groups would offer a reformist alternative to insurgency as a means of expressing disillusionment.

Four additional themes emerge from recent experience of National Programs, and these deserve consideration in future program design. First, program design must take implementation as the starting point. Second, accountability systems must be written into the DNA of each program. Third, programs should serve as platforms or "operating systems" that allow coordination and collaboration between different stakeholders. Fourth, attention should be given to how bottom-up and top-down approaches are to be synchronized across the village, district, province and national government structures.

More broadly, as a concrete mechanism that overcomes limitations upon existing aid approaches, focuses on building states, and breaks state-building into manageable and realistic tasks, National Programs deserve broader study for their potential to contribute to the stability and resilience in other contexts across the world.

↑ back to the top

---

[1] Ghani, A, Clare Lockhart and Michael Carnahan, "An Agenda for State-building in the Twenty-first Century. *The Fletcher Forum of World Affairs*, Vol. 30:1 Winter 2006, pp.101-123.

[2] The case of Haiti is particularly stark. See for example Buss, Terry F. and Adam Gardner, "Why Foreign Aid to Haiti Failed," National Academy of Public Administration, 2006.

[3] ISE continues to invite discussion and debate about the proposed ten functions of the state, which were previously identified in Ghani, A., Clare Lockhart and Michael Carnahan, "Closing the Sovereignty Gap," ODI Working Paper 253, 2005 and articulated in Ghani, A. and Clare Lockhart, "Fixing Failed States: A Framework for Rebuilding a Fractured World," OUP 2008.

[4] http://www.nspafghanistan.org/content/index_eng.html

[5] Upon completion of this program, the capitals of Afghanistan's Central Asian neighbors are no more than 32 hours' drive from the Gulf.

2. Independent Trend and Trajectory of the Future: Issues Impacting Social Regional Stability

2.1. Demographics

2.1.A. Demographic and Economic Trends: Implications for International Migration (Philip Martin)

Author and Organization: Dr. Philip Martin, Professor, University of California Davis, UC Comparative Immigration & Integration Program

Contact Information: martin@primal.ucdavis.edu

Abstract

About three percent of the world's 6.1 billion people were international migrants in 2000. Population growth is expected to slow between 2000 and 2050 in comparison to 1950–2000, but international migration is expected to rise as persisting demographic and economic inequalities that motivate migration interact with revolutions in communications and transportation that enable people to cross borders. The default policy option to manage what is sometimes deemed out-of-control migration, adjusting the rights of migrants, is unsatisfactory, prompting this review of longer term factors affecting migration patterns, including aging in industrial countries, rural-urban migration that spills over national borders, and the migration infrastructure of agents and networks that moves people. The paper concludes with an assessment of the likely effects of the 2008–09 recession on international migration.

Introduction

This paper outlines the major demographic and economic trends likely to affect international migration flows through 2050, with special emphasis on the effects of three factors—population aging, rural-urban movements, and the inequalities that encourage workers to move from lower to higher wage countries. The paper also reviews the "migration infrastructure," the institutions and actors ranging from recruiters, transportation firms, and those who benefit economically from employing and integrating migrants. Finally, the paper reviews the likely effects of the current recession on migration flows.

The half century between 1950 and 2000 was a period of extraordinarily fast population growth and urbanization—the global population rose by 1.5 times, the populations of many countries doubled, tripled, and quadrupled, and the urban share of world's population rose from 29 to 48 percent.[1] Population growth is expected to slow, with the world's fertility falling below the replacement level of 2.1 in 2025 as urbanization continues. Unlike the past 50 years, when almost all countries had baby booms and falling mortality, the hallmark of the next half century is expected to be unevenness in demographic change. Almost all of the world's approximately 200 countries are projected to have aging populations, while a quarter of the countries are expected to have shrinking populations, and another quarter of the countries are expected to continue growing at rates that will double their populations within 50 years.

The world's labor force is drawn from people 15 years of age and older. The major trends of the past 50 years have been the increasing share of married women with children who work for wages, later labor force entry due to less child labor and more education, and earlier retirement associated with better pensions. Over the next 50 years, these trends are expected to persist,

albeit with variations, including fewer persons self employed in agriculture and more persons employed for wages in cities. A major unknown is the age of retirement—will most workers continue to retire in their early 60s and rely on public and private resources for another 20 to 25 years, or will a rising share of workers continue working, keeping the total period of retirement closer to the current average of 15 to 20 years. In the U.S., the average age at retirement dropped from 70 in 1940 to 65 in 1970 and 62 in 2000, making the average number of years in retirement 10 years in 1940, 13 in 1970, and 18 in 2000.[2]

Differences in demographic trends between countries are likely to combine with economic inequalities to provide incentives to migrate within and between countries. Rural-urban migration within countries is poised to continue, but patterns of international migration are harder to predict, largely because regulating the entry and stay of non-citizens is a core attribute of national sovereignty jealously that is closely guarded by most governments. Public opinion in most countries opposes large-scale immigration, and the result of sparring between employers seeking additional workers, migrants seeking opportunity, and public opposition to migrants is a mélange of guest workers, irregular migration, and other forms of cross-border movement.

There are two extreme scenarios for international migration over the next several decades. One imagines the emergence of bilateral and regional free-movement regimes, such as the one between Australia and New Zealand or within the European Union, and eventually a global organization such as the World Trade Organization or a new World Migration Organization with the goal of fostering more movements to increase global economic efficiency. A WTO or WMO could also give national governments facing anti-migrant publics "cover" to liberalize labor flows over national borders, as when they say they do not want to allow the entry of migrants, but their international obligations require them to do so. The other scenario imagines that current levels of migration, which involve about three percent of the world's residents, will persist, and that national governments will remain restrictive toward migrants for reasons that range from fears of crime and terrorism to the adverse effects of migrants on native workers and productivity.

Most migration literature predicts and endorses regional and global liberalization of labor migration.[3,4,5]

The most likely scenario lies between these extremes, meaning there is likely to be a variety of migration systems reflecting the circumstances of the countries involved. Rentier societies such as Gulf oil exporters seem poised to continue to rely on migrant workers to fill most private sector jobs; most of these migrants are likely to rotate in and out of the country on two- and three-year contracts. More industrial countries are likely to join the four traditional immigration countries—Australia, Canada, New Zealand, and the US—to select a certain number of immigrants each year and to plan for their integration.

The great unknown factor is policy development in middle-income developing countries receiving migrants, ranging from Costa Rica to Argentina, Libya to South Africa, and Korea to Thailand. Will they follow the lead of industrial countries and plan for immigrants based on family, economic, or other criteria, including the return of descendants of previous emigrants? Alternatively, they could follow the model of Singapore, welcoming professionals to settle with their families while aiming to rotate less skilled workers in and out of the labor force and country.

National policy development is likely to be complicated because of links between rural-urban and international migration. Migration is generally a journey of hope for individuals wishing to improve their economic status and enjoy new opportunities. Rapid economic and job growth in urban areas of countries in the midst of rural-urban migration can provide opportunity for most migrants within borders, as in China and Turkey. Without opportunities within borders for those seeking opportunity, rural-urban migration can become international migration, as from Mexico to the U.S.

Many migrants face difficult trade-offs between economic and other freedoms—should they cross national borders for higher wages if they lose their right to engage in political activity to advocate for change in the circumstances in which they live and work? The front doors for legal immigrants open widest for those with the most education and skills, raising questions about whether remittances are sufficient compensation for a new brain drain. Finally, many developing countries that are richer than their neighbors are likely to learn the lessons of the 1960s, viz, that there is nothing more permanent than temporary workers, and that recruiting or tolerating migrant workers inevitably adds permanent residents to the population.

The world in 2008–09 is in the midst of the most severe recession in half a century, marked by shrinking economies and rising unemployment. Just before the recession spread around the world in 2008, the number of migrants and remittances were at record levels, and many analysts expected this growth to continue. Flows of labor from poorer to richer countries have slowed and in some cases reversed due to the recession, and it is not clear that migration and remittances will resume their upward trajectory during the economic recovery.

The current recession could affect migration patterns quite differently than past recessions for four major reasons.

First, this global recession began in industrial countries and spread to almost all countries, limiting the ability of migrants to shift from lagging to boom areas, as they could when high oil prices attracted migrants to the Gulf countries while doors closed to Western Europe.

Second, the effects of recession are most severe in cyclically and trade-sensitive sectors such as construction to manufacturing that hire large numbers of migrant workers and may not recover quickly. Some of the "migrant jobs" in manufacturing may not return in the recovery.

Third, most migrant-sending governments and development institutions have urged migrant-receiving governments to minimize disruptions to labor migration by stopping recruitment or forcing migrants to leave. (ILO DG Juan Somavia's December 18, 2008 International Migrants Day message asked migrant-receiving countries to "assess their labor market needs before resorting to general layoffs of migrant workers."[6] Employers rather than governments usually make hiring and layoff decisions; the message presumably argues against policies such as that adopted by the Malaysian government, which ordered migrants to be laid off before Malaysian workers.) Before the recession, it was widely assumed that migration and remittances would continue to increase. Current projections anticipate that remittances to developing countries will decline, but far less than foreign direct investment, which is expected to fall sharply. Some migrant-sending governments have appealed to migrant-receivers to continue accepting migrants to avoid political unrest among citizens unable to find foreign jobs.

The fourth issue is immigration, with the issuance of about 1.5 million visas a year to people allowed to settle and eventually naturalize in the traditional immigration countries of Australia,

Canada, New Zealand, and the U.S. (the U.S. accepts about three-fourths of these immigrants). Australia, Canada, and New Zealand select most of their immigrants on the basis of a point system that favors the admission of those most likely to be successful finding jobs, while the U.S. admits mostly foreigners who are joining settled family members. Family-based immigration to the U.S. is expected to be less sensitive to recession than economically motivated migration to countries that select newcomers based primarily on economic criteria.

Population Trends

The world's population, 6.8 billion in 2009, included 1.2 billion residents in "more developed" regions, and 5.4 billion or 82 percent in "less developed" regions. The world's population is projected to rise to 9.2 billion by 2050 in the median variant. Keilman[7] finds that UN projections of population are most reliable for 10-15 years into the future.[8] With almost all population growth projected to occur in what are now considered less developed regions, the share of global population in more developed countries is projected to fall from 18 percent today to 13 percent by 2050.

The world's population was about a billion in 1800 and rose to 2.5 billion in 1950. High birth rates and falling mortality rates explain the rise in the global population, and migration explains the sharp increases in the share of the global population in Latin America, North America, and Oceania.

Between 1950 and 2000, when the world's population increased from 2.5 billion to 6.1 billion, the major shifts in population weights by continent were the result of changes in fertility and mortality rather than large-scale migration. The most noticeable shift was the decline in the share of the world's people in Europe, down from 22 to 12 percent, and the rising share of the world's people in Africa, up from nine to 13 percent. This trend of an expanding Africa and shrinking Europe is expected to continue, so that by 2050, there is projected to be three Africans for every European.

Table 2.1.A-1. World Population by Continent, 1800, 1950, 2000, 2050 (Percent Shares)[9,10]

	1800	1950	2000	2050 (projected)
World (× million)	978	2,535	6,124	9,191
Africa (percent)	11	9	13	20
Asia	65	56	61	59
Europe	21	22	12	7
Latin America & Caribbean	3	7	9	9
Northern America	1	7	5	4
Oceania		1	1	1

The global population has been increasing by about 78 million a year, adding the equivalent of a Germany to the world population each year. The UN projects declining fertility, from the average 2.6 children per woman in 2005 to 2.05 by 2050. Fertility in the more developed regions in 2005 was 1.6, versus 4.6 in less developed regions. Fertility is expected to rise to 1.8 by 2050 in more developed regions. If fertility does not decline to below-replacement levels as projected, the global population would reach 10.8 billion by 2050.[11] A combination of below-replacement fertility and low immigration is expected to result in the shrinking of populations in a quarter of

the world's countries, including many European countries, Russia and Eastern European countries, and Japan and Korea.[12]

The UN groups people into three age groups—under 15, 15 to 59 and 60 and older. The UN projects population by age for 21 five-year age groups (e.g., 0–4, 5–9, 10–14), ending with 100+. The major changes expected in the world's population over the next four decades include a rising share of people 60 and older, and growing differences in age structures between continents. Globally, the major demographic trend is aging—10 percent of people were 60 or older in 2005, and their share of the global population is projected to more than double to 22 percent by 2050. Aging will be most pronounced in more developed countries, where the share of residents 60 or older is projected to rise from 20 percent in 2005 to 33 percent by 2050, so that there will be two persons 60 and older for each child 15 and younger.[13] In less-developed countries, by contrast, there will be far more elderly, but still about one child for each person 60 and older.

Table 2.1.A-2. World Population by Age Group, 2005 and 2050 (Percent Dist)[14]

	2005	**2005**	**2005**	**2050**	**2050**	**2050**
	0-14	15-59	60+	0-14	15-59	60+
World	28	61	10	20	58	22
More Developed	17	63	20	15	52	33
Less Developed	31	61	8	21	59	20
Africa	41	53	5	28	62	10
Asia	28	63	9	18	58	24
Europe	16	64	21	15	51	34
Latin America & Caribbean	30	61	9	18	58	24
Northern America	20	63	17	17	56	27
Oceania	25	61	14	18	57	25

Age differences between continents are most pronounced for the old and young. In 2005, there were more children under 15 than residents over 60 in every continent, including 8 children for every elderly adult in Africa. By 2050, there will be more people over 60 than under 15 in every continent except Africa, including two elderly adults for each child in Europe.

In four major labor-sending countries over the past half century, there was rapid population growth, led by the Philippines. The populations of these four countries are expected to continue increasing at a slower pace, but the population of the Philippines is projected to be greater than that of Mexico by 2050 because of higher fertility. If these projections prove accurate, one would expect emigration pressures due to demography to decline faster in Mexico than the Philippines.

Table 2.1.A-3. Population: Mexico, Morocco, Philippines, Turkey, 1950-2050 (in millions).[15]

	1950	***2000***	***2050***	***1950-00***	***2000-50***
Mexico	28	100	132	257%	32%
Morocco	9	29	43	222%	48%
Philippines	20	76	140	280%	84%

Turkey	21	68	98	224%	44%

Labor Force Trends

The International Labor Organization (ILO) labor force projections add labor force participation rates (LFPR) to the UN's population projections, meaning that the ILO projects an increase in the share of persons who are in the economically active population (EAP), defined as those who are employed or looking for work. Many countries do not report LFPRs, and the ILO estimates LFPRs for them. The ILO had LFPR data for 23 of the 191 countries for which projections were made; non-reporters are generally countries that are small and poor.[16]

The ILO estimates LFPRs for 11 age groups, from 15–19, 20–24, etc., to over 65, and for men and women. It finds that LFPRs behave as follows:

- Rise from younger to older age groups.
- Reach their maximum values for men and women between the ages of 25 and 55, when 90 percent or more of men are in the labor force (women's LFPRs are slightly lower than men's because of motherhood).
- Begin falling for workers after age 55.

The result is a hat-shaped distribution of LFPRs, rising steeply as young people enter the labor force, reaching a maximum during the primary working years, and falling with retirement.

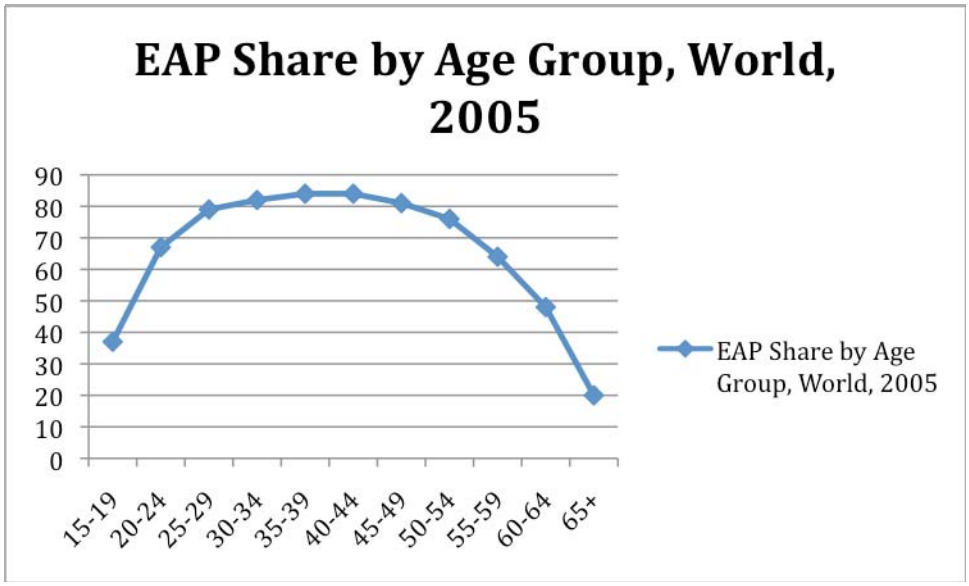


Figure 2.1.A-1. EAP Share by Age Group.[17]

The ILO's challenge is to apply appropriate LFPRs by age and sex to the correct base data—the number of persons in each age and sex group. It does this by country, as explained in volume 10 of the methodology. For example, in Mexico, male LFPRs averaged 90 percent for the 25–50 age groups, while female LFPRs averaged 40 percent for the 25–50 age groups. In 1991, about 34 percent of economically active men were in agriculture, 25 percent in industry, and 41 percent in services; the distribution of economically active Mexican women was 11 percent in agriculture, 19 percent in industry, and 70 percent in services.[18]

The ILO's most recent labor force projections cover the 1980-2020 period. The major trend at the global level is the gradual rise in the economically active share of the total population, from 43 percent in 1980 to 47 percent 2005 and 48 percent in 2020.

Table 2.1.A-4. EAP (millions).[19,20,21]

	1980	***1985***	***1990***	***1995***	***2000***		
World Pop	4,451	4,855	5,295	5,719	6,124		
World EAP	1,930	2,160	2,406	2,605	2,818		
Share	43%	44%	45%	46%	46%		
	2005	***2010***	***2015***	***2020***	***2025***		
World Pop	6,514	6,910	7,295	7,667	8,010		
World EAP	3,050	3,279	3,481	3,651	3,845		
Share	47%	47%	48%	48%	48%		
	2030	***2035***	***2040***	***2045***	***2050***		
World Pop	8,317	8,587	8,824	9,025	9,191		
World EAP	3,992	4,122	4,236	4,332	4,412		
Share	48%	48%	48%	48%	48%		
EAP Change	***1980-90***	***1990-00***	***2000-10***	***2010-20***	***2020-30***	***2030-40***	***2040-50***
World EAP	***25%***	***21%***	***17%***	***17%***	***9%***	***6%***	***4%***

If the economically active share of the global population remains at 48 percent between 2020 and 2050, the world's labor force will increase from about 3 billion in 2005 to 4.4 billion in 2050. However, the rate of labor force growth will slow. There was remarkably fast labor force growth in the 1980s and 1990s, when the global labor force expanded by over 20 percent as a result of the baby boom of the 1960s and 1970s, to a projected expansion of four percent in the decade of 2040–50.

In 2005, about 20 percent of the world's workers were in the more developed countries, and 80 percent were in the less developed countries. The labor force of the more developed countries is expected to remain at about 600 million if the overall LFPR remains at 48 percent, largely because of the stable populations of industrial countries. The ILO projects that the economically active share of populations in less developed countries will converge to the 48 percent of more developed countries, so the labor force of less developed countries will increase strictly with population after 2020. Thus, the labor force of more developed countries is projected to stop growing after 2020 because of the stable population, while the labor force of less developed countries continues rising at a slower rate and is still growing toward four billion in 2050.

Conclusions

The world's population is aging at a time when almost all population and labor force growth is concentrated in the 170 developing countries, where over half of residents live in rural areas.

Migration is a response to differences between areas, and millions of younger rural residents of developing nations are moving in search of higher wages and more opportunities. It is easier and cheaper for additional migrants to join those who have settled in urban areas within nations and abroad.

Inequalities that promote migration, coupled with communication and transportation revolutions that make it easier and cheaper to move, have prompted policy makers to turn to the instrument over which they have most control—rights—in an effort to manage migration. In western Europe during the 1990s, access to the asylum system was restricted by defining some countries as safe and requiring those seeking asylum to apply for refuge in the first "safe" country they reached, a policy change that reduced applications. In the US in the 1990s, the access of legal and unauthorized foreigners to means-tested federal assistance was reduced in an effort to head off a backlash against foreigners arriving "with their hands out," rather than seeking a "hand up" the ladder of economic opportunity.

Managing migration by adjusting the rights of migrants may become increasingly difficult in a world of persisting demographic and economic inequalities. For the first time in human history, sharp differences in demography are expected to exist across countries, with at least a quarter of the world's countries expected to have aging and shrinking populations. The question is whether these generally richer industrial countries will open border gates to migrants, whether they will raise the retirement ages, or whether they will take other steps in an effort to minimize migration.

Two extremes mark the spectrum of likely responses to slower population growth and aging. One imagines a world of bilateral, regional, and global migration agreements that make it easier for workers to cross national borders legally so that in a few decades, migrants will be accepted in workplaces with the same indifference as most consumers have about where their TVs, cars and other goods were made. The other extreme foresees relatively low levels of labor migration as aging societies that may "need" migrants resist the changes that accompany migration.

Reality is likely to lie somewhere between these extremes, making it difficult to assess the implications for human development. Migration is likely to remain a journey of hope for millions of people around the world, and most of those who move are likely to find the opportunities they are seeking. Receiving areas can benefit economically and culturally from the contributions of migrants, and migrant-areas-of-origin can benefit from remittances that reduce poverty and provide a new impetus for development. However, there are likely to be yawning gaps between the rights and protections laid out in international conventions and the realities faced by many migrants.

↑ Back to the top

---

[1] *World Urbanization Prospects,* 2007 revision, www.un.org/esa/population/publications/wup2007/2007wup.htm.
[2] Aaron, 1999.
[3] Ghosh, 2000.
[4] Bhagwati, 2005.
[5] Helton, 2003.
[6] www.ilo.org/public/english/bureau/dgo/speeches/somavia/2008/migrants.pdf
[7] Keilman, 1998, 15.
[8] UN, 2007, pxxi.
[9] UN, 1999. *The World at Six Billion*, Table 2, UN. 2007.

[10] *World Population Prospects*, 2006 revision.
[11] UN, 2007, pxxi.
[12] UN, 2007, pxxiii.
[13] UN, 2007, pxxiii.
[14] UN. 2007. *World Population Prospects*. 2006 revision (pxxiii).
[15] UN, 2007, Vol 1. *Medium variant*.
[16] ILO (2008, 11).
[17] World, 2005.
[18] http://laborsta.ilo.org/applv8/data/SSMe.html, p 120.
[19] *World*, 1980-2020.
[20] http://laborsta.ilo.org (48 % EAP after 2020).
[21] Population data from UN, 2007, medium variant.

2.1.B. Demography and Security (Jack Goldstone)

Author and Organization: Dr. Jack A. Goldstone, Professor, George Mason University
Contact Information: jgoldsto@gmu.edu

Five major demographic trends are likely to pose significant security challenges to the United States, Europe, Japan, and most other developed nations in the next two decades:

1. Disproportionate population growth in large and Muslim countries
2. Shrinking populations in Europe and the west Pacific Rim countries (Japan, S. Korea, Taiwan)
3. Sharply opposing age shifts between aging developed countries and youthful developing countries
4. Increased immigration from developing to developed countries
5. A shift in developing countries toward rapid and dominant urbanization

The security and conflict problems caused by population changes in the coming decades are not mainly due to overall population growth creating shortages of resources (although, with climate change looming, we cannot discount that possibility in some vulnerable nations). Rather, population distortions—in which populations grow too young, or too fast, or too urbanized—will make it difficult for prevailing economic and administrative institutions to maintain stable socialization and labor force absorption, leading to risks of mass migrations, government breakdowns, and international and civil conflicts and a reduced ability of the United States and its major allies to respond to these challenges.[1,2,3]

Big Emerging Markets and the World Economy

Countries are growing today for two major reasons: high population growth rates and demographic momentum. In some countries, mainly in Africa and the Middle East (as well as a few in Latin America and South Asia), birth rates remain much higher than mortality rates, so their population growth rates remain high, at more than 2.0 percent per year. In these countries—which include Afghanistan, the Democratic Republic of the Congo, Guatemala, Iraq, Jordan, Nepal, Saudi Arabia, Pakistan, and Yemen—the population is still doubling every generation (roughly every 30-35 years).[4]

In other countries such as China, India, and Indonesia, population growth rates have recently dropped substantially; in percentage terms, they are growing more slowly than they have in the past.[5] However, these countries already have such a large cohort of women of childbearing age that their populations continue to add significant numbers each year. In China, for example, although most couples have fewer than two children, zero population growth is still several decades away. While current growth rates have sunk to around 0.6 percent per year, China will add nearly 80 million people during each of the next two decades before its population peaks. India, though not quite as large as China today, is growing twice as fast, at 1.4 percent per year, and will add roughly 135 million people per decade for the next two decades. Even with a continued decline in their birth rates, these two countries alone are expected to add roughly 400 million people by 2025—more than today's entire populations of the United States, the United Kingdom, the Netherlands, and Belgium *combined*.

Table 2.1.B-1. The World's Largest Countries (Population in Thousands)
(Countries in **Bold** have large Muslim Populations)

2009		2025	
China	1,345,751	**China**	1,453,140
India	1,198,003	**India**	1,431,272
USA	314,659	USA	358,735
Indonesia	229,965	**Indonesia**	263,287
Brazil	193,734	**Pakistan**	246,286
Pakistan	180,808	Brazil	213,802
Bang'desh	162,221	**Nigeria**	210,057
Nigeria	154,729	**Bang'desh**	193,752
Russia	140,874	**Russia**	132,345
Japan	127,156	Mexico	123,366
Mexico	109,610	Japan	120,793
Philippines	91,983	**Ethiopia**	119,822
Vietnam	88,069	Philippines	117,270
Egypt	82,999	**Egypt**	104,970
Ethiopia	82,825	Vietnam	102,054
Germany	82,167	Congo, DR	98,123
Turkey	74,816	**Turkey**	87,364
Iran	74,196	**Iran**	87,134
Thailand	67,764	Germany	79,258
Congo, DR	66,020	Thailand	72,628
France	62,343	**Tanzania**	67,394
UK	61,565	UK	66,601
Italy	59,870	France	65,769
S. Africa	50,110	Italy	60,018
Myanmar	50,020	Myanmar	57,582

Most of the 20 largest countries in the world have modest growth rates but large demographic momentum and thus will make the largest contributions to total world population growth in the next 20 years. The fastest growing countries are generally smaller, but they are facing the largest burden of additional growth on a percentage basis (see Table 2.1.B-1). For the next several decades, global population growth will be concentrated in only a few regions and countries, mainly Muslim societies (almost the entire top half of Table 2.1.B-1) and huge states with

populations of 75 million or more. Most of the states that comprise Table 2.1.B-1 are also among the world's lower income countries. By contrast, population growth rates in Europe and Japan are already low, and in some cases they are negative. Therefore, the proportion of the world's population living in Muslim states, or in the very largest and very poorest states, will grow, and the proportion of the world's population living in developed countries will shrink. The major exception is the United States, which is expected to add 50 million people in the next 20 years—mostly due to recent and projected immigration of people born elsewhere.

Some countries with extremely rapid population growth are likely to manage it reasonably well due to sound management and strong economic growth (e.g., Kuwait and the United Arab Emirates). However, in a number of "flash points," the inability to integrate rapidly expanding populations into politics and the economy will lead to radical political mobilization among those angry at not attaining the level of prosperity reached by some of their neighbors. Similarly, some of the extremely large countries will probably manage their anticipated growth without conflicts. However, the sheer size of the population increases the large countries face in coming years, combined with their efforts to rapidly industrialize, means that many will also face a "tipping point," when uneven development leaves tens of millions of disadvantaged people to watch other millions reap the benefits of rapid growth. The disparities of economic fortune among classes, regions, or ethnic groups may become so great as to spark violent protests. Alternately, the migration of rural masses to urban and industrial centers could produce a social crisis. We cannot predict which countries will face such crises, as the crises are due to failed political leadership and administrative management more than population changes per se. But we can say that in many of the largest countries, governments will face exceptional challenges in meeting their populations' demands for both strong and equitable economic growth and sound political management.

We can say with certainty that these trends pose major dilemmas for the economic policy and development of the West, particularly Europe. In 2005, all of Europe comprised 731 million people, which is projected to shrink to just 664 million by 2050, while the rest of the world is projected to grow from 5.8 billion to 8.5 billion.[6] That is, in a single generation (the next 42 years), global population outside of Europe will increase by 2.7 billion, while Europe's population will decrease by about 67 million. The shrinking demographic weight of European countries puts them on the horns of a dilemma. If the economies of fast-growing developing countries do not catch up to those of the richer countries, then the standard of living enjoyed by the West will seem more elite and unfair than ever, fueling resentment of developing countries against the G-8. On the other hand, if economic growth in those countries does exceed that of the West so that living standards in poor countries or regions starts to approach those of rich countries or regions, then the combination of shrinking population and lagging economies will render the G-8 countries more and more irrelevant to the world economy. Greater resentment or greater irrelevance: certainly a difficult choice.

Europe's combined GDP in 2007 was US$14 trillion dollars (CIA, 2008).[7] Assuming GDP growth per capita of 2.5 percent per year and no net population growth, Europe's economy would increase by US$9 trillion (excluding inflation) by 2025. For Asia (excluding Japan), 2007 GDP was slightly larger, at US$18 trillion dollars (CIA, 2008). But due to modest growth in GDP per capita plus large population increases in most countries, total GDP is growing far more rapidly in this region. Iran and Pakistan achieved recent growth rates of 4 and 6 percent per year, respectively, while India and China were growing by 8-10 percent per year—and despite the

global economic downturn, both countries are expected to continue growing by 6-7 percent in 2009.[8,9] If Asia (excluding Japan) can sustain an overall growth rate of total GDP of 5 percent per year over the next 20 years, the increase in Asia's GDP would be US$30 trillion¸ or more than three times the total economic growth of Europe.

If Asian (excluding Japanese) GDP does not grow at 5 percent per year, living standards there will not catch up to those in Europe and Japan. Yet if Asian (excluding Japanese) GDP does grow at that pace, then given the size of the area, the preponderance of economic growth on the Eurasian continent will be occurring outside of Europe. Investment and innovation are likely to move increasingly to areas outside of Europe, further weakening Europe's economic strength and leadership. In other words, we are on the cusp of a global tipping point, in which East and South Asia come to eclipse Europe and Japan as major sources of global economic growth—a point made all the more sharply as Europe and Japan remained in recession in early 2009 while China is returning to rapid growth.

These demographic and economic changes also indicate that the military capacities of large developing countries will increase, while the ability of rich nations to put "boots on the ground" in conflict zones will diminish. Managing conflicts involving developing countries will become more difficult, and the effort will put more strain on developed countries' economies than before.

As the portion of the global economy contributed by the G-8 countries shrinks, countries such as China, India, Turkey, Brazil, Indonesia, and Mexico will become global economic powers. Admitting major regional powers into international governance bodies is vital if those organizations are to retain legitimacy. The November 2008 Summit on Financial Markets and the World Economy expanded the "G-group" to include these big emerging democratic economies —a trend that must continue if such efforts are truly going to grapple with the global economy.

Moreover, if the United States and Europe fail to support economic growth in the developing world, the rapidly increasing numbers of people in non-European and mainly Muslim countries is simply going to fuel ever-greater resentment of the West's position, exacerbating the problems of terrorism, smuggling, and illegal trafficking as the ways to "get ahead" and "get even." In short, the U.S. and Europe have no choice but to support and actively engage the fast-growing countries of the world, to improve relations with their populations, and to support and seek to share in their growth.

The Great Slowdown in Population Growth in High-Income Countries

During the next several decades, the population of most European countries, including Russia, Germany, Italy, Ukraine, Spain, Poland, Romania, the Czech Republic, and Hungary, will shrink substantially. This is due mainly to a sharp decline in the number of children per couple to well under 2.0 and in some cases under 1.5.[10] This slowdown will be accompanied by a rapid increase in the percentage of the population in higher age brackets, as the number of young children falls further behind the number of aging baby boomers. By 2050, the percentage of Japan's and Europe's population over age 60 is expected to double, ultimately comprising 35 to 40 percent of the total population.[11]

This pattern is highly novel and abnormal. Historically, population growth has stagnated on occasion or has been substantially reduced by major epidemics, with the cause being high mortality, especially among the young. Birth rates remained high, and when conditions were more propitious to growth, population increase resumed. In modern Europe, the United States,

Canada, and Japan, decreasing birth rates have precipitated population decline. Women are marrying later, if at all, and having fewer children. The result is an unprecedented aging of populations (less so in the United States), at the very same time that national economies can be expected to decline dramatically as a percentage of global GDP.

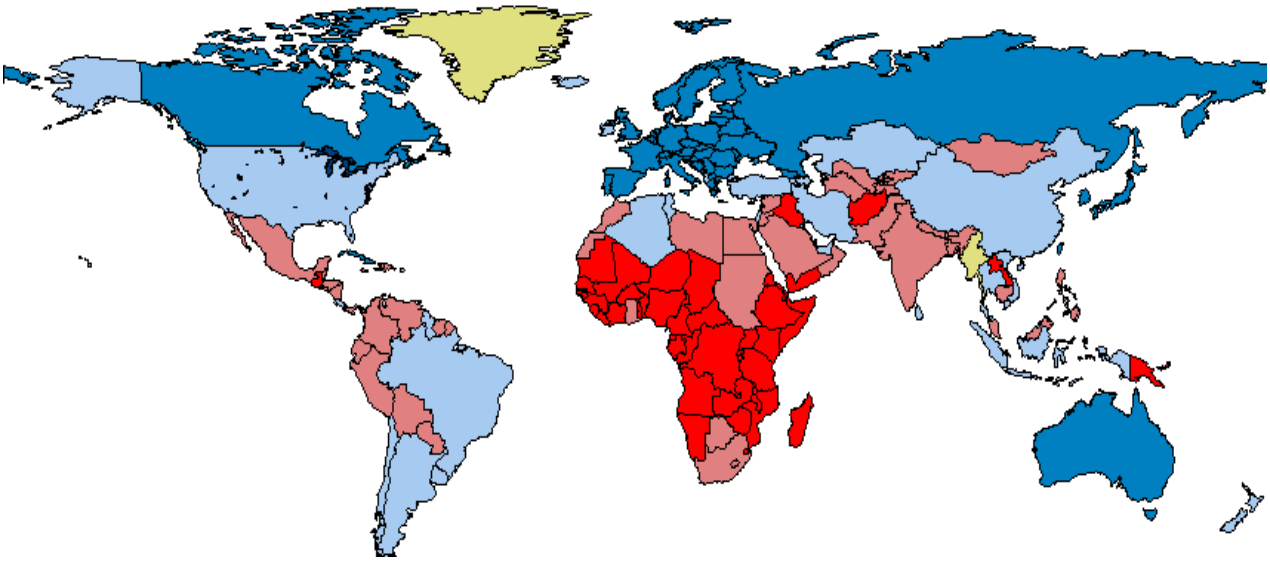
This slowdown in population growth has major implications for overall economic growth.[12] The economies of aging nations will not be stimulated by growing numbers of consumers and demand for housing. The capital growth generated by larger generations of young people approaching their peak earning years and saving for retirement will cease as well. Even if the growth of Europe's income per capita remained constant, its overall economic growth rate would be cut in half as the population declines over the next 30-50 years. An overall growth rate this small allows few margins for accumulation to invest for the future. As Benjamin Friedman[13] has argued, substantial growth rates allow more groups to share to some degree in growth, and provide social resources for a variety of services and investments. Overall growth rates below 2 percent per year, by contrast, allow for little redistribution or investment and tend to heighten social conflicts over such issues as pensions, migration, and labor/employer relations—situations we might see as the global economic downturn progresses.

Opposing Age Shifts and Migration Pressures

At the same time, the populations of much of the developing world will be tilted in the opposite direction, toward a larger percentage of youth. The youngest countries—all in the developing world—will have populations with only about 5 percent above age 60, but with nearly 50 percent under age 14.[14] While Europe and Japan will approach the mid-21st century with populations that are tilted toward the old, much of the developing world will have populations that are tilted toward the young (see map).

The obvious result of this imbalance is already taking place: a massive migration of young and working-age populations from the developing world to the developed world. Between 2000 and 2005, 2.6 million migrants moved each year to more developed countries from less developed regions.[15] Seeking new livelihood opportunities and entry-level jobs, young people are irresistibly drawn from high-youth-density regions to those with a lower percentage of youth; the OECD countries currently host 10 million foreign-born immigrants ages 15-24 and 55 million between ages 25-64.[16]

Yet this immigration—increasingly contentious in the developed world—is not the only consequence of this imbalance. To sustain their elderly populations, Europe, Japan, and North America will have to spend more money on health care and pension support. Whether active or ailing, the elderly population will need intensive medical procedures and medications necessary to sustain an active and healthy life into older ages—at a time when the domestic supply of new doctors and nurses will likely decline.



Red: 40+ Pink: 30-39 Light Blue: 20-29 Dark Blue: <20

Figure 2.1.B-1. Percentage of Population under 15 Years Old, 2005.

Keeping the elderly population at work is not a solution; older workers will generally not welcome entry-level work at entry-level wages, nor will they be likely to perform physically demanding work. Those gaps in the labor force will have to be filled by younger workers. Moreover, while older workers excel in experience and judgment, they do less thinking "outside the box." Path-breaking innovations in science and technology overwhelmingly come from those under age 45; countries with fewer and fewer younger workers will likely lose an edge in innovation, as well.

Developed nations can try to head off this impending growth slowdown in four ways.

1. **Improve Productivity.** First, developed nations can improve productivity by investing in technology, education, and innovation. An increase in productivity producing a 1 percent greater gain in output per capita per year would more than offset the change in population. Europe in particular should make it easier for individuals to start companies and use capital and labor flexibly to encourage entrepreneurial enterprises, which are the most important source of productivity-increasing growth.[17] Universities should seek increased support for training and research in the most technically important fields of biology, materials science, and engineering, and they should offer incentives to steer more students to the technical and engineering fields. Human capital must not be allowed to sit unused. In 2006, in the United States and Canada, roughly 63 percent of the population over age 16 was employed, while in the EU-15, only 52 percent of people over age 16 were employed. Although some European countries had workforce participation rates of 60 percent or more, France, Germany and Spain were at only 51-52 percent, and Italy was at 46 percent.[18] Increasing Europe's overall employment participation rates to North American or upper European levels would by itself offset the decline in its working age population for nearly a decade.

2. **Increase Immigration.** Secondly, countries could increase immigration and seek to raise immigrants' productivity and earnings to the average level as quickly as possible. While integrating and educating immigrants can take a generation or more, the United States, Australia, and Canada have enjoyed the benefits of making it easy for immigrants (especially skilled ones) to start businesses, acquire education, and move into the mainstream, such that the incomes of many immigrant groups exceeds the national norm. Even lower-skilled migrants can raise the overall productivity of a society if they work for lower wages than had previously been paid to non-migrants for similar work. Unfortunately, both in Europe and recently in the United States, debates on immigration have exposed the fear that immigration steals wealth from the native population. This pernicious view echoes the similarly mistaken idea that protecting trade by imposing high tariffs or blocking foreign investment will preserve the prosperity of a country. Migrants tend to self-select for entrepreneurial talent, ambition, and energy, and therefore they produce net gains for national economies that accept them.[19] A European country (or Japan) that has lost much of its own demographic momentum and energy can ill afford to exclude new generations, even if they come from abroad.
3. **Implement Pro-natal Policies.** A third way to head off this impending growth slowdown would be to pursue pro-natal policies that encourage larger families among the existing populations. However, it is not clear which policies would cause this result; demographers do not fully agree on the reasons underlying a baby boom. Unless societies start placing a higher worth on larger families than on expanding the consumption of consumer goods, small families will continue to be preferred. In richer countries, higher fertility is mainly found among more religious families, which is one of the factors accounting for much higher population growth in the United States than in Europe.[20] Short of a religious revival in Europe, a major increase in fertility and family size seems the least likely solution to the continent's demographic and economic decline.
4. **Encourage Older Migrants to Relocate to Developing Countries.** Fourth, and perhaps least discussed, encouraging a "reverse flow" of older migrants from developed to developing countries could create great benefits for both. If older migrants take their retirement along the southern coast of the Mediterranean, or in Latin America or Africa, it can greatly reduce the costs of their retirement. Of course, developing countries will need quality residential and medical facilities to ensure that these are desirable destinations. This effort could also counteract the constant drain of medical and nursing talent to rich developed countries. "Medical tourism" to many developing countries has already begun as residents of developed countries seek lower prices for medical procedures. Investment in facilities that will make long-term retirement attractive in cheaper locales will reduce the pension and medical cost burden for developed countries while channeling jobs and investment to developing countries with ample labor.

While Europe, the United States, and Japan will have older populations and many nearby developing countries will have young populations, the global population as a whole will nonetheless be heading for a relatively healthy age distribution of population. The most logical way to overcome the population distortions in varied regions will therefore be to ease the barriers to movement across borders to take advantage of the overall balance.

No doubt, a combination of all four methods will be required to offset the slowdown in population growth in high-income countries. However, we should recognize that one of the

biggest obstacles is the growing antagonism between the West and much of the Muslim world. The way forward for the West lies in greater openness and integration, increased investment in growth abroad, better integration of immigrant communities, and reduced barriers to emigration from fast-growing but youthful societies. None of this is possible with the high levels of fear, mistrust, and antagonism between the West and populations of many of the largest and fastest growing countries of the world. We must reach the degree of cooperation necessary to respond to the global population changes already in place for the next half-century. Much more than terrorism, these trends will affect the long-term prosperity of the developed but stagnating and rapidly aging populations of the West, as well as the fast-growing and extremely youthful population of the developing and largely Muslim nations.

An Increasingly Urban World

A hallmark of higher-income countries and regions throughout history has been the rise of large cities, from Xi'an to Rome to Constantinople (Istanbul) to Edo to London. Most low- and middle-income countries are now going through a transition from a situation in which most families were rural farm-producing families to a condition in which most families are urban consuming families. This creates new pressures on education, sanitation, energy supply, transportation, food storage, and food distribution. These pressures are greater in cities, and these demands fall chiefly on governments which can face opposition and even violence in the absence of jobs and adequate food supplies. Today, it is not so much the number of people per se, but the changing consumption and distribution patterns of populations (much of it linked to rising incomes and urbanization) that is putting pressure on global food and energy supplies.

For less-developed regions as a whole, the United Nations projects the percentage of urban population will increase from 42.7% in 2005 to 56% by 2030 and 67% by 2050. That is a fifty percent increase in the urban percentage, but since this increase is occurring in fast-growing countries, the total urban population is projected to more than double by mid-century. That projection entails the incredible figure of three billion additional urban residents in the less developed countries, added to the 2.3 billion who were there in 2005.

In sub-Saharan Africa, urbanization will be even more rapid, with the urban percentage expected to almost double, from 35% in 2005 to 67% in 2050, bringing about a tripling in total urban population from roughly 300 million today to over 1 billion by mid-century. In China and India, urban populations are also expected to grow explosively to 2050, increasing to 73% of total population in China (from roughly 40%) today and rising to 55% of total population in India (from under 30% today). Together, these two mega-populations are projected to increase their urban populations by over one billion people.[21]

While these projections may not be realized, it is vital to recognize that the "normal" urban percentage for developed countries is about 75% today, and it is logical to expect low- and middle-income countries to move toward this figure as they develop.

Since increasing employment and the efficiency of agriculture generally entails moving people off the land and into urban centers, it is reasonable to treat these projections as likely rather than as grim or abnormal.

Still, these levels of urbanization are likely to be reached at much lower levels of per capita income in most of the world than has been the case for the developed nations today. Today's more developed regions did not attain urbanization levels of 65% until 1970, and the United

States did not reach this level until 1950.[22] In fact, the rapid urbanization in today's developing world is most likely to occur in a manner similar to that of 19th century in Europe, when policing was inadequate, jobs were demanding, cyclical, and unregulated, and sanitation and education were limited. The result then was widespread labor strife, periodic urban rebellion, and even revolutions (1820s, 1830s, and 1848). Given that the developing world is facing a very similar situation of high numbers of youth, rapid development, high levels of urbanization, and relatively weak and often undemocratic governance, it is likely that systemic disorders will also arise unless efforts at global integration and support for fragile states can blunt those tendencies.

However, a major concern, again, is the continuing hostility between Muslim and western nations. Urban settings—even more than remote caves—offer excellent opportunities for recruitment and hiding of terrorist networks. The sheer proliferation of mosque and neighborhood networks, access to the Internet, easy transportation links, and concentration of targets mean that unless the proclivity of Jihadist groups to seek violence against the US and regimes allied to the West is sharply reduced, we can expect rapid urbanization to create increased opportunities for terrorist acts.

In addition, migration from countryside to city is only one step in what tend to become growing migration networks. In these networks, people move first from the countryside to regional urban centers, and then a portion of those move to national metropoles, major ports or industrial centers. This provides them with increased access to international migration networks. Increased urbanization in the developing world is thus likely to increase the flow of international migration from the developing to the developed world, again, just as the 19th century European pattern showed increased international migration to the Americas.

In short, the world of the first third of the twenty-first century will be quite different than the previous half century. The future will be a world of (1) fast-growing economies, (2) predominantly urbanized populations in the developing countries, (3) population stagnation, and (4) relative economic decline in Europe, Japan, and the United States. It will be a world in which the pressures for migration of young people to the aging societies of the West will become overwhelming, and in which the center of new investment and innovation shifts to the East. It will be a world in which civil and regional conflicts over the distribution of wealth and opportunities are likely to arise, yet where the aging and relatively smaller U.S. and NATO economies are less able to act independently to keep order in relatively larger transitional societies. Overall, and most importantly, it will be a world that cannot and will not be dominated by the power and economic weight of the Atlantic and west Pacific Rim economies, but rather by the mainland economies of Asia, the Middle East, and parts of Africa and Latin America. It will thus be a world in which expanded cooperation and deeper integration between the West and the developing (and especially Muslim) societies will be essential to maintaining order.

Conclusion: Embracing Globalization and New Global Institutions

As a greater portion of global economic growth and manpower shifts inevitably to Asia, Africa, and Latin America, the security and prosperity of the north Atlantic societies will depend on their ability to share in that growth, both by direct investment and economic linkages. At the same time, however, threats of local and regional conflicts in these high-manpower areas will confront North Atlantic societies who will have fewer young men and fewer free resources to support their military power given their needs to support pensions and health care for their older population.

These trends therefore require action in two major respects: first in regard to the major international institutions for military, financial, and diplomatic action, and second in regard to policies for international migration and law enforcement.

At this writing, the G-20 is replacing the G-8 as the international assembly that will arbitrate and coordinate global financial policies; that is the natural direction for future revision of the global institutional architecture. The global institutions established at the end of the second world war were designed to perpetuate the power of that war's victors: The United States, France, Britain, Russia, and China (first the Republic on Taiwan, now the People's Republic on the mainland). The World Bank, the International Monetary Fund, and of course the United Nations retain a structure giving these countries a dominant role, yet in the near future, a realistic appraisal of global economic and manpower strength would show a more diverse and more balanced roster: the United States, Russia, and China to be sure, but also the European Union as a whole (with Germany, Poland, Spain, and Italy as key players), India, Brazil, Indonesia, Pakistan, Turkey, Nigeria, South Africa, Mexico, and perhaps Iran, Ukraine, Vietnam, and others. Given the rapid shifts in economic and population growth expected in the years ahead, it seems it would be wise if international institutions had a more flexible leadership structure. Rather than a fixed set of countries filling key roles or staffing major institutions, it might be wise to stipulate that any country, or any regional group of countries that establishes a common leadership or representative faculty, which comprises 4% or more of the world's population or 4% or more of the world's economy should have a seat at the leader's table.

While such trends seem already underway in regard to global finance, a shift in the world's military alignments seems to be required, as well. If the major conflicts of the future are not likely to be between NATO and Russia, or between Australia and Indonesia, or between Bolivia and Brazil—in short, less likely to be international than domestic or regional conflicts brought on by state failures—then the major international military alliances need to have the capacity to intervene for humanitarian purposes and to provide order in failed states and regions affected by them. That is a requirement for relatively low-tech but trained military and civilian manpower that can be deployed in various regions of the world at an affordable cost. This requirement simply cannot be supplied by NATO for the extended future, or not as NATO is currently configured. This means it will be necessary either to expand NATO itself or replace it with a new global military alliance (with similar shared training and equipment to allow coordinated joint functions) geared to moving large numbers of military and civilian units into unstable areas and supporting them there. This will clearly require a partnership that spans regions and that includes both countries with large numbers of moderate-cost military and civilians available, and countries with logistic and financial resources to support them. More explicitly, this means incorporating countries such as Brazil, Indonesia, and possibly Pakistan, the Philippines, South Africa, and others into a joint structure. Right now, the United Nations' blue helmets draw on such countries precisely for such stabilization and peace-keeping missions. However, their resources are limited, and their joint training and coordination a work-in-progress. Moreover, they have the anomalous character of being deployed (and often commanded) by the decisions of the UN Security Council, which is dominated by the handful of WWII dominant nations but is manned mainly by troops from poorer countries who are often unrepresented on the Security Council. Given the likely prevalence of such missions in the future, driven by growing numbers of young men and urban expansion in developing economies affected by climate change and economic change, it would seem prudent to prepare the most effective forces for such missions by pooling the resources of diverse nations.

Finally, given the huge pressures for migration from poor to rich countries that will arise in the coming decades, it seems the choice faced by rich countries is whether to struggle against it or take measures that will make such migration more peaceful and productive. Aside from the fact that fighting such migration is unlikely to work, it will deny the rich countries the young workers they desperately need, and it will also deny poorer countries the safety-valve and means of training entrepreneurs that they desperately need. It would be more fruitful to invest in ways to integrate migrants, encourage circular migration (e.g., departure after a period of work), and even stimulate reverse migration to send retirees and medical-care and second career workers from the rich world to developing countries.

Yet it must be recognized that with growing migration and international economic flows will come opportunities for illicit trafficking, smuggling, and even terror operations. Therefore, the need for international cooperation in policing (not merely military actions), border patrol, and extradition will grow rapidly, as well. Again, new international arrangements for joint action and coordination will be necessary if the legitimate authorities are to keep pace with the globalization of criminal actions.

While population change poses major challenges, there is no reason to fear adverse consequences unless there is an attempt to willfully try to preserve a vanishing past in which a handful of dominant nations that dominated global growth since 1800 continue to regard themselves as primary and the rest of the world as secondary in power and importance. In fact, the population and economic growth of the developing world is a great opportunity for all people of the world to share in another century of economic growth and enriched cultural and technical diversity. Canadian tourists visiting the Brazilian or Indonesian rainforests in solar power vehicles and wearing natural/organic garments produced in African and south Asia while pondering their investments in new communication and entertainment technologies is the future we should seek. Enabling that future and avoiding risks of chaos will require new economic and military institutions that allow a closer meshing of peoples and economies while maintaining a framework for preserving order in a more globalized world.

↑ Back to the top

---

[1] Goldstone, Jack A. (2002). "Population and security: How demographic change can lead to violent conflict. *Columbia Journal of International Affairs* 56, 245-263.

[2] Cincotta, Richard, Robert Engelman, & Daniele Anastasion. (2003). *The Security Demographic*. Washington, DC: Population Action International.

[3] Leahy, Elizabeth, with Robert Engelman, Carolyn Gibb Vogel, Sarah Haddock, & Tod Preston. (2007). *The shape of things to come*. Washington, DC: Population Action International.

[4] UN Population Division. (2007). *World population prospects, 2006 revision* [Database]. Available online at http://esa.un.org/unpp/

[5] UN Population Division. (2007). *World population prospects, 2006 revision* [Database]. Available online at http://esa.un.org/unpp/

[6] UN Population Division. (2008). *World urbanization prospects: The 2007 revision population Database* [Database]. Available online at http://esa.un.org/unup/

[7] Central Intelligence Agency (CIA). (2008). *World factbook*. Washington, DC: U.S. Government Printing Office. Available online at https://www.cia.gov/library/publications/the-world-factbook/

[8] Central Intelligence Agency (CIA). (2008). *World factbook*. Washington, DC: U.S. Government Printing Office. Available online at https://www.cia.gov/library/publications/the-world-factbook/

[9] Economist Intelligence Unit. (2008, December 1). *Country Forecast December 2008 Updater*. Available online at www.eiu.com

[10] UN Population Division. (2007). *World population prospects, 2006 revision* [Database]. Available online at http://esa.un.org/unpp/

[11] Jackson, Richard, & Neil Howe, with Rebecca Strauss & Keisuke Nakashima. (2008). *The graying of the great powers: Demography and geopolitics in the 21st century*. Washington, DC: Center for Strategic and International Studies.

[12] Eberstadt, Nicholas. (2001, March/April). "The population implosion." *Foreign Policy*.

[13] Friedman, Benjamin. (2005). *The moral consequences of economic growth.* New York: Alfred Knopf.

[14] UN Population Division. (2007). *World population prospects, 2006 revision* [Database]. Available online at http://esa.un.org/unpp/

[15] UN Population Division. (2006, March). *International migration 2006*. New York: United Nations. Available online at http://www.un.org/esa/population/publications/2006Migration_Chart/Migration2006.pdf

[16] Longman, Philip. (2006, March/April). "The return of patriarchy." *Foreign Policy*. Organisation for Economic Cooperation and Development (OECD). (2008). OECD. *Statistical Extracts* [Database]. Available online at http://stats.oecd. org/wbos/index.aspx?lang=en

[17] Goldstone, Jack A. (2006). "Engineering, culture, innovation, and modern wealth creation." In Irene Johansson (Ed.), *Innovations and entrepreneurship in functional regions, Uddevalla Symposium 2005* (pp. 455-474). Trollhattan, Sweden: University West.

[18] United States Department of Labor (DOL). (2008). *Chartbook of international labor comparisons: The Americas, Asia, Europe.* Washington, DC: DOL.

[19] Simon, Julian L. (1999). *The economic consequences of immigration* (2nd ed). Ann Arbor, MI: University of Michigan Press.

[20] Longman, Philip. (2006, March/April). "The return of patriarchy." *Foreign Policy*. Organisation for Economic Cooperation and Development (OECD). (2008). OECD. *Statistical Extracts* [Database]. Available online at http://stats.oecd. org/wbos/index.aspx?lang=en

[21] UN Population Division. (2008). *World urbanization prospects: The 2007 revision population Database* [Database]. Available online at http://esa.un.org/unup/

[22] UN Population Division. (2008). *World urbanization prospects: The 2007 revision population Database* [Database]. Available online at http://esa.un.org/unup/

2.1.C. Demographic Security (Elizabeth Leahy Madsen)

Author and Organization: Elizabeth Leahy Madsen, Senior Research Associate, Population Action International
Contact Information: emadsen@popact.org

This paper reviews findings related to population and two aspects of stability—outbreaks of civil conflict and level of democratic governance—at the global scale. Because future stability will in part be shaped by the demographic trends that are already underway, this paper evaluates the implications of population projections and the assumptions that will be required to achieve them. The theme of the paper is that population influences security and development and is an important underlying variable in global stability because of its interactions with other factors. Research has demonstrated that countries with rapidly growing populations, and the youthful age structures they entail, are more likely to experience outbreaks of civil conflict and undemocratic regimes, in addition to other development challenges related to poverty, health and natural resources. These relationships are complex and should not be distilled to simple cause and effect; however, demography is a key component of future-oriented assessments and planning.

Although absolute population size has historically been considered an indicator of political power, current studies typically measure the impact of demographics through various categorizations of age structure—the relative size of age groups within a population—as well as through population growth rates and density. All countries' populations can be classified into one of four major age structure types based on their progression through the demographic transition: *very young, youthful, transitional, and mature.* The transition is the decades-long shift from high mortality and fertility rates, beginning with initial longer life expectancies, and later, smaller family size. The four-age structure types relate the share of a population comprised of children and young adults under the age of 30 to the share of older adults above age 60 and are termed very young, youthful, transitional and mature.[1]

Generally, countries with a very young age structure are those in which two-thirds or more of the population is younger than 30; in 2005, there were 64 such countries with a total population of more than one billion, including nearly all of sub-Saharan Africa, Afghanistan and Pakistan. In such countries, the youngest age groups comprise successively larger shares of the population, creating a pyramid-shaped age structure (see Yemen in Figure 2.1.C-1). Once a country's fertility rate declines below four children per woman, it is likely to reach the second category of age structures along the demographic transition, those termed youthful. The 31 countries in this category in 2005, with a combined population of nearly two billion, included India and most of North Africa. Countries in the middle of the demographic transition with a "transitional" age structure all have fertility rates lower than three children per woman, and between 45 and 60 percent of their population is comprised of young people under age 30. These 45 countries, including China, have a combined population of 2.4 billion. At the end of the demographic transition, countries with a mature age structure have fertility rates at the replacement level of 2.1 children per woman or below; less than 45 percent of their population is under age 30, while up to one-quarter of the population is comprised of older adults above 60. In 2005, the 56 countries with this type of age structure totaled 1.3 billion people and included nearly all of Europe.[2] The United States, with 42 percent of its population under age 30, had among the more youthful of the countries with "mature" age structures in 2005, given its relatively high replacement-level fertility rate,

which results in an age structure that is fairly balanced among children, youth and working-age adults (see Figure 2.1.C-1).

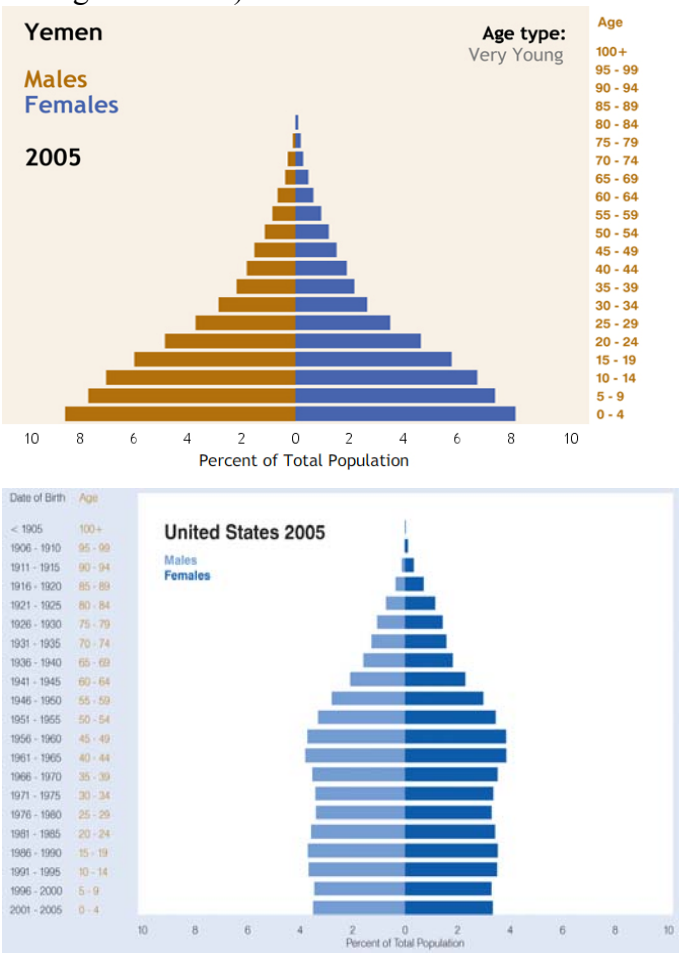


Figure 2.1.C-1. Age Structure Types along the Demographic Transition[3]

Population, Conflict and Governance

In an analysis of all new outbreaks of civil conflict with at least 25 battle deaths over the period between 1970 and 2007, countries with a very young age structure type (in which at least two-thirds of the population is younger than 30) were more than three times as likely to have experienced conflict than those with mature age structures that had completed the demographic transition.[4] Although there is a decline in the risk of conflict with each successive age structure type, the most profound drop in vulnerability lies between the first two categories of age structures, as shown in Figure 2.1.C-2. Countries with very young age structures are twice as likely to have experienced civil conflict as those with youthful age structures, demonstrating that some degree of fertility decline, even if fertility rates remain well above replacement level, may have broader ramifications.

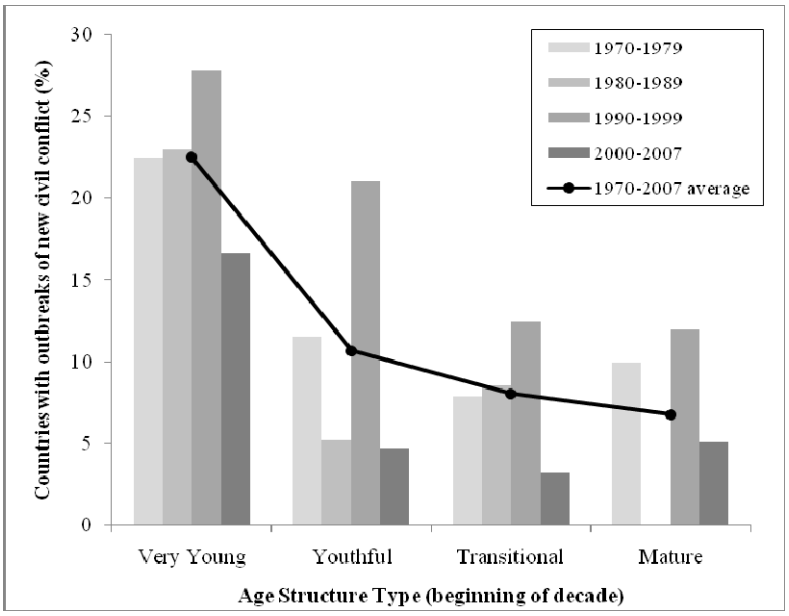


Figure 2.1.C-2. Risk of Civil Conflict by Age Structure Type[5]

These associations between age structure and conflict are verified and further quantified by empirical analysis. In a study controlling for level of development, regime type and previous incidence of conflict, countries in which the share of young people aged 15 to 24 exceeded 35 percent of the total adult population were 150 percent more likely to experience an outbreak of civil conflict between 1950 and 2000 than those with a more balanced age structure.[6] A young age structure is most strongly correlated with outbreaks of conflict in the case of countries with ongoing high fertility rates. Once fertility rates begin to decline and the demographic transition is underway, even though a country retains a youthful age structure for a few decades due to previous high fertility, outbreaks of conflict are less likely. Many researchers suggest that the link between population and conflict lies in competition over resources, especially economic resources, namely jobs. Unemployment, especially after education, or the ability to only eke out a meager living, enables resentment and hopelessness that can easily be drawn into uprisings or rebellions. The members of a "youth bulge" are not inherently dangerous, but without employment or the prospects of stability, they can be vulnerable for recruitment into a movement promising them income and perceived well-being, especially networks that attribute their troubles to the state or other groups.

Research is also underway on the relationship between age structure and governance. Cincotta finds that although a very young age structure creates "a social environment where the regime's legitimacy is strained and the political mobilization of young men is relatively easy," this volatile situation rarely results in sustained new democratic governments. Instead, autocracies and partial democracies are more likely in countries with a youthful age structure precisely because countries with such demographic profiles are also prone to instability and conflict; given the option, societies would choose the security of a heavy-handed ruler over the unknown threats of civil violence. Moreover, if states with a young age structure do attain liberal democracy, they are prone to backsliding or otherwise reverting towards authoritarianism. This analysis finds that once the "youth bulge" ratio of young adults aged 15 to 29, as a share of the working-age adult population, falls below 40 percent, countries are more likely to have regimes characterized by liberal democracy. After this demographic change, lower dependency ratios allow for greater

savings and investments in health and education, a more highly trained workforce, and economic growth, if governments invest in human capital and financial systems. In such situations, authoritarian leaders can be more willing to loosen political restrictions, or elites may create pressure for a change in government.[7]

In a historical analysis similar to that described above for civil conflict, the likelihood of a government being rated as a full democracy progressively increases as countries pass through the demographic transition. Between 1970 and 2007, 14 percent of countries with very young age structures and 80 percent of countries with mature age structures were classified as full democracies, on average, at the end of each decade, as shown in Figure 2.1.C-3.[8] This pattern continues with other measurements of governance: countries with very young and youthful age structures are also more likely to be characterized by weak institutional capacity, government corruption, poor regulatory quality, and fewer political freedoms and civil liberties. Countries that have passed through the demographic transition and achieved transitional or, in particular, mature age structures, are rated much more highly in terms of these governance measures.[9]

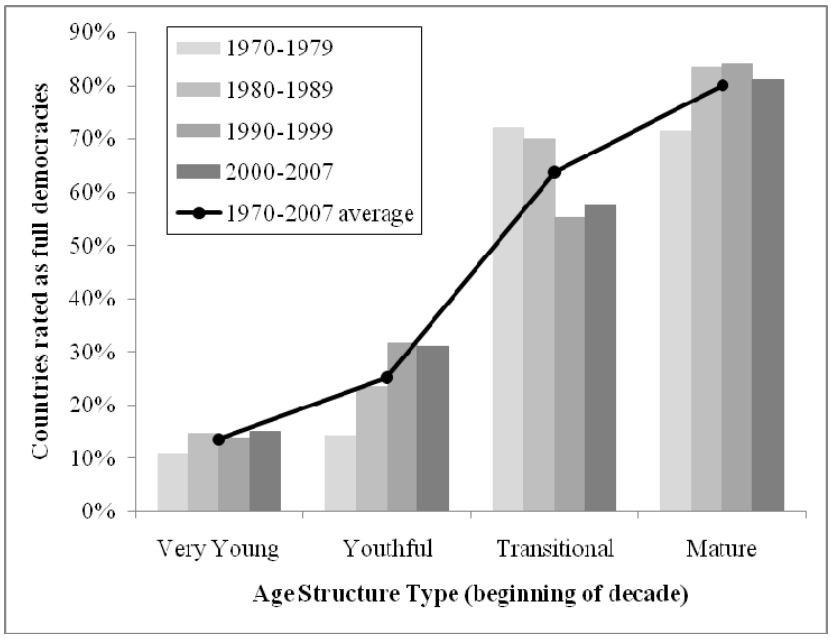


Figure 2.1.C-3. Likelihood of Democratic Governance by Age Structure Type[10]

It is clear that there is no direct causal relationship between population and conflict or population and democracy—no single demographic threshold that, once crossed, leaves a state doomed to upheaval or tyranny. Most countries with youthful, growing populations pass through decades free from outbreaks of conflict despite demographic pressures, and the ultimate successes and challenges to development depend on individual context. Researchers have widely reiterated that the presence of a youthful age structure is not sufficient to create conflict; nor does a mature age structure guarantee domestic peace and democracy. Rather, underlying social conditions that create grievance and make involvement in an insurgency a viable or even appealing option are necessary. In fact, because of the interplay of factors such as poverty, inequity, and institutional capacity, the relationship is more holistically framed as one between population and development.

The influence of age structure on development is driven by the three forces of demographic change (fertility, mortality and migration), each of which have their own distinct and variable drivers, and the ways in which demographic forces exacerbate or mitigate other structural forces. Considered alone, demography is unlikely to be the tipping point that leads a country into war or ensures its stability and the rule of law. However, the role of demography as an underlying force influences a state's path. Meanwhile, the elements that motivate demographic change are fairly well understood, and its future course can be fairly clearly outlined, yet demography is often neglected by policymakers.

Population Projections and Future Trajectories

The most notable characteristic of current population trends is their disparity, which has created a demographic divide that is unparalleled in history. This divide is most evident between the most- and least-developed countries of the world. Currently, most population trends—whether growth or decline—are caused by fertility rates. Mortality rates have declined everywhere, though they still remain high in some places. Still, because most children who are born today survive to adulthood, the number of children born per woman drives the pace at which populations grow. Mortality still plays a role, especially in cases like HIV/AIDS, where disease reached an epidemic scale. However, even in the south African countries that have been hardest hit by AIDS, birth rates are higher than death rates, and populations are growing. Migration also plays a role. About 3 percent of world's population lives in countries other than those in which they were born. In countries like the United States that receive a large share of migrants, this has made labor forces larger than they otherwise would be, and to a certain extent has offset the effect of population aging.

Because steady fertility declines have occurred in most regions of the developing world, and because industrialized countries face the emerging challenges of sustained low birth rates, it would be easy to surmise that population aging is inevitable. However, rapid population growth will remain the main demographic focus for the foreseeable future in a wide swath of the developing world, incorporating most of sub-Saharan Africa, as well as countries in southwest Asia such as Afghanistan, Pakistan and Yemen. For the entire sub-Saharan African region, fertility rates have declined less than 20 percent since 1960 and remain very high at 5.5 children per woman.[11] Fifty-eight percent of the world's people live in countries with fertility rates higher than replacement level, guaranteeing ongoing population growth for the long term. Within this group, almost one billion people, including 91 percent of sub-Saharan Africa, live in countries with fertility rates above four children per woman, a level that would lead their populations to double in less than 35 years. Meanwhile, 42 percent of the world's population resides in countries at the end of the demographic transition, with fertility rates at or below replacement level. According to United Nations population estimates, 17 countries are losing population. These countries are concentrated in Eastern Europe, but they also include Germany, Japan and Russia.

However, the scale of demographic change is also bifurcated across countries experiencing growth or aging. While the region with the highest pace of population decline, Eastern Europe, is projected to see its collective population shrink by seven percent between 2005 and 2025, the fastest-growing region, Eastern Africa, is projected to have a population increase of 63 percent in the same time period.[12] The challenges faced by such regions with youthful age structures and rapidly growing populations are sometimes recognized by government officials and the broader policy community. A report outlining guidelines for U.S. policy towards Africa has described

demographic trends as one of three "megachallenges" facing the continent because "they can create a context in which insecurity thrives."[13]

Demography is unique among the social sciences because current conditions reliably inform us of much that lies ahead. Most of the world's population twenty years hence is alive today. Along with assessments of the present situation, projections of the future of demographic trends are a second major component in policy design, particularly those projections occurring the relatively short term of the next 20 to 30 years. However, it is too often that policymakers and others presume that projections are sure predictions when in fact they are based on a series of uncertain assumptions. For this reason, the trends set into motion by current indicators and policies, not demographers' best projections, are the true determining forces of future populations.

The most commonly cited and complete set of demographic estimates and projections at the national level is revised biannually by the United Nations Population Division. The projections are grouped into variants based on assumptions related to fertility, mortality, and migration. Fertility carries the most weight in influencing current demographic trends, and the majority of the UN variants are associated with differing assumptions about fertility. The medium-fertility variant assumes that all countries' fertility rates will eventually converge at a universal level of 1.85 children per woman and begin movement toward that level virtually immediately, regardless of whether their current total fertility rate is far above or far below that figure.

These projections can seem optimistic for countries at both ends of the demographic transition due to their discrepancy with recent historical trends. Two-thirds of the countries in sub-Saharan Africa where survey data are available have had no significant decline in fertility rates in recent years.[14] Such stalls in fertility decline can wreak havoc with the projections, as was recently the case with Kenya. Between 1998 and 2003, Kenya's total fertility rate rose slightly from 4.7 to 4.8 children per woman. Because the UN assumed a declining fertility rate during that period, the projection of Kenya's total population in 2050 nearly doubled from 44 to 83 million.[15]

The assumption that countries will quickly undergo the rapid changes in fertility rates that would be needed to achieve a standard level is at odds with the paths many countries' demographics have followed dating back decades. Under the medium-fertility variant, Uganda's total fertility rate would decline to 4.8 children per woman by 2025. This would require that fertility decline by 28 percent over the next 20 years, a pace more than four times faster than that of the previous two decades. Although rapid fertility declines in short timeframes have occurred in several countries and may well do so in the future, Uganda is not foremost among the likely candidates for rapid demographic change.

The UN projections also assume an outright reversal of recent demographic trends in countries experiencing population aging. Between 1960 and 2005, the total fertility rate for the world's developed countries as a whole declined by 43 percent to 1.6 children per woman.[16] According to the UN projections, fertility in these countries is expected to *rise* 14 percent by 2050. China's administrative regions of Hong Kong and Macao, where fertility currently averages less than one child per woman, are the lowest fertility areas of the future, with their fertility projected to rise to almost 1.4.

The relationships between demographics, conflict and governance described above demonstrate that analysis of population projections and their assumptions is an important element of future stability planning. For the U.S. and many developed countries, the geographic focus for security and development efforts is shrinking. The National Intelligence Council notes that the number of

countries with youthful populations in the conflict-prone "arc of instability" is in decline, and those receiving the most attention in the near future are likely to be mostly in sub-Saharan Africa, as well as parts of the Middle East and southern Asia. This geographic focus aligns with the countries facing the youngest age structures, which compounds their vulnerability to instability now and in the future. Population projections assume that these countries will follow the path of others through the demographic transition and that their age structures will become more balanced over time. However, such trajectories are far from guaranteed and may in fact be overly optimistic. Demography is a powerful and dynamic force, but it is the aggregation of individual decisions, and it is subject to the policy decisions made today.

↑ Back to the top

---

[1] Leahy, E., et. al., *The Shape of Things to Come: Why Age Structure Matters to a Safer, More Equitable World,* Population Action International, Washington, DC, 2007.

[2] Data for age structure are calculated by the author using United Nations Population Division, *World Population Prospects: The 2008* Revision, United Nations Population Division, New York, 2009.

[3] Figure created by Population Action International. U.S. figure originally printed in Leahy et. al. 2007.

[4] These findings are based on the 2007 PAI report *The Shape of Things to Come,* updated to include the years 2000-2007. Data for conflict are drawn from the Uppsala Conflict Data Program/International Peace Research Institute *Armed Conflict Dataset* (Version 4, 2008).

[5] Figure generated by author using Population Action International age structure data and conflict data from Uppsala Conflict Data Program/International Peace Research Institute 2008.

[6] Urdal, H. "A Clash of Generations? Youth Bulges and Political Violence." *International Studies Quarterly* 50(3): 607-629, 2006.

[7] Cincotta, R. "Half a Chance: Youth Bulges and Transitions to Liberal Democracy." *Environmental Change and Security Program Report* 13:10-18, 2009

[8] Center for Systemic Peace. *Polity IV Project: Political Regime Characteristics and Transitions.* Available at http://www.systemicpeace.org/polity/polity4.htm.

[9] Freedom House. *Freedom in the World.* Available at http://www.freedomhouse.org/template.cfm?page=15; World Bank. *Governance Matters: Worldwide Governance Indicators.* Available at http://info.worldbank.org/governance/wgi/index.asp.

[10] Figure generated by author using Population Action International age structure data and governance data from Center for Systemic Peace.

[11] All data in this paragraph are derived from United Nations Population Division 2009.

[12] United Nations Population Division 2009. These projections assume significant changes in fertility in all countries toward a universal rate below replacement level.

[13] Gavin, M. "Africa's Looming Megachallenges." In Cooke, J. and J.S. Morrison (eds.). *U.S. Africa Policy Beyond the Bush Years,* Center for Strategic and International Studies, Washington, DC, 2009.

[14] Bongaarts, J. "Fertility Transitions in Developing Countries: Progress or Stagnation?" Working Paper No. 7. Population Council, New York, 2008.

[15] United Nations Population Division 2003 and United Nations Population Division 2005.

[16] United Nations Population Division 2009.

2.2. Economics

2.2.A. Banking Crises, Collective Protest and Rebellion (David Richards, Ronald Gelleny)

Authors and Organization: David L. Richards, Department of Political Science, University of Memphis and Ronald Gelleny, Dept of Political Science, University of Akron
Contact Information: drich1@memphis.edu; rgelleny@hotmail.com

Abstract

Both developed and developing countries have experienced turmoil associated with banking sector failures. However, while literature exists studying the economic costs of banking meltdowns, little scholarly attention has been devoted to examining the relationship between banking crises and domestic agitation (internal conflict). The failure to explore the impact of banking crises on domestic agitation risks excluding a key component in understanding the puzzle of domestic political stability. Examining a dataset of 125 countries for the years 1981 to 2000, we find banking crises to be systematically associated with greater levels of collective protest activities such as riots, anti-government demonstrations, and strikes.

Introduction

Regardless of their geographic regions, developed and developing countries have experienced economic and political turmoil associated with banking sector failures. Economists often acknowledge the existence of political pressures associated with the containment and solution to banking failures. However, analyses of corrective bank intervention policies routinely discount or downplay the potential political hazards and costs associated with such policies.[1,2,3,4,5] (Mahon and Corrales [2002] observed that while scholars were closely following [and predicting] Argentina's financial crisis, little speculation was devoted to the political consequences of the banking crisis that eventually led to the resignation of four presidents.[6]) Thus, while there is a growing body of literature studying the economic costs of banking meltdowns, little scholarly attention has been devoted to actually examining the relationship between banking crises and domestic agitation (internal conflict).

The authors believe that the failure to explore the impact of banking crises on domestic agitation risks excluding a key component in understanding the puzzle of domestic political stability. The financial sector plays the crucial role of dispersing capital to competing domestic interests, and the failure of domestic banks has significant repercussions on domestic economic and societal stability.[7,8,9] Furthermore, banks are generally considered more fragile and subject to failure than other firms.[10,11]

Consequently, owing to the banking sector's fragility and economic importance, some studies of domestic agitation may be seriously flawed due to the problem of misspecification. Furthermore, since financial investments (and by implication the accompanying financial institutional framework) are critical to the globalization process, it is important to understand the relationship between banking crises and domestic agitation. After all, domestic stability—political as well as economic—is widely considered to be a key element in attracting and maintaining foreign investment. Understanding the consequences of the relationship between the two variables can help us understand the tide of the globalization process.

This article attempts to address what we perceive to be a gap in the stability literature by empirically investigating the relationship between banking crises and domestic agitation. We perform our empirical examination on a cross-sectional time-series data set comprising 125 economically developed and developing countries of the world for the years 1981 through 2000. Rather than focusing on a single indicator of domestic agitation, we conduct a dimensional analysis of the concept and ultimately employ two separate indicators: collective protest and rebellion/internal war. This allows us to paint a broad picture of how banking crises affect domestic political stability. We find a country's level of banking crisis to be strongly associated with its level of collective protest activities, but not at all associated with its level of rebellion/internal war activities.

Banking Crises and the Economy

A banking crisis is a situation where a country's banking system has exhausted most of its banking capital and has become profoundly "unsound."[12] By and large, the soundness of a banking system is measured by its "ability to withstand adverse events."[13] More fully:

> *A sound banking system may be defined as one in which most banks (those accounting for most of the system's assets and liabilities) are solvent and are likely to remain so. Solvency is reflected in the positive net worth of a bank, as measured by the difference between the assets and liabilities (excluding capital and reserves) in its balance sheet. In other words, the distance between soundness and insolvency can be gauged in terms of capitalization, since net worth is equivalent to capital plus reserves.*[14]

Undercapitalized banks are thought more likely to be prone to political, economic or even nature-induced shocks and thus are more likely to collapse.

Banks play a unique role in an economy. The financial sector can be described as the "brain of the economy."[15]. It is the banking sector that directs scarce capital to competing domestic interests. Without a sound banking sector, macroeconomic stability can be difficult to achieve and, thereby, the country is more prone to economic crisis.[16,17,18]). (Clearly, the reverse is also true. That is, the banking system is much more likely to be stable if the macroeconomic condition of a country is sound [i.e., steady economic growth and low inflation rates].)[19] This is particularly true for less developed countries (LDCs), where banks often provide the vast majority of finance through the allocation of the nation's savings to creditworthy individuals and corporations. For most LDCs, alternative financial instruments (stocks and equities) are limited, and therefore indigenous economic actors must rely on domestic banks. Without available financial resources, domestic enterprises cannot access the capital required to maintain, let alone expand, levels of production. Declining levels of income and continued economic hardship are associated with falling production levels.[20,21] Solvent banks with a steady source of financial capital are essential for the process of late industrialization.[22,23]

Regardless of international economic standing, all countries that are exposed to a banking crisis experience negative economic costs. At a minimum, citizens experience some loss of wealth and a disruption to the supply of credit for investment projects.[24] Nonetheless, the depth of any particular crisis is dependent upon several factors, namely the size of the banking sector relative to the total wealth of the economy and the number of sound banks in the system.[25,26] For example, although the Savings and Loan crisis of the late 1980s eventually cost the United States economy well over $100 billion. The huge US economy, combined with the existence of a large number of healthy banks, was able to absorb the loss without experiencing a significant

economic downturn. However, LDCs have been especially vulnerable to banking crises due to the weakness of their domestic economies and the frailness of many financial institutions. For instance, the expense of bailing out the banking sector following the Mexican crisis of 1994–1995 has been estimated at 14.4 percent of Mexico's GDP for 1997, to be amortized over 30 years during the lifetime of the programmes. Moreover, in 1995, the Mexican economy slowed by 6.2 percent and inflation climbed to over 52 percent.[27] (It is also important to remember that Mexico was given an approximately US$40 billion lifeline from the US government. Most less developed countries [LDCs] cannot count on such support to weather economic turbulence.) Similarly, between 1988 and 1996, 20 African countries experienced banking crises, of which five had to spend more than 10 percent of their GDP to repair the damage.[28] Thus, banking crises impose economic strains on all countries, but they are particularly expensive to those nations already in a fragile economic condition.

Although it is widely accepted that banking crises impose significant costs on affected countries, it is not as easy to identify how or when banking crises occur.[29,30,31] There are, however, several key dangers that appear to be prevalent in many banking crises. For one, many banks have established connected lending patterns. That is, bank managers often form relationships with companies and/or government officials that they lend to at very reasonable interest rates. After some time, it becomes extremely difficult to deviate from this established pattern, regardless of the risk of the venture, thus inviting more bad loans to follow.[32,33] On August 11, 2001, Antonio Siba-Siba Macuacua, an economist hired to investigate the bank failure of Mozambique's Banco Austral, was found murdered at the bank's headquarters. Siba-Siba had begun a process of reducing the size of the workforce and had moved to recover loans that had been made predominately to members of the governing elite.[34] For example, South Korea is an example of the ill effects of "connected" borrowing. South Korean government officials applied considerable pressure on banks to support domestic projects and to funnel loans to large industrial conglomerates. In particular, bank loans were directed to the giant steel firm Hanbo, although the firm was failing to generate any profits by the mid-1990s. The result was that South Korean bank portfolios were dominated by non-performing loans. Hanbo alone is reported to have cost South Korean banks around US $7 billion.[35]

Additionally, many governments see the banking sector as another tool for it to carry out ill-founded industrial policy. In the 1980s the Zambian government set up a number of state-owned banks for the purpose of funding public sector companies. By the 1990s most were insolvent, with billions of dollars having been wasted.[36] Lastly, poor and/or fraudulent bank management can lead to bank failures.[37] The Bank of New England, for example, found itself buried in non-performing loans by the end of the 1980s due to overexposure of loans in the real estate market (that is, a lack of loan diversity). A study by the US Office of Comptroller of the Currency reviewed about 160 cases of national bank failures since 1979 and found that in most cases banks failed to have an adequate loan policy, or if a policy was in place, they failed to follow it.[38] (It is interesting to note that Claude Trichet, governor of France's Central Bank, was investigated for complicity in falsifying the 1992 accounts of the loss-making Credit Lyonnais. The eventual cost of bailing out Credit Lyonnais to the French taxpayers was approximately US $17 billion.)

Banking Crises and Domestic Agitation

What is the relationship between banking crises and domestic sociopolitical conditions? The most immediate impact of financial unsoundness is that banks become vulnerable to deposit runs. When citizens are threatened with the loss of their savings, they will scramble to recover their money before the banks can close their doors. That is, citizens demand full and immediate access to their financial funds during a financial crisis. The failure of citizens to recoup their savings can incite violent dissatisfaction with the government. As noted by former Uruguayan president Julio Maria Sanguinette, "The banking system will never take you to paradise, but it can bury you in hell in an afternoon."[39]

Moreover, the attempt to revitalize the banking sector is an extremely costly endeavor. Funds that could have been used for social programmes or job creation projects are instead funneled into repairing the banking sector and implementing policies designed to recover domestic and foreign investor confidence. The result is budget cutbacks that often eliminate subsidies and programmes for the middle and lower classes, thereby decreasing domestic political support for the government.[40] Furthermore, well-connected bankers often use their political influence to exempt themselves from intervention policies enacted to stabilize the banking crisis. Consequently, political leaders are hesitant to impose interventionist policies in a swift and decisive manner.[41,42] (If governments fail to impose interventionist policies that gain the confidence of bank depositors, the demise of a bank can quickly spill over to other banks, thereby worsening the financial sector.[43])

In the late 1990s and early 2000s, many countries had the misfortune of experiencing banking crises. In many of these cases, governments were faced with an angry citizenry demanding full access to their deposits. For instance, in December 2001, Argentina experienced economic and political turmoil because of a significant bank failure. In an attempt to end a depositor run on the banks, the government imposed a temporary suspension on access to bank accounts. Specifically, the government limited cash withdrawals to $250 per week for a period of 90 days, and overseas cash transfers were limited to $1000. The only exception to these caps was the allowance of the continued use of credit and debit cards.[44,45,46] The policy had the effect of enraging millions of middle-class and poor Argentines who lacked access to banking credit. Consequently, President Adolfo Rodriquez Saa's government was forced to resign after facing a large general strike organized by trade unions and countrywide violent street protests responding to the government's bank intervention policies.

Argentina was not the only South American country to experience political unrest owing to banking failures. In 1999, the Ecuadorian government was forced to close eight banks due to insolvency. (In this case, non-performing loans were caused by the widespread "El Nino" floods of the previous year.[47]) In order to limit a run on the banks, the government ordered the freezing of all bank deposits. In response, trade unions, indigenous organizations, and grassroots activists staged a general strike that crippled the country. To stop the ensuing social unrest, the president of the country imposed a state of emergency. In 2000, Nicaragua also experienced a banking crisis in which authorities limited access to deposits. Once again, this policy was met with angry and loud street protests. In the end, public pressure forced the government to guarantee all depositors full recovery of their savings.

Restructuring the banking sector often includes the loss of jobs as banks are streamlined to decrease financial losses. At the extreme, many banks may be closed, resulting in a permanent

loss of jobs in the financial sector. For example, the French government's restructuring of *Credit Lyonnais* included the closure of 73 branches at the cost of numerous jobs. Well over 2,000 employees protested in the streets of Paris over the job cuts.[48]

Long-term externalities may also accompany banking crises. For one, firms that require financing are negatively affected by the ensuing credit crunch that inevitably ensues. Simply, firms are often unable to obtain the necessary financial support to stay in business. The Nicaraguan government's decision to close the Bancafe bank in 2000 threatened the livelihood of farmers since it was one of the few banks that lent money to farmers and ranchers. Not surprisingly, hundreds of farmers protested in front of the bank's offices throughout the country. The financial meltdown in Southeast Asia in the late 1990s also provides an example of the adverse effects of a credit crunch. Because South Korean banks lacked the capacity to lend, the small- and medium-sized firms increasingly faced a credit shortage, thus adding to the economic woes of the country. The result was increased domestic agitation as unions and citizens increasingly protested the economic situation. As a consequence, a union rally protesting the economic crisis in May 1998 resulted in a violent riot.

A lack of government willingness to enact a reputable banking intervention policy can also contribute to domestic political and economic turmoil. For example, if influential banks are exempted from closure policies, the public trust in the overall intervention strategy may crumble. This is exactly what occurred in Indonesia in 1997, when the government allowed well-connected owners of banks to resume operations and the cost of solving the banking crisis continued to climb. (The failure to close insolvent banks often encourages banker managers to accept high-risk loans in hopes of high returns. Inevitably the bank continues to accrue significant losses, thereby increasing the costs of future financial bailouts.) As a result, disillusioned Indonesians lost confidence in the intervention plan and continued to withdraw funds from banks[49] Moreover, the political and economic conditions in Indonesia worsened as government corruption in the banking sector was exposed. Unable to establish political and economic stability, the government of President Suharto experienced severe domestic strain.[50] In an effort to end the run on banks and restore confidence in the economy, President Suharto decided to guarantee all bank deposits and liabilities. Although bank runs began to subside, anger at his government continued to grow.

In an effort to further shore up the financial sector and stabilize its balance of payment crisis, Indonesia turned to the International Monetary Fund (IMF) for financial help. In turn, the IMF demanded that the Suharto government reign in domestic spending. The chief opponents to Suharto tended to be labour unions, political rivals of the government, and students disenchanted with the corruption of the Suharto family and its political allies. In addition to protests and riots, Indonesian citizens targeted the ethnic Chinese community for revenge. The eventual outcome was the overthrow of the Suharto regime. (Similar events occurred in Malaysia and Thailand. The two respective governments failed to effectively monitor domestic banking crises, and economic growth stalled. In an effort to create domestic jobs [and perhaps to deflect attention from their inadequate handling of the banking crisis], both governments deported foreign workers. The end result was a series of massive protests and riots.)

Additionally, banking crises can often be associated with severe economic downturns.[51,52] This can happen when a banking crisis is accompanied with a currency crisis, thus causing a significant outflow of investment capital. (Even if governments guarantee deposits, this places an additional strain on government resources, particularly for LDCs, since the government is

responsible for all refunds. Furthermore, this policy invokes the issue of moral hazard. That is, if bankers know that depositors will not punish them, they have a greater incentive to act irresponsibly.) As a result, inflation is likely to increase rapidly, thus forcing the government to raise interest rates. Moreover, since there is no longer an ample source of credit available to domestic enterprises (or available only at high interest rates), economic growth can slow dramatically. Consequently, LDCs are often forced to appeal to the IMF for loans that require weighty decreases in public spending and other economic liberalization measures. (The elimination of public projects as a condition of IMF loans was met negatively by Mahathir Mohamad, the prime minister of Malaysia, since these projects were to supply a substantial number of jobs for the country.) All of these factors contribute to a substantial increase in unemployment rates. The overall costs of the crisis and the resulting adjustment programmes fall disproportionately on middle-to-low income households, causing decreased support for the prevailing government. In the mid-1990s, for instance, Venezuela's government assumed control of over 43 financial institutions, and by the end of the crisis, the government owned over 70 percent of the banking system.[53] To help stabilize the financial system and correct a balance of payments problem, the Venezuelan government turned to the IMF. The ensuing austerity programme was protested by thousands of public employees opposing the suspension of pay wages. Others protested the decision to eliminate subsidies to the poor. The result of the banking crisis was the decision by the government to temporarily suspend six constitutional guarantees.

The above discussion provides us with a testable hypothesis regarding the relationship between banking crises and domestic agitation: countries experiencing a banking crisis will encounter a contemporaneous increase in domestic agitation. A corollary of this hypothesis is that these effects may carry over into the future. Below we describe our research design and findings.

Data, Models and Estimation

Data and Models. For our empirical analyses, we employ a pooled cross-sectional time series data set comprised of 125 economically developed and developing countries for the years 1981 to 2000. This period is lower bound by the availability of our banking crisis data and upper bound by the availability of both our domestic agitation data and banking crisis data. The country-years included are those that were common to all source-datasets used in the analyses. Our data are representative of all geopolitical regions of the world, as well as all regime types. To examine the relationship between banking crises and domestic agitation, we examine both the contemporaneous and lagged relationship between a country's level of banking crisis and its level of two types of domestic agitation. The variables included in our models are discussed below, and each variable's expected relationship with domestic agitation is stated.

Domestic Agitation. This study differs from some other large cross-national empirical studies of contentious politics in that the chief concept of interest is not government-based action such as negative sanctions, [54,55,56,57,58,59] or human rights violations, [60,61,62,63] but rather, the chief concept is collective action taken by citizens against government. We are interested in what causes domestic agitation by citizens rather than what causes government repression of citizens.

In our previous discussions, we have used the term "domestic agitation" to represent a varied group of collective activities which citizens may employ to agitate against a government. To test our central hypothesis that banking crises may instigate domestic agitation, however, we must be able to distinguish between various types of such agitation. Gurr (1970), Hibbs (1973) and Lichbach and Gurr (1981) imply that the important sources of variation between agitation

activities are (1) the extent of organization, (2) the scope of citizen participation, and (3) the violent0non-violent nature of activities.[64,65,66] Gurr (1970) distinguishes between turmoil (mass, spontaneous unorganized violence) and revolution (consisting of conspiracy and internal war, both of which are organized violence, but differ in scope of citizen participation).[67] Hibbs (1973) differentiates between what he calls collective protest (riots, anti-government demonstrations and politicized strikes) and internal war (armed attacks, assassinations).[68] Similarly, Lichbach and Gurr (1981: 5) distinguish protest activities which are "short-lived challenges by members of associational groups" (demonstrations, strikes, riots) from protracted rebellion activities, which involve combat between rebels and governments.[69] (Others, such as Muller and Seligson [1987] take a unidimensional view. Concerned with the relationship between inequality and insurgency [political violence], they operationalize this type of agitation as "the death rate from domestic conflict per one million population."[70]) Essentially, these studies posit dimensions of domestic agitation.

Our domestic agitation data come from Arthur S. Banks's *Cross-National Time-Series Domestic Conflict Data*.[71] Banks provides annual event-count indicators of six types of agitation: assassinations, general strikes, guerrilla activities, riots, revolutions, and antigovernment demonstrations. Based on Hibbs (1973)[72] and on Lichbach and Gurr (1981),[73] we would expect to see two dimensions in our domestic agitation data. First, we would expect to see a "collective protest" dimension associated with activities such as riots, anti-government demonstrations, and strikes. Second, we would expect to see a "rebellion/internal war" dimension associated with activities such as guerrilla activities, assassinations, terrorism and revolution. According to Gurr (1970), we would expect to see three dimensions: (1) a "turmoil" dimension associated with mass spontaneous unorganized violence such as riots or large-scale demonstrations, (2) a "conspiracy" dimension associated with organized violence of limited participatory scope, such as assassinations and other very selective forms of terrorism, and (3) an "internal war" dimension characterized by organized mass violence such as all-out guerrilla warfare and revolution.[74]

Table 2.2.A-1 shows the results of a varimax principal component analysis used to determine how our data conforms to any of the aforementioned dimensional expectations. We find that the two-dimensional schemes of Hibbs (1973)[75] and Lichbach and Gurr (1981)[76] are clearly supported, while Gurr's (1970) three-dimensional scheme is not.[77] The analysis returned two identifiable factors, each of which clearly included three of our six indicators. Factor 1 includes riots, anti-government demonstrations and general strikes. This would be the "collective protest" dimension. Factor 2 includes guerrilla activity, revolutions and assassinations. This would be the "rebellion/internal war" dimension.

Table 2.2.A-1. Varimax Principal Component Analysis of Six Types of Domestic Agitation

Type of Agitation	***Factor 1*** ***collective protest***	***Factor 2*** ***rebellion/internal war***
Assassinations	.14	.58
General strikes	.61	.07
Guerrilla warfare	.10	.84
Riots	.86	.05
Revolutions	–.03	.79
Anti-government demonstrations	.86	.05

Thus, the analyses in this article will examine the relationship between banking crises and two types of domestic agitation, "collective protest" and "rebellion/internal war." The two indicators of domestic agitation used as dependent variables in this study were created by using principal components analysis to create factor scores from the raw event-count data. (Principal components analysis produces a variable that lends itself to regression analysis. By transforming a "given set of observed variables into another set of variables," principal components analysis helps achieve "economy of representation" with the objective of accounting for as much variance in the data as possible. For a discussion on how this economy is achieved, see Kim and Mueller (1978).[78] Two scores were generated for each country-year, one representing the level of collective protest activities, and one indicating the level of rebellion/internal war activities. For use in our analyses, we log the factor scores for both domestic agitation indicators.

Bank Crises

The IMF and the World Bank have developed an ordinal indicator measuring the level of banking system crises in IMF member-states, 1980–2000.[79,80] This measure uses solvency as a proxy for the degree to which a banking system is sound (or lacking crisis). The indicator is coded as indicated in Table 2.2.A-2 below:

Table 2.2.A-2. Ordinal Indicators for Banking System Crises in IMF Member-states.

Score	***Indication***
(0)	No crisis
(1)	Significant0non-systemic problem
(2)	Crisis/systemic

Lindgren et al. (1996: 20)[81] explain that, following the lead of Sundararajan and Balino (1991),[82] cases "where there were runs or other substantial portfolio shifts, collapses of financial firms, or massive government interventions [are referred to as] crises. Extensive unsoundness short of a crisis is termed significant." While there is some deal of subjectivity in this measure, it is the only one that we have found to exist, and to its credit, the country-specific case evidence provided along with these rankings is considerable.

The costs of banking crises vary considerably and are not restricted to any particular geographic region. For example, in 1997 Nigeria experienced financial distress that accounted for about 4 percent of banking systems assets.[83] However, between 1980 and 1982, Argentina experienced a catastrophic systemic crisis in the financial sector. In fact, around 170 financial institutions experienced central bank intervention and/or were dissolved. In total, the financial cost to Argentina amounted to about 55 percent of its gross domestic product.[84,85] Countries that experience crisis/systemic financial problems are experiencing capital exhaustion. Countries that are facing significant0non-systemic problems continue to retain a majority of banks that are solvent.

We use this ordinal measure as our indicator of a country's level of banking crisis. The data are taken from the International Monetary Fund's *Bank Soundness*[86] and the World Bank's *Episodes of Systemic and Borderline Financial Crise*s.[87]

Control Variables. We include indicators in our analyses that account for some of the competing hypotheses regarding the causes of domestic agitation.

Democracy

The first alternative hypothesis is that regime type may affect citizen propensity towards agitated behaviour. We would expect a different relationship between the extent to which a regime is democratized and both of our indicators of domestic agitation. Most empirical studies that examine the relationship between democracy and political agitation or repression assume a linear association. Yet many scholars argue that a curvilinear relationship exists between the two variables. Tilly (1978) asserts that we would expect democracies to be more able and willing to be facilitative to powerful contenders engaging in larger-scale activities, possibly short-circuiting revolution (and a necessarily linear relationship between democracy and rebellion/internal war actions).[88] Also, newly established democracies face a myriad of obstacles to overcome. These include the formation of new political institutions, the elimination of state restrictions on individual behaviour and social inhibitions, and the perplexity regarding standards of morals.[89] The initiation of free elections also requires politicians to compete for votes. Huntington (1997) argues that it is relatively easy for inexperienced politicians to solicit public support by appealing to ethnic, religious and moral differences, thereby exacerbating social, ethnic and religious conflict.[90] (The former Soviet Union and Yugoslavian Republic provide examples of this problem.)

Furthermore, newly formed democracies are responsible for solving the problems inherited from their predecessors. These often include economic stagnation, corruption, an inept bureaucracy and public mistrust.[91,92,93,94] If the new democratic government fails to successfully address these problems, public anxiety and unrest are likely to erupt. Only when democracy is fully consolidated (that is, it becomes the "only game in town") does it play a wholly stabilizing role in the domestic political environment.[95,96,97] At this stage, individuals and political organizations rely on the formal rules of conduct and sanctions to vie for political power, and no major political group seriously attempts to overthrow the democratic regime.[98,99] Thus, we might reasonably expect to see a nonlinear relationship between democracy and domestic agitation that takes an upside-down "U" shape. To test for this non-linear relationship, we express democracy as a quadratic polynomial.

Our measure of democracy is a 21-point ordinal regime-type indicator from the *Polity IV: Political Regime Characteristics and Transitions, 1800–2002* dataset.[100] This measure is constructed by subtracting the ordinal Polity measure of a country's level of institutionalized autocracy from the ordinal measure of a country's level of institutionalized democracy. This results in an ordinal variable ranging from –10 (strongly autocratic) to +10 (strongly democratic). Since we needed to create a polynomial for use in our analyses, it was necessary to modify the original scale of the original regime-type variable. The fact that the variable had both negative and positive values created a potential problem for substantive interpretation after squaring scale values. Thus, we shifted the original scale to 0 through 20 before creating the polynomial used in our analyses. We expect the linear democracy indicator to manifest a positive relationship with domestic agitation, and the polynomial indicator to manifest a negative relationship with domestic agitation.

Level of Economic Development

We include an indicator to control for the alternative hypothesis that a country's level of economic development may influence the propensity of its citizenry to engage in domestic agitation. As previously discussed, countries with stronger economies are more likely to be able

to afford and withstand a banking crisis. Furthermore, Gurr (1994: 359) points out that at least in terms of ethno-political conflict (such as in our Indonesian or Malaysian examples), it is easy to specify why poverty may cause more frequent and more severe conflict, as "systemic poverty means limited state capacity: substantial concessions to communal contenders therefore are prohibitively costly ... and conflicts over power and material issues tend to be seen by all contenders in zero-sum terms."[101] Lichbach (1989: 464) notes that while "some dimensions of economic inequality might be related to some dimensions of conflict under some conditions ... no empirical work has so far established this claim."[102] Thus, this indicator is of interest.

Like Cingranelli and Richards (1999), Mitchell and McCormick (1988), Poe et al. (1999) and Richards et al. (2001), we use the logged per capita value of a state's gross domestic product (GDP) (purchasing power parity) as our indicator of economic development.[103,104,105,106] The fact that, by definition, GDP per capita does not take into account distributional issues has not gone unnoticed here. However, statistical examinations employing other measures, such as the Physical Quality of Life Index (PQLI) and the Human Development Index (HDI), do not produce results different from those produced using GDP per capita in constant dollars. Thus, we use the logged value of GDP per capita (purchasing power parity), as it is a more widely (spatially and temporally) and readily available indicator than either PQLI or HDI. Data were gathered from the World Bank's *World Development Indicators 2004* dataset.[107] We would expect the level of domestic agitation to rise as the level of economic development falls.

Repression

Our next alternative hypothesis is that the level of government repression in a society may be related to the propensity of citizens to engage in domestic agitation. That is, does repression lead to turmoil? Jackson et al. (1978)[108] and Lichbach and Gurr (1981)[109] have both posited that increased repression leads to increased rebellion. Indeed, in a longitudinal analysis of Chile, Davis and Ward (1990: 468) find evidence that "rebellion is (primarily) a response to government repressiveness in the form of deaths from political violence."[110]

While many agree that some relationship exists between repression and turmoil, there is some debate surrounding the nature of this relationship. Many have specified this relationship as linear, while others have made the case that the relationship is actually more complex than that. Gartner and Regan (1996: 285) note the following:

> *One widely held view [of the interaction between opposition groups and a government] posits that as government coercion increases political dissent will increase accordingly, though at some point, where government coercion becomes particularly oppressive, the marginal effect of the continued coercion will generate a decrease in overt political dissent.*[111]

This non-linear approach suggests that an upside-down "U"-shaped relationship exists between repression and domestic agitation such that repression, up to some point, will produce greater turmoil until this coercion becomes so overwhelming that it actually decreases levels of dissent. (Gartner and Regan [1996] do not test this exact relationship; rather, they are concerned with levels of regime coercion relative to opposition demand levels, given international and domestic costs associated with this ratio. They do, however, make a strong general case that the relationship between coercion and protest is in some way nonlinear.)[112]

This approach assumes that repression works. That is, if a regime bears down hard enough on an opposition by, say, eliminating its leaders and jailing its rank-and-file, agitated opposition will, at some point, deteriorate. A tabular analysis of our cross-national data strongly indicates that such a nonlinear relationship is a possibility, and certainly, many states with high equilibrium levels of repression have very low levels of domestic agitation. As Gartner and Regan state, a linear relationship between coercion and repression would lead to an ever-increasing game of "tit-for-tat" culminating in civil war.[113]

To control for this nonlinear alternative hypothesis, we include an indicator of repression from Arthur S. Banks' *Cross-National Time-Series Domestic Conflict Data* that counts the number of "[s]ystematic elimination[s] by jailing or execution of political opposition within the ranks of the regime or the opposition."[114] In our models, we include a quadratic polynomial to account for nonlinearity. (In our analyses, we replaced Banks' measure with the CIRI physical integrity rights index [Cingranelli and Richards, 2005][115] as a simple sensitivity analysis. Our results were substantively unchanged.) We expect the linear repression indicator to manifest a positive relationship with domestic agitation, and the polynomial indicator to manifest a negative relationship with domestic agitation.

Discrimination

Our final alternative hypothesis posits that high degrees of economic and political discrimination may lead to turmoil. This hypothesis is based in the framework of relative deprivation.[116,117,118,119,120] The basic logic of relative deprivation is straightforward. It assumes people compare their condition, either as individuals or members of a group (or both), with the conditions of others. To the extent that one's own condition, or that of one's group, is inferior or degraded relative to that of others, one is likely to become dissatisfied/frustrated. Incorporating the frustration-aggression model of behaviour, relative deprivation theory asserts that the more frustrated one becomes at this perceived lack of status relative to others, the more likely one is to become aggressive or to engage in violence. There is a strong parallel to the hypothesis that as economic inequality grows in society, so does the risk of violent behavior.[121,122,123,124]

We focus on what is known as fraternal relative deprivation, whereby group status is the mode of comparison.[125,126]. That is, to the point one's group (however self-identified) is seen as degraded relative to other groups, the more potential for dissatisfaction and perhaps, ultimately, violence. We work from the assumption that individuals as members of groups are cognizant of their status relative to other groups in society. It follows, importantly, that they are then aware of discrimination against them (fraternal relative deprivation) and that this discrimination can lead to violent behaviour as a result.

The results of empirical tests of both relative deprivation and economic inequality as wellsprings of violence have been famously mixed.[127,128] Canache (1996) notes, however, that this is largely the result of flawed research designs and data not suited for these tests.[129] Our data on discrimination against groups in society come from the Minorities at Risk (MAR) Project.[130,131] (We also tried using a country's GINI coefficient score as a measure of economic inequality and/or relative deprivation,[132,133,134] but because of the cross-regional nature of our study, even the most comprehensive source of GINI data [Version 2.0a of the UNU/WIDER World Income Inequality Database {WIID}] significantly reduced our N for analysis. Indicators from the World Bank's World Development Indicators database of percentage of population at the top and bottom quintiles of income were subject to even greater missing data problems.[135] Our analyses

conducted using GINI had an overall N of 625, and while there was no substantive effect on our relationship of chief theoretical interest [banking crises and domestic turmoil]. The GINI measure itself was statistically insignificant.) The MAR Project defines a "minority at risk" as "an ethnopolitical group (non-state communal group) that: collectively suffers, or benefits from, systematic discriminatory treatment vis-à-vis other groups in a society; and0or collectively mobilizes in defense or promotion of its self-defined interests."[136] Our indicator of fraternal relative deprivation additively combines the MAR indices of both a country's mean level of political discrimination and economic discrimination against groups. The MAR indicators describe "the role of public policy and social practice in maintaining or redressing political inequalities."[137] Our indicator is continuous and ranges from zero (no such discrimination) to eight (a high level of discrimination). We would expect domestic agitation to increase along with increased discrimination.

Estimation Technique

Because our data are in pooled cross-section time-series form (with significantly more spatial units than temporal units), we employ the generalized estimation equation (GEE) estimation technique with robust standard errors. The GEE approach was developed to extend "generalized linear models (GLMs) to a regression setting with correlated observations within subjects," making it very attractive for use with panel data.[138] When "the primary question of interest is ... comparison across groups or subpopulations," GEE is a more appropriate estimation technique than "conditional" models (e.g., fixed-effects).[139] Further, GEE allows researchers to specify correlation patterns within clusters. We specify an exchangeable correlation structure, where covariance is assumed to be equal within any given individual cluster. This correlation structure corresponds to a random-effects model.

Findings

In the majority of cases in our data (approximately 61percent), no banking system problem was present. Approximately 16 percent of our cases were coded as having significant banking problems, while approximately 22 percent of the cases were identified as having full-blown banking crises. Thus, roughly 38 percent of our cases had some banking problem. It was mentioned earlier that the creators of the banking crisis indicator acknowledge that there is some degree of subjectivity involved in differentiating between a "significant banking problem" and a "banking crisis." Were these categories independent for the purposes of our analysis, we might expect to see a statistically significant difference in the mean level of domestic agitation. That is, we would see a statistically significant different mean level of agitation between "no banking crisis," "significant problem" and "banking crisis" cases. Using a difference-of means test we find that in the collective protest context, the "significant problem" and "banking crises" categories are not statistically independent, and that in the rebellion/internal war context, they are barely independent.

This might indicate that these data really represent only two categories of information, "no banking crisis" or "banking crisis," and not three. To test this possibility, we created a dichotomous version of the banking crisis variable, leaving "no banking crisis" as one category, and combining the other two categories into a single "banking crisis" category. However, when employed in the regression analyses reported below, this indicator returned substantively identical results compared to the original ordinal indicator.

Table 2.2.A-3 presents the results of two regressions, examining the contemporaneous relationship between a country's level of banking crisis and its level of two types of domestic agitation. Chi-squared tests show that both models are statistically significant from their null counterparts at the .00 level of significance. We see that a country's level of banking crisis is a statistically significant predictor of its level of collective protest activities, but not its rebellion/internal war activities. We find a weak positive linear relationship between collective protest activities and democracy, although there is no evidence for a nonlinear relationship. No statistically reliable relationship is found between democracy and rebellion/internal war.

Table 2.2.A-3. GEE Estimates of the Contemporaneous Relationship between Level of Banking Crisis and Two Types of Domestic Agitation.

Level of Banking Crisis	***Type of Domestic Agitation***	
	Collective Protest	***Rebellion/Internal War***
Level of Banking Crisis	.10**	-.00
	(.04)	(.03)
Level of Democracy	.05**	.04
	(.02)	(.04)
Level of Democracy Squared	-.00	-.00
	(.00)	(.00)
Level of Economic Development	.04	-.18**
	(.04)	(.05)
Repression	.90**	1.13**
	(.41)	(.39)
Repression Squared	-.31*	-.36**
	(.16)	(.16)
Discrimination	.06**	.08**
	(.02)	(.02)
Constant	-.82**	1.04**
	(.36)	(.43)
N	1525	1521
Prob > Chi-squared	.000	.000

Figures in parentheses are heteroskedasticity-corrected standard errors.

*$p \leq .10$ **$p \leq .05$

In line with our a *priori* expectations, we found a negative relationship between level of economic development and context of rebellion/internal war activities. That is, from these results, we would expect levels of rebellion/internal war activities to fall when a country's level of economic development rises, and vice versa. No statistically reliable relationship was found between level of economic development and collective protest. In both models, we found evidence of a nonlinear relationship between domestic agitation activities and repression such that domestic agitation seems to be greatest at middle values of repression.

Finally, Table 2.2.A-3 demonstrates a statistically reliable and positive relationship between discrimination and both types of domestic agitation activities. That is, increases in political and economic discrimination against minorities at risk should be expected to be accompanied by increased levels of collective protest and rebellion/internal war activities. The relationship is slightly stronger in the case of rebellion/internal war activities.

The models used for the analyses in Table 2.2.A-4 are similar to those used in the previous analysis, except that they take into account the fact that the effects of banking crises may extend to a following year. Thus, along with our indicator of the level of banking crisis at time (t), we also employ an indicator of the level of banking crisis at time (t-1). The Kendall's Tau estimation of the association between level of banking crisis at time (t) and time (t-1) is 0.77, indicating less-than-perfect correlation and thus, the presence of some variation between the two time points. (See Friedrich [1982][140] or Jaccard, Turrissi and Wan [1990][141] for a discussion on the effects of less-than-perfect multicollinearity among indicators in a regression. Our assumption is that less-than-perfect correlation is not a problem.) As in Table 2.2.A-3, chi-squared tests show that both models are statistically significant from their null counterparts at the .00 level of significance.

In Table 2.2.A-4, we first see that a country's level of banking crisis at time (t) is only reliably associated with domestic agitation in the context of collective protest activities. In neither context of domestic agitation, however, is a state's level of banking crisis at time (t-1) a significant indicator of domestic agitation at time (t). In terms of significance and direction of relationship, other variables perform as in the non-lagged models.

Looking at our raw data, we find different combinations of collective protest activities to be associated with similar levels of changes in the level of banking crises, although continued/increased demonstrations seem to be a mainstay. For example, Senegal in 1982 to 1983 went from no crisis to severe crisis and saw an increase in demonstrations. Given the same increase in the level of banking crisis, Bangladesh in 1987 to 1988 saw increases in riots and demonstrations and a continued level of strikes. Lebanon in 1987 to 1988 saw an increase in riots. Given the establishment here of a banking crisis-collective protest relationship at the macro level of analysis, future studies, perhaps region-specific ones, would do well to examine collective protest activities in a disaggregated fashion.

Table 2.2.A-4. GEE Estimates of the Lagged (t-1) Relationship between Level of Banking Crisis and Two Types of Domestic Agitation

Level of Banking Crisis	**Type of Domestic Agitation**	
	Collective Protest	**Rebellion/Internal War**
Level of Banking Crisis	.10**	-.03
	(.04)	(.03)
Banking Crisis (T-1)	.01	.04
	(.03)	(.03)
Level of Democracy	.04*	.04
	(.02)	(.04)
Level of Democracy Squared	-.00	-.00
	(.00)	(.00)
Level of Economic Development	.03	-.18**
	(.04)	(.05)
Repression	-.90**	-1.14**
	(.41)	(.39)
Repression Squared	-.31*	-.36**
	(.16)	(.16)
Discrimination	.06**	.08**
	(.02)	(.02)
Constant	-.77**	1.04**

Table 2.2.A-4. GEE Estimates of the Lagged (t-1) Relationship between Level of Banking Crisis and Two Types of Domestic Agitation

Level of Banking Crisis	**Type of Domestic Agitation**	
	Collective Protest	**Rebellion/Internal War**
	(.35)	(.43)
N	1514	1510
Prob > Chi-squared	.000	.000

Figures in parentheses are heteroskedasticity-corrected standard errors.
$*p \leq .10$ *$**p \leq .05$*

Conclusion

The findings of this study have important implications for understanding the puzzle of domestic political stability. We posited that banking crises appear to be a sufficient factor in causing domestic agitation, as citizens protest against both the direct negative effects of a banking crisis itself (such as the loss of access to savings and credit), and against a government's handling of an economy adversely affected by a banking crisis. These protests are often manifested as food riots, labour strikes, and anti-austerity demonstrations.

Using a cross-national data set comprised of 125 countries for the years 1981 through 2000, we tested our hypothesis that banking crises are a significant factor in fomenting domestic agitation. We found a contemporaneous relationship between a country's level of banking crisis and its level of collective protest activities. The relationship was such that we would reliably expect to see collective protest in countries experiencing a banking crisis (the worse the crisis, the more domestic agitation). This finding is supported by our case examples, where we almost exclusively saw collective protest activities, such as strikes, anti-government demonstrations and riots, directly resulting from the effects of banking crises on citizens. Thus, the impact of a banking crisis on political stability was found to be immediate and damaging. The silver lining here may be that such ill effects appear to be relatively short lived. This is supported by the finding that the lagged banking crisis variable demonstrated no significant association with collective protest or rebellion/internal war.

Although there was a considerable amount of rebellion/internal war activity in our sample, our analyses seem to indicate that the effects of banking crises on types of collective action are bounded. On average, the effect of a banking crisis appears to stop short of being related to widely participatory and organized violence against a government. Instead, banking crises seem to be reliably associated with immediate and intense protests by adversely affected individuals.

Certainly no less interesting was the support we found for a relationship between fraternal relative deprivation and amounts of both types of domestic agitation such that greater relative deprivation is associated with greater domestic agitation. In addition, lending support to many who have so surmised, our findings suggest that nascent democracy is a seemingly accommodating context for collective protest activities. We found a negative relationship between level of economic development and rebellion/internal war activity. This relationship was expected, as economic growth is expected to alleviate the plight of poverty and positions countries to better withstand the cost of banking crises, thereby lowering the likelihood of citizens participating in rebellion/internal war.

We found some support for those who posit nonlinear relationships between domestic agitation and repression. In particular, we found possible empirical support for Gartner and Regan's (1996) supposition that while low-to-moderate levels of repression may spur domestic agitation; agitation will actually decline in the face of higher levels of repression.[142] This is an interesting finding and certainly deserves future consideration.

Our findings should be of some interest to those seeking to identify the causes of banking crises themselves. Economists have yet to generate any uniform agreement regarding the specific causes of banking crises. While there is agreement that a particular set of conditions may be present for one particular crisis, it is by no means accepted that these conditions will be present at other financial crises. Certainly, the discovery of a reliable association between banking crises and collective action activities makes discovering the causes of banking crises all the more important.

To summarize, we found a country's level of banking crisis to be reliably associated with its level of collective protest activities. Given the frequency of banking crises and the fragility and importance of the banking sector, this is a significant finding. A government experiencing a banking crisis faces not only certain economic costs, but also should probably expect to face social disruptions such as strikes, riots and antigovernment demonstrations. Moreover, our findings add to the ongoing discussion regarding the economic globalization process. The results of the study clearly suggest that banking crises impose economic and political costs on domestic governments. To reestablish a stable political environment, including the settlement of strikes and anti-government demonstrations, governments may be required to limit participation in international markets. While firm agreement on this will require further study, it is clear that the interconnection between economic and political frontiers adds greater pressure to better understand the determinants and consequences of banking crises; not only to avoid economic distress, but also to avoid the pain of social agitation.

↑ Back to the top

---

[1] Enoch, Charles. 2000. "Interventions in Banks During Banking Crises: The Experience of Indonesia." *IMF Policy Discussion Paper 00/2*. Washington, DC: International Monetary Fund.

[2] Frydl, Edward J. and Marc Quintyn. 2000. "The Benefits and Costs of Intervening in Banking Crises." *IMF Working Paper 00/147*. Washington, DC: The International Monetary Fund.

[3] Garcia, Gillian. 1997. "Protecting Bank Deposits." *Economic Issues*. Washington, DC: International Monetary Fund.

[4] Kaufman, George G. and Steven A. Seelig. 2001. "Post Resolution Treatment of Depositors at Failed Banks: Implications for the Severity of Banking Crises, Systemic Risk, and Too-Big-To-Fail." *IMF Working Paper 01/83*.Washington, DC: International Monetary Fund.

[5] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.

[6] Mahon, James E. and Javier Corrales. 2002. "Pegged for Failure? Argentina's Crisis." *Current History* 101: 72–75.

[7] Carstens, Agustin, Daniel Hardy and Ceyla Pazarbasioglu. 2004. "Avoiding Banking Crises in Latin America." *Financial Development* 41: 30–3.

[8] Hardy, Daniel C. 1998. "Are Banking Crises Predictable?" *Finance and Development* 35 (December): 38–41.

[9] Stiglitz, Joseph E. 2003. *Globalization and its Discontents*. New York: W.W. Norton & Company, Inc.

[10] Rennhack, Robert. 2000. "Banking Supervision." *Finance and Development* 37: 9–15.

[11] Kaufman, 2000.
[12] Caprio, Gerard and Daniela Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." http://econ.worldbank.org/view.php?id=23456 (February 20, 2006).
[13] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[14] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[15] Stiglitz, Joseph E. 2003. *Globalization and its Discontents*. New York: W.W. Norton & Company, Inc.
[16] *The Economist*. 1997. "The end of the miracle?" November 29, 21–23.
[17] Eichengreen, Barry. 1997. "The Tyranny of the Financial Markets." *Current History* 96: 377–82.
[18] Hardy, 1999.
[19] Knight, Malcolm. 1999. "Developing and Transition Countries Confront Financial Globalization." *Finance and Development* 36: 32–5.
[20] Kaufman, George. 1996. "Banking Failures, Systemic Risk, and Bank Regulation." *Cato Journal* 16: 17–45.
[21] Hardy, 1999.
[22] Garcia, Gillian and Carl-Hohan Lindgren. 1998. "Averting the Crash: Economic Institutions and Financial Sector Soundness." *SIAS Review* 33: 20–32.
[23] Singh, Ajit and Bruce A.Weisse. 1998. "Emerging Stock Markets, Portfolio Capital Flows and Long-term Economic Growth: Micro and Macroeconomic Perspectives." *World Development* 26: 607–22.
[24] Hardy, Daniel C. 1998. "Are Banking Crises Predictable?" *Finance and Development* 35 (December): 38–41.
[25] Enoch, Charles. 2000. "Interventions in Banks during Banking Crises: The Experience of Indonesia." *IMF Policy Discussion Paper 00/2*. Washington, DC: International Monetary Fund.
[26] Frydl, Edward J. and Marc Quintyn. 2000. "The Benefits and Costs of Intervening in Banking Crises." *IMF Working Paper 00/147*. Washington, DC: The International Monetary Fund.
[27] Martinez, Guillermo Ortiz. 1998. "What Lessons Does the Mexican Crisis Hold for Recovery in Asia?" *Finance and Development* 35 (June): 6–9.
[28] Valencia, Matthew. 1997. "Fragile, handle with care." *The Economist*, April 10, 1–40.
[29] Carstens, Agustin, Daniel Hardy and Ceyla Pazarbasioglu. 2004. "Avoiding Banking Crises in Latin America." *Financial Development* 41: 30–3.
[30] Ouattara, Alassane. 1998. "Globalization, Lessons from the Asian Crisis and Central Bank Policies." Paper presented at the Reunion des Gouverneurs des Banques Centrales des Pays, Ottawa.
[31] Schwartz, Anna J. 1998. "International Financial Crises: Myths and Realities." *The Cato Journal* 17: 251–56.
[32] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[33] Lindgren, Carl-John, Gillian Garcia and Matthew I. Saal. 1999*. Financial Sector Crisis and Restructuring: Lessons from Asia*.Washington, DC: International Monetary Fund.
[34] Hanlon, Joseph. 2001. "Corruption Busting Bank Chairman Assassinated." *Southern African Regional Poverty Network.* http://www.sarpn.org.za/countryPovertyPapers/Mozambique/Assassin/index.php (April 1, 2006).
[35] Valencia, Matthew. 1997. "Fragile, handle with care." *The Economist*, April 10, 1–40.
[36] Valencia, Matthew. 1997. "Fragile, handle with care." *The Economist*, April 10, 1–40.
[37] Madura, Jeff. 1998. *Financial Markets and Institutions*. Cincinnati: South-Western College Publishing.
[38] Madura, Jeff. 1998. *Financial Markets and Institutions*. Cincinnati: South-Western College Publishing.

[39] Carstens, Agustin, Daniel Hardy and Ceyla Pazarbasioglu. 2004. "Avoiding Banking Crises in Latin America." *Financial Development* 41: 30–3.

[40] Mydans, Seth. 1998. "Newly Jobless—and Newly Angry—Threaten Indonesia's Stability." The *New York Times*, January 29, A3.

[41] Enoch, Charles. 2000. "Interventions in Banks During Banking Crises: The Experience of Indonesia." *IMF Policy Discussion Paper 00/2*. Washington, DC: International Monetary Fund.

[42] Kaufman, George G. and Steven A. Seelig. 2001. "Post Resolution Treatment of Depositors at Failed Banks: Implications for the Severity of Banking Crises, Systemic Risk, and Too-Big-To-Fail." *IMF Working Paper 01/83*.Washington, DC: International Monetary Fund.

[43] Garcia, Gillian. 1999. "Deposit Insurance: A Survey of Actual and Best Practices." *IMF Working Paper 99/54*. Washington, DC: International Monetary Fund.

[44] *The Economist*. 2002. "Between the creditors and the streets." January 5, 29–30.

[45] Keesings. 2001. *Keesings Record of World Events*. Bethesda, MD: Keesings Worldwide LLC.

[46] Pearlstein, Steven. 2001. "Argentina's Financial Quandary Is No Surprise." *Washington Post*, December 21, A38.

[47] Keesings. 2000. *Keesings Record of World Events*. Bethesda, MD: Keesings Worldwide LLC.

[48] *The Economist*. 1998. "The last bailout of Credit Lyonnais." May 23, 67.

[49] Enoch, Charles. 2000. "Interventions in Banks During Banking Crises: The Experience of Indonesia." *IMF Policy Discussion Paper 00/2*. Washington, DC: International Monetary Fund.

[50] Uchitelle, Louis. 1998. "Economists Blame Short-Term Loans for Asian Crisis." The *New York Times*, January 8, D4.

[51] Eichengreen, Barry. 1997. "The Tyranny of the Financial Markets." *Current History* 96: 377–82.

[52] Valencia, Matthew. 1997. "Fragile, handle with care." *The Economist*, April 10, 1–40.

[53] Keesings. 1995. *Keesings Record of World Events*. Bethesda, MD: Keesings Worldwide LLC.

[54] Davenport, Christian. 1995. "Multi-Dimensional Threat Perception and State Repression: An Inquiry into Why States Apply Negative Sanctions." *American Journal of Political Science* 39: 683–713.

[55] Davenport, Christian. 1996. "The Weight of the Past: Exploring Lagged Determinants of Political Repression." *Political Research Quarterly* 49: 377–403.

[56] Davenport, Christian. 1998. "Liberalizing Event or Lethal Episode?: An Empirical Assessment of How National Elections Affect the Suppression of Political and Civil Liberties." *Social Science Quarterly* 79: 321–41.

[57] Francisco, Ronald A. 1996. "Coercion and Protest: An Empirical Test in Two Democratic States." *American Journal of Political Science* 40: 1179–1204.

[58] King, John C. 1998. "Repression, Domestic Threat, and Interactions in Argentina and Chile." *Journal of Political and Military Sociology* 26: 191–211.

[59] Ziegenghagen, Eduard. 1986. *The Regulation of Political Conflict*. New York: Praeger.

[60] Cingranelli, David L. and David L. Richards. 1999. "Respect for Human Rights after the End of the Cold War." *Journal of Peace Research* 36: 511–34.

[61] Mitchell, Neil J. and James M. McCormick. 1988. "Economic and Political Explanations of Human Rights Violations." *World Politics* 40: 476–98.

[62] Poe, Steven C., C. Neal Tate and Linda Camp-Keith. 1999. "Repression of the Human Right to Personal Integrity Revisited: A Global Cross-national study Covering The Years 1976–1993." *International Studies Quarterly* 43: 291–313.

[63] Richards, David L., Ronald D. Gelleny and David H. Sacko. 2001. "Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries." *International Studies Quarterly* 45: 219–239.

[64] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.

[65] Hibbs, Douglas A. 1973. *Mass Political Violence: A Cross-National Causal Analysis*. New York: John Wiley.

[66] Lichbach, Mark Irving and Ted Robert Gurr. 1981. "The Conflict Process: A Formal Model." *Journal of Conflict Resolution* 25: 3–29.
[67] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
[68] Hibbs, Douglas A. 1973. *Mass Political Violence: A Cross-National Causal Analysis*. New York: John Wiley.
[69] Lichbach, Mark Irving and Ted Robert Gurr. 1981. "The Conflict Process: A Formal Model." *Journal of Conflict Resolution* 25: 3–29.
[70] Muller, Edward N. and Mitchell A. Seligson. 1987. "Inequality and Insurgency." *American Political Science Review* 81: 425–52.
[71] Banks, Arthur S. 2003. *Cross-National Time-Series Data Archive*. http://www.databanks.sitehosting.net (April 1, 2006).
[72] Hibbs, Douglas A. 1973. *Mass Political Violence: A Cross-National Causal Analysis*. New York: John Wiley.
[73] Lichbach, Mark Irving and Ted Robert Gurr. 1981. "The Conflict Process: A Formal Model." *Journal of Conflict Resolution* 25: 3–29.
[74] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
[75] Hibbs, Douglas A. 1973. *Mass Political Violence: A Cross-National Causal Analysis*. New York: John Wiley.
[76] Lichbach, Mark Irving and Ted Robert Gurr. 1981. "The Conflict Process: A Formal Model." *Journal of Conflict Resolution* 25: 3–29.
[77] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
[78] Kim, Jae-On and CharlesW. Mueller. 1978. *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills: Sage Publications.
[79] Caprio, Gerard and Daniela Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." http://econ.worldbank.org/view.php?id=23456 (February 20, 2006).
[80] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[81] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[82] Sundararajan, V. and Tomas J.T. Balino. 1991. "Issues in Recent Banking Crises." In *Banking Crises: Cases and Issues*, eds. V. Sundararajan and Tomas J.T. Balino. Washington, DC: International Monetary Fund.
[83] Caprio, Gerard and Daniela Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." http://econ.worldbank.org/view.php?id=23456 (February 20, 2006).
[84] Caprio, Gerard and Daniela Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." http://econ.worldbank.org/view.php?id=23456 (February 20, 2006).
[85] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[86] Lindgren, Carl-Johan, Gillian Garcia and Matthew I. Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, DC: International Monetary Fund.
[87] Caprio, Gerard and Daniela Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." http://econ.worldbank.org/view.php?id=23456 (February 20, 2006).
[88] Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, MA: Addison-Wesley.
[89] Huntington, Samuel P. 1997. "Democracy for the Long Haul." In *Consolidating the Third Wave Democracies*, ed. Larry Diamond. Baltimore: The Johns Hopkins University Press.
[90] Huntington, Samuel P. 1997. "Democracy for the Long Haul." In *Consolidating the Third Wave Democracies*, ed. Larry Diamond. Baltimore: The Johns Hopkins University Press.
[91] Narayan, Deepa. 2000. *Voices of the Poor: Can Anyone Hear Us?* Washington, DC: The World Bank.

[92] Rose, Richard, Doh C. Shin and Neil Munro. 1999. "Tensions Between the Democratic Ideal and Reality: South Korea." In *Critical Citizens: Global Support for Democratic Governance*, ed. Pippa Norris. Oxford: Oxford University Press.
[93] Snow, Donald M. 1997. *Distant Thunder*. London: M.E. Sharpe, Inc.
[94] Wiarda, Howard J. 2004. *Political Development in Emerging Nations: Is There Still a Third World?* Belmont: Thomson/Wadsworth.
[95] Diamond, Larry. 1999. *Developing Democracy: Toward Consolidation*. Baltimore: The Johns Hopkins University Press.
[96] Newton, Kenneth. 1999. "Social and Political Trust in Established Democracies." In *Critical Citizens: Global Support for Democratic Governance*, ed. Pippa Norris. Oxford: Oxford University Press.
[97] Snow, Donald M. 1997. *Distant Thunder*. London: M.E. Sharpe, Inc.
[98] Linz, Juan J. and Alfred Stepan. 1997. "Toward Consolidated Democracies." In *Consolidating the Third Wave Democracies,* ed. Larry Diamond. Baltimore: The Johns Hopkins University Press.
[99] Newton, Kenneth. 1999. "Social and Political Trust in Established Democracies." In *Critical Citizens: Global Support for Democratic Governance*, ed. Pippa Norris. Oxford: Oxford University Press.
[100] Marshall, Monty G. and Keith Jaggers. 2003. Polity IV: *Political Regime Characteristics and Transitions*, 1800–2002. http://www.cidcm.umd.edu/inscr/polity (April 1, 2006).
[101] Gurr, Ted Robert. 1994. "Peoples Against States: Ethnopolitical Conflict and the Changing World System." *International Studies Quarterly* 38: 347–77.
[102] Lichbach, Mark Irving. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict?' Studies." *World Politics* 41: 431–70.
[103] Cingranelli, David L. and David L. Richards. 1999. "Respect for Human Rights after the End of the Cold War." *Journal of Peace Research* 36: 511–34.
[104] Mitchell, Neil J. and James M. McCormick. 1988. "Economic and Political Explanations of Human Rights Violations." *World Politics* 40: 476–98.
[105] Poe, Steven C., C. Neal Tate and Linda Camp-Keith. 1999. "Repression of the Human Right to Personal Integrity Revisited: A Global Cross-national study Covering The Years 1976–1993." *International Studies Quarterly* 43: 291–313.
[106] Richards, David L., Ronald D. Gelleny and David H. Sacko. 2001. "Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries." *International Studies Quarterly* 45: 219–239.
[107] World Bank. 2004. *World Development Indicators on CD-ROM*. Washington, DC: World Bank.
[108] Jackson, Steven, Bruce Russett, Duncan Snidal and Donald Sylvan. 1978. "Conflict and Coercion in Dependent States." *Journal of Conflict Resolution* 22: 627–57.
[109] Lichbach, Mark Irving and Ted Robert Gurr. 1981. "The Conflict Process: A Formal Model." *Journal of Conflict Resolution* 25: 3–29.
[110] Davis, David R. and Michael D. Ward. 1990. "They Dance Alone: Deaths and the Disappeared in Contemporary Chile." *Journal of Conflict Resolution* 34: 449–75.
[111] Gartner, Scott Sigmund and Patrick M. Regan. 1996. "Threat and Repression: The Non- Linear Relationship Between Government and Opposition Violence." *Journal of Peace Research* 33: 273–87.
[112] Gartner, Scott Sigmund and Patrick M. Regan. 1996. "Threat and Repression: The Non- Linear Relationship Between Government and Opposition Violence." *Journal of Peace Research* 33: 273–87.
[113] Gartner, Scott Sigmund and Patrick M. Regan. 1996. "Threat and Repression: The Non- Linear Relationship Between Government and Opposition Violence." *Journal of Peace Research* 33: 273–87.
[114] Banks, Arthur S. 2003. *Cross-National Time-Series Data Archive*. http://www.databanks.sitehosting.net (April 1, 2006).
[115] Cingranelli, David L. and David L. Richards. 2005. *The CIRI Human Rights Dataset, Version 2005.10.12*. http://www.humanrightsdata.org (April 1, 2006).
[116] Crosby, Faye. 1976. "A Model of Egoistical Relative Deprivation." *Psychological Review* 83: 85–113.

[117] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
[118] Hibbs, Douglas. 1974. *Industrial Conflict in Advanced Industrial Societies*. Cambridge, MA: MIT Center for International Studies.
[119] Runciman, Walter Garrison. 1966. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth Century England*. Berkeley: University of California Press.
[120] Stouffer et al., 1949.
[121] Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
[122] Hey, John D. and Peter J. Lambert. 1980. "Relative Deprivation and the Gini Coefficient: Comment." *The Quarterly Journal of Economics* 95.3: 567–73.
[123] Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, MA: Addison-Wesley.
[124] Yitzhaki, Shlomo. 1979. "Relative Deprivation and the Gini Coefficient." *The Quarterly Journal of Economics* 93: 321–24.
[125] Crosby, Faye. 1979. "Relative Deprivation Revisited: A Response to Miller, Bolce, and Halligan." *American Political Science Review* 73: 103–12.
[126] Runciman, Walter Garrison. 1966. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth Century England*. Berkeley: University of California Press.
[127] Canache, Damarys. 1996. "Looking out My Back Door: The Neighborhood Context and Perceptions of Relative Deprivation." *Political Research Quarterly* 49: 547–571.
[128] Lichbach, Mark Irving. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict?' Studies." *World Politics* 41: 431–70.
[129] Canache, Damarys. 1996. "Looking out My Back Door: The Neighborhood Context and Perceptions of Relative Deprivation." *Political Research Quarterly* 49: 547–571.
[130] Davenport, Christian. 2003. "Minorities At Risk: Dataset Users Manual 030703." *The Minorities At Risk (MAR) Project*. Center for International Development and Conflict Management. University of Maryland, Baltimore. http://www.cidcm.umd.edu/inscr/mar/margene/MAR-codebook_040903.doc (April 1, 2006).
[131] Gurr, Ted Robert. 2000. *Peoples Versus States: Minorities at Risk in the New Century*. Washington, DC: United States Institute of Peace Press.
[132] Hey, John D. and Peter J. Lambert. 1980. "Relative Deprivation and the Gini Coefficient: Comment." *The Quarterly Journal of Economics* 95.3: 567–73.
[133] Lichbach, Mark Irving. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict?' Studies." *World Politics* 41: 431–70.
[134] Yitzhaki, Shlomo. 1979. "Relative Deprivation and the Gini Coefficient." *The Quarterly Journal of Economics* 93: 321–24.
[135] World Bank. 2004. *World Development Indicators on CD-ROM*. Washington, DC: World Bank.
[136] Davenport, Christian. 2003. "Minorities At Risk: Dataset Users Manual 030703." *The Minorities At Risk (MAR) Project*. Center for International Development and Conflict Management. University of Maryland, Baltimore. http://www.cidcm.umd.edu/inscr/mar/margene/MAR-codebook_040903.doc (April 1, 2006).
[137] Davenport, Christian. 2003. "Minorities At Risk: Dataset Users Manual 030703." *The Minorities At Risk (MAR) Project*. Center for International Development and Conflict Management. University of Maryland, Baltimore. http://www.cidcm.umd.edu/inscr/mar/margene/MAR-codebook_040903.doc (April 1, 2006).
[138] Horton, Nicholas J. and Stuart R. Lipsitz. 1999. "Review of Software to Fit Generalized Estimating Equation Regression Models." *The American Statistician* 53: 160–69.
[139] Zorn, Christopher J. W. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications." *American Journal of Political Science* 45: 470– 90.
[140] Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26: 797–833.

[141] Jaccard, James, Robert Turrisi and Choi K. Wan. 1990. *Interaction Effects in Multiple Regression*. Newbury Park, California: Sage Publications.

[142] Gartner, Scott Sigmund and Patrick M. Regan. 1996. "Threat and Repression: The Non-Linear Relationship Between Government and Opposition Violence." *Journal of Peace Research* 33: 273–87.

2.2.B. Economics and Political Instability within the Global Economic Crisis (Peter Steen)

Author and Organization: Peter Steen, Analytic Services, Inc., in support of the Deputy Directorate for Strategy and Policy, Joint Staff

Contact Information: steenpm@js.pentagon.mil; peter.steen@anser.org

Introduction

This paper provides an overview of work completed at the Joint Staff J-5, seeking to provide a strategic, national-level understanding on the interplay between economics and political instability within the context of the ongoing Global Economic Crisis. The paper begins with key definitions. From that baseline, a brief discussion of methods provides insight into how the Joint Staff arrived at a variety of conclusions. Conclusions are delineated based on completed analysis. The next section delineates a model for linking economics with national security. The paper ends with a brief discussion of future analysis.

Grounds for Analysis and Definitions

The Deputy Directorate for Strategy and Policy at the Joint Staff drafts key strategy and policy documents for the Chairman. In order to accomplish this, an encompassing and inclusive understanding of the context within which military strategy and policy operates must be maintained. As Lehman Brothers failed in late 2008, it became clear that the economic context within which the U.S. military operates was undergoing fundamental turmoil. What would that turmoil mean to the U.S. military? It was necessary to define the most important aspect of the problems before they could be investigated and their significance to the military could be determined.

Definitions

There were several key aspects of this effort: economics, political instability, U.S. national security, and the global economic crisis.

Economics encompasses the creation, exchange and non-authoritative distribution of value within a society. This definition includes the production and trade of goods (material or non-material) and services, as well as the distribution of income based, at least in part, on that production.

Global Economic Crisis is the drop in global production, trade, and financial flows, as well as the turmoil in financial markets that has been called the "Great Recession." The fall of Lehman Brothers has been pointed to as the culmination of behind-the-scenes trends that led to the drastic financial turmoil and economic negative growth characterizing the last quarter of 2008 and the first quarter of 2009.

Political instability describes unexpected change in the society's system of authoritative allocation of value. This general definition can be clarified through the identification of two types of political instability. Type 1 is unexpected changes in the political system that occur within and do not result in changes to the rule set of the political system. Type 2 is unexpected change that occurs outside of the rule set or legal framework, often this is extra-legal, usually, violent, political change.

U.S. national security is defined as accomplishing the goals articulated within the U.S. National Security Strategy.

Methods

The Deputy Directorate for Strategy and Policy within the Joint Staff has significant background knowledge and data in quantitative and qualitative research. The objective was to (1) gather information on what books/experts *said* about the driving question above, (2) empirically test these assertions, and (3) ground findings in empirical analysis.

A "mixed methods" research agenda was implemented for this effort. First, experts and key theoretical propositions within the relevant literature were identified. Second, statistical analyses were conducted, including the adoption of both hypothesis testing and exploratory data analysis approaches. Third, experts were interviewed. Fourth, further statistical analyses were conducted.

Economics and Political Instability

First, the Political Instability Task Force (PITF) provided key assistance. Goldstone et al (2005) measured political instability as one or more of one of the following instances of state failure: "Revolutionary Wars, Ethnic Wars, Adversary Regime Changes, and Genocides and Politicides."[1] The PITF tested hundreds of economic variables against this very strict definition of political instability and found "most economic, demographic, geographic, and political variables do *not* have consistent and statistically significant effects on the risk of instability onset." Infant mortality was found to be a predictor of the onset of political instability 2 years out. No other economic variable was found to be a useful predictor of political instability.

Second, we went "back to the drawing board." Strategy and Policy leadership changed its focus away from explaining the onset of instability toward uncovering the potential for political instability. Therefore, the research agenda changed to *index creation*. The agenda was simple: review the literature and identify candidate measures that could be used to identify the potential for political instability. This loose, exploratory data analysis agenda allowed for the identification of over 90 variables, over half of which were economic. These variables were then logistically regressed against the countries' PITF scores (0 or 1) four years into the future. Our goal was to have a useful tool for assessing the potential for political instability. This had to be balanced against practical, statistical considerations, all of which resulted in the 4-year time frame of this work. No economic measures were found to be significant in predicting the PITF score. The initial statistical efforts yielded a fundamental finding: there was no statistical support for the assertion that economic measures can forecast political instability four years from the date of measurement. (For example, no economic measure in 2001 could be found to forecast political instability in 2005). The statistical analysis revealed that the World Governance Indicators (WGIs) were predictors.[2] An index was constructed using the WGIs.

Third, a meeting was held with Dr. Jack Goldstone and others in an effort to get a richer understanding as to the nature of political instability.[3] These interviews resulted in a new conceptualization of political instability as described above: Type 1 (unexpected, peaceful change within the system), and Type 2 (unexpected and violent change to the system), as described above in the "Definitions" section. With that new conceptualization in mind, we returned to the statistics.

Fourth, the research approach turned back to statistical analysis. Could empirical support be found for the assertion that economic variables explain changes in the different types of political

instability? The dependent variable had been changed. Before, the PITF dichotomous measure offered no gradation in instability. In order to operationalize the new two-type concept of political instability discussed above, statistical analysis was modified to include three values for every country of the world. Countries could be highly unstable as measured by the PITF data, countries could be unstable as measured by an irregular regime change in the Archigos Dataset[4] or non-violent "campaign,"[5] or countries could be stable (by not being in either of the first two categories). Leveraging the previous work, because of time constraints and data availability, research focused on 11 possible economic predictors of political instability:

- Commercial openness
- GDP per capita
- Inflation
- Foreign direct investment
- Infant mortality
- Agricultural imports
- Agricultural exports
- Oil imports
- Oil exports
- Metal exports
- Diversification

Preliminary findings were that only infant mortality and commercial openness were significant predictors of the new operationalization of political instability four years out.

The four research steps involved a variety of qualitative and quantitative analyses. In this analysis, infant mortality, as well as commercial openness, can be used to forecast the likelihood of political instability four years from the date of measurement. As infant mortality improves, data indicate that the likelihood of political instability to declines. As commercial openness improves, data indicate that the likelihood of political instability declines. These findings were based on *correlation*; the analysis could be the basis for further study that could identify causation.

Economics and U.S. National Security

The current 2006 National Security Strategy (NSS)[6] states the following: "Promoting free and fair trade has long been a bedrock tenet of American foreign policy. Greater economic freedom is ultimately inseparable from political liberty." It calls for several initiatives relating to economics: (1) "opening markets and integrating developing countries," (2) "opening, integrating, and diversifying energy markets to ensure energy independence," (3) "Reforming the International Financial System to ensure Stability and Growth." Even before the economic crisis that became so evident in late 2008, U.S. national security included an economic aspect.

As the economic crisis developed, senior leaders stated the importance of the economic crisis in disrupting the context within which U.S. national security is executed.[7] How might the economic crisis actually affect U.S. national security, beyond impacting the above initiatives?

Various groups within and outside the Defense Department were convened in order to properly answer this question. The figure shown below was developed to delineate the ways in which economic instability could affect U.S. national security.

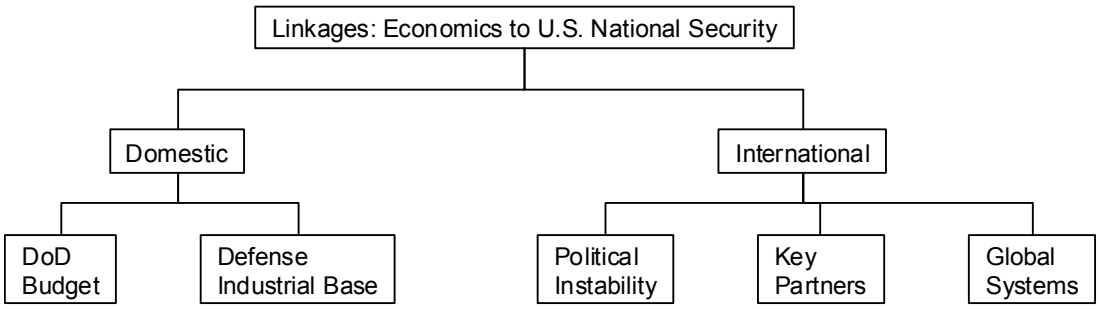


Figure 2.2.B-1. Economic Instability's Impact on U.S. National Security.

This model illustrates the relationships between economic and political instability. Economics can be traced through a variety of avenues of effect on U.S. national security. The political instability box depicts the proposition that political instability could lead to changes in countries that were friends and allies of the U.S. Such an unexpected change could hinder or even counter accomplishment of the NSS objectives. Conversely, political instability in a country that is currently hindering or countering accomplishment of NSS objectives could be a positive. Either way, political instability matters from a national security perspective. The left side of Figure 2.2.B-1 shows the domestic avenues of effect. Shrinking GDP indicates shrinking tax bases and increased difficulty in non-deficit financing of defense. As the Defense Industrial Base is integrated into the greater economy, and as the greater economy suffers, suppliers and other firms that are important players in defense may be hurt or even rendered insolvent due to the market shocks outside of defense.

The international avenue of effect has at least three lanes. Political instability has already been discussed. Changing economic fortunes can affect the wills and/or capabilities of key international partners to carry out operations in support of shared objectives. Finally (this is the least developed of the lanes) global systems were considered. Economics affect the systems within which international interactions occur. For example, economic strength can be articulated through institutions such as the International Monetary Fund (IMF). As economic strength is redistributed, demands grow to change the IMF's governance to reflect the new realities.

Future Analysis

This paper describes the ground covered by the Joint Staff J-5 Deputy Directorate for Strategy and Policy, as it explored the relationship between economics and political instability. The Joint Staff discovered a variety of players who provided key insight into the question. The most important player in this field is clearly the Political Instability Task Force (PITF), which provided very useful data and analyses for policy and strategy development that were key to our analysis.

The way forward is to move out on two tracks. First, different techniques must be examined and tested to uncover possible relationships between economics and political instability. This area of work encompasses two main efforts: research of the modeling of economic volatility (rather than static economic measures) and political instability, (2) research of more current measures of economics. Unfortunately, the standard measures of economy, such as GDP per capita, are released well after the year in question. Very often data are already two years' old when they are finalized (2006 data becoming available in 2008). This makes the data "vintage" and inhibits

forecasting. Investigations are being conducted to determine what economic data might be available that are less "vintage" to allow for better forecasting.

Second, the concept of political instability must be differentiated into a more nuanced variable showing "shades of gray." Dichotomous measures of political instability offer very useful insights into state failure, but they inhibit identification of political stability trends within the "stable" and "unstable" categories. Methods are being reviewed for index creation, and the underlying dynamics within this complex concept of political instability are being examined. As the conceptualization changes to reflect the requirements of Joint Staff strategy and policy development, the operationalization will inevitably alter, as well.

This paper reports on nearly a year of work at the Joint Staff, attempting to inform military strategy and policy development through systematic analysis of the relationship between economics and political instability. This first year was the first step in a long journey. Lessons and insights from the other authors in this volume will be key in keeping the Joint Staff informed as development and revision of national level strategy and policy continue.

↑ Back to the top

---

[1] Goldstone, Jack, Robert Bates, Ted Robert Gurr, Michael Lustik, Monty Marshall, Jay Ulfelder, and Mark Woodward. "A Global Forecasting Model of Political Instability." APSA Paper, 2005.

[2] http://info.worldbank.org/governance/wgi/pdf/wgidataset.xls.

[3] SDD interviewed over 20 experts. Particularly important were the 10 March 2009 interview of Jack Goldstone (George Mason), the 23 March 2009 interview of Joe Hewitt (U of MD), multiple thinkers at the Center for a New America Security 9 October 2008, and multiple sessions with the Defense Business Board.

[4] Gleditsch, Kristian. "Archigos: A New Database of Political Leaders." Paper presented at the annual meeting of the International Studies Association, Hilton Hawaiian Village, Honolulu, Hawaii, Mar 05, 2005. 2009-05-26 (http://www.allacademic.com/meta/p71997_index.html).

[5] Chenoweth, Erica and Maria Stephan. 2008. "Why Civil Resistance Works: the Strategic Logic of Nonviolent Conflict," *International Security* 33(1). (http://echenoweth.faculty.wesleyan.edu/research-and-data/).

[6] http://georgewbush-whitehouse.archives.gov/nsc/nss/2006/.

[7] CNN quotes Dennis Blair speaking to Congress: "The primary near-term security concern of the United States is the global economic crisis and its geopolitical implications."

2.3. Water

2.3.A. Water and Security: Cause War or Help Community Building? (Jerome Delli Priscoli)

Author and Organization: Dr. Jerome Delli Priscoli, Senior Advisor, IWR –USACE, Governor of WWC, Editor in Chief, *Water Policy*: Institute for Water Resources, U.S. Army Corps of Engineer
Contact Information: Jerome.DelliPriscoli@usace.army.mil; priscoli@erols.com

Introduction

Our news has been filled with statements like "water as the next oil," "water wars," and "water as the liquid of the 21st century." For example:

> *"The only matter that could take Egypt to war again is water."(Anwar Saddat, 1979)*
>
> *"The next war in the Middle East will be fought over water, not politics."(Boutros Boutros Ghali, 1985)*
>
> *"The wars of the next century will be about water."(Ismail Seageldin, VP World Bank, 1995)*

But are these remarks more misleading than illuminating? In 1995, the Israeli Defense Forces analyst responsible for long-term planning during the 1982 invasion of Lebanon noted the following:

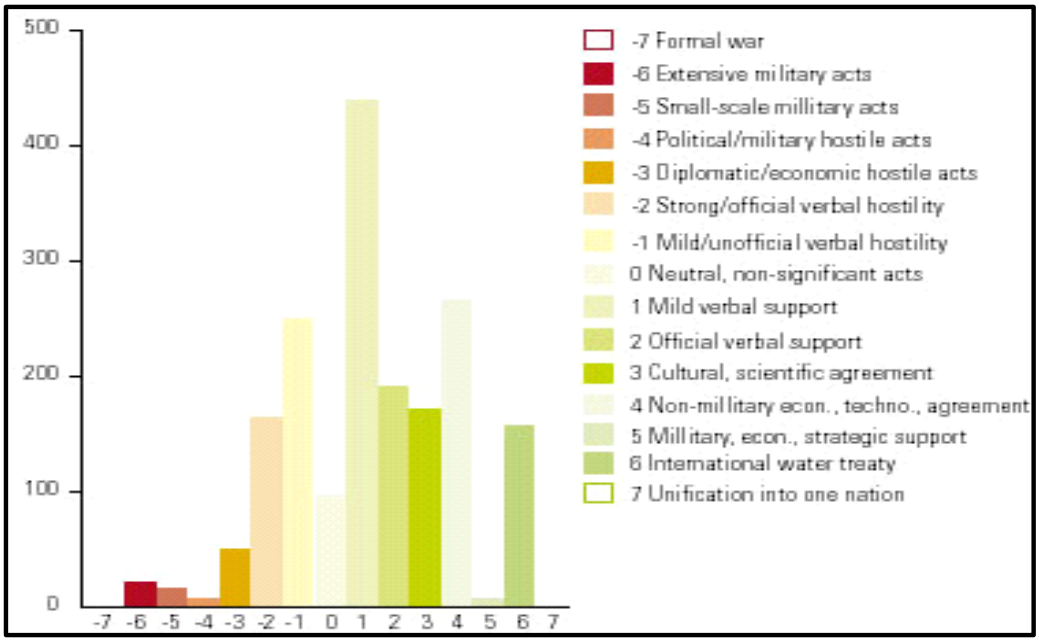
> *Why go to war over water? For the price of one week's fighting, you could build five desalination plants. No loss of life, no international pressure, and a reliable supply you don't have to defend in hostile territory.*[1]

Even some of the commentators themselves seem to be changing. For example:

> *"Fierce competition for fresh water may well become a source of conflict and wars in the future."(Kofi Annan, March 2001).*
>
> *"But the water problems of our world need not be only a cause of tension; they can also be a catalyst for cooperation...If we work together, a secure and sustainable water future can be ours."(Kofi Annan, January 2002).*

If one sees water as a resource to be fought over in terms of allocations of limited flows, as many in our security community now seem to, it is easy to come to the conclusion that water wars will result. However, the history of water resource management's role in the development of civilizations has been very different, especially at the macro social levels. That history of water resource management has been one of learning (albeit often painfully) to jointly create and share benefits beyond the river through use of water.

Oregon State University (OSU) data indicate that for 1,831conflicting and cooperative water interactions over the last fifty years, 7 disputes involved violence, 507 conflictive events have occurred, 200 were treaties signed, and there were 1,228 cooperative events.



Fifty years of 1,831conflicting and cooperative interactions over the last fifty years.
- 7 disputes involved violence; 507 conflictive events have occurred
- 200 treaties signed;1,228 cooperative events.

Figure 2.3.A-1. Water Vector of Cooperation vs. Source of Conflict.[2]

Strategic arguments do not point to water as cause of war. There are rich shared interests around water. Institutions can be developed so that they are resilient and can change with new needs. Economically, it makes little sense to fight over water, and there are increasing trends toward cooperative behavior, as noted below.[3]

- **Technical information beginning to bond vs.. separate**
- **Price for control over river = cooperation**
- **Opportunity costs for not cooperating growing**
- **New actors/new claims that cross jurisdictions**
- **Austerity requires multipliers**
- **Events of 1990's in world water community**
- **Science showing cooperation as root to evolution and security**
- **Functional and second track diplomacy**
- **Accept facilitator vs.. expert dictator role**
- **Technologies now cheap and facilitative**
- Others

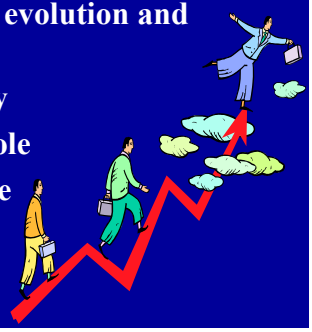


Figure 2.3.A-2. Trends toward Cooperation.

There are, however, river basins at risk, especially if they are left unattended with poor water management. Forty percent of the world's population lives on shared basins, which comprise more than 50% of the Earth's landmass. The OSU database identifies 16 basins at risk, including 51 nations on five continents, 8 of which are in Africa, and 6 of which are in Asia.

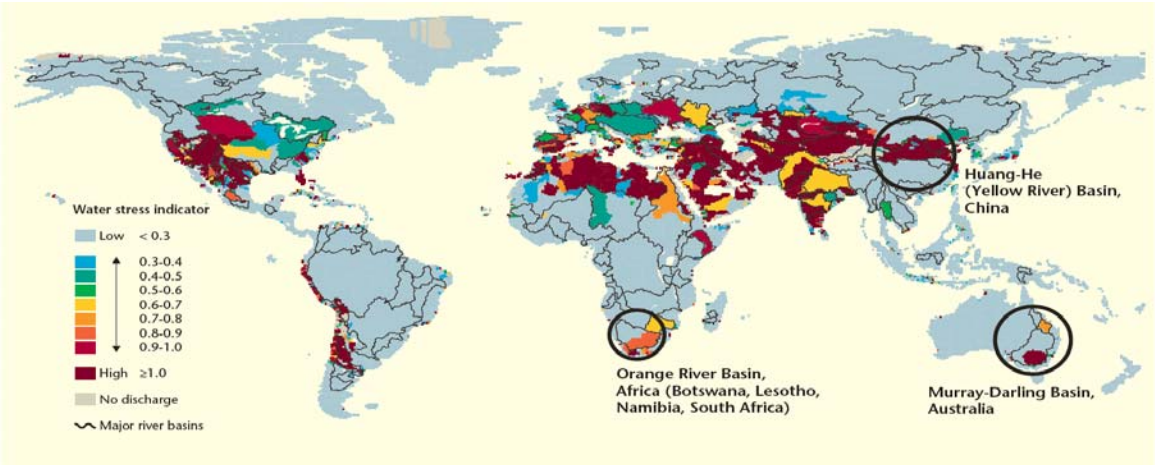


Figure 2.3.A-3. Water Stress Level of Major River Basins, around 2002.[4]

The likelihood of conflict in basins rises as the rate of change within the basin exceeds the institutional capacity to absorb that change. The basin characteristics which enhance resilience to change include the following:

- International and intra-state, cross-jurisdictional agreements and institutions (e.g., River Basin Organizations)[5]
- A history of collaborative projects
- Generally positive political relations
- Higher levels of economic development

The basin characteristics which indicate vulnerability include the following:

- Rapid environmental change
- Rapid population growth or asymmetric economic growth
- Major unilateral development projects
- The absence of institutional and/or organizational capacity
- Generally hostile relations

The world's renewable freshwater is in lakes, surface runoff and in groundwater and equals 40,000 cubic kilometers available/year; 10% is withdrawn, 5% is consumed, and the remainder goes to stream and aquifers.[6] The key problems are (1) managing the time and place of water, (2) managing uncertainty, and (3) aligning people with water. A large portion of water is located where there are no people, in areas such as the Amazon, Canada, and Alaska. Security issues arise in cases where a large portion of water arrives where people live in very short periods of time (e.g., in Asia, more the 2/3 of people live in areas where 80% rainfall arrives in 25% /year)[7]. A large portion of the remainder of water is degraded, and a high portion of available water is used (e.g., in arid areas that is 90%++). Overall, the patterns of water use are roughly 70% for agriculture, 20% for industry, and 10% for municipal.[8] Most of the world is struggling to

determine how to realign these uses to fit our changing needs that stem from changing demographics.

The water crisis is mainly one of distribution of water, knowledge, and resources and is not one of absolute scarcity. However, the water and security debate is driven by our notions of scarcity. Water is rarely the cause of war or large-scale social violence. However, such focus dominates the security and water debates.

Consequently, the most salient aspects of water are often passed over in the debate, including the water's powerful role in the following:

- Building the social community
- Generating wealth through established preconditions of economic activities
- Convening adversaries and providing common language for joint and creative dialogue
- Practically integrating diverse interests and values
- Providing a principle tool for preventive diplomacy and for building cultures of cooperation, if not peace

Strategy: Using the "Small s" Water Actions, to Achieve the Ends of the "Big S," Security

War and large-scale violence are what we might call the "Big S security" or the traditional concerns of the security community. Within this framework, water is seen as independent variable, the lack of which can be seen as a cause of war, a tool of war, a trigger for large social violence, and as a cause of eco-shocks leading to social unrest.

Our English dictionaries define *security* as, "freedom from danger, from fear or anxiety, from want or deprivation." Security is the basis of humanity's historical management of water, of taking measures (1) to assure that we have good water in the right quantity at the proper time and place, (2) to predict floods, (3) to impound water for droughts, and (4) to use water to help us generate wealth and avoid deprivation. This is what we might call the "small s security." Within this framework, water is seen more as a means to achieving important social ends. This "small s security" framework is the key to understanding water and its role in civilizations and to see how water relates to today's security issues.

Early humans in ancient Mesopotamia, China, Egypt, and other parts of the world sought to predict floods and droughts and to manage such events and their potential impacts on people. Herodotus ascribes the beginning of geometry in ancient Egypt to trying to predict floods. As humans focused on water they worked to create irrigation systems and other water works. These water actions effectively helped humans to adapt to the uncertainties of nature. In so doing, they reduced their fear and anxieties about these exigencies of nature. They changed from the fatalism of being victims to a sense of creating and molding their societies. As Jacob Bronowski notes in the beginning of his *Ascent of Man*, "humans are not just the perfectly adapted one cell being, they have made their home." This illustrates how "small s security" water actions are effective adaptive behavior.

To the degree that humans have enhanced their personal sense of security and reduced internal fears from the fatalism of droughts and floods, they began to create. They invented languages, they freed up time to invest in developing tools and systems, and they started to create and mold their homes. The definition of *security* as freedom from fear anxiety, want and deprivation, has been enhanced. All rich civilizations have invested their social capital in actions to help achieve the sense of managing such uncertainties as a precursor to growth and prosperity. When such

efforts deteriorated, so too did societies, and the same is true today, as evidenced in poor African countries.

The notion that the lack of water can cause war should be reconsidered. Rather than considering how water distributions or flows might lead to war, we should identify strategic regions and ask how investment in water resources management and investments (as a means) might help achieve strategic ends to ensuring the availability of water. This approach allows for water investment to serve as a means to achieving security rather than as an end itself that drives security. Water actions are key societal adaption tools. Investments in the "small s" become critical to enhancing and ensuring the "large S." How does this work? Figure 3.3.A-4 illustrates the process.

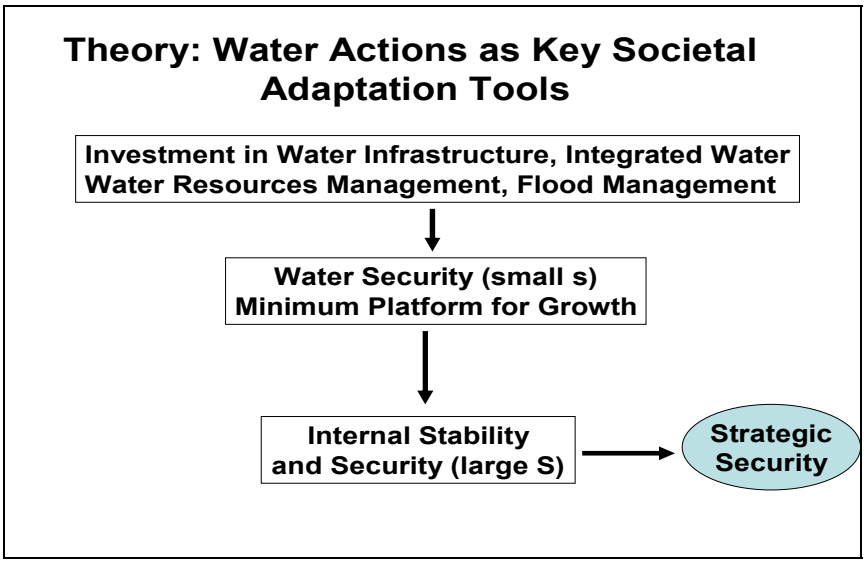


Figure 2.3.A-4. Water Actions as Key Societal Adaptation Tools.

The climate change environmental shocks projected to impact society are primarily water related (e.g.,) droughts, floods, tsunamis, and rising sea-levels. This puts projections of these environmental shocks and ways to deal with them in the historical purview of water resource management. Both behavioral and hard infrastructure water actions are prime societal means to adapt to and manage the uncertainties of change. Rather than longer term mitigation measures, these more immediate adaptations are the keys to social resilience and thus keys to achieving the small "s" of security. To the degree that they achieve stability, they contribute to larger sense of security through a reduced sense of vulnerability, therefore contributing to the "large s" security.

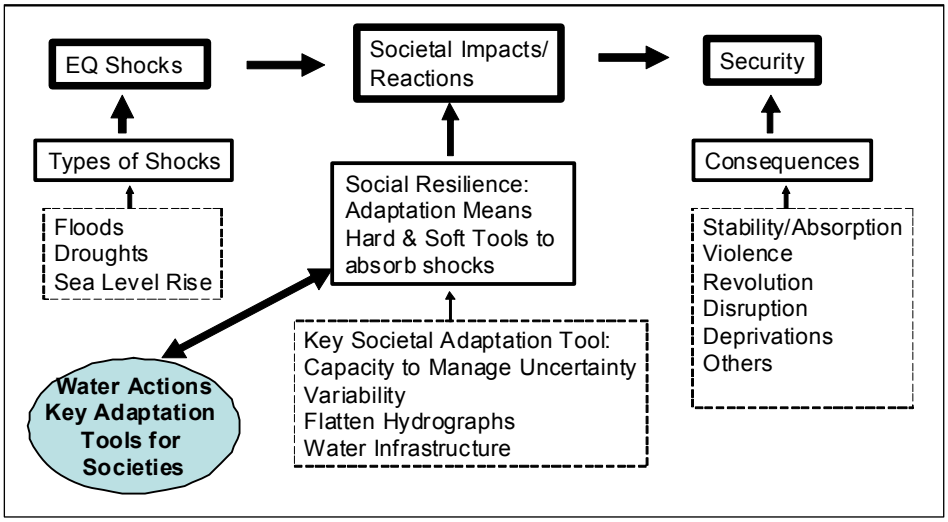


Figure 2.3.A-5. Environment/Water Actions as Adaptation Tools and Keys to Societal Security/Stability.

Water and Security Today

The gross domestic products (GDPs) of Zimbabwe, Ethiopia, Kenya and Mozambique all seem to vary with rainfall. In fact, the variations in GDP due to the inability to deal with variations in rainfall (the peaks and valleys of the hydrograph or of floods and droughts) might account for almost 25% of variations in these countries' GDPs. Societies that do not work to attenuate the impact of flood and drought occurrences have little chance to develop, and often become fatalistic, accepting and actually coming to accept the fate of being wiped out and having to start over again every several years. Fear and the "small s security are pervasive in this setting.

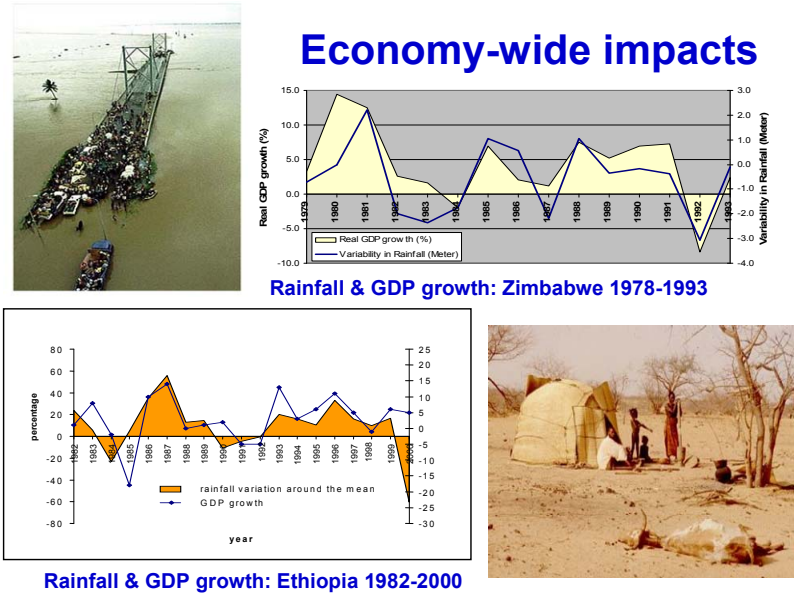
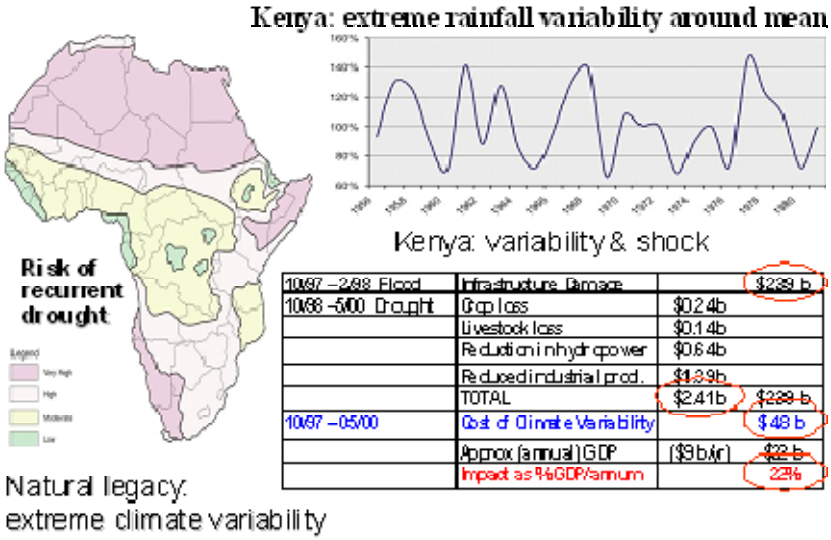


Figure 2.3.A-6. Economy-wide Impacts of Rainfall.

As psychologists point out, this becomes a cycle of thinking that itself further defeats development—a deterministic fatalism. This mindset often results in pockets of poverty, and no money moves in their directions. These pockets of poverty are then vulnerable to forces conspiring to "big S Security" actions of violence and disruption.

What Does This Mean to Water and Security?

First, one may notice that flattening out the peaks and valleys of the curves is what hydrology is all about. For the last 500 years that is how and why humans—in the broad social sense—became engineers: to build more certainty into an uncertain environment, to manage uncertainty so as to build stable expectations, and to create time for investment in other areas, which in turn creates more resources, which leads to more innovation.



10/97 – 2/98 Flood	Infrastructure Damage		$2.39 b
10/98 – 5/00 Drought	Crop loss	$0.24b	
	Livestock loss	$0.14b	
	Reduction in hydropower	$0.64b	
	Reduced industrial prod.	$1.39b	
	TOTAL	$2.41b	$2.39 b
10/97 – 05/00	Cost of Climate Variability		$4.8 b
	Approx (annual) GDP	($9 b/yr)	$32 b
	Impact as %GDP/annum		22%

Figure 2.3.A-7. Climate Variability Undermines Growth.

Second, where there are droughts there are floods. They are opposite sides of the same coin, and they coexist with the poverty and the poorest of the poor.

Third, we might posit a hypothesis: there is a minimum platform of water resources infrastructure and institutions necessary to achieve "water security."[9] Water security is essential for sustained economic growth and poverty eradication. Water security is achieved through balancing the productive use of water with managing vulnerabilities to water's destructive power; to balancing access to water with living within acceptable levels of risks from unpredicted events.

	Actual		Projection				
			Before the Floods		After the Floods		
	1998	1999	2000	2001	2000	2001	
Real GDP (ann. Growth rate)	12.0	9.0	7.0	7.2	5.4	7.9	-23%
Inflation (ann. average, %)	0.6	2.0	6.6	5.0	9.5	5.0	+44%
External current account:							
Before grants	-20.5	-31.7	-23.0	-15.7	-31.5	-18.4	
After grants	-12.4	-21.5	-16.3	-9.1	-19.7	-11.0	
Fiscal Balance:							
Before grants	-10.7	-12.1	-12.1	-10.7	-16.0	-11.5	
After grants	-2.4	-1.2	-5.2	-4.4	-7.0	-5.1	
Memorandum: GDP (Mt billion)	46,134	52,913	60,177	67,790	61,471	69,673	

Source: Staff estimates, IMF and Government of Mozambique

Figure 2.3.A-8. Rainfall Affects Growth: The Case of Mozambique's Year 2000 Floods.

From Fighting over Allocation of Flows to Jointly Creating Benefits from Water Use: Water Security is More Than Access

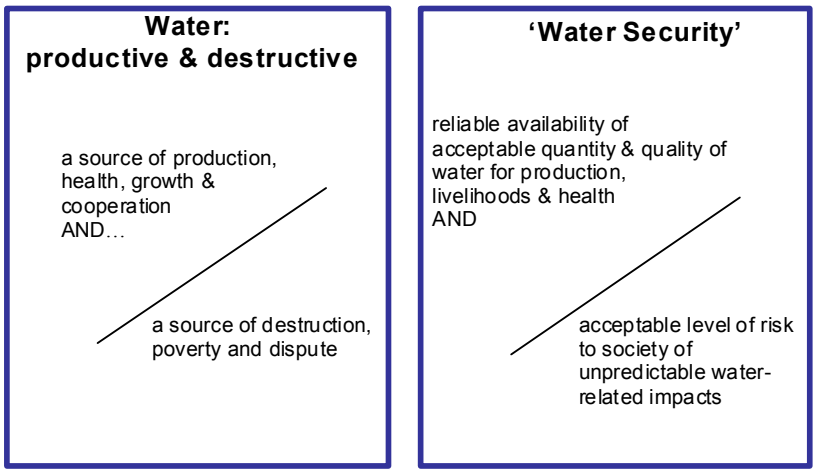


Figure 2.3.A-9. Water Security.

Throughout history, water has been a venue for dialog often where dialog is prohibited, and this continues today. The question is, why? Ensuring water security involves more than ensuring access. Water's many uses include irrigation, flood damage reduction, drought prevention, ecological flows, hydropower, energy coolant, navigation, and recreation. These multiple purposes or uses of water allow for joint creation of benefits from the instead of fighting over allocation of flows. Achieving this realization is a key to understanding water as a venue of dialog.

Throughout history, water actions have been a means to achieving other macro social goals. In this manner, water has often expanded the negotiating pie instead of requiring reallocation of the limited pie. This means that decisions were really about relative and not absolute scarcity. This history fits well with new paradigms of negotiations and with interest- and needs-based approaches to negotiations. In this sense, water is seen as humanity's proving ground for building a community rather than as a generator of war.

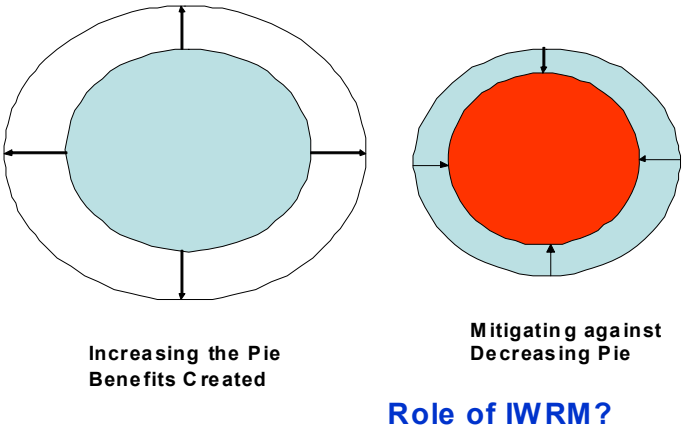


Figure 2.3.A-10. Negotiating Arenas: Benefits - Interests

This is not new. Implementing multiple uses of the same water to create joint benefits has been the hallmark of water management in the wealthy or developed world. As the figure below shows, it was central to nation building in North America,

1808: Gallatin Report

Waterways to be used for:
•Building Political Unity and Nation
•National Defense
•Economic Development

Marshal Gibbons Ogden

1824: GIBBONS VS. OGDEN

(Estbl. Federal Powers vs. States)
Claims are said to be repugnant–
1st. To that clause in the constitution which authorizes Congress to regulate commerce.
2d. To that which authorizes Congress to promote the progress of science and useful arts.

1920's -"308" Reports: **Congress Authorizes USACE do Comp. assessments of all major rivers of the US**

Figure. 2.3.A-11. Water Ways and Establishing National Federal Interventions over Interstate Issues.

This approach is known as integrated water resources management (IWRM).[10] Water can serve many purposes: hydropower, flood control, navigation, water supply for cities and farms, ecological maintenance, recreation, and others. USACE has been a means to achieve IRWM in North America, with the potential of providing these services around the world. USACE has pioneered IRWM concepts on the Columbia, Missouri, Mississippi, Ohio and other rivers.

Successful implementation of IRWM is an essential key to unleashing development, creating a platform for growth in various economic and social areas. IRWM has been a prerequisite for successful development in the US, and this development can clearly being seen as a prerequisite for supporting and maintaining social stability, as well as reducing violence in the third world. IRWM requires water infrastructure investment that brings multipurpose uses and options to decision makers.

The reapportionment or reconfiguration of purpose for which the current U.S. water infrastructure is used in the U.S. and in other wealthy countries will be a central focus as we adapt to climate changes, ecological changes and changing demographic patterns; likewise, it will be key to maintaining national security.

Conclusions: A Strategic Argument

The strategic aspects of water seem to lend themselves to finding means for cooperation. Evidence abounds. Look at the Middle East Table Talks, during which parties were at war, the Multilateral Water Talks in the Middle East, the World Bank and US interventions in the *Indus River Treaty* of the 1950s and its current arbitration, the Recent Istanbul World Water Forum, and the meetings of the riparian of T-E River, and others.

Southern Africa is a region with prime strategic importance to the U.S. Southern Africa is dependent on water. Investment in the "small s security" in Southern Africa could pay huge dividends to the "big S security" and ultimately to US security interests. However, the water system in this region is hugely complex. The roles of international rivers are a chief element of a region's security complex, but this role in Southern Africa is as yet largely unexplored. Threats to economic security derive from the role of water as a foundation for the economic growth and prosperity of a given state. International river basins form an important element of the Southern African Regional Security Complex.

The horn of Africa is clearly an area of strategic interest to the US. Opportunities, as opposed to only difficulties, abound in the Nile Basin. There are major potential win-win benefits from

cooperative development. For example, in power production/trade, 90% hydropower potential is undeveloped, and 85% of the population is un-served. In food production and trade, 60% of irrigable land is un-irrigated. Multipurpose storage potential is huge, but there is very little being implemented despite very high rainfall variability. Environmental sustainability could be enhanced on watersheds, soils, wetlands, and lakes. It is important to note that, in conflict prevention, reduced tensions promote integration.

Today the old Red Dead Canal is once again alive. New reports now present the project as a significant means for cooperation in the volatile region, whereas the focus of such reports has historically been focused on the associations between water and war. The famed Israeli billionaire Tshuva stated, "This is the only way to get out of the cycle of violence and the dead end in the area,…..jobs and prosperity,… would moderate Arabs in the region and give them an alternative to violence and terror…….peace will be made not by peace agreements but by making cooperation and goodwill among the peoples of the region. The Valley of Peace will provide a solution for generations to come."

Environmental communities are also engaged in this dialogue. Israeli environmentalists accused the World Bank of stacking the timetable of the Red Sea-Dead Sea Canal project discussions in their favor. Many conferences, reports, and policies are implemented which hinder meaningful dialog and the relationships necessary to support the approaches outlined.

To move down the strategic path suggested, the rhetoric of the wealthy vs. poor dialog should be changed. Often prescriptions from the wealthy nations are borne from unconscious assumptions embedded in the water experiences gained in those situations. What may seem to be obviously correct may not be what others in different cultures and situations need. U.S. investments in water infrastructure actually helped build a regionally divided nation into one; they transformed major regions from poverty to productivity and affluence. The Pacific Northwest transitioned from deep poverty in the 1930s to a region with one of the largest GDPs in the world. The same was true and even more dramatic in the southeast due to the Tennessee Valley Authority, an organization that now oversees some of the richest soil in the US. Parts of this region were once so poor agriculturally or were inundated with floods to the point that little could grow. TVA is now the second largest tax payer in the southern U.S.

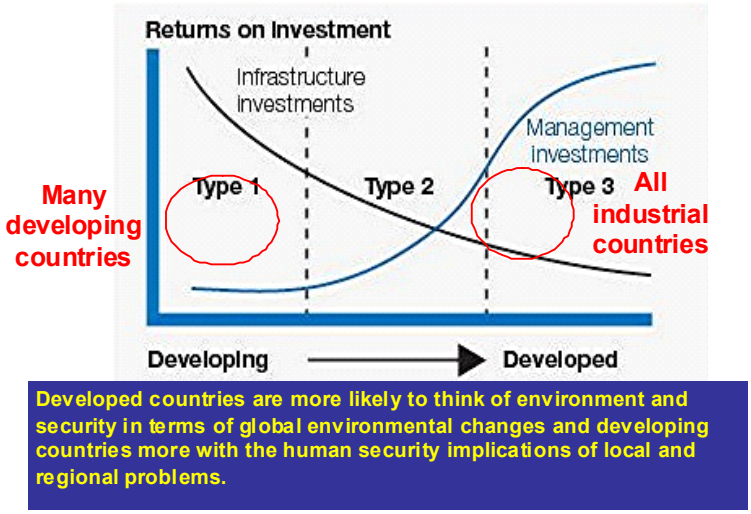


Figure 2.3.A-12. Relations between Developed and Developing Worlds.

Figure 2.3.A-12 shows that type 1 countries (poor/developing) are spending much higher percentages on infrastructure investment than on management investments. Type 3 (the US and other wealthy nations) spend more on management investments and less on infrastructure. While type 1 countries can learn from type 3 countries about conservation or other measures, they can also

learn about how type 3 countries invested and used water. Then type 3 countries can outline the unanticipated costs over time and help design measures to allow investment while mitigating or avoiding unwanted effects through sound design.

Ultimately, the creation of platforms for growth is essential to breaking the fatalism and poverty throughout the world, enhancing the focus on the "small s security," which in turn will pay subsequent dividends to the "big S Security." Those platforms for growth depend greatly on water actions and investments. Water actions, in turn, are directly related to "small s security," and wide dispersions of the "small s security" are a key ingredient to for achieving goals of the "big S security."

In Africa the US could fund and convene major discussions on key river basins that represent strategic areas of concern to the US and could conduct studies on how IRWM might increase "small s security," consequently paying dividends to "big S Security." outlined. Such discussions would place the US in a leadership role in an area where it really has comparative advantage, and it could also produce new coalitions and partnerships of investments.

Priority areas of U.S. security concerns could be as follows:

- How can water actions be used as a means to achieve security ends in each priority area?
- How can we work to prevent and reduce vulnerability to disasters?
- How can be used to create a platform for growth while limiting mitigation cost to the environment?

↑ Back to the top

---

[1] Jerome Delli Priscoli and Aaron T. Wolf, Managing and Transforming Water Conflicts, Cambridge University Press, New York, 2009, Chapter 2 pages 9-28.

[2] Wolf, A; Yoffie, S. Giordano, M. Forthcoming. *International Waters: Identifying Basins at Risk.* Corvallis, University of Oregon.

[3] Delli Priscoli and Wolf, op. cit., Chapter 2 and page 3.

[4] Smakhtin, Revenga, and Döll, 2004.

[5] Delli Priscoli and Wolf, op. cit., Appendix B

[6] The United Nations, Water in a Changing World, UNESCO Publishing, Paris, 2009, Chapter 2 pages 80-95.

[7] Delli Priscoli and Wolf, op. cit., page xxii

[8] United Nations, op. cit., Chapter 2.

[9] Sink or Swim? Water Security for Growth and Development, David Grey and Claudia W. Sadoff, *Water Policy*, June 2007, page 17.

[10] Presentation by Jerome Delli Priscoli, "Water and Security," Johns Hopkins University (SAIS) November 18, 2008, Washington, D.C.

2.3.B. Water Resources and Stability (J. Rolf Olsen, Kathleen White, Julie Kiang, D. Phil Turnipseed, Levi Brekke, David Raff, Roger Pulwarty, Robert Webb)

Authors and Organizations: J. Rolf Olsen and Kathleen D. White, US Army Corps of Engineers; Julie E. Kiang and D. Phil Turnipseed, U.S. Geological Survey; Levi D. Brekke and David A. Raff , US Bureau of Reclamation; and Roger S. Pulwarty and Robert Webb, National Oceanic and Atmospheric Administration

Contact Information: J.Rolf.Olsen@usace.army.mil, Kathleen.D.White@usace.army.mil, jkiang@usgs.gov, pturnip@usgs.gov, LBrekke@usbr.gov; DRAFF@do.usbr.gov; Roger.Pulwarty@noaa.gov, Robert.S.Webb@noaa.gov

Abstract

Water is a fundamental requirement for ecosystems, human health, and economic growth, and consequently the management of this resource is closely linked to social stability. Water resources have been managed to provide water for irrigated agriculture, municipal, rural and industrial water supply, energy, and other uses. However, water resources are stressed by multiple factors, including population growth and demographic changes, urbanization, land use changes, water pollution, climate change, and inadequate financing and investment. Droughts and floods can cause catastrophic losses for poor countries. Integrated water resources management (IWRM) is an effective approach for addressing water resources problems because it follows a comprehensive approach that integrates multiple objectives for all stakeholders over the relevant spatial scale and time horizon. An IWRM approach includes a framework for evaluation of potential impacts to society and ecosystems that could contribute to social instability.

Introduction

Water is critical to the well being of humans and ecosystems. According to the United Nations, "Adequate water in both quantity and quality underpins health and basic quality of life."[1] This paper describes why water is essential to social stability. Water resources affect ecosystems, human health, and economic development. Hydrologic extremes can cause major social disruptions. The paper also describes the multiple factors that can impact and stress water resources. Because of these interactions, solutions to water resources problems should follow a comprehensive approach that integrates multiple objectives across the proper spatial and temporal scales. All relevant stakeholders should participate in the decision making process.

The Need for Water

Water is essential for plant, animal, and human life. Water availability is a limiting factor for primary productivity of terrestrial ecosystems because living things require water to grow and survive. Water supports terrestrial ecosystems and the services they provide people: agriculture for food production, forests for firewood and construction materials, and grasslands for grazing animal habitat. Water is the medium of life for aquatic ecosystems. Lakes, streams and rivers support fisheries which are a major component of the diet in many countries.

Clean water is also essential to human health. According to the World Health Organization, "Adequate quantities of safe water for consumption and its use to promote hygiene are complementary measures for protecting health." [2] Waterborne diseases include cholera and

typhoid, dysentery, hepatitis A and diarrhea. Water-washed diseases occur when people have inadequate water to keep their hands, bodies and domestic environments clean.[3]

Water resources availability and management are one determinant of a country's growth.[4] "Societies do not become wealthy first and then invest in water management; they find ways to manage water and risk first, which then leads to wealth."[5] Irrigation can improve agricultural productivity. Safe drinking water and sanitation cut the loss of life from a number of illnesses, so that health resources can be spent on other conditions. Water also must be provided for energy and industrial production.

Managed Water Resources. Water provides for multiple social uses. One of the major uses of water is for irrigated agriculture. More than 80% of agriculture depends only on adequate rain and is subject to the variability of climate. Only 18% of agricultural land is irrigated, but the irrigated land produces half of the world's grain supply.[6] About 70% of all freshwater withdrawals go to irrigated agriculture.[7]

Another important use for water is for human drinking water in both urban and rural communities. In urban areas, water supply infrastructure is necessary to supply water to the population. Water supply in rural areas may be simpler but is still necessary.[8] The Second World Water Forum and the United Nations Millennium Declaration endorsed a target of reducing by one-half the proportion of people without sustainable access to adequate quantities of affordable and safe water by 2015. However, 18% of the world population (1.1 billion people) does not have access to improved water supply.[9] A related issue is access to sanitation facilities. An estimated 2.4 billion people lack access to improved sanitation.[10] The areas lagging behind achieving improvements in sustainable water supply and sanitation include sub-Saharan Africa and low-income Arab states,[11] which are regions of concern for social stability.

Water is important for energy production and electrical generation. Water not only provides hydropower, but it also is used as cooling water for thermal power plants. Rivers are used by both developed and developing countries for transportation of both people and freight. Aquatic systems may compete for water with human needs. Ecosystems require water, but ecosystems also provide multiple benefits for human and social well being, including food and fiber.

Conflicting Water Uses. There is often a perceived conflict between providing water for human needs and supporting aquatic ecosystems. There are different stakeholders for different uses. For example, urban dwellers benefit from lower electric costs, while fishermen depend on a productive aquatic habitat. An example of the potential tradeoff between hydroelectricity and fisheries is the Pak Mun Dam in Northeast Thailand. Pak Mun Dam is a run-of-the-river hydropower plant on the Mun River just upstream from its confluence with the Mekong. Even though a fish ladder was built, fish species that migrate and depend on the river rapids are seriously affected because the migration route is blocked in the beginning of the rainy season. Before dam construction, local communities depended on river fisheries for food and income. After the dam construction, fishermen reported a 50-100% decline in their fish catch. Households that depended on fisheries upstream of the dam declined from 95.6% to 66.7%. Many villagers have been unable to find alternative means of livelihood, and many have migrated to the cities looking for wage labor. Compensation was provided to the fishermen, but it was not sufficient to replace their loss of livelihood.[12]

Hydrological Extremes and Social Stability

Precipitation varies across time and space. Hydrological extremes of droughts and floods can have major impacts on human society and ecosystems. Economic losses from global disasters, including floods, have been increasing in recent decades.[13] The cost of disasters as a percentage of gross domestic product (GDP) is much higher in poor countries than in developed countries.[14]

Droughts. Droughts are episodic dry periods caused by a period of abnormally low precipitation. Meteorological drought is generally defined by a deficit of precipitation and/or increased evaporation. Hydrological drought is concerned with reduced surface and/or subsurface water supply, including streamflow, reservoir and lake levels, as well as groundwater levels. Agricultural drought deals with the impact of moisture deficiency on agricultural production.

Droughts have a wide range of social impacts. Drought can reduce the available water supply in groundwater and reservoir levels. Reduced streamflow can degrade water quality since pollutants become less diluted. Drought reduces crop yields in rain-fed agriculture, and it also reduces the available supply for irrigated fields. Droughts provide favorable conditions for some insects that further destroy crops, such as locusts. Dry vegetation also increases the potential for wildfires. Some fish species face increased mortality as lake and river levels fall,[15] affecting those who depend on the fish for their food or livelihood. Droughts can have especially serious consequences for poor countries. Droughts and the resulting famine in Ethiopia killed about 400,000 people in the 1970s and about 1 million in the 1980s. Millions more were displaced.[16]

Floods. Too much water also causes major social impacts. Flooding is the most destructive natural disaster. One study found that flooding accounts for two-thirds of the people affected by natural disasters each year.[17] In addition to causing deaths and damages to settlements, floods can damage agricultural production, threaten water supplies, lead to outbreaks of waterborne disease, and damage terrestrial and aquatic habitat.[18] Hurricane Mitch caused massive destruction in Central America, with economic losses in Honduras estimated as being about 69% of their annual GDP in 1998.[19] In Bangladesh, a flood in 1998 inundated over two-thirds of the country, and over 10% of the estimated rice production was lost.[20]

Stresses and Changes Affecting Water Resources

While water resources affect society in numerous ways, water resources are, in turn, affected by many dynamic processes, including societal change. Population growth increases the demand for water. Society's vulnerability to floods depends on where people live and increases when people move to coastal areas and flood plains. Changes in land use patterns and land cover can cause changes in runoff patterns and the hydrologic cycle. Ground-water supply can be depleted. Water pollution threatens human health and aquatic ecosystems. Natural resources can be overharvested.

Population Changes. The world's population tripled during the 20th century.[21] From 1990 to 2000, the population increased by 15% to over 6 billion people.[22] Larger population and related economic growth increase the demand for water. Estimated water withdrawals have increased by seven times in the 20th century.[23] Vorosmarty and others[24] evaluated the vulnerability of water supply in 2025 by estimating the ratio of water withdrawals to discharges. They concluded "that impending global-scale changes in population and economic development over the next 25 years will dictate the future relation between water supply and demand to a much greater degree than will changes in mean climate."

Population and economic growth have also occurred in areas that are subject to flooding and coastal storms. Analyses of long-term records of disaster losses indicate that population growth and more economic development in vulnerable locations are the principal factors responsible for increasing losses from hurricanes and floods.[25]

Urbanization. The growth rate of the world's urban population has exceeded the general population growth rate. In 2000, 47% of the world's population lives in cities, up from 43.5% in 1990.[26] The increase in urban population has placed a strain on water supply and wastewater treatment infrastructure in urban environments.[27] Urbanization also affects the hydrologic cycle. Roofs, paved roads and parking lots are generally impervious and prevent precipitation from infiltrating into the ground. The amount of runoff generally increases as the percentage of impervious area increases. In urban areas, runoff is often quickly concentrated into open channels and storm sewers. Peak flow rates tend to increase in urban areas due to higher runoff volume and higher velocity.[28]

Land Use Changes. Land use changes affect the hydrology of a watershed. Afforestation and deforestation are the largest land cover changes in terms of area.[29] The amount of infiltration is generally higher in forested land cover compared to grass, agricultural, or urban land cover. Plants absorb water, which is then transpired through the plant and eventually escapes through the leaf pores into the atmosphere. The overall effect of forest is a reduction of runoff from storms compared to grass or crop land.[30] Afforestation can reduce both annual runoff and flood discharges. Deforestation has the opposite effect, reducing interception of precipitation, evaporation, and transpiration, and subsequently increasing runoff. Deforestation may have worsened the impact of flooding in the Yangtze River basin in China and in Central America from Hurricane Mitch, both in 1998.[31]

Agricultural practices can alter the hydrologic cycle and water quality. For example, increased irrigation can cause or exacerbate problems with groundwater depletion. Runoff from agricultural fields can increase nutrient loads in rivers, leading to hypoxia in downstream lakes and estuaries. Loss of wetlands reduces storage in the watershed and affects runoff patterns.

Groundwater Depletion. In many parts of the world, groundwater is a major source of water supply. Groundwater extraction data are highly uncertain, but most of the pumped groundwater is used for agriculture, with domestic uses being second. Areas irrigated with groundwater have substantially higher agricultural yield than regions irrigated from surface supplies,[32] as groundwater is typically a more stable source of water during drought.

Prior to development, a groundwater system is in equilibrium, and inflow equals outflow. Groundwater is recharged by infiltration from precipitation and surface water. It discharges naturally to springs or as base flow in streams. When groundwater is pumped, however, the total discharge may no longer equal the recharge, resulting in declining groundwater levels.[33]

Extraction rates exceed recharge rates in a number of areas in the world, including several countries in the Middle East. These countries include Saudi Arabia, United Arab Emirates, Libya, Egypt, Algeria, Pakistan and Mauritania.[34] As groundwater levels decline, the water must be pumped from deeper levels up to the surface, increasing energy costs. If the levels decline below the bottom of the existing pump, the pump must be lowered or the well must be deepened. Yields decline and the cost of the water increases. Groundwater pumping reduces the base flow for streams and can cause land subsidence. In coastal areas, groundwater pumping can cause saltwater to migrate inland, resulting in saltwater intrusion into the water supply.[35]

Water Pollution. Water has been used for disposing of domestic, agricultural and industrial waste. Deteriorating water quality caused by this pollution is a worsening problem that threatens human health and aquatic ecosystems. Insufficient sanitation infrastructure is one source of pollution. Nutrient loading is another problem. Humans have doubled the amount of reactive nitrogen in streams. High levels of nitrogen contribute to eutrophication of freshwater and coastal marine ecosystems and acidification of freshwater and terrestrial ecosystems.[36]

Ecosystem Changes. Ecosystems face multiple stressors, in addition to concerns with water quantity and water quality. One problem is overharvesting, especially overfishing. At least one quarter of commercial fish stocks are overharvested.[37] Invasive species (introduced either accidentally or intentionally) compete with native species and upset the balance of native ecosystems. Land use changes such as deforestation and destruction of wetlands cause habitat fragmentation. Species may then be limited in their ability to migrate if they are under pressure due to other changes in their habitat.

Climate Change. The best available scientific evidence indicates that the Earth is warming. Climate change has the potential to affect many sectors in water resources.[38] The Intergovernmental Panel on Climate Change (IPCC) says with high confidence that "observed warming over several decades has been linked to changes in the large-scale hydrological cycle."[39] The IPCC projects that the frequency of heavy precipitation events is very likely to increase, while the proportion of land in drought at any time is also likely to increase. Sea level is expected to rise due to warming ocean temperatures, as well as melting glaciers and ice sheets. A higher sea level will increase the potential for flooding of people living in low-lying coastal areas. Global warming may affect agriculture and food production and could impact water availability. Climate change may aggravate the effects of other stresses on aquatic ecosystems.[40]

Financing and Investment in Water Resources. Water resources management costs money. New water supply storage, municipal water infrastructure, and sanitation facilities require capital investments. Water management functions such as flood risk management, environmental protection, monitoring, administration, planning and research need funding for labor. Infrastructure deteriorates over time and must be repaired and rehabilitated. Structures need to be maintained to stay functional, and deferred maintenance accelerates deterioration. Adequate financing often limits water resources management. One study placed the world's gap in annual capital requirements for water supply and wastewater services at $67.5 to $115.2 billion .[41] Water resources management will likely continue to be constrained by a lack of financial resources.

Integrated Water Resources Management (IWRM)

Water management is often fragmented among sectors. It is often controlled by top-down institutions, "the legitimacy and effectiveness of which have increasingly been questioned." [42] IWRM is being advocated as an alternative approach to water management.

IWRM can be defined in various ways,[43,44] but the approach usually requires the use of a comprehensive systems approach applied throughout a watershed rather than focusing on a single project or sector.[45,46] Integration should be performed across space, objectives, institutions, and time.[47] An appropriate spatial region should be defined for the problem. For surface water, this region is usually a basin or watershed, but the bounds of an aquifer should be included when conjunctive use of ground and surface water is considered. The vertical bounds may range from the lithosphere to the atmosphere. Integrated water resources management must integrate multiple objectives. For example, the need for water supply storage and hydropower must be balanced with the need to sustain aquatic environments and fisheries. To be most effective, IWRM should follow a real participatory approach in which all stakeholders are part of a collaborative decision process. IWRM should also integrate across temporal scales to ensure that ecosystems, economics, and society are sustainable over the long term.

IWRM can include some of the following specific activities.

- Land management and planning can be integrated with water management and planning.
- Conjunctive use of groundwater and surface water can be considered.
- Planners can evaluate alternatives for a range of possible future conditions, including climate change, to account for future uncertainty.
- Truly integrative water planning should evaluate ecosystem and social impacts in addition to economic development.
- IWRM should also consider the needs of all segments of a society, including groups that may be disadvantaged.

Although IWRM may be a desired paradigm for water management, there are certainly impediments to its implementation. Many societies lack a participatory process that brings together all stakeholders. Groups in power may not want to share decision making with other stakeholders. Lack of adequate funding may be another roadblock to implementation of IWRM.

Conclusion

Water is intrinsically tied to a nation's economic development, human health and wellbeing, and the conditions of ecosystems. Hydrologic extremes of droughts and floods can lead to severe social disruptions. Stability operations need to recognize water's important role. Water resources could be impacted by multiple changes in the future, including climate change, population increases, urbanization, land use changes, and depletion of resources. Water resources management is likely to be chronically underfunded. A comprehensive framework is needed to understand the potential impact of these stresses on water and social stability. The IWRM provides a framework that considers the sustainability of water resources over time. A holistic approach recognizes that water management decisions can impact many sectors affecting environmental and social wellbeing. IWRM can evaluate the potential for social instability from water resources stresses and the potential alternatives that could be employed by water managers.

↑ Back to the top

[1] "Status Report on IWRM and Water Efficiency Plans for CSD16," *United Nations-Water*, 2008. available at http://www.un.org/esa/dsd/resources/res_publsdt_wat.shtml

[2] "Global Water Supply and Sanitation Assessment 2000 Report." World Health Organization (WHO) and United Nations Children's Fund, 2000.

[3] "Global Water Supply and Sanitation Assessment 2000 Report." World Health Organization (WHO) and United Nations Children's Fund, 2000.

[4] "The United Nations World Water Development Report 3: Water in a Changing World." *World Water Assessment Programme* (WWAP), 2009. Paris: UNESCO, and London: *Earthscan*, available at http://www.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml

[5] "The United Nations World Water Development Report 3: Water in a Changing World." *World Water Assessment Programme* (WWAP), 2009. Paris: UNESCO, and London: *Earthscan*, available at http://www.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml

[6] Bates, B.C., Z.W. Kundzewicz, S. Wu and J.P. Palutikof, Eds., 2008: "Technical Paper of the Intergovernmental Panel on Climate Change," *Climate Change and Water.* IPCC Secretariat, Geneva, 210 pp.

[7] "The United Nations World Water Development Report 3: Water in a Changing World." *World Water Assessment Programme* (WWAP), 2009. Paris: UNESCO, and London: *Earthscan*, available at http://www.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml

[8] Helweg, Otto J., 2000. "Water for a Growing Population Water Supply and Groundwater Issues in Developing Countries," *International Water Resources Association*, 25 (1), pp. 33-39.

[9] "Global Water Supply and Sanitation Assessment 2000 Report."World Health Organization (WHO) and United Nations Children's Fund, 2000.

[10] "Global Water Supply and Sanitation Assessment 2000 Report." World Health Organization (WHO) and United Nations Children's Fund, 2000.

[11] "The United Nations World Water Development Report 3: Water in a Changing World." World Water Assessment Programme (WWAP), 2009. Paris: UNESCO, and London: *Earthscan*, available at http://www.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml

[12] "Pak Mun Dam Mekong River Basin Thailand." World Commission on Dams (WCD), 2000.

[13] Höppe, P. and R.A. Pielke, Jr. (eds.), 2006. "Understanding and Attributing Trends and Projections," Final Workshop Report. Workshop on Climate Change and Disaster Losses. Hohenkammer, Germany, 25-26 May, Available at http://sciencepolicy.colorado.edu/sparc/research/projects/extreme_events/munich_workshop/full_workshop_report.pdf

[14] Delli Priscoli, J., and A. T. Wolf. 2009. "Managing and Transforming Water Conflicts." *International Hydrology Series*. Cambridge, UK: Cambridge University Press.

[15] Cooley, Heather, 2006. "Floods and Droughts," in Peter Gleick, *The World's Water 2006-2007*, Island Press.

[16] United Nations Environment Programme (UNEP), 2002. *Africa Environment Outlook Past, present and future perspectives*, available at http://hqweb.unep.org/dewa/Africa/publications/AEO-1/index.htm

[17] United Nations (UN), 2004.Guidelines for Reducing Flood Losses, available at www.unisdr.org/eng/library/Literature/8689.pdf

[18] Cooley, Heather, 2006. "Floods and Droughts," in Peter Gleick, *The World's Water 2006-2007*, Island Press.

[19]"Guidelines for Reducing Flood Losses," United Nations (UN), 2004, available at www.unisdr.org/eng/library/Literature/8689.pdf

[20] Del Ninno, Carlo, Paul A Dorosh, Lisa C. Smith, and Dilip K. Roy, 2001. "The 1998 floods in Bangladesh: disaster impacts, household coping strategies," International Food Policy Research Institute.

[21] "Integrated Water Resources Management," TAC Background Paper No. 4. Global Water Partnership (GWP), Technical Advisory Committee (TAC), 2000.
[22] "Global Water Supply and Sanitation Assessment 2000 Report." World Health Organization (WHO) and United Nations Children's Fund, 2000.
[23] "Integrated Water Resources Management," TAC Background Paper No. 4. Global Water Partnership (GWP), Technical Advisory Committee (TAC), 2000.
[24] Vorosmarty, C.J., P. Green, J. Salisbury, R.B. Lammers, 2000. "Global Water Resources: Vulnerability from Climate Change and Population Growth," *Science* 289: 284-287.
[25] Höppe, P. and R.A. Pielke, Jr. (eds.), 2006. "Understanding and Attributing Trends and Projections," Final Workshop Report.Workshop on Climate Change and Disaster Losses. Hohenkammer, Germany, 25-26 May, Available at http://sciencepolicy.colorado.edu/sparc/research/projects/extreme_events/munich_workshop/full_workshop_report.pdf
[26] "Global Water Supply and Sanitation Assessment 2000 Report." World Health Organization (WHO) and United Nations Children's Fund, 2000.
[27] Helweg, Otto J., 2000. "Water for a Growing Population Water Supply and Groundwater Issues in Developing Countries," *International Water Resources Association*, 25 (1), pp. 33-39.
[28] Urbonas, B.R. and L.A. Roesner, 1993. "Hydrologic Design for Urban Drainage and Flood Control." *Handbook of Hydrology*. David R. Maidment, ed. New York: McGraw-Hill, Inc. pp. 28.1-28.52.
[29] Calder, Ian R., 1993. "Hydrologic Effects of Land-Use Change." *Handbook of Hydrology*. David R. Maidment, ed. New York: McGraw-Hill, Inc. pp. 13.1-13.50.
[30] Calder, Ian R., 1993. "Hydrologic Effects of Land-Use Change." *Handbook of Hydrology*. David R. Maidment, ed. New York: McGraw-Hill, Inc. pp. 13.1-13.50.
[31] United Nations (UN), 2004.Guidelines for Reducing Flood Losses, available at www.unisdr.org/eng/library/Literature/8689.pdf
[32] Moench, Marcus, 2004. "Groundwater: The Challenge of Monitoring and Management," in Peter Gleick, *The World's Water 2004-2005*, Island Press.
[33] Bartolino, J.R. and W.L. Cunningham, 2003. "Ground-Water Depletion across the Nation." *U.S Geological Survey Fact Sheet* 103-03, November 2003, available at http://pubs.usgs.gov/fs/fs-103-03/
[34] Moench, Marcus, 2004. "Groundwater: The Challenge of Monitoring and Management," in Peter Gleick, *The World's Water 2004-2005*, Island Press.
[35] Bartolino, J.R. and W.L. Cunningham, 2003. "Ground-Water Depletion across the Nation." *U.S Geological Survey Fact Sheet* 103-03, November 2003, available at http://pubs.usgs.gov/fs/fs-103-03/
[36] "Millennium Ecosystem Assessment," 2005. *Ecosystems and Human Well-Being*, Island Press, Washington, DC.
[37] "Millennium Ecosystem Assessment," 2005. *Ecosystems and Human Well-Being*, Island Press, Washington, DC.
[38] Brekke, L.D., Kiang, J.E., Olsen, J.R., Pulwarty, R.S., Raff, D.A., Turnipseed, D.P., Webb, R.S., and White, K.D., 2009, "Climate change and water resources management—A federal perspective." *U.S. Geological Survey Circular* 1331, 65 p. (Also available online at http://pubs.usgs.gov/circ/1331/.)
[39] Bates, B.C., Z.W. Kundzewicz, S. Wu and J.P. Palutikof, Eds., 2008: "Climate Change and Water." Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp.
[40] Bates, B.C., Z.W. Kundzewicz, S. Wu and J.P. Palutikof, Eds., 2008: "Climate Change and Water." Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp.
[41] "The United Nations World Water Development Report 3: Water in a Changing World." *World Water Assessment Programme* (WWAP), 2009. Paris: UNESCO, and London: Earthscan, available at http://www.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml
[42] "TAC Background Paper No. 4," Integrated Water Resources Management Global Water Partnership (GWP), Technical Advisory Committee (TAC) 2000.

[43] Jonch-Clause, T, 2004. "Integrated Water Resources Management (IWRM) and Water Efficiency Plans by 2005: Why, What and How? Global Water Partnership," TEC Report No. 10, Elanders Infologistics Väst AB: Sweden. 45pp.

[44] Biswas, A.K., 2004. "Integrated water resources management: a reassessment." *Water International*, 29(2):249-256.

[45] Bauwer, H., 2000. "Integrated water management: emerging issues and challenges." *Agricultural Water Management* 45(3): 217-228.

[46] Thomas, J-S and B. Durham, "2003. Integrated water resources management: looking at the whole picture." *Desalinization* 156: 21-28.

[47] Cardwell, H., R. Cole, L. Cartwright, and L. Martin, 2006. "Integrated Water Resources Management: Definitions and Conceptual Musings," Pages 8-17 in B. Hooper (Editor). "Integrated Water Resources Management: New Governance, Tools, and Challenges—Selected International Perspectives." *Journal of Contemporary Water Research & Education.* Issue 135.

2.3.C. Water Security and Scarcity: Potential Destabilization in Western Afghanistan (Alex Dehgan, Laura Jean Palmer-Moloney)

Authors and Organizations: Dr. Alex Dehgan, Senior Advisor, Office of the Science and Technology Advisor to the Secretary, Department Of State;
Dr. Jean Palmer-Moloney, Senior Research Geographer, U.S. Army Corps of Engineers, Engineer Research & Development Center, Alexandria, VA (formerly with National Geospatial-Intelligence Agency)
Contact Information: DehganAO@state.gov; laura.j.palmer-moloney@usace.army.mil

Summary

Iran sees current Afghan stabilization and reconstruction efforts--agricultural development and water withdrawal, diversion, and containment projects--in the middle and lower Helmand River, the Farah River, and the Hari Rud watersheds of Afghanistan--as undermining water security in eastern Iran. The concern is particularly relevant in the region known as Sistan and Baluchestan, the most desolate, the poorest, and the most unstable of Iran's provinces. This is Iran's Achilles Heel, and Iran is sensitive to the region's stability. Accordingly, U.S. & Afghan efforts to harness or divert water from these watersheds without resolution of the water dispute between Afghanistan and Iran – one that has endured since 1870 – has encouraged Iran to adopt a mixed strategy of destabilization and cooperation in Afghanistan. Iranian destabilization activities include support for the Taliban and direct action against water diversion targets. Though population growth and consumption patterns are at the root of short-term water challenges in both Iran and Afghanistan, changes in climate are expected to exacerbate the situation (Alcamo et al., 2007; Asian Development Bank, 2009; Zimmerman et al., 2009). Projected climate models of the region predict that less water will flow through these watersheds due to decreased rainfall, and more water will be lost to evaporation due to increased temperatures. This will further undermine stability and increase tensions within countries and between them. Failure to address water concerns has the potential to increase tensions between both countries explosively, slowing or even preventing stability gains in Afghanistan's Western provinces. However, handling water concerns adroitly could encourage closer cooperation on stability and development in Afghanistan.

Background

Five of Afghanistan's six major river basins extend beyond its borders and are a source of tension with its neighbors. Accordingly, any water-based development projects in Afghanistan must take into account the equities of Afghanistan's neighbors and the rights and obligations imposed on Afghanistan by international law, bilateral treaties, and state practice on the use of water from shared basins. At the same time, Afghanistan's immediate future economic growth will occur through agriculture, which requires increasing water use, improving irrigation and distribution systems, and increasing water efficiency. Water is necessary for Afghanistan's development, security, and stability.

The Helmand River is the longest (1,150 km) and among the most strategic of the country's watersheds. It rises in the Hindu Kush Mountains and crosses south-west through the desert of Dashti Margo, where it, along with the Farah River, feed into the Sistan Depression at the Afghan-Iranian border, a large complex of wetlands, lakes and lagoons, internationally-recognized as a haven for wildlife. Water of the Helmand River and its tributaries is used

extensively for irrigation in the middle and lower reaches of the watershed, in Kandahar, Helmand, and Nimruz provinces of Afghanistan, and in the Sistan and Baluchestan province of Iran, part of the Persian granary. The Helmand River Basin is inhabited by more than 7 million people.

The Helmand River is a renewable water resource, but it is vulnerable to overuse and seasonal exhaustion. Though the river and its tributaries are fed annually by melting snow from the highlands of central Afghanistan (Favre et al., 2004), recent below-average snowfall and reduced water in the snow pack have led to reduced water in the overall drainage system. Two major dams, constructed circa 1950 (the Kajakai dam and the Arghandab dam), help control seasonal flooding and release water into the river during the seasonally dry times (FAO, 1962; US Army Geospatial Center, 2009; US Army Engineer Research and Development Center, 2002). Predictable year-round flow of surface water is not guaranteed, and groundwater drawn from wells is the traditional and most reliable source of water for communities in the area. US forces bombed the Kajakai dam on Halloween, 2001, and started reconstruction of its 3 turbines in 2004.

Water Insecurity in Iran & Afghanistan

Iran's agricultural sector accounts for about one-fifth of its GDP, employs one-third of its workforce, and is necessary for the country's goals of achieving food absence. Although developments of agriculture and food self-sufficiency are important goals of the Iranian government, strong population growth in recent decades has left Iran unable to grow enough food to feed its population. The processes of urbanization, industrialization and the development of irrigated agriculture to support population growth have raised the demand for water, and simultaneously reduced the supply. This increased demand has drawn down Iran's aquifers. More than 90 percent of the country's renewable water resources are used in agriculture, but because of low irrigation & transport efficiency, about 50-60% of the water is lost during the process.

Although only ten percent of Iran's water resources originate outside the country, those regions dependent on transboundary water resources are highly susceptible to its scarcity. In 1999, the Taliban-led Afghan government reduced the flow of the Helmand River to Iran by 98 percent at the peak of tension with Iran and in the midst of a 5-year drought. As a result, Iran's Hamun Lake dried up and sandstorms buried more than a hundred villages and destroyed farmland. Iran is dependent on 80% of the Helmand's current downstream flow for agriculture. This region is economically important to Iran, having frequently been referred to in the historical documents as the "bread basket" of Central Asia. However, Sistan's traditional economic fortune has diminished as a result of the diminishing water supplies from the Helmand. Iran has similar concerns over the planned Salma Dam of the Hari Rud River, which forms the border between Iran, Turkmenistan, and Afghanistan. Iran perceives the construction of the Salma dam, like current NATO/USAID efforts to rehabilitate the Kajakai dam on the Helmand River, as a direct security threat.

In Afghanistan, at least 97 percent of the water being withdrawn from the Helmand Basin is used in the agriculture sector. Moreover, 90% of the Afghan population is dependent on natural resources management – primarily animal husbandry and agriculture (Batson, 2008). Future agricultural development—a key portion of Obama's strategy to rapidly increase economic activity as a bulwark against the insurgency—will depend on having sufficient water resources

for expansion (Wegerich, 2009). Although Afghanistan has theoretically sufficient water for its needs, it is hampered by poor infrastructure damaged by 30 years of war. Population growth and climate change also further increasingly undermine Afghanistan's water security.

Predicted Changes to Water Availability

Proposed population growth, increased agricultural expansion, and climate change conditions are likely to exacerbate existing tension with Iran over dams and other water control features on the Helmand River. The most recent climate projections from the Intergovernmental Panel on Climate Change, coupled with downscaled modeling of the IPCC global predictions for the region, suggest that water stress will continue to increase in southeast Iran and southwest Afghanistan (Alcamo et al., 2007). As the population in the lower Helmand watershed grows, it will place more demands on water supply for domestic consumption, business, and food production. Current reconstruction efforts in Afghanistan – particularly in the agricultural sector – are proceeding haphazardly, with no understanding of the total impact such activities will have on the water flow to Iran and other countries, or on the existing aquifers in Afghanistan. As seen in examples from other arid regions, i.e., Iraq (Lightfoot, 2009) and Pakistan (Mustafa, 2007), when new groundwater wells are drilled and water is pumped from deep aquifers, the water table can fall rapidly in some places, at least disrupting water availability and at worst causing entire villages to be abandoned (Lightfoot, 2009; Warner et al., 2009). There is insufficient systematic monitoring of ground water tables, salinization, water flow, snow pack, and rainfall to understand how water flow and availability is changing in the Helmand basin. Similar problems exist for the other Afghan river basins. Moreover, a 30-year gap in Afghanistan's hydrometeorological data (Williams-Sether, 2008; Whitney, 2006) further undermines Afghanistan's ability to manage its water resources wisely.

The decreased water flow in the region not only affects the water table, but also increases desertification. Sandstorms that scoured southwestern Afghanistan in early June-Sep 2003 - called the worst in living memory by residents of the area - buried villages, filled waterways, destroyed crops, and killed livestock. Though locally strong wind (known as the Wind of 120 Days) occurs annually (Whitney, 2006), wind-generated sandstorms persisted longer than expected in the recent years of drought, creating a huge environmental problem for the residents of this region. Most of the windblown dust originated in the Sistan Basin wetlands in the beds of the dried out hamuns (lakes). Persistent drought conditions there, coupled with increased irrigation drawn from the Helmand River, have quickly turned these wetlands into arid salt pans (IRIN 2007, UNEP, 2003 and 2006).

Negotiations between Iran and Afghanistan

Although the headwaters of the Helmand watershed is within Afghanistan's territorial boundaries, Afghanistan's rights to the use of that water are limited under international law, state practice, and long-standing water sharing agreements with Iran. Under customary international law, no riparian state has the exclusive right to the use of a transboundary river; instead each state's sovereign right to develop a river is limited by the equal rights of other states to develop. The Goldsmid Arbitration Award in 1872 gave equal shares of the water in the lower reach of the river to both Afghanistan and Iran. A key portion of the award language provides:

> "It is, moreover, to be well understood that no works are to be carried out on either side calculated to interfere with the requisite supply of water for irrigation on both banks of Hirmand."

After the Helmand River changed its course in 1896 and after a severe drought in 1902, the two governments attempted through arbitration to reach agreements in 1905 and again in 1938, but were unsuccessful. On September 7, 1950, the Afghan and Iranian governments signed the "Terms of Reference of the Helmand River Delta Commission and an interpretative statement relative thereto, agreed by Conferees of Afghanistan and Iran" and established the Helmand River Delta Commission, with the assistance of the US Department of State. This Commission was designed to elaborate on the technical methods for sharing waters of the Helmand River, and was to be composed of three technical experts from disinterested states, with powers of recommendation only. The 1951 recommendation of the Commission was rejected by both countries and the dispute over the division of water continued.

Building of the Kajakai Dam in 1953 increased in-stream flow during the dry season but reduced the flood waters on which pastoralists were dependent for fertilization. Removing water from the river system upstream from the historical point-of-measurement (at Kamal Khan) complicated water sharing dialogue. Iranian Prime Minister Amir Abbas Hoveida and Afghan Prime Minister Mohammad Musa Shafiq signed an accord in 1973 that determined the specific amount of water that should flow into Iran: 22 cubic meters per second--with an option for Iran to purchase an additional 4 cubic meters per second in "normal" water years. However, this agreement was never fully implemented due to the Afghan coup d'état, the Soviet invasion of Afghanistan, the Islamic revolution, and tensions between the Wahhabist Taliban and the Shia Tehran government. Although Iran and Afghanistan have held discussions with the Karzai government under the framework of the 1973 agreement, they have been inconclusive to date.

Current Status

The Iranians have been eager to negotiate the status of the Helmand, Farah, and Hari Rud rivers in Afghanistan. They have tried to get Afghanistan to sign agreements over water, but the Afghans have continued to demur claiming they lack expertise and capacity to participate in transboundary water negotiation (EastWest Institute, 2009). The Afghans are correct in noting that there are huge data gaps about local hydrology and water use that would need to be filled before negotiations were conducted, but at the same time, Afghanistan is, with US assistance, seeking to use and capture an increasing share of the Helmand, Farah, and Hari Rud rivers through agricultural projects and new dam construction and revitalization without conducting a comprehensive study on how the net effect of these individuals projects would affect overall water flow and availability (USAID Water Atlas, 2008). These projects have raised concerns in Tehran over the geopolitical use of water. Furthermore, international donors are unlikely to be willing to fund needed infrastructure in Afghanistan if there is no agreement governing water use between the two countries.

For Iran, this is a serious security issue. Baluchistan is one of Iran's most volatile regions, and Iran is sensitive to fears of foreign interference seeking to destabilize the Baluch population. Since 2004, Iran's Sistan and Baluchistan province has witnessed regular attacks, bombings, and kidnappings at the hands of Jundullah, an ethnic Sunni Baluch group, which Iran has accused the United States of supporting (the most recent attack was on Oct. 15, 2009). Moreover, Baluchistan remains Iran's poorest and least educated region, and the population already feels disenfranchised by Tehran, isolation made worse by the blatantly fabricated results of Iran's June elections. Iran's reported support for different Taliban factions in Western Afghanistan could be potentially linked to its concerns over its water security, and Iran will likely continue to

destabilize Western Afghanistan if the water dispute is not resolved. In March 2009, Afghan security forces found a cache of Iranian-made explosives and ammunition around the Bakhashabad Dam in Farah Province. The 2.2 million U.S. dollar construction project will substantially bolster the water and power supply to Afghan communities, but at the cost of reducing the flow of water into neighboring Iran.

Discussion & Strategy

Afghanistan cannot proceed to build new dams, restore old dams, and take an increasingly greater fraction of shared groundwater and surface water resources from the Helmand, Farah, and Hari Rud rivers in Afghanistan without serious consequences if it does not equitably address the transboundary water issues with Iran. Moreover, international donors will not fund these projects if an agreement with Iran, the downstream riparian, is not reached. As Afghanistan claims that it lacks the capacity and data to enter into any such negotiations, the US should be focused on increasing capacity on the technical level, creating monitoring systems that are transparent and shared, and building mechanisms for technical cooperation to the benefit of both countries. The proposed strategy implements this approach.

1. Develop transparent and shared monitoring systems and databases for Afghanistan's transboundary river basins to improve the scientific basis for any agreement. The major technological and knowledge deficit in the water sector restricts prospects for efficient management and use of water resources and hinders the policy development process (Homayoonnezhad et al., 2008; Meybeck et al., 2004; Peyman et al., 2006; Rahimzadeh et al., 2008; and Van Beek et al., 2008). Decision-makers need to know the quantity and quality of water in transboundary river basins before moving forward with upstream projects that will affect in-stream flow as well as ground water levels. Accordingly, the international community—possibly under the direction of UNESCO, UNDP, or UNEP —should continue to help Afghanistan put in stream gauges, map glacier melt, record changes in precipitation and snow fall, and monitor water table depth and inflows and outflows to underground aquifers associated with those river basins. Based on these data, dynamic models of water flow for each transboundary basin should be developed to predict the potential impacts of agricultural projects, population growth, climate change, and infrastructure projects on downstream water flow and quantity. Development of this knowledge base would allow the parties to negotiate water sharing agreements that are more sensitive to and cognizant of seasonal, interannual, and long-term changes in water flow in Afghanistan and throughout the region. These data should be available on-line in close to real time, and shared with surrounding countries. The monitoring stations and infrastructure should be opened to inspections by participating countries, as well as observer states and organizations. The international community should look beyond aid delivery at just the national level and incorporate regional water strategies into policy development and aid programming, and work with the UN Assistance Mission to Afghanistan to improve coordination of water-related donor aid.

2. Identify historical hydrometeorological baseline data to better understand changes in water availability in Afghanistan's river systems. Afghanistan's 30 years of civil war have interrupted data collection on the amount of water moving through its river basins since the agreements negotiated in the 1950s-1970s (Helmand River Delta Commission Report, 1951).

Discussions with scientists at the National Center for Atmosphere Research suggest that Afghanistan's river systems probably have less water in the last fifty years due to global warming, similar to patterns found in other transboundary river systems in the region, and may see further decreases in the future. There are ways to develop proxies for past changes in water distribution and flow, including looking at past records of well depth, time series remote sensing imaging of glaciers, snow cover, and rivers, and pre-existing river gauges. The United States should assist Afghanistan in establishing historical baseline information on water flow – such as obtaining the Fairchild Aerial Survey dataset and sharing commercial remotely sensed images of its critical watersheds. Establishing the extent of previous water flow is important to adequately predict future changes, and to adjust previous water sharing agreements in light of current realities.

3. Create a Transboundary Commission between Afghanistan & Iran to manage transboundary water resources. The Transboundary Commission should serve as a formal body to negotiate water sharing agreements, collect and share data, and building confidence and capacity on both sides. It should have as its initial mandate the creation of a technical working group of 3 representatives from disinterested countries (agreed upon by Iran and Afghanistan) to review the entire course of the Helmand and to assess existing and planned infrastructure, agriculture production, and population density and growth. The technical working group would oversee data collection, monitoring, and modeling of the Helmand River Basin (including the Farah River), assess future changes, and make recommendations.
4. Build capacity for water management and increase irrigation efficiency. The United States, U.N. and international organizations and agencies, and partner states, should build capacity in Iran and Afghanistan to collect, interpret, model, and maintain the water databases, to better understand international law governing shared water resources, and to build capacity for negotiations (Chavoshian et al., 2005; McMurray et al., 2005). Technical officials from these states should study to other regional water sharing commissions, such as the Mekong River Commission, Colorado River Commission and the Nile River Basin Commission.
5. Build an international support network for negotiations on water. The United States is no longer a disinterested country, nor necessarily a trusted partner, for either party to the dispute. Accordingly, the US should continue efforts to build an international support network for resolution of the transboundary water dispute. The US should push for expansion of the existing Water Sector Group – which includes the United Nations Assistance Mission in Afghanistan, the Canadian International Development Agency, the United States Army Corps of Engineers, the European Commission, the German Development Association (GTZ), USAID, and the Dutch Embassy to include other potential donors and implementing agencies. These may include the UNDP, UNEP, OSCE, World Bank, and Asian Development Bank, as well as key NGOs such as the Wildlife Conservation Society, which work in both countries.

Conclusion

Afghanistan's plans for upgrading and developing its water infrastructure on each of its major river basins are aimed at exploiting the irrigation and energy potential of its rivers and at mitigating floods. While crucial to the social and economic development of Afghanistan, these plans will also impact transboundary water flow and as a result Afghanistan's relations with its neighbors. Failure to address water concerns in the middle and lower Helmand River, the Farah River, and the Hari Rud watersheds can have the potential to increase tensions between Iran and

Afghanistan explosively, slowing or preventing stability gains in Afghanistan's West, and interfering with the success of important development projects in agriculture and energy generation. Moreover, Iran increasingly sees threats to its water security as threats to its own security, and will seek to destabilize the region, possibly through support for its natural adversaries, such as the Taliban. However, the adroit handling of water concerns could encourage closer cooperation with Iran on stability and development in Afghanistan and potentially even create a framework for U.S. cooperation with Iran.

Appendix 1: Projected Changes in Global Water Stress

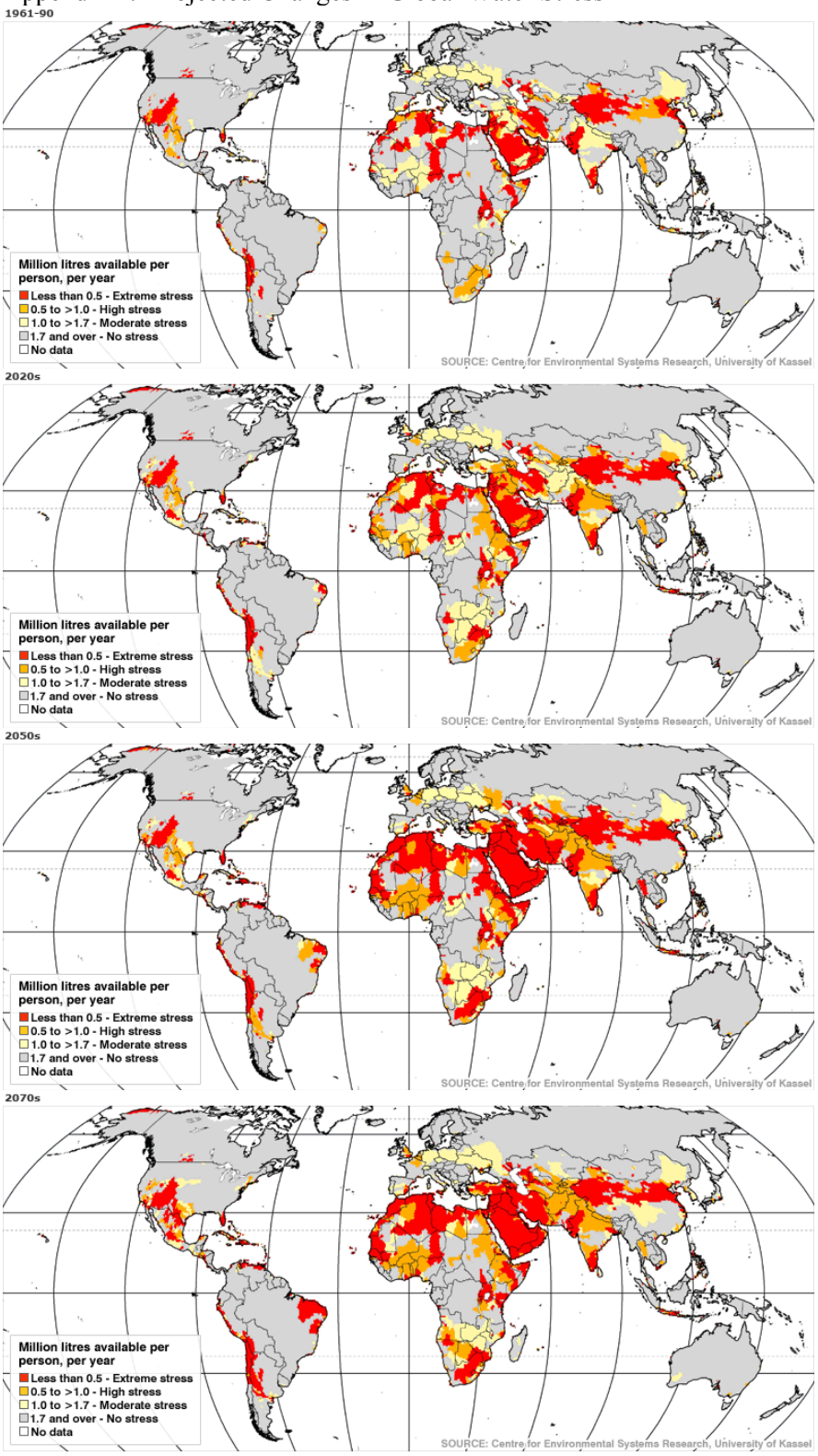


Figure 1: Projected changes in global water stress. Source: Alcamo, J. et al. 2007. "Future long-term changes in global water resources driven by socio-economic and climatic changes," Journal of Hydrological Sciences 52(2).

Appendix 2.

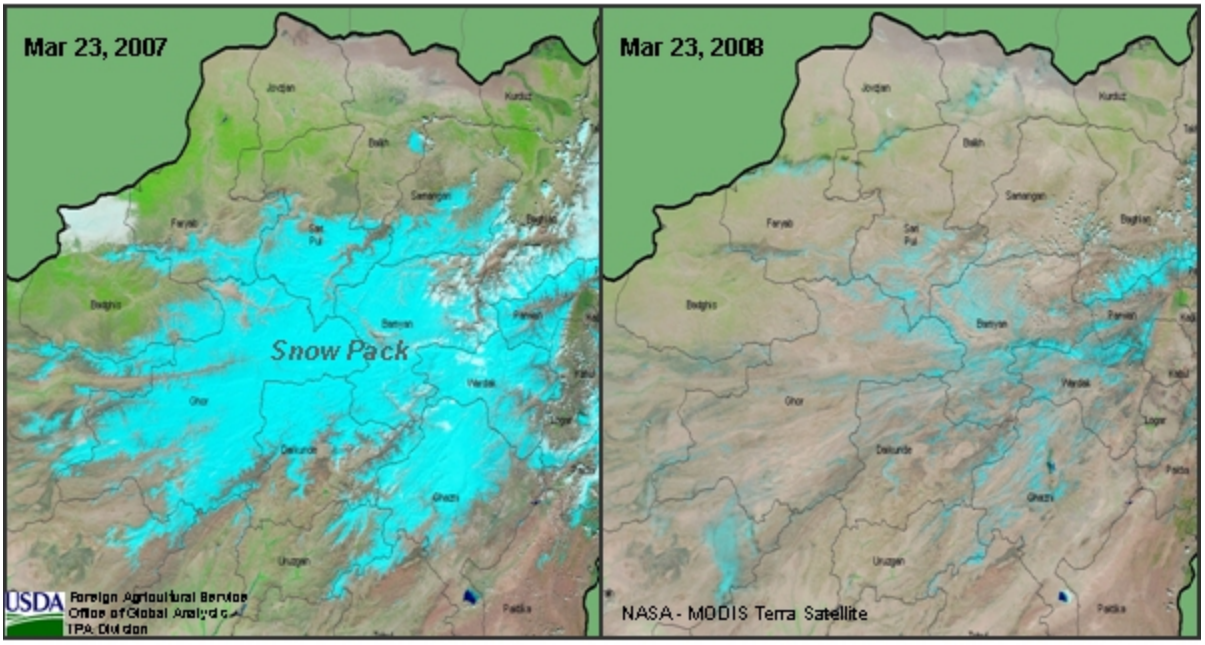


Figure 2. Central Afghanistan Winter Snow Pack Comparison (Though 2 years of data do not indicate a "trend," these examples illustrate the wide range of coverage difference. N.B.: These images do NOT address the amount of water that is in the snow pack.)

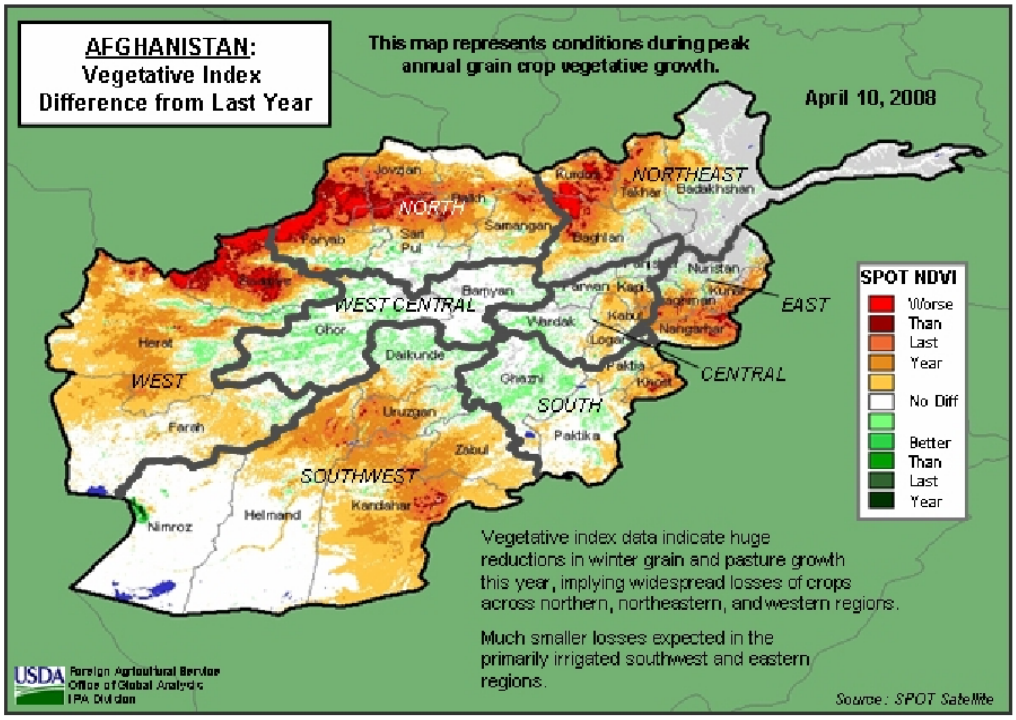


Figure 3. Change in crop production from 2007-2008. (Note "better than last year" color in Nimruz Province. If water is redirected upstream, less will be available for irrigation in this potentially productive area.)

Source for Figures 2 and 3: USDA Foreign Agricultural Service, Commodity Intelligence Report. "AFGHANISTAN: Severe Drought Causes Major Decline in 2008/09 Wheat Production" August 12, 2008.

↑ Back to the top

References:

Alcamo, J., M. Florke, and M. Marker. 2007. "Future long-term changes in global water resources driven by socio-economic and climatic changes," in Hydrological Sciences Journal, Vol. 52, Issue 2, (April), pgs. 247-275.

Arbitral Opinion of General Sir Frederic Goldsmid, arbitrator for adjustment and settlement of the rights of Iran and Afghanistan with regard to land and water problems of the Seistan, 1872. Unpublished manuscript. (When we have received permission to share the original document, access information will be provided.)

Asian Development Bank (ADB) 2009. Climate change threatens water, food security of 1.6 billion South Asians. Available from Relief Web, http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-7VH7TW?OpenDocument [Accessed 06 Sept 2009].

Award of Arbitration by Col. I. H. McMahon, Special Envoy of H. M. the King of England and Emperor of India and Arbitrator for the delimitation of the frontier between Iran and Afghanistan and the water questions in Seistan, 1905. Unpublished manuscript. (When we have received permission to share the original document, access information will be provided.)

Bates, B., Z.W. Kundzewicz, W. Shaohong, and J. Palutikof (eds.) 2008. Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp.

Batson, D. 2008. Registering the Human Terrain: A Valuation of Cadastre. Washington, DC: National Defense Intelligence College.

Chavoshian, S. A., K. Takeuchi, and S. Funada. 2005. "An Overview to Transboundary and Shared Water Resources Management in Iran, Technical Challenges and Solutions," in Role of Water Science in Transboundary River Basin Management. (volume editor and publisher not given. Place of publication given as "Thailand.")

Collins, M., CMIP Modelling Groups (BMRC (Australia), CCC (Canada), CCSR/NIES (Japan), CERFACS (France), CSIRO (Australia), MPI (Germany), GFDL (USA), GISS (USA), IAP (China), INM (Russia), LMD (France), MRI (Japan), NCAR (USA), NRL (USA), Hadley Center (UK), and YNU (South Korea)). 2005. "El Nino or La Nina-like climate change?" in Climate Dynamics, vol 24, pgs. 89-104 (DOI 10.1007/s00382-004-0478x)

Dept of State Information Memorandum. 2009. "Afghan Trans-boundary/Riparian Rights Dialog—time for Action," from ECON- Tom Engle, Counselor to Anthony Wayne.

East West Institute. 2009. Symposium "Alternative Futures for Afghanistan and the Stability of Southwest Asia: Improving Regional cooperation on Water. Session 4: The Helmand River Basin and the Harirud and Murghab Rivers," 25 June 2009, EastWest Institute Brussels Centre, Brussels, Belgium. See http://www.ewi.info/water-security-southwest-asia [Accessed 06 Sept 2009]

Food and Agriculture Organization (FAO) of the United Nations. 1962. Survey of Land and Water Resources: AFGHANISTAN General Report. New York: United Nations Special Fund.

Favre, R. and G. M. Kamal. 2004. Watershed Atlas of Afghanistan: 1st Edition Working Document for Planners. AIMS and FAO supported publication.

(need to get web address!)

Heath, R. 1983. Basic ground-water hydrology: U.S. Geological Survey Water-Supply Paper 2220, 86 p.

Helmand River Delta Commission Report and Recommendation on the Distribution of Helmand Water, 1951. (This information came from the unpublished draft manuscript of The Hirmand River Dispute.)

Homayoonnezhad, P. I., Amirian, and I. Piri. 2008. "Investigation on Water Quality of Zabol Chahnimeh Reservoirs, The Effective Means for Development of Water Management of Sistan and Baloochestan Province, Iran," in world Applied Sciences Journal, vol. 5, issue 3, pgs. 378-382.

IRIN NEWS.ORG. 30 July 2007. AFGHANISTAN: Environmental crisis looms as conflict goes on. United Nations Office for the Coordination of Humanitarian Affairs. Available from http://www.globalsecurity.org/military/library/news/2007/07/mil-070730-irin01.htm [Accessed 06 Sept 2009].

Lightfoot, D. 2009. Survey of Infiltration Karez in Northern Iraq: History and Current Status of Underground Aqueducts. UNESCO publication. Accessed 08 November 2009 at http://www.iauiraq.org/reports/UNESCO_Karez_survey_report_FINAL.pdf

McMurray, J.C. and d. Tarlock. 2005. "The Law of Later-Developing Riparian States: The Case of Afghanistan," in NYU Environmental Law Journal, vol. 12, pgs. 711-763.

Meybeck, M. and C. Vorosmarty. 2004. "The Integrity of River and Drainage Basin Systems: Challenges from Environmental Changes," in Kabat, P., et al. Eds., Vegetation, Water, Humans, and the Climate: A New Perspective on an Interactive System. Global Environmental Action International Conference, Tokyo, Oct 14-16.

Mustafa, D. and M. U. Qazi. 2007. "Transition from Karez to Tubewell Irrigation: Development, Modernization, and Social Capital in Balochistan, Pakistan," in World Development, vol. 35, number 10, pp. 1796-1813.

Peyman, D. A. And M. Tajrishy. 2006. "Estimation of Free Water Evaporation from hamun Wetlands Using Satellite Imagery," Available from http://ewrc.sharif.edu/pdf_folder/01%20Arasteh.pdf [Accessed 06 Sept 2009].

Rahimzadeh, F., A. Asgari, and E. Fattahi. 2008. "Variability of extreme temperature and precipitation in Iran during recent decades," in International Journal of Climtology, DOI: 10.1002/joc.1739.

USDA Foreign Agricultural Service, Commodity Intelligence Report. 2008. "AFGHANISTAN: Severe Drought Causes Major Decline in 2008/09 Wheat Production" August 12, 2008

UNEP. 29 January 2003. UNEP report chronicles environmental damage of the Afghanistan conflict. Available from http://new.unep.org/Documents.Multilingual/Default.asp?DocumentID=277&ArticleID=3201&l=en [Accessed 17 April 2009].

UNEP Post-Conflict Branch. 2006. History of Environmental Change in the Sistan Basin Based on Satellite Image Analysis: 1976-2005. Nairobi, Kenya: UNEP/PCoB.

USAID. 2008. Water Atlas (Afghanistan), sections 19-20 on middle and lower Helmand and section 21 on Sistan watershed contain discharge data from 1978.

U. S. Army Engineer Research and Development Center. 2002. Water Resources Areal Appraisal of Afghanistan. Alexandria, VA: USACE Topographic Engineering Center.

(This publication notes that the water table may have wide seasonal fluctuations. Recommends "extensive reconnaissance and exploratory drilling are necessary to locate zones of maximum yield and best quality water.")

U.S. Army Geospatial Center. 2009. Provincial Water Resources Data Summary of Helmand (Province), Afghanistan. Alexandria, VA: USACE AGC. Local Reality: Afghanistan Research and Evaluation Unit Issues Paper Series. Available from www.areu.org.af [Accessed 25 May 2009].

Whitney, J. W. 2006. Geology, water, and wind in the lower Helmand Basin, southern Afghanistan: U.S. Geological Survey Scientific Investigation Report 2006-5182, 40 p.

Williams-Seth, T. 2008. Streamflow Characteristics of Streams in the Helmand Basin, Afghanistan. USGS Fact Sheet 2008-3059.

Zimmerman, J. B., J. R. Mihelcic, and J. Smith. 2009. "Global Stressors on Water Quality and Quantity," in Environmental Science & Technology, (June 15), pgs. 4247-4254.

(This publication only covers the middle Helmand River in Helmand Province.)

Van Beek, E., B. Bozorgy, Z. Vekerdy, and K. Meijer. 2008. "Limits to agricultural growth in the Sistan closed Inland Delta, Iran," in Irrigation Drainage Systems, vol 22, pgs. 131-143.

Warner, K., C. Ehrhart, A. deSherbinin, S. Adamo, and T. Chai-Onn. 2009. In Search of Shelter: Mapping the Effects of Climate Change on Human Migration and Displacement. Available from http://ciesin.columbia.edu/publications.html [Accessed 20 July 2009]

Wegerich, Kai. Water Strategy Meets

2.4. Changing of Climate and State Weather

2.4.A. Maintaining Geopolitical and Social Stability throughout a Global Economy in an Era of Climate Change (James Diaz)

Author and Organization: Dr. James Diaz, Professor of Public Health and Preventive Medicine, Director, Environmental and Occupational Health Sciences, Schools of Public Health and Medicine, Louisiana State University
Contact Information: jdiaz@lsuhsc.edu

Introduction

In its 2007 Fourth Assessment Report, the Intergovernmental Panel on Climate Change (IPCC) concluded that mean land surface temperatures had increased by 0.74° C every decade since.[1] Thus, there now appears to be uniform agreement in the international scientific community that the earth is warming from a variety of climatic effects. Although the exact contributions of climatic effects on global warming are unknown, the major climatic effects include long-term cyclical weather patterns and, even more importantly, manmade greenhouse gas emissions and their cascading effects on melting polar icecaps and rising sea levels.

Most reports to date of the public health impact of global warming have featured heat stroke deaths following heat waves, drowning deaths in floods and tsunamis, and mosquito-borne disease outbreaks following prolonged drought-flood cycles. As nighttime temperatures warm and glaciers and permafrost retreat into the highlands, the geographic distribution ranges of mosquito and tick-borne diseases have now extended to higher altitudes and to new regions in formerly disease-free areas with competent insect vectors. The ultimate effects of global warming on rainfall, drought, tropical cyclone (hurricane), and tsunami activity will have far reaching human health and public safety impacts; not only on weather-related infectious disease outbreaks, but also on world food production and supplies, access to clean drinking water, immigration, urban relocation, and civil strife. The ultimate challenge for the US will be to assume leadership in maintaining geopolitical and social stability throughout the global economy in an era of climate change.

For starting targets, the World Health Organization (WHO) has now specifically identified and ranked the five major health consequences of climate change that need to be addressed to preserve the global economy and to reduce worldwide civil strife, including piracy, coups d'états, sectarian and tribal violence. The WHO's five major health consequences of climate change include (1) the vulnerability of food supplies to droughts, floods, over-fishing, and poor agricultural practices, such as soil conservation, with consequent malnutrition, famine, and refugee relocation, (2) waterborne infectious disease outbreaks, especially cholera in Africa and Bangladesh, following prolonged drought, and poisoning from contaminated water wells (usually by arsenic in Bangladesh), (3) diarrheal diseases following more frequent drought-flood cycles, (4) heat waves in mega-cities with high thermal masses (also known as urban "heat-islands") with increasing heat-related morbidity and mortality in the elderly with chronic cardiovascular diseases and in persons with ozone-sensitive asthma, and (5) vector-borne disease outbreaks, especially malaria, Chikungunya virus, and dengue fever, following more frequent drought-flood cycles, in expanded geographic distributions.

How Good are the Prediction Models?

Since the 1980s, the outcomes of climatic change on human health have received more attention, climate-change-associated infectious disease outbreaks have been described and investigated, and several climate prediction models have been developed by nations and intergovernmental panels. Recently, these global climate models have become even more comprehensive and include more detailed analyses at finer geospatial resolutions of ocean temperature fluctuations, ice cap and sea ice thickness, carbon cycle dynamics, greenhouse gas emissions, outgoing long wave radiation effects (as surrogates for drought conditions), and vegetation indices (as surrogates for rainfall). Some of the areas of convergence and divergence among global climate change models are compared in Table 2.4.A-1.[2,3]

Table 2.4.A-1. The Predictive Models of Climate Change

Climate Changes	Intergovernmental Panel on Climate Change-Fourth Assessment Report	British Hadley Centre for Climate Prediction and Research	Canadian Centre for Climate Modeling and Analysis
Temperature	***Uneven Warming***	***Slower Warming***	***Faster Warming***
	Troposphere > Stratosphere		
	Land > Sea		
	Higher > Lower Latitudes		
	Night > Day		
Drought	***Regional Increase***	***Regional Increase***	***Widespread Increase***
	Eastern Hemisphere > Western Hemisphere	Eastern Hemisphere > Western Hemisphere	Eastern Hemisphere > Western Hemisphere
	Southern Hemisphere > Northern Hemisphere		
Precipitation	***Uneven Increase***	***Regional Increase***	***Limited Increase***
	Rainfall > Snowfall	Rainfall > Snowfall	Rainfall > Snowfall
	Coastal > Interior		
	Northern Hemisphere > Southern Hemisphere		
	Western Hemisphere > Eastern Hemisphere		
Tropical Cyclones (Hurricanes)	***Increasing***	***Increasing***	***Increasing***
	Both Atlantic and Pacific Oceans	Atlantic Ocean > Pacific Ocean	Pacific Ocean > Atlantic Ocean

Demographics and Human Behaviors in a Global Economy

Besides climatic and ecosystem changes, many other factors will influence the onset of emerging and re-emerging infectious disease outbreaks, including population demographics, the globalization of trade and travel, the globalization and industrialization of the world's food supplies, new vector-pathogen-host relationships, human health behaviors, and human host susceptibilities.

Today's Population Demographics

The nations that were the first to organize around large cities, such as London, Paris, New York, and Tokyo, are the most developed nations today, offering more opportunities and services at less cost than dispersed rural areas. Today more than 50% of the world's populations of over 6 billion people are urban dwellers, and more than 90% of the world's population lives within less than 100 miles of seacoasts.[4] Soon the five largest mega-cities will be Tokyo, Mumbai, Dhaka, Sao Paolo, Mexico City, and Delhi, each with over 20 million residents.[5] The greatest threats modern mega-cities face now (and will face into the foreseeable future) come from global warming: a finite supply of fossil fuels and freshwater, terrorism, infectious disease pandemics, and natural and man-made disasters. Coastal city dwellers in Galveston and New Orleans, for example, may have to relocate to higher ground to avoid hurricane storm and tidal surges and combine with more inland cities, such as Houston and Baton Rouge, for improved protection and greater economic opportunity. Terrorists will continue to exploit available technologies, including abandoned nuclear warheads and bio-weaponized microbes, in order to attack person-dense targets of opportunity (in hotels as in the incident in Mumbai on November 26, 2008, in high-rise office buildings as in the September 11, 2001, World Trade Center attacks, in airports, train, and subway stations as in the Tokyo and Madrid incidents, as well as in stadiums, water treatment facilities, chemical and nuclear power plants). Just as the Victorians of London acclimatized to epidemic threats from cholera in the 1850s, modern city residents must acclimatize to modern-day threats of natural and man-made disasters.

Nevertheless, new or returning infectious disease threats will remain ever present. According to the WHO's World Malaria Report, 2005, 3.2 billion people live in malaria-endemic regions in 107 countries and territories, and there are between 350 million to 500 million cases worldwide per year, with 840,000 to 1.2 million deaths from malaria annually.[6] Today, malaria is most often imported into the industrialized world by an infected migrant worker or air traveler, and imported malaria cases are now increasing and expected to continue increasing. A ranking of the most common lethal infectious diseases in the world today include malaria, tuberculosis (TB), human immunodeficiency virus (HIV, the AIDS pathogen), and influenza. Malaria is now hyperendemic in sub-Saharan Africa and Asia, HIV and TB are pandemic worldwide, and influenza re-assorts its genes with animals and man periodically to threaten recurring pandemics of avian flu (H5NI avian-origin influenza A) and swine flu or novel influenza A (H1N1).[7]

The Vulnerability of a Globalized Food Supply: The Melamine Tragedy

In early 2007, the US Food and Drug Administration (FDA) ordered a national recall of most commercial pet foods following more than a dozen deaths and hundreds of cases of kidney-related illnesses in domestic cats and dogs. This was the largest food recall in US history. A subsequent trace-back investigation revealed that the wheat flour in the pet foods was intentionally contaminated in China with melamine, a protein look-alike used in fertilizers and plastics, and a related chemical, cyanuric acid, known to cause kidney stones in animals.[8] Despite the recall and precise identification of the nephrotoxins, the Chinese continued to add melamine and cyanuric acid to food products, specifically watered-down infant formulas, and other dairy products, to boost their protein contents and market prices.[9] In September 2008, the Chinese reported that over 50,000 infants and children had been hospitalized in China after consuming melamine-contaminated dairy products, with more than a dozen deaths from kidney stones and kidney failure. In November 2008, the FDA detected melamine and/or cyanuric acid (the combination of both is the most toxic) in at least four US food products; and the Director of

the FDA Center for Food Safety and Applied Nutrition made the following pronouncements: (1) the Chinese dairy products were intentionally adulterated with high levels of melamine and cyanuric acid not ordinarily present in legitimately produced dairy products, (2) melamine could leach out from formula can liners, (3) cyanuric acid was often used to sanitize equipment used in the production of infant formula, and (4) further testing and monitoring of infant dairy products were not planned due to an inability to detect low adverse effect levels. Consumer-watchdog groups have noted that these lowest adverse effect levels, especially those resulting from the combination of toxicants, have yet to be determined. Further domestic US and in-country testing and monitoring of Chinese production practices would appear to be indicated.[10]

The melamine tragedy highlighted the serious hazards of the globalized food chain and underscored the inability of the FDA to protect the US food supply in a global economy. The FDA record on assuring the safety of imported drugs is no better given the recent deaths and allergic reactions to contaminated Chinese-manufactured heparin administered to thousands of US patients for temporary anticoagulation during cardiovascular interventions, open heart surgery, and intravenous infusions.

The Globalization of Trade and Travel: Sea Trade

Infected arthropod vectors, infected animal hosts such as rodents and exotic pets, and even virulent microbes travel well at sea, especially in hot and humid ship cargo holds and in cargo containers. Recent outbreaks of cholera in Ecuador and Peru have followed increased shipping trade with Southeast Asia, where *Vibrio cholerae* is endemic in coastal estuaries. Traceback investigations have demonstrated that container vessels pump contaminated saltwater ballast into their hulls in their home ports for smoother transoceanic sailing and then discharge ballast prior to unloading in distant ports with warming estuaries.[11] Such practices can effectively import cholera bacteria, marine viruses, and even harmful algae to new, warming marine ecosystems and fisheries, causing microbial and algal toxin contamination of shellfish beds and regional marine fisheries.[12]

Monkeypox Arrives by Sea, and Cowpox Will Too

Recently the exotic pet trade imported monkeypox to the US in infected pet rodents. In 2003, 53 cases of monkeypox were reported among three Midwestern US states—Illinois, Indiana, and Wisconsin.[13] Fortunately there were no deaths, but 14 patients were hospitalized for supportive treatment, including one child with encephalitis. The monkeypox virus, a smallpox-like orthomyxovirus, was first isolated in the Congo River basin of West Africa in 1968. After an incubation period of 1-2 weeks, monkeypox is characterized by an onset of headache, fever, fatigue, and backache, followed by a characteristic rash. The monkeypox rash is similar to smallpox and cowpox, with evolving macules, vesicles, and pustules that crust over and heal with 14-21 days. Unlike smallpox, pronounced lymphadenopathy (swollen local lymph nodes) is usually present, and complications may include pneumonia and encephalitis. Although person-to-person transmission may occur, infection is usually transmitted by contact with contaminated animals.[14] The case fatality rate (CFR) ranges from 1-10%, with higher CFRs in young children.[15]

The source of the monkeypox virus introduced into the US was later identified as infected giant Gambian rats imported from Ghana to Texas and sold to an Illinois pet distributor, who housed the animals together prior to sale to pet shop owners and others. Although the monkeypox virus is endemic in West African river basins, including the Gambia and Congo, the warming, humid

ecosystem of the massive Mississippi River drainage basin in the Midwestern US clearly supported the transmission of the virus from imported rodents to domestic rodents, especially prairie dogs and squirrels, and to rabbits. The close contact between pet owners and their pets permitted the transmission of the zoonotic infection, for which there is no specific treatment, to humans. If sentinel animal surveillance and disease control initiatives fail, monkeypox, cowpox, and other exotic pox viruses will become unwelcomed but firmly established zoonosis in the US as a direct result of relaxed free trade regulations, the exotic pet trade, domestic animal importation, and a warming, welcoming ecosystem.

This Vector Came by Sea, but its Pathogens will Come by Air

The worldwide spread of the Asian tiger mosquito, *Aedes albopictus*, by imported tire shipments on container ships from Southeast Asia and Japan, has introduced a new secondary (to *Aedes aegypti*) vector for dengue virus in endemic regions. In addition, *A. albopictus* has genetically adapted itself to be a competent new vector for Chikungunya virus and has caused Chikungunya virus epidemics in India and the Indian Ocean Islands, very popular island destination resorts for wealthy Europeans. More recently, co-infections with dengue and Chikungunya viruses have been reported from Africa and India. Chikungunya-infected air travelers returning to Europe from Chikungunya-endemic regions have imported a new viral pathogen for *A. albopictus* to transmit locally. Imported Chikungunya has been reported in France, Italy, and the UK; and an outbreak of autochthonous or locally transmitted (by *A. albopictus*) has been reported in Italy. Like West Nile virus, which arrived in the US by air from Africa or the Middle East in 1999, it is simply a matter of time before Chikungunya virus (which is more likely to cause symptomatic disease than West Nile virus) arrives in the US by air.

Figure 2.4.A-1. The female *Aedes albopictus*, or Asian tiger mosquito.[16]

The Globalization of Trade and Travel: Air Travel

Many models of climate change and vector-pathogen relationships now predict a significant expansion in potential malaria transmission cycles in the next few decades, with some studies predicting a 16–25% increase in person-months of exposure in malaria-endemic areas of Africa.[17] Accessible airline connections now permit infected individuals to travel anywhere in the world in less than 24 hours, delivering human reservoirs of malaria, dengue, West Nile virus, and Chikungunya virus to new temperate areas for autochthonous or local transmission by new and adaptable mosquito vectors who are often recent air- or sea-arrivals themselves.

The most common reasons for malaria to occur in the industrialized nations of North America and Europe, where malaria was once endemic, are related to international air travel in a warmer and wetter climate and include airport malaria and, more significantly, imported malaria.

Airport malaria is defined as the intercontinental transfer of malaria through the introduction of an infective anopheline mosquito vector into a nonendemic-disease area with a changing ecosystem that supports the vector-pathogen relationship. The malaria-infected mosquito vector is a new arrival on an international flight from a malaria-endemic region. Airport malaria is transmitted by the bite of an infected tropical anopheline mosquito within the vicinity of an international airport, usually a few miles or even less. Airport malaria has been reported in the US and in Europe (France, UK), but it is much less common than imported malaria, the most common cause of malaria in the developed world today.[18]

Imported malaria is defined as the intercontinental transfer of malaria by the movement of an infected or parasitemic person with malaria to a nonendemic-disease area with locally competent anopheline vectors in a welcoming ecosystem.[19] Climate change has now expanded the geographic distribution of malaria-endemic regions worldwide and has extended the length of seasonal malaria transmission cycles in endemic regions, so more arrivals of malaria-carrying mosquitoes and malaria-infected travelers are anticipated. The greatest public health threats that imported malaria-infected mosquitoes and patients with malaria pose to non-malarious regions include (1) the re-introduction of *Plasmodium* species (especially *Plasmodium vivax*, the causative agent of often misdiagnosed relapsing malaria, in the US and Europe) into regions with competent anopheline vectors, and (2) the re-establishment of locally transmitted or autochthonous malaria by local anopheline vectors.

Prevention and control strategies for airport, imported, and autochthonous or locally transmitted malaria should include (1) early case definition, case confirmation, and treatment, (2) strengthened vector surveillance to detect the potential for autochthonous transmission, and (3) drainage of potential mosquito breeding and egg-laying surface water sites. Although the relationships among infected vector importation, index case immigration, reclaimed disease ecosystems, and malaria transmission are complex, future attempts to control and eradicate airport and imported malaria should be based on an understanding of disease transmission mechanisms and an appreciation that climate and ecosystem changes can support re-emerging local mosquito-borne infectious diseases in non-endemic areas, especially malaria, dengue, Chikungunya virus, and West Nile virus. The most capable mosquito vectors of malaria and other infectious diseases are compared by geographic distribution and infectious disease transmission in Table 2.4.A-2.

Table 2.4.A-2. The Mosquito Vectors of Major Medical Importance.

Mosquito Genera	Infectious Diseases Transmitted	Geographic Distribution Ranges	Causative Microbial Agents	Classification of Causative Agents
Anopheles spp.	Malaria	Africa, Asia, Central America, South America	*Plasmodium falciparum, P. vivax, P. ovale, P. malariae*	Protozoan parasites
Anopheles spp.	Bancroftian filariasis	Southeast Asia	*Wuchereria bancrofti*	Filarial worms causing lymphatic filariasis
	Brugian filariasis	Southeast Asia	*Brugia malayi*	
	Timorian filariasis	Timor, Indonesia	*Brugia timori*	
Anopheles spp.	O'nyong nyong fever	Africa	Alphavirus	Togaviruses
Aedes spp.	Yellow fever	Africa, Latin America	Flavivirus	Flaviviruses
	Dengue fever	Africa, Asia, Latin America	Flaviviruses DEN 1-4	Flaviviruses
	Chikungunya fever	Africa, Asia	Alphavirus	Togaviruses
	Eastern equine encephalitis	Eastern & South-eastern USA	Alphavirus	Togaviruses
	Ross River fever	Australia, Papua New Guinea	Alphavirus	Togaviruses
	California encephalitis	Western USA	Bunyavirus	Bunyaviruses
	LaCrosse encephalitis	Midwestern USA	Bunyavirus	Bunyaviruses
	Rift Valley fever	Africa	Phlebovirus	Bunyaviruses
Culex spp.	Western equine encephalitis	Western US, South America	Alphavirus	Togaviridae
	Venezuelan equine encephalitis	Central & South America	Alphavirus	Togaviruses
	Japanese encephalitis	Far East, Southeast Asia, Indonesia	Flavivirus	Flaviviruses
	St. Louis encephalitis	USA	Flavivirus	Flaviviruses
	West Nile virus	Africa, Middle East, East Europe, North America	Flavivirus	Flaviviruses

Table 2.4.A-2. The Mosquito Vectors of Major Medical Importance.

Mosquito Genera	Infectious Diseases Transmitted	Geographic Distribution Ranges	Causative Microbial Agents	Classification of Causative Agents
Mansonia spp.	Brugian filariasis	Southeast Asia	*Brugia malayi*	Filarial worm causing lymphatic filariasis
Psorophora spp.	Venezuelan equine encephalitis	Central & South America	Alphavirus	Togaviruses
Haemophagus spp.	Yellow fever	Africa, South America	Flavivirus	Flaviviruses

In addition to malaria and flaviviral infections, sudden acute respiratory failure (SARS) and tuberculosis (TB) have also been imported into North America by infected air travelers; and avian-origin influenza A (H5NI) and swine-origin novel influenza A (H1NI) have been imported into North America by land (swine flu), sea (avian flu), and air (avian flu, swine flu).

The Most Predictable Pathogens and their Food Water, and Insect Transmission Vectors

The food, water, and vector-borne microbial agents that will cause climate change-induced infectious disease outbreaks have now independently identified themselves as (1) having low infective doses and high virulence, (2) being environmentally stable, and (3) and being salt, chlorine, and iodine-resistant. Water- and food-borne diseases will include gastroenteritis caused by the coccidian protozoans (*Cryptosporidium parvum*, *Cyclospora cayetanensis*, *Isospora belli*), hepatitis A virus, astroviruses and caliciviruses, *Giardia lamblia*, the non-cholera marine *Vibrios* (*V. vulnificus*, *V. parahaemolyticus*), *Campylobacter jejuni*, *Shigella* species, and the typhoidal and non-typhoidal *Salmonella* species.[20]

Campylobacter jejuni and many *Salmonella* species are natural commensals that live in the gastrointestinal tracts of birds, amphibians, and reptiles. Following natural disasters, man may be exposed to these bacteria following the consumption of contaminated drinking water or food exposed to animal or human carriers, or to contaminated fomites such as cutting boards and carving knives. Rarely, human beings can become chronic asymptomatic carriers and fecal super-shedders of *S. typhi*. These bacteria are very heat-sensitive and can be killed by boiling drinking water and thoroughly cooking meat, especially poultry.

Following natural disasters, particularly floods and hurricanes, reptiles are often displaced from their dens and can contaminate food preparation surfaces with non-typhoidal *Salmonella* species, particularly outdoor surfaces on which fresh fruits are later sliced for eating or squeezed or tapped for juices. On the other hand, the *Vibrio* bacteria, including *V. cholera*, are natural seawater dwelling organisms that often contaminate human drinking water reservoirs following storm surges, and can also contaminate open wounds. The non-cholera-causing *Vibrios* (*V. parahaemolyticus*, *V. vulnificus*) are much more ubiquitous and virulent pathogens than are the cholera-causing *Vibrios*, and require treatment with intravenous rehydration and antibiotics. In addition to dysentery, they can cause systemic septicemia and liver failure with high case fatality rates, especially in patients with pre-existing liver diseases such as alcoholic cirrhosis, hepatitis, and hemochromatosis (high iron stores in the liver and heart).

Outbreaks of plague caused by *Yersinia pestis* and transmitted from rodents to humans by fleas will continue to occur in regions where the organism is endemic in prairie dogs and other

rodents, such as the US West and the Southern Himalayas. *Leptospirosis* caused by the *spirochete*, *Leptospira interrogans*, and transmitted to man via rodent urine-contaminated surface waters will occur in coastal and other waterside municipalities worldwide. The large rodent populations in these cities will be frequently displaced by heavy rainfall, flash flooding, tidal and storm surges. Pulmonary hantavirus syndrome, another rodent zoonosis transmitted by rodent fecal-aerosols, will occur in dry areas such as the US Four Corners (shared state boundaries of Arizona, Colorado, New Mexico, and Utah) following heavy rainfalls that force domestic deer mice from their underground burrows and into vacant homes, garages, and crawl spaces. All of the tick-transmitted diseases, including spirochetal diseases such as Lyme disease, babesiosis, ehrlichiosis, anaplasmosis, and the tick-borne viral *encephalidides*, will shift their geographic distributions northward by several latitudes and their altitude distributions higher by tens of meters.

New Vector-pathogen Relationships and Potential Epidemic Disease Threats

Emerging infectious diseases may be defined as those that have recently appeared in human populations, have expanded their ranges of distribution, or have threaten to increase their prevalences and distribution ranges in the near future.[21] Today most emerging infectious diseases arise first in the natural environment as zoonoses, such as Hantavirus pulmonary syndrome, avian-origin influenza A (H5N1), swine-origin novel influenza A (H1N1), and SARS.[22] The arthropod-borne emerging infectious diseases often adopt a competent new insect vector such as *Aedes albopictus* for the transmission of dengue and Chikungunya viruses. On the other hand, re-emerging infectious diseases may be defined as those whose pathogenesis, clinical manifestations, and treatment strategies are well know, but they have re-emerged as public health threats, often with increased antimicrobial drug resistances (e.g., multi-drug resistant and extensively drug-resistant TB, nosocomial [or hospital-acquired] and community-acquired methicillin-resistant *Staphylococcus aureus* [MRSA], and nosocomical *Clostridium difficile* and vancomycin-resistant *enterococcal* [VRE] infections).

New Vector-pathogen Relationships

Dengue is a serious re-emerging threat for several reasons. Although an effective live vaccine is available for yellow fever, there is no vaccine to prevent dengue, a yellow fever-like flavivirus that is also transmitted by *Aedes* species mosquitoes. A dengue vaccine has proven to be very difficult to develop for several reasons: (1) the 4 dengue serotypes dictate a polyvalent vaccine, like the influenza vaccine, (2) a dengue vaccine must provide immunity against all 4 flaviviral serotypes (DEN 1-4) at once by stimulating effective neutralizing antibodies, (3) these neutralizing antibodies must not cross-react and activate T-cells causing the cytokine reactions characteristic of frequently fatal dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS), and (4) multiple vaccinations every few years will be required in order to achieve long-lasting immunity against all 4 serotypes. Dengue viruses are now endemic along the porous US-Mexico border and have caused DHF outbreaks on both sides of the border and one autochthonous case of DHF in Brownsville, Texas, in 2005.[23]

Although yellow fever and dengue viruses have historically been confined to the tropics and transmitted preferentially by *Aedes aegypti*, a new secondary *Aedes* vector for dengue, *A. albopictus*, the Asian tiger mosquito, has now expanded its range globally in a warming ecosystem and is a competent vector of dengue viruses. The WHO considers dengue to be one of the world's most important re-emerging infectious diseases, with 50 million to 100 million cases

annually, 0.5 million hospitalizations that often require blood product transfusions, and 22,000 deaths annually, mostly in children. Even though the first dengue infection may be mild, the second could be lethal from either DHF or DSS, even years later. Since there are no vaccines or specific drug treatments for dengue, and since local *A. aegypti* and *A. albopictus* mosquitoes are capable of transmitting dengue in the US, dengue poses a significant threat to the US, and a safe quadrivalent vaccine and better mosquito vector control along the US-Mexico border are needed now.

Vaccine Preparedness

The US was not prepared for the avian flu pandemic, which did not occur, much less the swine flu pandemic, which did. The avian-origin influenza (H5N1A) virus, or the avian flu virus, emerged as a new cause of human disease in Chinese poultry workers more than a dozen years ago and has now caused nearly 400 cases worldwide in poultry workers or their close contacts, with more than half of the cases being fatal. This virulent influenza strain has now been found in birds throughout the world except in the US. Although avian flu cases occurred in 14 countries across Asia and Africa, the much-feared pandemic spread to the US never occurred, and this finally led the US government to halt production of an avian flu vaccine.

New influenza vaccines for potentially pandemic strains are very time-consuming and expensive to produce for several reasons including: (1) enough virus has to be isolated from human cases and grown in seasonally available embryonated eggs, (2) the virus has to be attenuated or made less virulent through many passages through eggs just to produce live, whole virus vaccines (like the oral polio vaccine, which is no longer in use in the US) which could cause human disease, (3) live attenuated virions then have to be inactivated and separated into their component H (*hemagglutinin*) and N (*neuraminidase*) protein antigens to make a safer, immunologically effective subvirion vaccine, and (4) the new vaccines need to be tested in both adults and children before approval for widespread use. Since vaccines for novel influenza A (H1NI) or swine flu are still undergoing trials in adults and children, the most effective immediate strategies to control a swine flu pandemic in the US include identifying and isolating suspected cases and treating confirmed cases with the only effective neuraminidase inhibitor the virus is susceptible to, oseltamivir (Tamiflu®). Unfortunately, novel influenza A (H1N1) is resistant to the only other neuraminidase inhibitor currently FDA-approved to treat influenza, zanimivir (Relenza®).

An outbreak of H1N1 swine-origin influenza A, now known as novel influenza A (H1N1), began in southern Mexico in March 2009, and by May 2009 it had spread to 22 nations, with more than 1900 cases worldwide and the highest case fatality rates in Mexico (29 deaths).[24] The first novel influenza A (H1N1) cases were confirmed in the US in April 2009; and by August 2009, the cumulative number of US infections had reached 1 million, well in advance of the traditional wintertime influenza season.[25] By August 2009, the US Centers for Disease Control and Prevention (CDC) deployed 25% of the Strategic National Stockpile (SNS) of oseltamivir (Tamiflu®) to all U.S. states in response to the pandemic. In addition, the time-consuming process of new vaccine development was initiated to develop an effective vaccine against the novel influenza A (H1N1) virus. The swine flu pandemic spread more rapidly across the world than the avian flu outbreak of 1997-2004. Fortunately, there have been lower case fatality rates as new cases have moved away from the epicenter of the pandemic in Mexico.[26]

The Globalization and Industrialization of the World's Vaccine and Drug Supplies

The US will need to develop new vaccines for influenza containing H antigens from potentially pandemic strains and to restore oseltamivir stockpiles. New vaccine development will require embryonated eggs during seasonal shortages in the US. The manufacture of oseltamivir will require its key ingredient, shikimic acid from the Chinese star anise plant. Both embryonated eggs and shikimic acid would likely be imported from China. China is among the world's largest exporter of embryonated eggs for vaccine development and poultry farms, and a major exporter of shikimic acid for oseltamivir manufacture.

In December 2008, WHO inspectors detected melamine in eggs from chickens fed melamine-contaminated feed.[27] The consumption of melamine-contaminated chicken and unembryonated eggs could result in nephrotoxic exposures similar to those that followed the consumption of adulterated infant formula and dairy products. The use of melamine-adulterated embryonated eggs for influenza vaccine production could alter the safety and immunological efficacy of inactivated subvirion vaccines. Today, many drugs and vaccines, like foods, are imported from major manufacturers in Bangladesh, China, and India, where there is less regulatory oversight of production practices and intentional adulteration occurs to boost profits. Without greater control and oversight of its food, drug, and vaccine supplies, the US will remain vulnerable to food and drug shortages, to intentional food and drug adulteration, and to agricultural and biological terrorism. The consequences could be disastrous and include inadequate or contaminated meat and dairy supplies, and inadequate stockpiles of safe drugs and vaccines to fight pandemics or to contain intentional biological weapon exposures, particularly from smallpox or similar pox-viruses (cowpox, camel pox).

Air Quality, Water Quality, and Drinking Water Rights

Drier and hotter summers following milder winters with less snowmelt will empty reservoirs and will fuel lightning-ignited and intentionally set brush and forest fires, destroying timberlands, displacing residents, and causing increased smoke-inhalation injuries and deaths. Greater tropospheric trapping of primary pollutants from fossil fuel consumption and secondary pollutants from photochemical-smog reactions cause the continuing deterioration of air quality and visibility. Air pollution leads to increased seasonal allergic asthma and other chronic respiratory diseases, results in defoliate alpine forests, and contaminates surface freshwater with acid rain. Annual US hospital admission rates for acute asthma in children increased from 19 per 10,000 children in 1979 to 35 per 10,000 children in 2001.

Droughts typically provoke water-rights disputes between upstream and downstream populations. Litigation over water rights has started between California and other western-slope US states, among Alabama, Florida and Georgia, and between the US and Mexico. Droughts also cause crops to fail, triggering malnutrition, famine, and refugee relocation in developing nations. Tidal flooding will increase the salinity of water tables, coastal marshes, and alluvial deltas, destroying shellfish beds, contaminating freshwater aquifers, and rendering deltas unfit for agriculture.

Water-centered Disasters: Increasing Hurricane Activity

Water-centered disasters such as floods, rainstorm-spawned tornadoes, tropical storms, tropical cyclones or hurricanes, and earthquake-spawned tidal waves or tsunamis now occur more often than geological disasters. They occur more commonly in the Americas, Asia, and Africa, and

they often cause more deaths by drowning than other disasters. Water-centered disasters also cause more serious injuries than deaths overall, primarily through falls and falling objects, motor vehicle accidents, carbon monoxide poisoning from home generator misuse, and chain saw injuries during debris clean-up. Increased endemic infectious disease risks, especially acute gastrointestinal illnesses and mass population relocations are more common after water-centered disasters than after geological disasters such as earthquakes and volcanic eruptions.

A *storm surge* is a massive wall of seawater 50 to 100 miles wide that sweeps across the coastline near hurricane or tsunami landfall. Storm surge is the greatest threat to human life during hurricanes and tsunamis. A *storm tide* is the combination of the storm surge and the astronomical high tide and is the greatest threat to property during hurricanes and tsunamis. The storm tide sweeps along coastlines at hurricane or tsunami landfall for much longer distances (over 100 miles) and penetrates much deeper inland than the storm surge.

During the 5 weeks after hurricane Andrew, the proportional morbidity from injuries decreased, proportional morbidity from respiratory illnesses increased, and proportional morbidity from diarrheal diseases remained stable. No infectious disease outbreaks occurred. Investigators concluded that injuries and exacerbations of chronic diseases were the most common causes of morbidity immediately after hurricanes, and that infectious diseases were uncommon, presented weeks after hurricanes, and were not associated with communicable disease outbreaks.

Atmospheric and climate scientists have now been able to measure the impact of sea surface temperatures on the intensification and direction of tropical storms and hurricanes (tropical cyclones) using a worldwide network of ocean temperature buoys. Other continuous measurements in assessing ENSO activity and tropical cyclone intensification and landfall prediction have also been identified and include ocean subsurface temperatures, thermocline depths and thicknesses, and storm steering current barometric pressures.[28] Armed with these new measurements, forecasters will be better able to make earlier and more accurate predictions of hurricane intensity and landfall; permitting local authorities to order timely population evacuations prior to landfall.[29]

Earth-centered Disasters: Increasing Volcanic and Tectonic Activity

Global tectonic and volcanic activities have increased significantly since 1880 due to a combination of coastal tectonic plate undermining by the seawater-weighted Northern Pacific plate and wintertime sinking of the North Pole from ice melt. As winter water and ice weights increase, these seasonal geophysical changes will be accentuated, exerting greater influences on tidal surges, tectonic plate movements, and geothermal activities, especially in the Northern Hemisphere and along the Pacific Rim. These powerful, combined geophysical forces will generate more molten magma in the earth's mantle, which expands to uncork dormant volcanoes, as in Alaska, or to burst through thinning ice packs, as in Iceland. Case fatality rates for earth-centered natural disasters—such as earthquakes, earthquake-spawned tsunamis, and volcanic eruptions—often significantly exceed case fatalities from more predictable, water-centered disasters such as tropical cyclones and floods.

Conclusions and Recommendations

International agencies and climate modelers agree on the following conclusions:

1. The world's climate is changing.
2. Human activities are contributing to the change.

3. The earth is warming and the pace of warming is increasing.
4. Weather patterns have become more extreme.
5. All biological systems are responding to global warming, both on land and in the oceans.

Free trade, economic globalization, and the ease and speed of air and sea travel and commerce have offered several unique opportunities to genetically adaptable pathogens during this period of climate change, including the following:

1. New pathogen distribution systems
2. New animal hosts for incubation
3. New vectors of disease transmission
4. New warming and welcoming ecosystems for pathogen reproduction
5. More disease-susceptible human hosts for pathogens to infect

Several components of climate change, particularly warming temperatures and more frequent drought-rainy season cycles, have supported the success of new vector-pathogen relationships, as in airport and imported malaria transmission. Some arthropod vectors, particularly mosquitoes and ticks, have been given selective advantages by climate change, free trade, international air travel, and increasing outdoor contact with susceptible human hosts during prolonged insect breeding and feeding seasons. Pathogens such as the monkeypox and West Nile viruses that have been introduced from tropical regions have found new host animal reservoirs in warming ecosystems in industrialized nations north of the Equator.[30] In short, more un-welcomed tropical pathogen imports may be anticipated and transmitted by competent local insect vectors.

The governments of both developed and developing nations have several binding responsibilities to their citizens including:

1. To inculcate an informed environmental consciousness
2. To protect their citizens and economies from foreign exploitation
3. To recognize and control endemic and pandemic infectious diseases
4. To prevent infectious disease transmission by patient education, vector control, and vaccination
5. To practice responsible labor and trade practices in order to assure the safest workplaces and to produce the safest foods, drugs, and goods for export in a one-world economy

Additional Recommended References

1. St. Louis ME, Hess JJ. Climate change. Impacts on and implications for global health. Am J Prev Med 2008; 35: 527-38.
2. Okie S. A new attack on malaria. N Engl J Med 2008; 358: 2425-8.
3. Charrel, R., de Lamballerie, X., Raoult, D. Chikungunya outbreaks—the globalization of vectorborne diseases. New Engl J Med 2007; 356: 769-71.
4. Anyamba, A., Chretien, J.P., Small, J., Tucker, C.J., Linthicum, K.J. Developing global climate anomalies suggest potential disease risks for 2006-2007. Intl J Health Geographics 2006; 5: 60-3.
5. deLamballerie X, Leroy E, Charrel RN, et al. Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come? J Virology 2008; 5: 33.
6. Panning M, Grywna K, von Esbroek, Emmerich P, Drosten C. Chikungunya fever in travelers returning to Europe from the Indian Ocean region, 2006. Emerg Infect Dis 2008; 14: 416-22.

7. Hochedez P, Canestri A, Guihot A, et al. Management of travelers with fever and exanthema, notably dengue and Chikungunya infections. Am J Trop Med Hyg 2008; 78: 710-3.
8. Morens, D, Fauci, AS. Dengue and hemorrhagic fever: a potential threat to public health in the United States. JAMA 2008; 299: 214-6.
9. Monath TP. Dengue and yellow fever-Challenges for the development and use of vaccines. N Engl J Med 2007; 357: 2222-5.
10. Dengue hemorrhagic fever-US-Mexico border, 2005. MMWR 2007; 56: 785-9.
11. Voelker R. Effects of West Nile virus may persist. JAMA 2008; 299: 2135-6.
12. Whitfield D, Curtis CF, White CB, et al. Two cases of *falciparum* malaria acquired in Britain. Brit Med J 1984; 289: 1607-9.
13. Giacomini T, Mouchet J, Mathieu P, Petithory JC. Study of 6 cases of malaria acquired near Roissy-Charles-de-Gaulle in 1994. Bulletin de l'Academie Nationale de Medecine 1995; 179: 335-51.
14. Rodger AJ, Cooke GS, Ord R, Sutherand CJ, Pasvol G. Cluster of *falciparum* malaria cases in UK airport. Emerg Infect Dis 2008; 8: 1284-6.
15. Romi R, Sabatinelli G, Majori G. Malaria epidemiological situation in Italy and evaluation of malaria incidence in Italian travelers. J Travel Med 2001; 8: 6-11.
16. Mascarello M, Allegranzi B, Angheben A, et al. Imported malaria in adults and children: Epidemiological and clinical characteristics of 380 consecutive cases observed in Verona, Italy. J Travel Med 2008; 15: 226-36.
17. Transmission of *Plasmodium vivax* malaria-San Diego, County, California, 1988 and 1989. MMWR; 39: 91-4.
18. Thierfelder C, Schill C, Hatz C, Nüesch R. Trends in imported malaria to Basel, Switzerland. J Travel Med 2008; 15: 432-6.
19. Kruse, H., Kirkemo, A.M., Handeland, K. Wildlife as a source of zoonotic infections. Emerg Infect Dis 2004; 10: 2067-72.
20. Chomel, B., Belotto, A., Meslin, F.X. Wildlife, exotic pets, and emerging zoonoses. Emerg Infect Dis 2007; 13: 6-11.
21. Marano N, Arguin PM, Pappaioanou M. Impact of globalization and animal trade on infectious disease ecology. Emerg Infect Dis 2007; 13:1807-1809.
22. Süss J, Klaus C, Gerstengarbe FW, et al. What makes ticks tick? Climate change, ticks, and tickborne diseases. J Travel Med 2008; 15 (1): 39-45.
23. Freedman DO, Weld LH, Kozarsky PE, et al. Spectrum of disease and relation to place of exposure among ill returned travelers. N Engl J Med 2006; 354: 119-130.
24. Jensenius M, Parola P, Raoult D. Threats to international travelers posed by tickborne diseases. Travel Med Infect Dis 2006; 4: 4-13.
25. Steere AC. Lyme borreliosis in 2005, 30 years after initial observations in Lyme, Connecticut. Wein Klin Wochenschr 2006; 118: 21-22.
26. Parola P, Raoult D. Tickborne bacterial diseases emerging in Europe. Clin Microbial Infect Dis 2001; 7 (2): 80-83.
27. Lyme disease-United States, 1994. MMWR 1995; 44: 459-462.
28. Jensenius M, Fournier P-E, Raoult D. Rickettsioses and the international traveler. Clin Infect Dis 2004; 39: 1493-1499.
29. Parola P. Tickborne rickettsial diseases: Emerging risks in Europe. Comp Immunol Microbiol Infect Dis 2004; 27 (5): 297-304.

30. Raoult D, Fournier P-E, Fenoccar F, et al. *Rickettsia africae*, a tickborne pathogen in travelers to sub-Saharan Africa. N Engl J Med 2001; 344: 1504-1510.

↑ Back to the top

---

[1] 1975 (IPCC, Fourth Assessment Report. Climate Change: Synthesis Report. AR4 SYR-Summary for Policymakers, 2007; 2007: pp 1-23)
[2] IPCC, Fourth Assessment Report, op. cit.
[3] Smith DM, Cusack C, Coleman AW, et al. Improved surface temperature prediction for the coming decade from a global climate model. Science 2007; 317: 766-9
[4] Patel RB, Burke TF. Urbanization-An emerging humanitarian disaster. N Engl J Med 2009; 361: 141-3
[5] Lewis M. Megacities of the future. Forbes, 2007.
[6] World Health Organization-Malaria. In: Malaria. International Travel and Health. WHO, Geneva, 2005: 132-51
[7] Morens DM, Taubenberger JK, Fauci AS. The persistent legacy of the 1918 influenza virus. N Engl J Med 2009; 361: 225-7
[8] Ingelfinger JR. Melamine and the global implications of food contamination. N Engl J Med 2008; 359: 2745-8
[9] Ingelfinger JR, op. cit.
[10] Ingelfinger JR, op. cit.
[11] Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis 1995; 1: 7-15
[12] Morse SS, op.cit.
[13] Multistate outbreak of monkeypox-Illinois, Indiana, and Wisconsin. MMWR 2003; 52: 537-40)
[14] MMWR 2003, op. cit.
[15] MMWR 2003, op. cit.
[16] CDC Public Health Image Library (PHIL) Image Number 4735. Available at http://www.cdc.gov . No copyright permission required.
[17] WHO, Malaria Report 2005, op. cit.
[18] Malaria surveillance-United States, 2007. MMWR 2009; 58: SS-2
[19] Malaria surveillance-United States, 2007. MMWR 2009; 58: SS-2
[20] Surveillance for foodborne disease outbreaks-United States, 2006. MMWR 2009; 58: 609-15
[21] Morse SS, op. cit.
[22] Morse SS, op. cit.
[23] Dengue hemorrhagic fever—US-Mexico Border, 2005. MMWR 2005; 56: 785-9
[24] Influenza A activity-United States, April-August 2009. MMWR 2009; 56: 1009-12
[25] MMWR 2009, op. cit.
[26] MMWR 2009, op. cit.
[27] Kuehn BM. Melamine scandals highlight hazards of increasingly globalized food chain. J Am Med Assoc 2009; 301: 473-5
[28] IPCC Fourth Assessment Report, op. cit.
[29] IPCC Fourth Assessment Report, op. cit.; Smith DM et al, 2007, op.cit.
[30] Marano N, Pappaioanou M. Historical links between human and animal health. Emerg Infect Dis 2004; 12: 2065-6

2.4.B. The Arctic Region: Prospects for a Great Game or International Cooperation (Kenneth Yalowitz, Ross Virginia)

Authors and Organization: Kenneth S. Yalowitz, Director, Dickey Center for International Understanding, Dartmouth College; Myers; Ross A. Virginia, Family Professor and Director, Dickey Center Institute of Arctic Studies, Dartmouth College
Contact Information: kenneth.s.yalowitz@dartmouth.edu; ross.a.virginia@dartmouth.edu

The pace of ecological, political, social, and economic change in the Arctic region is accelerating due to a warming climate. As the Arctic Ocean summer sea ice thins and disappears, new opportunities and uncertainties are emerging over access to energy reserves, shipping routes, and fisheries. The traditional way of life is changing significantly for indigenous peoples and many Arctic residents who strive to be full partners in developing strategies to adapt to climate change. While science is providing detailed information about climate change and the responses of Arctic ecosystems to new weather and oceanographic conditions, models of the Arctic have consistently underestimated the rate of change, leaving policy makers facing important strategic decisions about the future of the Arctic sooner than expected.

Imbedded in this complexity are concerns that a less harsh Arctic environment will facilitate exploitation of Arctic resources and trigger a new "great game" of international power politics involving the principal Arctic coastal states—the United States, Russia, and Canada. This paper evaluates the prospects for contrasting outcomes in the Arctic region: a return to international power politics as states seek to claim Arctic energy, extend continental shelves, and enforce their will through military means versus the emergence of increased international cooperation around environmental protection and sustainable development.[1]

Science and Policy

Climate science is unambiguous in its assessment—our planet is a set of closely linked physical systems of land, water and atmosphere in which the Arctic is integral. Global warming is affecting the world's weather, and sea ice in the Arctic is melting and thinning rapidly, declining 15% every decade over the recent past. Moreover, the pace of change is accelerating.[2]

As scientific understanding of climate change advances, policy makers are inundated with crisis management decisions; specific climate change issues related to the Arctic must compete for their attention. Exacerbating this problem is the communication gap that exists between scientists and policy makers. Scientists need to communicate more effectively and accurately with policy makers and develop better qualitative and quantitative models with which to inform them. And policy makers must provide scientists with the necessary resources to develop useful data and models.

The traditional ecological knowledge of indigenous peoples is an underutilized asset for evaluating and understanding the processes contributing to rapid environmental change. Indigenous communities are central stakeholders in scientific assessment and policy development for a sustainable Arctic. Western researchers need to engage more with indigenous communities and organizations to build bonds of communication and trust, so together they can develop a body of scientific information that values indigenous knowledge on climate change and adaptation, and integrates it into scientific data.

Economics, Resources, and Development

The Arctic region has abundant natural resources—gas, oil, diamonds, coal, iron ore, gold, zinc, nickel, and others. Intensive development of non-renewable resources has been ongoing for four decades. In the short run, in terms of shipping, globalization of the Arctic and the condition of the global economy are the biggest issues controlling rates of activity, not climate change. Large numbers of ships—more than 6,000 annually—are already in the Arctic for tourism, fishing, and transport, mostly in the summer.

There has been great public fanfare about the opening up of northern sea routes and the Northwest Passage for longer time periods due to sea ice melt, offering cheaper transportation between Asia and Europe. In fact, there is much uncertainty among shipping experts about whether or under what conditions these routes will prove economical. Even using accelerated climate warming projections, sea ice will remain in place for 6-8 months a year, and potentially dangerous floating ice will remain year round, necessitating the use of ships designed to meet the rigors of the Arctic Ocean environment and ice breakers for winter shipping routes. Nevertheless, there are important safety and environmental issues. Arctic shipping today is regulated by voluntary guidelines and inconsistent ad hoc governance regulations. There is a pressing need for stable, rule-based regimes based on the United Nations Convention on the Law of the Sea (UNCLOS) and the International Maritime Organization (IMO) rules for international marine coordination in the Arctic on safety, emissions, and infrastructure issues. The Arctic Marine Shipping Assessment 2009 Report of the Arctic Council is an important step forward.

Mining, extraction and energy exploitation in the Arctic has been going on for some time. The largest zinc mine in the world (Red Dog Mine in northwestern Alaska) and the largest nickel mine in the world (Norilsk Mine in northern Siberia) are located there, and prospects are good for development of gold, diamonds, and concentrated ores in Greenland and elsewhere. Climate change affects the timing of mining, extraction and energy exploitation projects in the Arctic and creates uncertainties in forward planning. Melting ice may improve marine transportation opportunities, but melting permafrost will decrease the duration of potential overland travel in many regions. But the most important factors affecting short-term economic development are price and cost—not climate change—and here the current severe economic downturn is having a major effect. Charters for large ore carriers dropped to one percent of normal rates shortly after the global downturn,[3] and new or expanded Arctic projects are being delayed by the paralysis in global credit markets.

The main economic prizes in the Arctic are oil, gas, and mineral resources. The primary reserves belong to Russia, and the major exploration activity also is theirs. Recent estimates from the U.S. Geological Survey indicate that about 30% of the remaining world reserves of natural gas and some 10% of the oil are in the Arctic region.[4] To date, unresolved issues involving demarcation of sea beds under UNCLOS are not a major issue in the pace of energy development; rather, the key factors are the cost of development and the unpredictable price cycle of oil and gas. Offshore projects are the most costly and the most environmentally dangerous, and most of the known offshore reserves of oil and gas are within national Exclusive Economic Zones (EEZs), which extend 200 nautical miles from the coastline. Therefore, immediate prospects for interstate conflict over oil and gas reserves appear small.

Concerns for energy security will ultimately push development of Arctic oil and gas, but the financial crisis and the slowing of world economic activity is depressing energy prices and

consumption, making it very difficult to predict the pace and extent of development. As a result, there likely will not be pell-mell development in the North over the near term. This pause due to the slowdown in the global economy gives the international community the chance to develop environmental and development rules of the game, an opportunity that should not be lost.

Political and Security Issues

During the Cold War, when U.S. and Soviet nuclear submarines patrolled under the North Pole and bombers flew over the region, the Arctic was a security flashpoint. Today, following the breakup of the Soviet Union, the Arctic is largely disassociated from great power politics. Will that remain the case?[5]

The countries with military/security interests and naval capacity in the Arctic are Russia, Canada, Norway, Denmark, and the United States. Russia's activities could be disruptive to the region if its recent focus on politics and territorial claims retain priority over increased attention to science and international cooperation. The driving factors may be Russian prestige, identity, and image, all of which converge on borders and territorial claims. Russians do agree that environmental security is an important Arctic issue and that pollution from outside the Arctic must be reduced. They are not interested in a new international fault line over the Arctic.

Canada has never had significant military capability in the Arctic, but over the past two years it has conducted operations there to build capacity and presence. This reflects an overall increase in Canadian interest in the Arctic manifested also in development of Canadian International Polar Year (IPY) activities, mapping for development purposes, construction of new icebreakers, and opening new research stations. Canada is looking beyond the current economic downturn, toward exploitation of gas hydrates twenty years in the future, for example. Canada has made great progress in settling Northern land claims with First Nation groups. The key issue with the U.S. is whether the Northwest Passage sea route is considered to be Canadian internal water or if it is considered an international strait as claimed by the U.S. This dispute is "frozen" for now by mutual agreement.

Release of the revised *U.S. Arctic Regional Policy* occurred in January 2009.[6] The previous policy, written in 1994, listed meeting post-Cold War national security and defense needs as a U.S. goal in the Arctic. The new U.S. statement reiterates this goal. "The United States is an Arctic nation, with varied and compelling interests in the region," with "broad and fundamental national security interests . . . and is prepared to operate either independently or in conjunction with other states to safeguard these interests." However, the declining condition and capacity of the U.S. ice breaker fleet is a sign of disinterest on Washington's part in military or security threats emanating in the region from current disputed issues. It is also a serious limitation for U.S. science in the Polar Regions.

Anticipated territorial claims under UNCLOS leave large portions of the Arctic Ocean unclaimed, and some smaller zones have the potential for overlapping claims. For the past thirty years, unresolved claims between Russia and Norway have not prevented cooperation in managing fish stocks in the Barents Sea. States should be able to adjudicate overlapping claims through UNCLOS. Another possible approach is joint management or the establishment of a commons over the disputed areas. The need for large-scale ecosystem-based management regimes to protect the integrity of the Arctic Ocean is receiving increasing attention, including proposals for an Arctic treaty or park in order to manage and protect the Arctic Ocean as an international commons.

Security concerns and issues do not seem to be the pressing factor driving Arctic policy. There are no large geo-political fault lines, and no resource wars are anticipated. The May 28, 2008, *Illulissat Declaration of the Five Arctic Coastal States*[7] commits the states to resolving boundary and extended continental shelf claims issues under UNCLOS.

Governance and Institutions

Until now, international Arctic governance has been led by the Arctic Council, an organization of eight Arctic states, six permanent Participating Parties (indigenous peoples' organizations), and observers. The Council and its bodies have played an important role in focusing attention on environmental and climate-related issues; however, the Council has no enforcement mechanism, and security and political issues are not within its purview.

Governance within the Arctic presents unique challenges.[8] Unlike Antarctica, the Arctic is not a continent, but it has permanent residents, considerable natural resources, and a high degree of developed industrial activity. While Antarctica is burdened with territorial disputes, most territorial claims in the Arctic are not overlapping. Unlike Antarctica, the Arctic has strategic importance and was a venue for Cold War competition. These are some of the reasons why it is unlikely that an Antarctic-like treaty that provides for complete demilitarization, freezing of territorial claims, and freedom of scientific research will be concluded for the Arctic.

What are the views of the major stakeholders on this multitude of issues? Canada wants a stable, rule-based regime in the Arctic within which it can promote economic and social development, protect the environment, improve and develop governance, and exercise its Arctic sovereignty. Canada's dispute with the U.S. over the Northwest Passage, as well as the U.S. dispute with Russia over Northern Sea Route (NSR) straits, concerns their legal status—whether they are internal or international passageways—not ownership or sovereignty. The sides now agree to disagree. For Canada, the Ilulissat Declaration and UNCLOS represent the core of Canadian policy regarding navigation and access to ocean resources.

Russia's governance viewpoint seems close to that of Canada; that is, the Northern Sea Route is internal Russian waters. They also believe that the current international regimes are adequate for solving problems. Russia's declared policy is strongly supportive of the Arctic Council, but its actions are not always consistent with this policy. Russia does not endorse the idea of a new international treaty on the Arctic and believes that there is little support for it. Russia believes that UNCLOS is very important and is critical of the U.S. failure to ratify it.

There is broad support in the Obama administration for ratifying UNCLOS. They are expending considerable effort to deal internationally with climate change issues, especially engaging with India and China. The U.S. strongly supports the Arctic Council, but it is open, in principle, to consider modifications to the Council's structure and programs.

Conclusions

The Arctic strongly influences world climatic conditions, and in turn it is deeply affected by climate change and other outside factors, principally industrial pollution.

The North will be at the center of debate over international energy and environmental security issues. Competition for mineral resources such as zinc, gold, diamonds, and even fresh water will become more intense. While climate change influences the timing of Arctic resource development, the primary factors are price and cost. Regional social and economic instability are likely to increase if international agreement cannot be reached to reduce greenhouse gas

emissions. Economic activity, mainly mining and energy, which has been steadily increasing, now is affected by the overall economic decline, as well as the drop in price and demand for fuels and commodities. Even with the melting of sea ice, northern waterways will remain risky due to floating ice and undeveloped infrastructure. The biggest risks are safety at sea and environmental contamination of the water. Organizations such as the IMO and the Arctic Council must face up to these challenges.

Improved communication on many levels is a necessity, especially between scientists and policymakers. Scientists declare that climate change is causing major problems in the North, but they do not specify starting points for action, priorities, or clear steps to take in response. More targeted research is needed, as well as basic research regarding where the abrupt turning points may occur and when. On the other hand, policymakers must provide better guidance and sustained funding to scientists if they expect to obtain useful data and models.

In the Arctic, where there are no fundamental security issues, "great game" language should not be applicable. States are not yet defining their interests in zero sum terms. No geo-political fault lines reminiscent of the Cold War are evident. There are, however, some U.S.-Russian military tensions that bear watching. The U.S., Canada, and Russia all approach the Arctic in a state-centric fashion, so the applicable political model for this region is that of traditional interstate negotiations as typified by (1) the *Ilulissat Declaration of the five Arctic States*, and (2) management of border-resources issues under UNCLOS. Climate change, rather than security risks, is producing environmental issues that require new or bolstered institutional responses.

For the foreseeable future, the Arctic Council will be the principal body of governance in the North, and it assumes a key responsibility in helping to ensure regional environmental security. The Council has had a non-permanent secretariat funded by Norway for six years. With no political/security issues on the horizon, the Council is best situated to deal with issues resulting from globalization of the Arctic such as increased shipping, environmental security, etc. There is a broadly felt need for high-level policy attention to Arctic issues, and a permanent Arctic Council secretariat could provide that focus and develop recommendations for addressing functional issues.

The environment and the management of natural resources are the most pressing security issues in the North. Large-scale damage to the Arctic environment from transportation accidents, energy development, fishing, tourism, and the long-range transport of pollutants from the South pose greater immediate threats to the ecosystem and national priorities than classic security issues.

↑ Back to the top

---

[1] This paper is generally based on the outcome of a conference held at Dartmouth College, December 1-3, 2008. For a summary and conclusions see the published report: Yalowitz, K. S., J. F. Collins, and R. A. Virginia. 2009. *The Arctic Climate Change and Security Policy Conference—Final Report and Findings.* Carnegie Endowment for International Peace, Washington D.C., 36 pgs.

[2] For scientific assessments of climate change in the Arctic see, Arctic Climate Impact Assessment—Scientific Report, 2005, Cambridge University Press. For the most recent U.S. regional climate assessment that includes information on Alaska see, Global Climate Change Impacts in the United States, 2009, a report of the U.S. Global Change Research Program (USGCRP)

[3] Yalowitz, p. 13

[4] Gautier, G. L. K. J. Bird, R. R. Charpentier et al., 2009. Assessment of undiscovered oil in the Arctic. Science 324: 1175-1179, DOI: 10.1126/science.1169467

[5] See Borgerson, S. G. 2008. Arctic meltdown: The economic and security implications of global warming. Foreign Affairs, March/April 2008.

[6] National Security Presidential Directive and Homeland Security Presidential Directive, January 12, 2009, NATIONAL SECURITY PRESIDENTIAL DIRECTIVE/NSPD – 66, HOMELAND SECURITY PRESIDENTIAL DIRECTIVE/HSPD -- 25

[7] Available at www.oceanlaw.org/downloads/arctic/Ilulissat_Declaration.pdf

[8] Young, O. R. 2009. Whither the Arctic? Conflict or cooperation in the circumpolar North. Polar Record 45:73-82.

2.4.C. Changing Climate Impacts to Water Resources: Implications for Stability (Kathleen White, J.Rolf Olsen, Levi Brekke, David Raff, Roger Pulwarty, Robert Webb)

Authors and Organizations: Kathleen D. White and J. Rolf Olsen, US Army Corps of Engineers;
Levi D. Brekke and David A. Raff, U.S. Bureau of Reclamation;
Roger S. Pulwarty and Robert Webb, National Oceanic and Atmospheric Administration

Contact Information: Kathleen.D.White@usace.army.mil J.Rolf.Olsen@usace.army.mil, LBrekke@usbr.gov; DRAFF@do.usbr.gov; Roger.Pulwarty@noaa.gov, Robert.S.Webb@noaa.gov

Abstract

Climate encompasses change and variability by definition. The occurrence of a changing and variable climate has been observed in decades of environmental monitoring from a multitude of federal, state, and local data collection networks throughout the world, and it has been documented in the paleoclimate record. Although regional in range, primary indicators of recent global climate change as seen by water resources managers are increases in temperature, as well as the frequency and intensity of heavy downpours. Both obvious and subtle impacts are to the entire hydrologic cycle and are manifested as changes in the supply and demand for water. Impacts include the spatial and temporal occurrence and distribution of precipitation, rainfall intensity and duration, form (frozen or not), changes in the timing and extent of melting and runoff, as well as evapotranspiration and evaporation. Just as changing hydrologic conditions can have both adverse and positive impacts, so too can they provide both challenges and opportunities to water resources managers. Since water resources are a fundamental requirement for human health and economic growth, they are also a fundamental driver of political stability or instability in various regions of the world. This security threat was recognized during an April 2007 United Nations Security Council debate on impacts of climate change to peace and security. This paper discusses climate change in the context of water resources management, including potential implications for instability. Water resources managers are already accustomed to dealing with changes in land use, demographics, and social values. These water management techniques can provide a potential reservoir of resilience for operations in the face of climate change if water managers are prepared to act effectively in a timely fashion. Embedding a strategy of proactive and cooperative water management into planning can enhance stability by mitigating the potential for increased conflict over water in a changing climate.

Introduction

To date, the Intergovernmental Panel on Climate Change (IPCC) has produced four assessment reports which explore in detail the physical science basis of climate change, observed and projected climate impacts and vulnerabilities, and climate change adaptation and mitigation strategies, along with special reports on such issues as emissions scenarios.[1]

In the synthesis report for the Fourth Assessment, IPCC[2] reports that "Warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level." More recently, a technical paper on climate change and water[3] indicated "that freshwater resources are vulnerable and have the potential to be strongly impacted by climate change."

Stern addresses the impacts of global climate change, stating that "Climate change threatens the basic elements of life for people around the world – access to water, food, health, and use of land and the environment."[4]

There is a long history of conflict related to water and especially to access to fresh water.[5,6,7] In 1993 Gleick concluded that "Water already contributes to conflicts among nations, and future conflicts over water are increasingly likely."[8] Rising sea levels may impact water supplies and increase lowland flooding, with the potential for forced migration of large numbers of people.[9,10] In 2007, the United Nations Security Council 2007[11] noted that island nations are facing climate change impacts that are "… no less threatening than the dangers guns and bombs posed to large nations." Schwartz and Randall concluded that climate change could result in increased conflict threatening U.S. national security,[12] and the 55 delegations participating in the initial United Nations Security Council (2007) debate on the impacts of climate change raised issues potentially impacting global peace and security.

This paper reviews potential climate change impacts to water resources management and examines vulnerabilities and risks associated with potential conflicts. Considering water resources decision-making in the face of climate change as a class of "Wicked Problems"[13] provides a framework for collaborative solutions based on scenario planning. Scenario planning in water resources management can enhance preparedness, thus providing a reservoir of resilience that can mitigate adverse consequences. The paper next addresses climate change adaptation and mitigation strategies as they relate to stability. Given the critical nature of water resources in conflict and security discussions, the paper identifies a pivotal role that water managers may play in developing and implementing strong early actions that could reduce and alleviate the potential for conflict within the U.S., with U.S. bordering countries, or international conflicts between other countries.

Climate Change Impacts to Water Resources

Climate change poses numerous challenges to water resources managers because it affects fundamental drivers of the hydrological cycle as it circulates between the atmosphere, the earth surface, and oceans. Important drivers include precipitation, evaporation, condensation, and wind. Recently, the four major Federal agencies in the United States that manage water resources and water resources data and information collaborated on a report providing a comprehensive assessment of approaches to climate variability and change in water resources management. These four agencies, two termed "operating agencies" (the US Army Corps of Engineers [USACE] and the Bureau of Reclamation [Reclamation]) and two termed "science agencies" (the US Geological Survey [USGS] and the National Oceanic and Atmospheric Administration [NOAA]) published USGS Circular 1331 "Climate Change and Water Resources Management: A Federal Perspective" in February 2009.[14] The report provides a foundation upon which future agency policies, methods, and processes will be based. Although geared toward the U.S., the findings of this report are applicable to global considerations of climate change as it impacts water resources. Significant issues not addressed by Brekke et al include increased drought conditions in arid areas, the impacts of glacial melt on water supply, and the benefits of ecosystems in a changing climate. The key findings of Brekke et al related to climate change impacts to water resources are summarized as follows:

1. The best available scientific evidence based on observations from long-term hydrometeorological monitoring networks indicates that climate change is occurring, although the effects differ regionally.
2. Climate change *could affect all sectors of water resources management* [emphasis added], since it may require changed design and operational assumptions about resource supplies, system demands or performance requirements, and operational constraints. The assumption of temporal stationarity (defined by Milly et al [2008] as "the idea that natural systems fluctuate within an unchanging envelope of variability"[15]) in hydro climatic variables should be evaluated along with all other assumptions.
3. Climate change is but one of many challenges facing water resource managers. A holistic approach to water resources management includes all significant drivers of change, as shown in Table 2.4.C-1.

USGS Circular 1331 stresses the importance of long-term monitoring networks (Key Point 4) at locations that describe the climate signal (Key Point 5), and the need for research and monitoring to fill knowledge gaps required to improve water resources management (Key Point 12). Water resources management in many regions of the world, at a minimum, will likely be affected by the changes listed in Table 1 thus the emphasis on monitoring and development of new knowledge. In essence, these findings underscore those of Stern (2009), who points out that, despite the atmospheric drivers of climate change, it is the resulting changes in water on the land surface, in the form of "storms, droughts, floods, rising sea levels," that is the primary source of climate change damages and hence drives climate change vulnerabilities.[16]

Table 2.4.C-1. Potential Environmental and Socioeconomic Changes Affecting Water Resources Management Decision-making.[17,18]

Climate-induced	***Non-climate-induced***
Changes in rain-snow partitioning	Vertical land movement
Changes in precipitation intensity, frequency, sequencing	Population growth
Changes in air and water temperature	Changes in land use
Changes in snow melt volume and timing	Changes in social values
Changes in glacier melt	Changes in water use
Accelerated sea level rise	Changes in sediment budgets
Changes in sea-surface temperature	Changes in economic conditions
Changes in wind and wave patterns	Resources available to maintain and improve infrastructure
Changes in tropical and extra-tropical storm intensity and frequency	
Changes in drought frequency and intensity	
Changes in sediment budgets	
Altered wild-land fire conditions	
Changes resulting from climate adaptation measures	
Changes resulting from mitigation measures	

Vulnerabilities to Climate Change and Conflict

According to IPCC, vulnerability is defined as "the degree to which a system is susceptible to, and unable to cope with, adverse effects of climate change, including climate variability and extremes. Vulnerability is a function of the character, magnitude, and rate of climate change and variation to which a system is exposed, its sensitivity, and its adaptive capacity."[19] Detailed descriptions of global vulnerabilities to climate change are provided in the IPCC Fourth Assessment Working Group II Report[20] and the IPCC technical paper on climate and water.[21] Brekke et al (2009) identify the following climate changes as being of most concern to water resources managers in the U.S.:

- Temperature increases that are expected to change the mix of precipitation toward more rain and less snow, impact evaporation and transpiration processes, and potentially change water temperatures and ocean circulation to alter the intensity and frequency of coastal storms
- Precipitation changes, including regional and seasonal changes as well as changes in the intensity and frequency and extremes of precipitation.[22]

These changes, together with those identified in Table 2.4.C-1, can impact water availability, demand, and quality; the performance of flood and coastal storm risk reduction, stormwater, and wastewater infrastructure; wild-land fires; ecosystem functioning; navigation; and energy production and demand. [23] Furthermore, the impacts of these climate changes on water can be amplified or can be suppressed by natural variability. The changes and their impacts will be unevenly distributed across the globe. It is well known that the relationship between global average temperature changes and regional climate changes is very uncertain, particularly for precipitation. Stern (2006) suggests potential impacts associated with varying levels of atmospheric levels of GHG (expressed as CO_{2e} or carbon-equivalents) at stabilization.[24] As shown in Figure 2.4.C-1, there is clearly a range of uncertainty in the temperatures projected for CO_{2e} stabilization levels. Nevertheless, it is possible to describe potential impacts expected at different levels of warming. This figure shows potential impacts in different sectors and regions based on scientific literature available at that time.

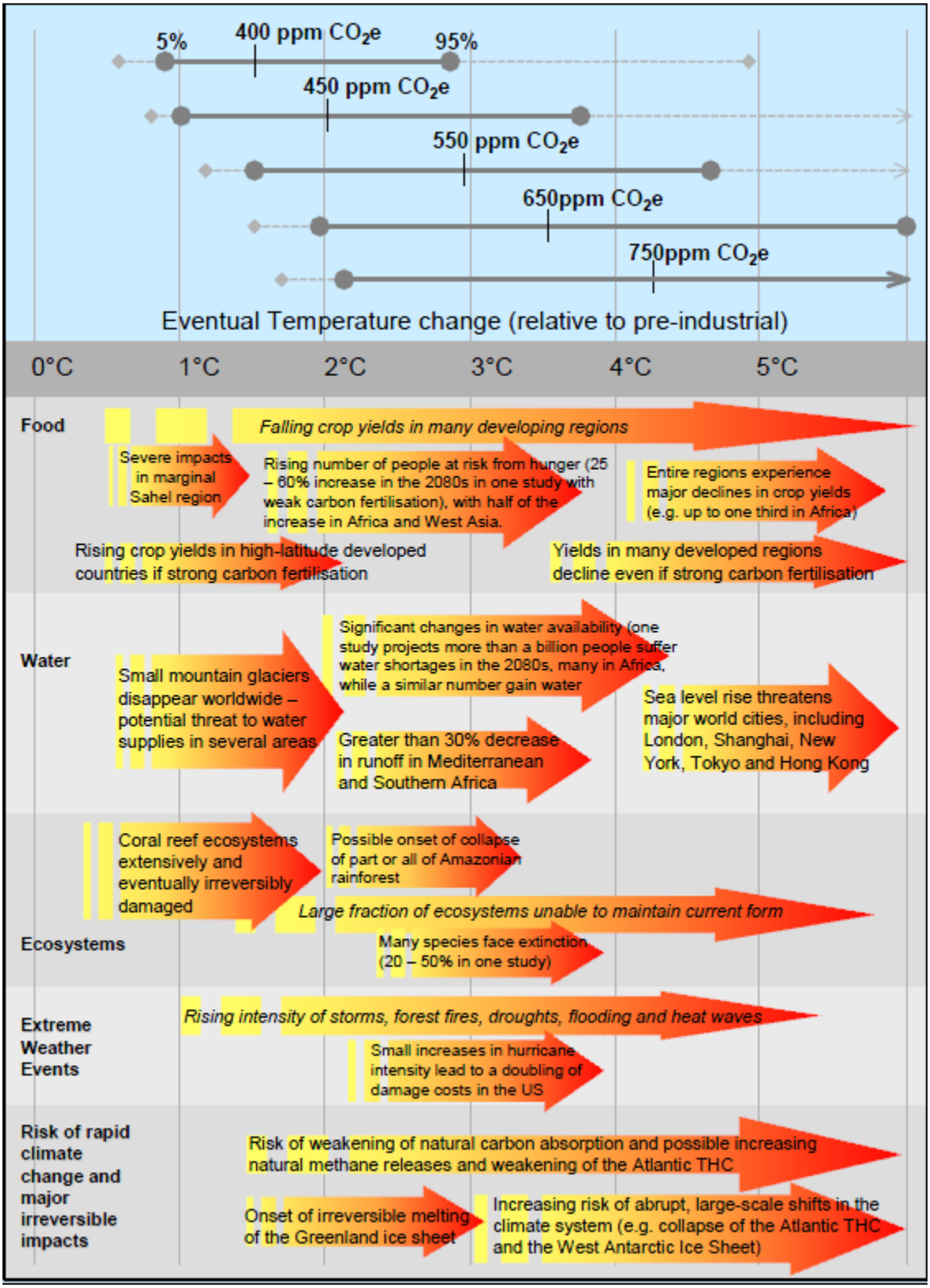


Figure 2.4.C-1. Potential Climate Change Impacts Resulting from a Range of CO_{2e} Stabilization Levels. (Confidence limits [5% and 95%] and mean of the 50th percentile shown for expected average global temperatures.) [25]

Considering the historical evidence that water and water resources management systems have been used as both offensive and defensive weapons in conflicts throughout the world,[26,27] a correlative approach would predict increased conflict due to increased social, political, and economic stress induced by climate change.[28] The Stern Review also suggested that increased water supply vulnerabilities in a changing climate could increase the potential for conflict.[29] Critical regions where vulnerabilities could result in increased conflict can be identified by a variety of screening methods.[30,31] For example, Alcamo and Henrichs (2002) identified critical regions (defined by them as those whose water resources have higher sensitivity to global change than other regions) for a variety of socio-economic scenarios that include climate change (~0.2°C rise in global air temperature per decade).[32] For all scenarios, they identified central Mexico, the Middle East, large parts of Indian sub-continent, and stretches of the North African coast as being at risk for increased water stress. On the other hand, Trondalen (2009) identifies demographics and land use as factors more critical than climate change in the Middle-East, and supports Key Point 3 above, saying that "the impact of climate change on water security should be seen in a broader water development and conflict context."[33] Within the U.S., the Department of Interior Water 2025 initiative was implemented to promote innovative approaches to minimize conflicts between water users through better management and more efficient use. The Western Governors' Association advocated for the creation of the National Integrated Drought Information System (NIDIS) to move the country toward a proactive approach to avoid conflicts and minimize the damage caused by droughts.

However, future changes in conflict due to climate change are not fully understood and may not be inevitable.[34,35,36] Wolf et al state that "the record of acute conflict over international water resources is overwhelmed by the record of cooperation," and they found no pattern of impact by climate on water disputes. [37] More efficient water resources management could also result in decreased conflict or help to avoid conflict. For example, Hendrix and Glaser conclude that if agricultural methods in sub-Saharan Africa were to change to methods that are not so rainfall-dependent, predicted conflict could be mitigated.[38] Nordas and Gleditsch also note that "the study of climate change and conflict needs to balance the positive and negative effects of climate change as well as the effects of various strategies of adaptation."[39] Ravnborg has emphasized that "water-related conflicts are caused by the way in which water and its uses are governed" rather than just the lack of water.[40] Since development, management, and delivery of water are the primary tools for implementing water governance,[41] an opportunity exists for water resource management practices to decrease the potential for conflict over water as climate changes through activities such as enhanced planning for sustainability.

Climate Change and Water Resources: Solving "Wicked Problems"

Developing plausible solutions to predict the likelihood of conflict will require resolving the multiple and often conflicting objectives of various stakeholders. These objectives are often highly complex and difficult to model, even before climate change consideration. Solutions to these problems could enhance or reduce stability in the face of changing climate on water resources. The term "wicked problems" was coined by Rittel and Webber to describe problems that are complex, resistant to solution, and have both numerous and conflicting intervention points and outcomes.[42] Traditional problem solving methods such as gathering more data, attempting to define the problem more clearly, and breaking the problem into smaller parts are not sufficient to manage wicked problems.[43] For problems of low complexity (which Kahane describes as those whose causes and effects are close in space and time, involve process that are

familiar and predictable, and align with common assumptions, values, rationales and objectives), the traditional problem-solving processes that are piecemeal, backward looking, and authoritarian, can be successful. Solving highly complex problems (i.e., those whose causes and effects are far apart in space and time, are unfamiliar and unpredictable, and for which there is a low degree of common assumptions, values, rationales, and objectives) requires processes that are "systemic, emergent and participatory."[44] These processes will require new methods of communicating risks, inclusion of social as well as physical science in problem approaches, and problem-solving processes that include parallel approaches rather than simple sequential steps.

Water policy, water resources management, and climate change adaptation have been identified as "wicked problems."[45,46,47] Freeman cites water policy problems in particular as: "…integrating [two] very different types of knowledge, of working across several socio-political units of analysis simultaneously, and of better organizing water as a common resource."[48] However, solutions can occur when the problem-solving process explicitly addresses the social components. As noted by Wolf et al, "Overall, shared interests, human creativity and institutional capacity along a waterway seem to consistently ameliorate water's conflict-inducing characteristics. Furthermore, once cooperative water regimes are established through treaties, they turn out to be impressively resilient over time, even when between otherwise hostile riparians."[49] Climate change may exacerbate existing stressors in a system, increasing complexity even when a wicked problem has achieved some level of stability. In particular, climate change poses a wicked problem with respect to stationarity, which is a fundamental assumption underlying past water resources development.[50] This assumption allowing backward-looking problem-solving must be replaced with a more dynamic view of the physical processes important in water resources development that is compatible with climate change, one that addresses nonlinear solutions and confounding and synergistic stresses and shocks.[51]

The water resources decision-making environment (e.g., social values, knowledge, economic factors) is changing along with climate. This evolution is an important step in beginning to characterize wicked problems and develop a range of possible solutions. Just as structured scenarios—combining quantitative and qualitative information in developing potential future conditions—can increase resilience to unpredictable or surprise results in social-ecological systems,[52,53] they may also be useful in wicked problems related to water resources and climate change. As pointed out by Shoemaker and Mavaddat, scenarios are especially useful in cases when important factors cannot be formally modeled (e.g., the effects of new regulations on development, the potential for climate adaptation measures to reduce vulnerability to climate change).[54] They also describe scenario planning as a useful tool to reveal and organize uncertainties in the type of complex and volatile environment that exists for water resources managers now and in the future. Development of alternative solutions to multiple scenarios provides an opportunity to evaluate the robustness of alternatives to a range of future conditions. The use of scenarios also allows us to recognize significant changes and rapidly adapt to them,[55] and to support robust water resources decision-making. Finally, scenario planning is useful in cases where our actions can change the future, as in the case of climate adaptation and mitigation.

Climate Change Adaptation and Mitigation

Climate change adaptation related to water resources can be defined generally as water resource management adjustments or changes in water resources management decision-making to enhance resilience or reduce vulnerability to observed or expected changes in climate. Climate

change mitigation can be defined generally as water resources management adjustments or changes to reduce the potential effects of global warming, through reduced demand for greenhouse gas-intensive goods and services, increased efficiency, implementation of more low-carbon technologies, and otherwise reducing fossil fuel emissions. The potential role of water resources managers to enhance social, economic, and security stability in the face of climate change is further supported considering that adaptation and mitigation are both required to potentially reduce climate change risks.[56]

Brekke et al (2009) identified water resources climate change adaptation options in the categories of operational, demand management, and infrastructure changes, and emphasized the value of adaptive management based on plausible future scenarios.[57] They suggest that a robust evaluation of planning alternatives is possible through developing climate scenarios with a wide variety of hydroclimatic conditions, which can be based on integrating paleoclimate information and stochastic modeling.[58] (This topic is also discussed in Circular 1331 Key Points 9 and 10):

9. Adopting [climate change adaptation in the form of operational, management, or infrastructure] alternatives that perform well over a wide range of future scenarios could improve system flexibility. Water resources planning and management requires an appreciation of existing and potential future uses of water resources, particularly when public health and safety are involved.
10. Adaptive management is a [collaborative] approach where decisions are made sequentially over time and allows adjustments to be made as more information is known. This approach may be useful [as a climate change adaptation measure] in dealing with the additional uncertainty introduced by potential climate change.[59]

The conclusion of the Stern Review regarding climate change mitigation was that "the benefits of strong, early action considerably outweigh the costs."[60] In the same vein, Trondalen concluded that when climate change aggravates existing water stressors, early action—especially cooperative frameworks between countries that include monitoring and appropriate use of the precautionary principle—is required.[61] (The precautionary principle has many formulations, but it can generally be described as the concept that managers or regulators "should take steps to protect against potential harms, even if causal chains are unclear and even if we do not know that those harms will come to fruition."[62]) Given the estimated 400 water-related treaties,[63] there are ample existing frameworks to serve as models for cooperative frameworks that improve stability. (See the Transboundary Freshwater Dispute Database at http://www.transboundarywaters.orst.edu/database/ for up-to-date information). Kundzewicz reports that water resources managers in the Netherlands, UK and Germany include climate change impacts explicitly in flood risk management and design.[64] Stern also identified improved data collection and sharing that reduces inefficient use of water as methods to reduce conflict, further indicating roles that water resources managers could play to improve stability.[65] Focusing on the southwestern U.S., Rajagopalan et al examined the impact of a number of climate scenarios, management strategies and variety of potential demands on Colorado River water storage over the next 50 years.[66] While the risks for drought-induced reservoir depletions were projected to increase rapidly after 2026, these risks were shown to be reduced significantly through preemptive management strategies. Furthermore, this study suggests that the impact of these management strategies would be maximized through implementation well before the rapid rise in risk in 2026.

Water resources managers globally have developed resilient technologies and institutional systems that balance competing needs for ecosystem functioning in addition to water availability, demand, and quality; flood and coastal storm risk reduction; navigation; and energy production and demand. Water resources monitoring, data collection, and data dissemination systems are highly automated, involve both in-situ and remote sensors, and are scalable on a global level. Wolf et al proposed that understanding of future water-related conflict could benefit from an approach focusing on "new technologies for negotiation and management; globalization, privatization and the WTO [World Trade Organization]; the geopolitics of desalination; and the changing sources of water and the changing nature of conflict."[67] All of these are within the sphere of water resources management. The examples presented illustrate the increasing opportunity that water resources managers, if adequately prepared, are uniquely positioned to provide both adaptive capacity through operational, demand management, and infrastructure changes[68] and mitigation capacity through innovative power generation and emission control strategies. Given the central role of water management in the implementation of the political, social, economic and administrative systems for water governance, the water management community must continue to implement approaches such as enhanced planning for sustainability that minimize the conflict associated with climate-induced changes in water availability.

Conclusions

Climate variability and change have and will result in changes to the hydrologic cycle that impact water availability and quality. These changes and the resulting climate risk vulnerabilities have the potential to increase conflicts and destabilize already stressed societies. However, increased water conflicts are not inevitable; water resources managers already have skills to deal with balancing competing uses that could help to mitigate conflict. Approaching water resources issues impacted by climate change from a "wicked problems" view can help identify where adaptation and mitigation measures may result in potential conflicting outcomes. Water resources managers are well positioned to develop and implement adaptively managed solutions to achieve positive outcomes in multiple stress environments when managing for multiple objectives. Drawing on the experience of the water resources and the watershed management communities, applying this approach early in the intervention process is critical to securing and maintaining stability. This approach allows for managing risks proactively rather reacting to prevailing crises and conflicts as climate changes.

↑ Back to the top

---

[1] Intergovernmental Panel on Climate Change (2000) "Special report on emissions scenarios." Nakicenovic, N., and Swart, R., eds.: Cambridge, United Kingdom, Cambridge University Press. (Also available online at http://www.grida.no/climate/ipcc/emission/.)

[2] Intergovernmental Panel on Climate Change (2007a) "Summary for Policymakers, of Climate change 2007—The physical science basis." in Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., and Miller, H.L., eds., Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change: Cambridge, United Kingdom, Cambridge University Press. (Also available online at http://www.ipcc.ch/ipccreports/ar4-wg1.htm.)

[3] Intergovernmental Panel on Climate Change (2008) Climate Change and Water. Intergovernmental Panel on Climate Change, 28th Session, April 9–10, 2008, Budapest, Hungary. (Also available online at http://www.ipcc.ch/meetings/session28/doc13.pdf.)

[4] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[5] Wolf, A.T., A Kramer, A. Carius, G.D. Dabelko (2005) Chapter 5, "Managing Water Conflict and Cooperation." In M. Renner, et al. (Eds.), *State of the World 2005: Redefining global security* (pp. 80-95) Worldwatch Institute: Washington, DC.

[6] Gleick, P.H. (1993) "Water and Conflict: Fresh Water Resources and International Security." *International Security*, Vol. 18, No. 1, pp. 79-112.

[7] Gleick, P.H. (2008) "Water Conflict Chronology as of November, 2008." Data from the Pacific Institute for Studies in Development, Environment, and Security database on Water and Conflict, available at http://www.worldwater.org/conflictchronology.pdf.

[8] Gleick, P.H. (1993) "Water and Conflict: Fresh Water Resources and International Security." *International Security*, Vol. 18, No. 1, pp. 79-112.

[9] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[10] Reuveny, R. (2007) "Climate change-induced migration and violent conflict." *Political Geography*, Volume 26, Issue 6, Climate Change and Conflict, pp. 656-673.

[11] See http://www.un.org/News/Press/docs/2007/sc9000.doc.htm

[12] Schwartz, P. and D. Randall (2003) *An Abrupt Climate Change Scenario and Its Implications for United States National Security*. Environmental Media Services: Washington DC (available at http://www.accc.gv.at/pdf/pentagon_climate_change.pdf).

[13] Rittel, H.W.J and M.M Webber (1973) Rittel, Horst, and Melvin Webber; "Dilemmas in a General Theory of Planning," *Policy Sciences*, Vol. 4, pp. 155-169. Reprinted in N. Cross (ed.), *Developments in Design Methodology*, J. Wiley & Sons: Chichester, UK, 1984, pp. 135-144.

[14] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) "Climate change and water resources management: A federal perspective." *USGS Circular 1331*, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[15] Milly, P.C.D, J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer (2008) "Stationarity Is Dead: Whither Water Management?" *Science* 319 (5863) p. 573.

[16] Stern, N. (2009) "The Global Deal: Climate Change and the Creation of a New Era of Progress and Prosperity." *Public Affairs*: New York.

[17] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) Climate change and water resources management: A federal perspective. *USGS Circular 1331*, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[18] Nicholls, R.J., P.P. Wong, V. Burkett, C.D. Woodroffe, and J. Hay (2008) "Climate change and coastal vulnerability assessment—Scenarios for integrated assessment": *Sustainability Science*, v. 3, no. 1, p. 89–102.

[19] Intergovernmental Panel on Climate Change (2007b) "Climate change 2007—Impacts, adaptation and vulnerability." *in* Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., and Hanson, C.E., eds., Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK. (Also available online at http://www.ipcc.ch/ipccreports/ar4-wg2.htm).

[20] Intergovernmental Panel on Climate Change (2007b) "Climate change 2007—Impacts, adaptation and vulnerability." *in* Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., and Hanson, C.E., eds., Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK. (Also available online at http://www.ipcc.ch/ipccreports/ar4-wg2.htm).

[21] Intergovernmental Panel on Climate Change (2008) *Climate Change and Water*. Intergovernmental Panel on Climate Change, 28th Session, April 9–10, 2008, Budapest, Hungary. (Also available online at http://www.ipcc.ch/meetings/session28/doc13.pdf.)

[22] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) "Climate change and water resources management: A federal perspective." *USGS Circular 1331*, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[23] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) Climate change and water resources management: A federal perspective. *USGS Circular 1331*, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[24] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[25] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[26] Gleick, P.H. (1993) "Water and Conflict: Fresh Water Resources and International Security." *International Security*, Vol. 18, No. 1, pp. 79-112.

[27] Gleick, P.H. (2008) "Water Conflict Chronology as of November, 2008." Data from the Pacific Institute for Studies in Development, Environment, and Security database on Water and Conflict, available at http://www.worldwater.org/conflictchronology.pdf.

[28]Barnett, J. and W.N. Adger (2007) "Climate change, human security and violent conflict." *Political Geography*, Vol. 26, pp. 639-655.

[29] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[30] Hendrix, C.S. and S.M. Glaser (2007). "Trends and triggers: Climate, climate change and civil conflict in Sub-Saharan Africa." *Political Geography*, Vol. 26, pp. 696-715.

[31] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[32] Alcamo, J. and T. Henrichs (2002) "Critical regions: A model-based estimation of world water resources sensitive to global changes." *Aquatic Sciences*, Vol. 64, No. 4, pp. 352-362.

[33] Trondalen, J.M. (2009) "Climate Changes, Water Security and Possible Remedies for the Middle East." The United Nations World Water Assessment Programme Side Publications Series Scientific Paper. United Nations Educational, Scientific, and Cultural Organization: Paris, France.

[34] Stern, N. (2009) "The Global Deal: Climate Change and the Creation of a New Era of Progress and Prosperity." *Public Affairs*: New York.

[35] Nordas, R. and N.P. Gleditsch (2007) "Climate change and conflict." *Political Geography*, Vol. 26, pp. 627-638.

[36] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[37] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[38] Hendrix, C.S. and S.M. Glaser (2007). "Trends and triggers: Climate, climate change and civil conflict in Sub-Saharan Africa." *Political Geography*, Vol. 26, pp. 696-715.

[39] Nordas, R. and N.P. Gleditsch (2007) "Climate change and conflict." *Political Geography*, Vol. 26, pp. 627-638.

[40] Ravnborg, H.M. (2004) Water and Conflict - Conflict prevention and mitigation in water resources management. *DIIS Report* 2004:2, 99 p.

[41] Rogers, P. and A. Hall. (2003), Effective Water Governance., TEC Background Papers No. 7. Global Water Partnership/SIDA, Stockholm.

[42] Rittel, H.W.J and M.M Webber (1973) Rittel, Horst, and Melvin Webber; "Dilemmas in a General Theory of Planning," *Policy Sciences*, Vol. 4, pp. 155-169. [Reprinted in N. Cross (ed.), Developments in Design Methodology, J. Wiley & Sons: Chichester, UK, 1984, pp. 135-144.]

[43] Camillus, J.C. (2008) "Strategy as a Wicked Problem." *Harvard Business Review*, May: 99-131.

Carpenter, S.R., P.L. Pingali, E.M. Bennett, and M.B. Zurek (eds) (2005) "Ecosystems and human well-being: Scenarios, Volume 2." *Findings of the Scenarios Working Group, Millennium Ecosystem Assessment*. Island Press: Washington, DC. Available at http://www.millenniumassessment.org/en/Scenarios.aspx.

[44] Kahane, A. (2007) *Solving Tough Problems*. Berry-Koehler Publishers, Inc: San Francisco.

[45] Freeman, D.M. (2000) "Wicked water problems: Sociology and local water organizations in addressing water resources policy." *Journal of the American Water Resources Association* 36 (3): 483–491.

[46] Mills, B. (ed.) (2008) "SERA North: Economics of Weather, Climate, and Climate Change." Synthesis of a meeting held 21-22 February, Waterloo, Canada. Adaptation and Impacts Research Division, Environment Canada: Waterloo, Canada.

[47] White, K.D., A. Dornstauder, S.Durden, and J. C. Shockley (2009) "Systems Approach to Water Resources Decision Making in the US Army Corps of Engineers." Proceedings, 2009 AWWA Water Resources Symposium: "Managing in Times of Change and Uncertainty" [Portland, Oregon, January 25-27 2009]. Electronic proceedings.

[48] Freeman, D.M. (2000) "Wicked water problems: Sociology and local water organizations in addressing water resources policy." *Journal of the American Water Resources Association* 36 (3): 483–491.

[49] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[50] Milly, P.C.D, J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer (2008) "Stationarity Is Dead: Whither Water Management?" *Science* **319** (5863) p. 573.

[51] Folke, C., S. Carpenter, T. Elmqvist, L. Gunderson, C.S. Holling, and B. Walker (2002) "Resilience and sustainable development: Building adaptive capacity in a world of transformations." *Ambio*, 31(15): 437-440.

[52] Folke, C., S. Carpenter, T. Elmqvist, L. Gunderson, C.S. Holling, and B. Walker (2002) "Resilience and sustainable development: Building adaptive capacity in a world of transformations." *Ambio*, 31(15): 437-440.

[53] Carpenter, S.R., P.L. Pingali, E.M. Bennett, and M.B. Zurek (eds) (2005) "Ecosystems and human well-being: Scenarios, Volume 2." *Findings of the Scenarios Working Group, Millennium Ecosystem Assessment*. Island Press: Washington, DC. Available at http://www.millenniumassessment.org/en/Scenarios.aspx.

[54] Shoemaker, P.J. and V.M. Mavaddat (2000) "Scenario planning for disruptive technologies." in Day, G.S., P.J. Shoemaker, and R.E. Guenther (eds) *Wharton on Managing Emerging Technologies.* John Wiley & Sons, Inc: Hoboken, NJ.

[55] Kahane, A. (2007) *Solving Tough Problems*. Berry-Koehler Publishers, Inc: San Francisco.

Kundzewicz, Z. W. (2008) "Flood risk and vulnerability in the changing climate." *Annals of Warsaw University of Life Sciences - SGGW. Land Reclamation*, Vol. 39, pp. 21-31.

[56] IPCC 2007d.

[57] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) Climate change *and water resources management: A federal perspective*. USGS Circular 1331, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[58] Prairie, J., Nowak, K. Rajagopalan, B., Lall, U. and Fulp, T., (2008) "A stochastic nonparametric approach for streamflow generation combining observational and paleo reconstructed data." *Water Resources Research*, 44, W06423, doi:10.1029/2007WR006684.

[59] USGS Circular 1331, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

[60] Stern, N. (2006) Stern Review on the Economics of Climate Change. Cambridge University Press: Cambridge, UK, p. 56.

[61] Trondalen, J.M. (2009) "Climate Changes, Water Security and Possible Remedies for the Middle East." The United Nations World Water Assessment Programme Side Publications Series Scientific Paper. United Nations Educational, Scientific, and Cultural Organization: Paris, France.

[62] Sunstein, C.R. (2005) "Laws of Fear: Beyond the Precautionary Principle." University Press: Cambridge, England.

[63] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[64] Kundzewicz, Z. W. (2008) "Flood risk and vulnerability in the changing climate." *Annals of Warsaw University of Life Sciences - SGGW. Land Reclamation*, Vol. 39, pp. 21-31.

[65] Stern, N. (2006) *Stern Review on the Economics of Climate Change*. Cambridge University Press: Cambridge, UK, p. 56.

[66] Rajagopalan, B., Nowak, K., Prairie, J., Hoerling, M., Harding, B., Barsugli, J., Ray, A., and B. Udall (2009). "Water Supply Risk on the Colorado River: Can Management Mitigate?" *Water Resources Research* 45.

[67] Wolf, A.T. S.B. Yoffe, and M. Giordano (2003) "International waters: identifying basins at risk." *Water Policy*, Vol. 5, pp. 29-60.

[68] Brekke, L., J.E. Kiang, J.R. Olsen, D. R. Pulwarty, D. Raff, D.P. Turnipseed, R.S. Webb, and K.D. White (2009) "Climate change and water resources management: A federal perspective." *USGS Circular 1331*, USGS: Reston, VA (http://pubs.usgs.gov/circ/1331/).

Energy

2.4.D. Energy, Africa, and Civil Conflict: What Does the Future Hold? (Richard Stoll)

Author and Organization: Richard J. Stoll, Department of Political Science, Rice University
Contact Information: stoll@rice.edu

Introduction

Traditionally United States foreign policy has not focused on Africa (particularly Sub-Saharan Africa). However, this is likely to change in the future. While there are several reasons for this, one is that Africa will become increasingly important as a source of resources—including oil—for the U.S. At the same time, Africa is rife with civil conflict. Such conflicts create significant humanitarian problems and cause the deaths of millions. It can also be difficult for states undergoing civil conflict to export resources. For both humanitarian and economic reasons, the United States and other countries will have to pay more attention to Africa particularly in the following areas: (1) how to prevent the outbreak of civil conflict, (2) if a decision is made to intervene in a civil conflict, how to do so successfully, and (3) how to prevent these conflicts from recurring.

This paper notes a few select characteristics of Africa, including a brief assessment of the minimal attention from the U.S. to the region and the increasing importance of Africa to the U.S. and other countries. The paper also addresses the prevalence of civil conflict in the region, including a consideration of whether civil conflicts in Africa have different causes or characteristics from civil conflicts in other regions. Finally, the paper considers what we know about civil wars that could be used to reduce their occurrence, duration, and recurrence in Africa.

This paper makes use of findings and conclusions from systematic research into civil conflict. Some of this research is specifically about conflict in Africa, but most of it comes from studies of civil conflict throughout the world. Below is a discussion of why the non-Africa findings presented here are relevant.

The United States and Oil

It is no surprise to learn that the United States is heavily dependent on oil. According to the Department of Energy, oil currently supplies about 40 percent of our total energy demand. Oil is particularly important in transportation, supplying over 99 percent of the fuel the U.S. uses in cars and trucks.[1] While there has been much talk about "energy independence,"[2] the truth of the matter is that it is highly unlikely that the drastic changes will be undertaken that would be necessary to make this happen. If there is no drastic change in U.S. energy usage, then from 2010 to 2030, U.S. oil consumption is projected to rise from 22.2 million barrels of oil per day to 27.6 million barrels per day, an increase of about 24 percent. At the global level, the projected increase in oil consumption is almost 29 percent. This growth will be led by China, with an anticipated increase in its oil consumption of 72 percent.[3]

If the demand for oil is forecast to rise by 29 percent, what are the prospects on the supply side? Unfortunately, over the same time period, oil production is only expected to rise by about 23.5 percent.[4] While all projections are subject to error, it is prudent to assume that there the demand

for oil will increase and accommodating this increase will be difficult. Even though Africa's oil production (excluding North Africa) will be a little less than 9 percent of total global production, this is not a trivial amount because the world will need every possible source of supply to meet its demands.

The United States and Africa

Africa has never been a high priority region for the United States. In the wake of the Cold War (with the end of the superpower rivalry for influence), attention declined even more. For example, in 1998 Africa received only 13.3 percent of U.S. foreign aid.[5] While this proportion is about the same as Africa's proportion of the world population, it does not take into account the vast development gap between Africa and the rest of the world. Taking the world as a whole, gross domestic product (GDP) per capita in 1998 (current dollars) was $5025. In Sub-Saharan Africa, the figure was $717,[6] or about 14 percent of the global value. This comparison demonstrates that Africa (particularly Sub-Saharan Africa) is desperately poor.

However, Africa has begun to receive more attention from the United States. In 2008, due primarily to President Bush's AIDS initiative, Africa's share of U.S. foreign aid has risen to almost 29 percent.[7] Furthermore, the U.S. military has established a new unified combatant command for Africa (AFRICOM). Both of these efforts came about in response to serious problems in Africa: the first because of the dramatic spread of AIDS throughout Africa, and the second due to the increased concern that failed (and failing) states in Africa would provide safe haven for terrorist groups such as Al-Qaeda.

Humanitarian concerns and counter terrorism operations are two important reasons for the United States to devote more attention to Africa. But there are others. The National Intelligence Council believes that *by 2015*, about one-quarter of the oil in the U.S. market will come from Africa: Angola, Nigeria, Congo-Brazzaville, Gabon, Cameroon, and Equatorial Guinea.[8]

Civil Conflict in Africa

Africa has a significant problem with civil (internal) conflict. The University of Uppsala and the International Peace Research Institute in Oslo, Norway, maintain one of the most widely used datasets on internal conflict.[9] They code an internal conflict if at least 25 people are killed in a year. If we tally each year in each country in Africa that meets this standard, between 1946 and 2006 there were 397 country-years of internal conflict in Africa. This is a large number. Civil conflict has a number of negative effects. First and foremost is that civil conflict kills, wounds, and displaces people. And many of these people are not combatants but are simply in the wrong place at the wrong time. Second, infrastructure is damaged. Third, the economy is damaged. Some parts of it are destroyed, while others are distorted. Since oil is so valuable, during a civil conflict the oil industry of the country is a prime target. The government wants to protect, insurgents want to control it, and others simply want to hijack it. Consider what has happened in Iraq during the insurgency. It is estimated that in 2007, $2 billion of the oil production of the Baiji refinery (located 110 miles north of Baghdad) was siphoned off by the black market.[10]

Between 1960 and 2000, nearly 20 African countries (40 percent of Sub-Saharan Africa) experienced at least one year of civil conflict.[11] Does this prevalence indicate that civil conflicts in Africa are fundamentally different from civil conflicts in other regions? This is an important question because, if Africa is different, then we cannot apply what we know from general studies of civil conflict.

Researchers who have systematically examined this question have concluded that the causes of civil conflict Africa are no different than the causes of civil conflict elsewhere.[12] The most important factors associated with the outbreak of civil conflict are as follows:

1. Low levels of economic development
2. Low rates of economic growth
3. Dependence on primary commodity exports (note this relationship is curvilinear; the chances of civil conflict increase with dependence, but beyond a point, the chances of civil conflict decrease)
4. Failed political institutions

The reason that Africa has such a high incidence of civil conflict is not because it is unique but because it has a number of countries with the conditions that are associated with the outbreak of civil conflict. From this we can conclude that (1) we can use the knowledge we have gain from the study of civil conflict in general to predict and understand civil conflict in Africa, and (2) (more importantly) as countries in Africa develop economically, this should change the conditions listed above and reduce the frequency of civil conflict in that continent.

What Should the U.S. Do? Strategy Choices

The U.S. cannot (and should not) take on the task of intervening in every potential or ongoing civil conflict in Africa. Interventions can be costly and take significant amounts of time; they should not be undertaken lightly. However, if a decision is made to intervene, what should the United States do to prevent the outbreak of civil conflict, to stop it once it begins, or to prevent it from reoccurring?

Preventing the Outbreak of Civil Conflict

The most effective long-term strategy to prevent the outbreak of civil conflict is to promote economic development. This will have a direct impact on the first three factors associated with the outbreak of civil conflict listed above. It may also have an indirect impact on the fourth factor. Economic development and democracy are tightly intertwined; a more developed economy may facilitate the development of stronger political institutions.

It is beyond the scope of this paper to offer an extensive set of suggestions. (This is best left up to those who specialize in the study of economic development). The work of Hernando De Soto[13] demonstrates that if the system of laws and regulation in a country make property rights insecure, people will in effect withdraw from the economy and live and work "off the grid." The result is that significant amounts of capital become unavailable to the regular economy. That is, it is not enough to have rule of law, it must be *reasonable* rule of law to provide secure property rights for all individuals. This will increase the amount of capital available in the country. Second, development efforts such as those being implemented in Iraq and Afghanistan are often hampered by corruption. The same can be said in Africa, and efforts to reduce corruption can help speed development. Finally, the Bush-initiated effort to combat AIDS in Africa should be seen not only as a humanitarian and health initiative, but also a key element of the effort to speed development by saving the lives of people who are critical to moving these countries forward.

Intervention in Civil Conflict

Once a civil conflict starts, should the United States intervene? If it is in our interest to intervene, what should be done? That depends critically on the goals of the United States. Is the goal to see one side win the war, or is the goal to stop the conflict? This is a critical matter because the

systematic research on intervention in civil conflicts makes it very clear that the factors that are associated with success in winning a war are not the same as those associated with stopping a war.

Winning the Conflict. An external military intervention for either side in a civil conflict increases its chances of victory. Military intervention for either side also increases the chance of a negotiated settlement. The effects are stronger when intervening for the rebels than for the government.[14] However, military intervention by outsiders also lengthens the duration of the civil conflict,[15] and of course the longer the conflict goes on, the greater the costs. If the United States wishes to have one side prevail in a civil conflict, intervention can help its choice to prevail but it is a costly option.

Stopping the Conflict. There are other circumstances in which the goal is to bring the conflict to an end as quickly as possible. If the goal is to get the fighting stopped, then non-military means can be used (and in some circumstances are preferred). Diplomacy can be an effective tool to bring the conflict to an end, but timing is important. The relationship between success of diplomacy and duration is curvilinear. Diplomacy is not likely to be successful at the beginning of the conflict, but as time goes on, it is more likely to be successful. However at some point, the efficacy of diplomacy starts to decline once again.[16]

Keeping the Peace

One of the most tragic aspects of civil conflicts is that they are likely to reoccur. If one examines civil wars between 1944 and 1997, more than half of them broke out more than once.[17] So what can be done to reduce the chances of a civil conflict resuming? First of all, the same factors that are associated with the outbreak of civil conflict (see above) are associated with the recurrence of civil conflict. Efforts to facilitate economic development in Africa will also serve to reduce the recurrence of civil conflict. While this may not bear fruit in the short run, it is nevertheless an important element in the effort to reduce the chances of civil conflict occurring or recurring.

What else can be done? One answer is peacebuilding. Peacebuilding "is an attempt, after a peace has been negotiated or imposed, to address the sources of current hostility and build local capacities for conflict resolution."[18] While there are no guarantees of success, peacebuilding is more successful if it is a large-scale operation conducted by the United Nations and involve high levels of financial assistance, rather than a unilateral effort by a single country. Thus, in most cases, the United States should not consider undertaking peacebuilding efforts alone.

Conclusions

Historically, the United States has devoted little attention to Africa, but the situation is changing. The Bush administration dramatically increased funding to combat AIDS globally, and this significantly increased the amount of U.S. aid to Africa. The U.S. concern that failed or weak states in Africa could become willing or unwilling hosts to terrorists or other groups that threaten U.S. interests (e.g., pirates) has increased the amount of military attention to the continent. And as time goes on, there will be at least one additional reason from the United States to devote attention to Africa: we will become more dependent on that continent for energy supplies.

Since African states became independent, many have experienced serious civil conflict. These conflicts create many serious problems, including a decline in oil exports. Of course oil is not the only reason the U.S. should be concerned about civil conflict in Africa. No one wants to see

another Rwanda happen. However, whether one believes that the U.S. should care about Africa for humanitarian reasons or for the energy security of the United States (or both), we must face the fact that Africa is very susceptible to civil conflict.

While it should not be the policy of the United States to intervene in all civil conflicts in Africa, it is important that we take advantage of what scholars have learned about civil conflict and apply it (as well as other forms of knowledge, insights, and experience) to help inform our policy choices. Based on the systematic study of civil conflicts, we would do well to keep the following in mind:

1. Despite the high frequency of civil conflicts in Africa, there is no reason to believe that the causes of African civil conflicts are fundamentally different that the causes of civil conflict in other regions. Research demonstrates that the fundamental causes of civil conflict in Africa are the same as the causes of civil conflict elsewhere.
2. The fundamental factors that drive the outbreak of civil conflict in Africa stem from low levels of development. An important strategy to prevent civil conflict (as well as to improve the lives of those living in Africa) is to facilitate the process of development. Note that programs to reduce the ravages of AIDS are an integral part of this effort. Also, the U.S. should promote the reasonable rule of law to keep capital from disappearing from the economies of African states.
3. If the United States favors one side over the other in a civil conflict, direct (military) intervention can help that side win. Military intervention is typically more helpful for the insurgent side than for the government side. But military intervention on either side in a civil conflict is likely to make it last longer and drive up the costs of the war.
4. If the goal of the United States is to end the conflict, then diplomacy can play an important role. However, timing is important. Success is less likely if diplomacy takes place too early or too late in the conflict.
5. It is difficult to keep a civil conflict from reigniting. The same factors that are associated with civil conflict (low levels of development) also help to restart it. Therefore, efforts to aid economic development of a country will also reduce the chances of civil conflict restarting.
6. Robust multilateral peacebuilding efforts run by the United Nations can also reduce the chances of a civil conflict restarting.

The next few decades will provide a number of challenges for American foreign policy. Energy security is one of the most important of these. As we move forward, our energy security will become more closely tied to events in Africa. Given the high frequency of civil conflict in Africa, the United States has to consider its options to prevent, stop, and reduce the recurrence of civil conflict in that continent.

↑ Back to the top

---

[1] Department of Energy. 2009. Energy Sources: Oil. http://www.energy.gov/energysources/oil.htm.

[2] Woolley, John and Gerhard Peters. 2009. Radio Address About the National Energy Crisis. The American Presidency Project. http://www.presidency.ucsb.edu/ws/index.php?pid=4208&st=energy&st1=independence

[3] Energy Information Administration. 2009. Oil (Petroleum) Consumption. Forecasts & Analyses. http://www.eia.doe.gov/oiaf/ieo/pdf/ieoreftab_4.pdf.

[4] Energy Information Administration. 2009a. International Energy Outlook 2009. Report #:DOE/EIA-0484(2009). Table G1 World Total Liquids Production by Region and Country, Reference Case. http://www.eia.doe.gov/oiaf/ieo/index.html

[5] Tarnoff, Curt and Marian Lawson. 2009. Foreign Aid: An Introduction to U.S. Programs and Policy. Congressional Research Service. R40213, p. 15.

[6] United Nations Statistics Division – National Accounts. Available at: http://unstats.un.org/unsd/snaama/dnllist.asp

[7] Tarnoff, Curt and Marian Lawson. 2009. Foreign Aid: An Introduction to U.S. Programs and Policy. Congressional Research Service. R40213, p. 15.

[8] Garrett, Laurie. 2005. HIV and National Security: Where Are the Links? Council on Foreign Relations. Available at: http://www.cfr.org/publication/8256/hiv_and_national_security.html.

[9] Harbom, Lotta. 2008. UCDP/PRIO Armed Conflict Dataset Codebook. Version 4-2008. Uppsala Conflict Data Program/ International Peace Research Institute, Oslo.

[10] Oppel, Richard A. 2008. Five Years In; Iraq Insurgency Runs on Stolen Oil Profits. New York Times March 16, 2008, Section A, Page 1

[11] Elbadawi, Ibrahim, and Nicholas Sambanis. 2000. "Why Are There So Many Civil Wars in Africa? Understanding and Preventing Violent Conflict." *Journal of African Economies* 9(3), p. 244

[12] See Collier, Paul, and Anke Hoeffler. 2002. "On the Incidence of Civil War in Africa." *Journal of Conflict Resolution* 46(1): 13-28, and Elbadawi, Ibrahim, and Nicholas Sambanis. 2000. "Why Are There So Many Civil Wars in Africa? Understanding and Preventing Violent Conflict." *Journal of African Economies* 9(3): 244-269.

[13] Soto, Hernando De. 2003. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. Hew York. Basic Books.

[14] Balch-Lindsay, Dylan, Andrew J. Enterline, and Kyle A. Joyce. 2008. "Third-Party Intervention and the Civil War Process." *Journal of Peace Research* 45(3):345-363.

[15] Elbadawi, Ibrahim, and Nicholas Sambanis. 2000. "External Interventions and the Duration of Civil Wars." Policy Research Working Paper 2433. The World Bank Development Research Group Public Economics.

[16] Regan, Patrick M., and Aysegul Aydin. 2006. "Diplomacy and Other Forms of Intervention in Civil Wars." *Journal of Conflict Resolution* 50(5):736-756.

[17] Quinn, J. Michael, T. David Mason, and Mehmet Gurses. 2007. "Sustaining the Peace: Determinants of Civil War Recurrence." *International Interactions* 33(2). p. 167.

[18] Doyle, Michael W., and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *The American Political Science Review* 94(4), p. 779.

2.4.E. Energy: A National and Global Issue (Douglas Arent)

Author and Organization: Douglas J. Arent, National Renewable Energy Laboratory
Contact Information: Douglas.Arent@nrel.gov

Executive Summary

Global energy use has grown more than 20-fold over the past two centuries. The primary determinants of energy use are population, economic activity per capita, and energy use per dollar of economic activity or "energy intensity." In the 20th century, the world has experienced a four-fold growth in population, an explosion in economic development, and a gradually decreasing trend in energy intensity. Energy use in individual countries has evolved in phases related to economic development, starting from a take-off in industrialization and personal mobility, to slowing industrialization, and then to a transition to services.

Projected trends for the 21st century in a business-as-usual scenario are (1) an average annual population growth of 0.9 percent, (2) growth in economic activity per capita of 1.4 percent, and (3) decrease in energy intensity of 0.8 percent. These trends imply an average annual increase in energy use of 1.5 percent, or a doubling by mid-century from current levels of about 400 quadrillion British thermal units (BTUs) and a tripling from current levels by the end of the 21st century. A large fraction of the global growth in energy use by mid-century is expected to be in industrializing economies, particularly China and India.[1]

While the world's endowment of energy resources is sufficient to meet anticipated demand through the century, existing fossil energy supplies will become increasingly strained, with anticipated economic and geopolitical consequences, as well as continued undesirable consequences for the environment. New energy pathways are needed. These pathways must balance increasingly complex policy goals of accessibility, environmental concerns, and geopolitical issues, and they must consider affordability in the U.S. and throughout the world. Forces are driving fundamental changes in both the transportation and electric power sectors, where technologies that have allowed tremendous economic and industrial development over the past century are increasingly less desirable when viewed in light of the evolving geopolitical and environmental framework. Continued reliance on geographically concentrated oil and natural gas will threaten international energy security. The continued growth in energy demand and continued use of fossil fuel energy to meet that demand will pose various issues: the steady depletion of fossil fuels and the consequent increase in fossil fuel prices and price volatility, reliance on oil and natural gas from a few unstable regions, and increased concentrations of carbon dioxide in the atmosphere. Alternative energy pathways, including those that significantly reduce geopolitical concentrations of risk, utilize indigenous resources of each country in environmentally benign ways, and promote both energy and economic diversity, are experiencing rapid global market growth, with more than $150B in transactions in 2007 and 2008. The global energy sector is enormous, representing many trillions of investment, and processing more than 40,000 gallons of liquid fuels per second, with capital intensive requirements and long-lived infrastructure. Transitions should be expected to take multiple decades and will require careful navigation of an increasingly complex set of global interdependencies.

World Energy History

Global energy use has grown by a factor of more than 20 over the past 200 years. This 20-fold increase has far exceeded the 6-fold growth in world population. Figure 2.5.B-1 shows the increase in primary energy use and in population, for the developing world and the (now) industrialized world.

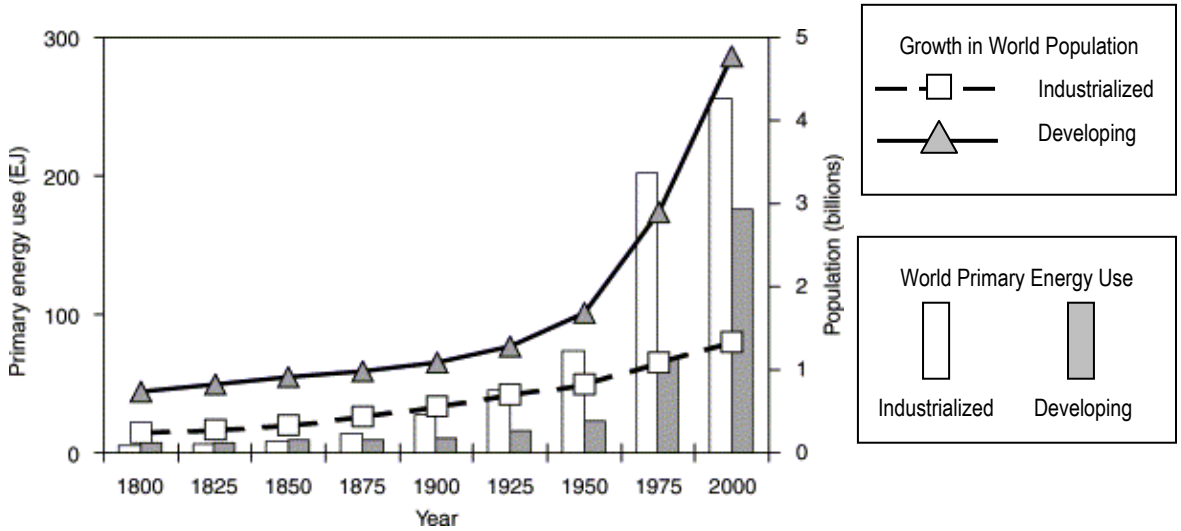


Figure 2.5.B-1, Increase in Primary Energy Use and Population for Developing and Industrialized World.[2]
EJ: exajoules or 10^{18} joules (1.055 exajoules equal 1 quadrillion BTUs)

Growth by Energy Type and Drivers of Major Transitions. There have been several drivers of this significant increase in primary energy use per capita. In the early 19th century, the world's population was still overwhelmingly rural. Dominant energy converters were human labor and draft animals. With industrial changes starting in mid-century, steam engines and turbines began to replace people and draft animals as the dominant energy converters. The expanding trade in fossil fuels, first coal and later oil and natural gas, overcame energy supply limitations imposed by locally available renewable energy. Populations shifted to cities, and manufacturing activities displaced agriculture as the primary source of economic output. Changes in economic activity resulted in increased transportation of both goods and people by steamboats and railways.

In the 20th century, an explosion in economic development led to rapid growth in energy usage. Diffusion of automobiles greatly expanded oil use in transportation. Consumer preferences for cleaner and more convenient energy services led to a transition from direct use of coal to direct use of oil and gas and then to the use of electricity. Figure 2.5.B-2 shows the growth in world primary energy consumption by source from 1970 to 2004.

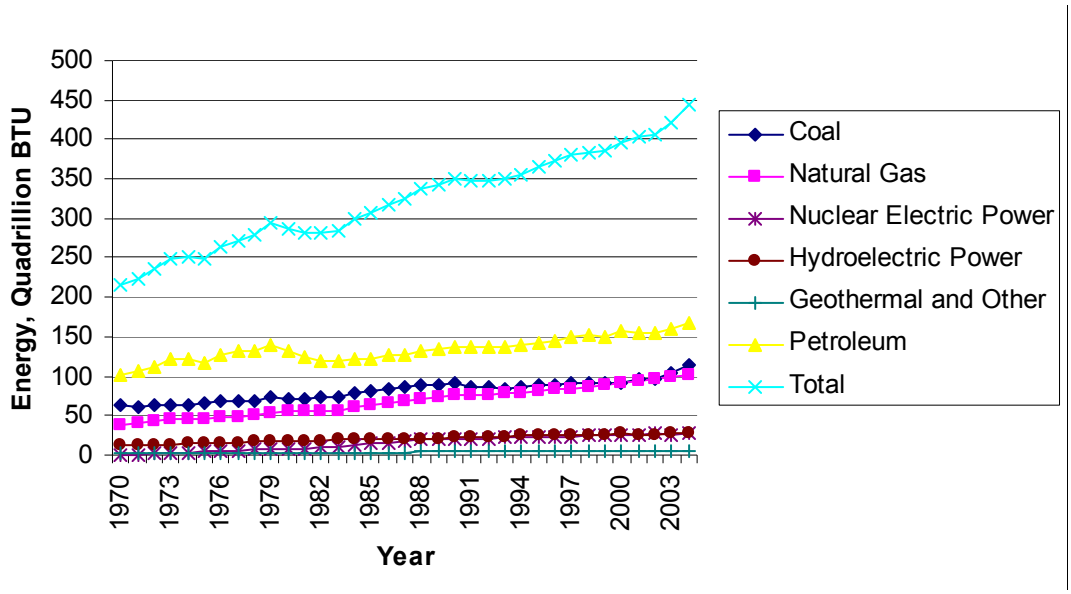


Figure 2.5.B-2. World Energy Consumption by Source, 1970-2004.[3]

Scale of the World Energy Challenge. The scale of the global energy system today is enormous. As detailed above, annual consumption of around 400 Quads of energy includes worldwide consumption of nearly 80mbpd (million barrels per day) of petroleum products (approximately 40,000 gallons per second!), an installed electric generation capacity of around 4300 TW, and continued reliance on traditional biomass for a significant (e.g., 1–2 billion) population. Estimates of the installed capital (while very approximate) range from $20–100 trillion. Over the next 20 years, the International Energy Agency (IEA) estimates that $20–50 trillion will be required to meet growing demand and to decarbonize the world energy sector to address climate change.

There are three principle driving factors for energy consumption: population, economic growth, and energy intensity of the economy. Algebraically,

Energy consumption = Population × GDP/pop × Energy/GDP.

A concise articulation was put forth by Hoffert et al, including implications for atmospheric CO_2.[4] Hoffert outlines historic and projected world population growth, GDP growth (globally) and energy consumption as a function of GDP. His key findings are shown in Figure 2.5.B-3 below. The world population was ~ 6.1 billion in 2001 and is projected to increase by 0.9%/yr to ~ 9.4 billion by 2050. World per capita GDP was ~$7,500 per capita in 2001 and is projected to increase at the historical average rate of 1.4%/yr to ~$15,000 per capita by 2050. Including an estimated continued decrease in global energy intensity of 0.8%/yr, the world energy consumption rate would grow at a rate of 1.5%/yr, from a level of 400 Quads today to around 800 Quads in 2050, and to triple to 1,200 Quads by 2100.

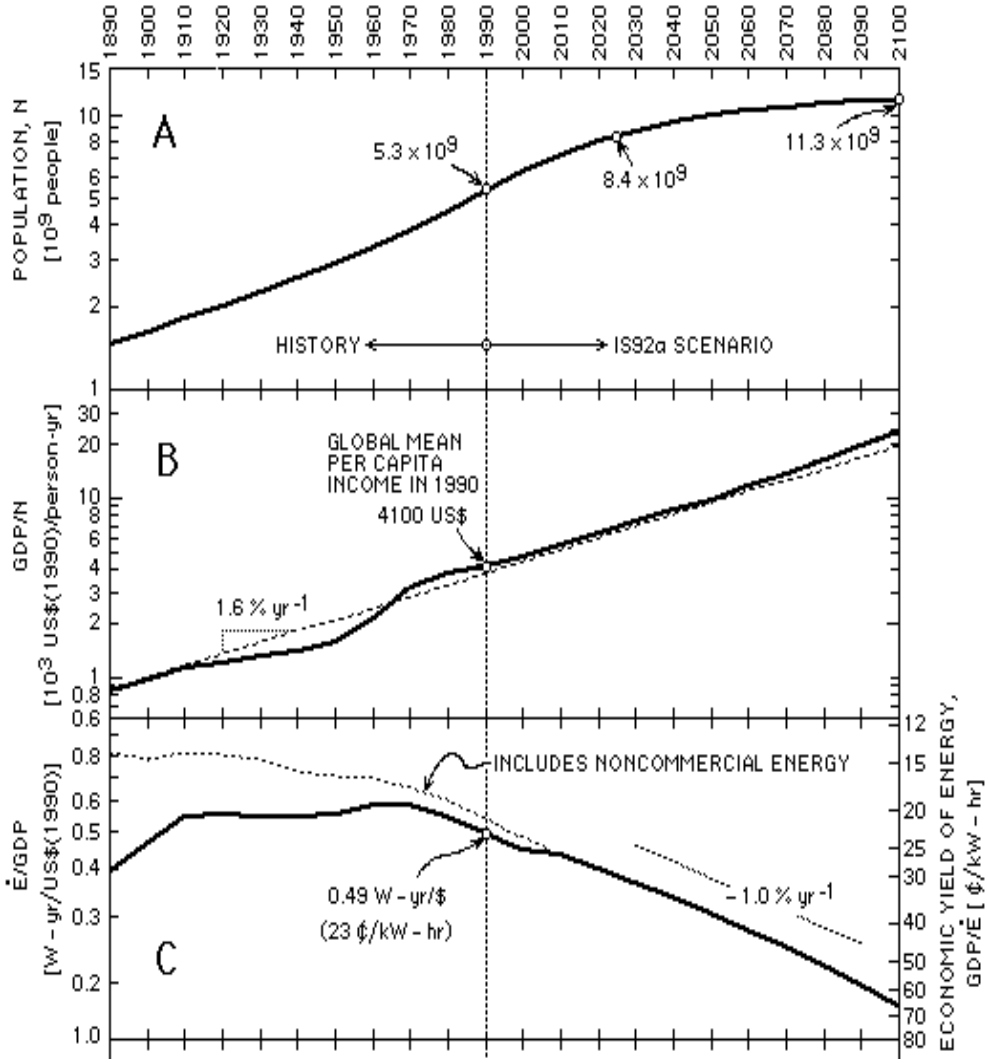


Figure 2.5.B-3: Historical and Projected World Population, Economic Growth and Energy Intensity for Scenario IS92a. [5]

World Energy Future. Multiple agencies and institutions publish scenarios of possible world energy futures. Scenarios from the Energy Information Agency (EIA) and IEA are most frequently referenced. Utilizing similar techniques and models, other institutions publish alternative scenarios based on "materially different" assumptions. One is outlined here for comparison, recognizing that many others are also available for further comparison.[6] An alternative to the future energy scenarios developed by the EIA and IEA was recently published by Greenpeace in a report titled "Energy Revolution: A Sustainable World Energy Outlook."[7] The report compares the authors' "energy [r]evolution scenario" to a reference scenario based on the *IEA World Energy Outlook 2004*. These scenarios are not intended to predict the future, but rather to simply describe two potential development paths out of a broad range of possible futures.

The reference scenario is based on a business-as-usual approach, taking into account only existing policies. "The assumptions include, for example, continuing progress in electricity and gas market reforms, the liberalization of cross border energy trade and recent policies designed

to combat environmental pollution. The reference scenario does not include additional policies to reduce greenhouse gas emissions."[8] As the IEA's scenario ends in 2030, it has been extrapolated to provide a baseline for comparison with the energy [r]evolution scenario.

The energy [r]evolution scenario illustrates the efforts and actions required to achieve the ambitious objectives of limiting the increase in global temperature to under 2° Celsius by 2050 and to phase out nuclear energy. "To achieve these targets, the scenario is characterized by significant efforts to fully exploit the large potential for energy efficiency. At the same time, cost-effective renewable energy sources are accessed for heat and electricity generation, as well as, the production of biofuels."[9] The energy [r]evolution scenario uses the same parameters for population and GDP growth as the reference scenario.

Figures 2.5.B-4 and 2.5.B-5 provide a summary of global primary energy consumption under the two scenarios. Compared to the reference scenario, in the energy [r]evolution scenario, overall energy demand will be reduced by almost 50 percent in 2050. Around half of the remaining demand will be covered by renewable energy sources. In the figures below, efficiency represents reduction *compared to* the reference scenario. *Note that because of the "efficiency method" used for the calculation of primary energy consumption (which postulates that the amount of electricity generation from hydro, wind, solar and geothermal energy equals the primary energy consumption), the share of renewables *seems* to be lower than their actual importance as energy suppliers.

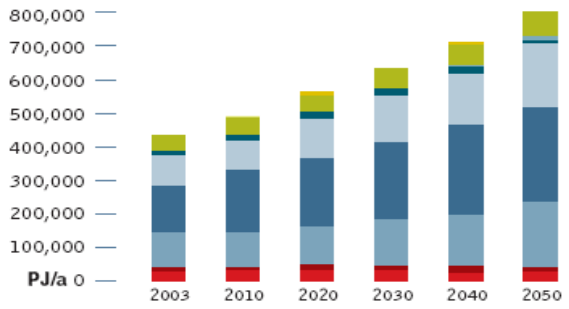


Figure 2.5.B-4 Development of Global Primary Energy Consumption under the Reference Scenario.

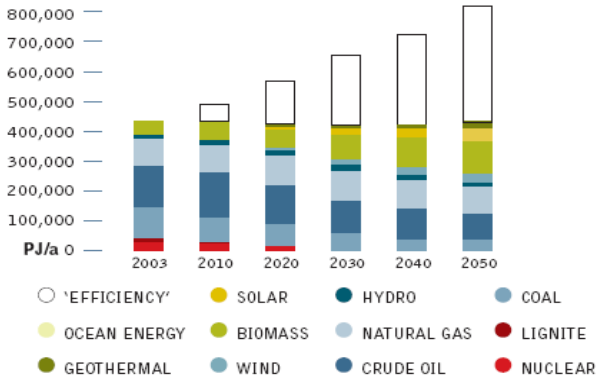


Figure 2.5.B-5. Development of Global Primary Energy Consumption under the Energy Revolution Scenario.

PJ = petajoules or 10^{15} joules (1055 petajoules equal 1 quadrillion BTUs)[10]

U.S. Energy History. The sources of energy in the United States have gone through two major transitions, from primarily wood to coal between 1850 and 1920, and from coal to oil and natural gas between 1920 and 1970. Non-fossil energy sources—nuclear and renewables—have contributed less than 20 percent of annual energy use since 1920. Renewable sources other than hydro-electric power, such as biomass, geothermal, wind, and solar, account for less than 2.5 percent of current annual energy use. According to the EIA, petroleum use in the U.S. is split

among different sectors: 65% transport, 25% "industrial/petrochemical" 4% residential, 1.5% commercial, and 1.5% electric generation.[1]

Energy intensity in the United States economy has continued to decrease, driven by energy policy, technology development and diffusion, and economic trends such as shifting of energy intensive activities to other nations. Between 1980 and 1985, the energy intensity of the economy decreased by almost 20 percent due to energy policies and technological developments spurred by the oil embargo of the 1970s. It declined yet another 20 percent by 2005.

Energy as a Part of the U.S. Economy. The primary determinants of energy use are population, economic activity per capita, and energy use per dollar of economic activity, or "energy intensity." The product of economic activity per capita and energy intensity is energy use per capita. The drivers of trends in energy intensity are energy policies, technological development and diffusion, and economic trends such as shifting of energy intensive activities to other nations. Energy intensity is expected to continue decreasing from 2005 to 2030 at a rate similar to that from 1985 onwards.

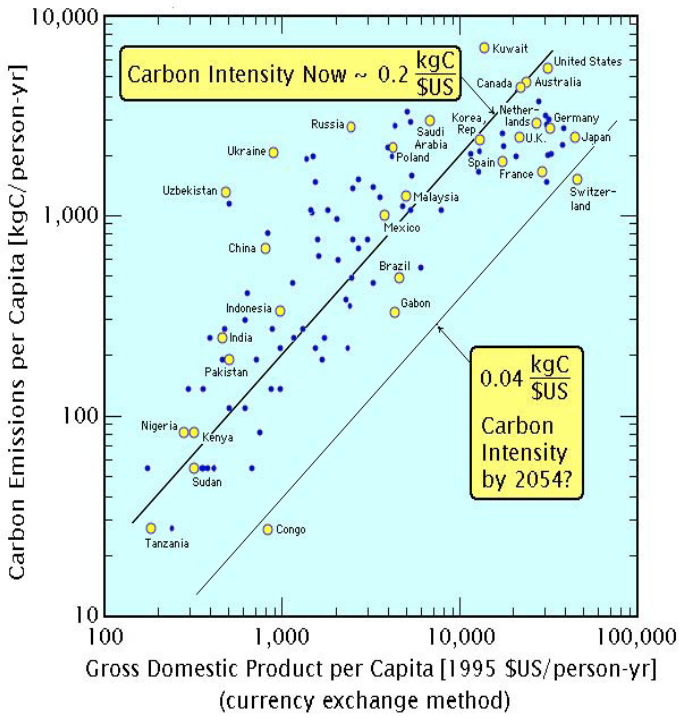


Figure 2.5.B-6: Economic Development and Carbon Intensity is Tightly Correlated with Energy Consumption.

U.S. energy use per capita has remained roughly constant for the last 20 years, with increases in economic activity per capita offsetting decreases in energy intensity. In their "business-as-usual" scenario, the EIA projects a slight increase in U.S. energy use per capita for the years 2005 to 2030 driven by demographic and geographic changes. These changes include an aging population, a predicted migration of the population from the North to the South, and projected increases in income.

While energy use per capita in the U.S. and other industrialized countries has been growing slowly in recent years, it has been growing more quickly in many other countries. The U.S. remains the world's leading energy consumer on a per capita basis and in aggregate. Figure 2.5.B-6 shows the relationship over time between energy use per capita and economic development.[11]

Transitions in the economic and social development of countries and societies are marked by changes in the energy system as they climb the "energy ladder."

[1] EIA Annual Energy Outlook. Available at www.eia.doe.gov

The first and most important step is substituting commercial for traditional fuels. Perhaps one billion people worldwide still rely mainly on biomass fuels and animal wastes. After that there is a strong—but constantly changing—relationship between income and energy demand. When per capita GDP (on a purchasing power parity basis) reaches some:

- *$3,000 – demand explodes as industrialization and personal mobility take off*
- *$10,000 – demand slows as the main spurt of industrialization is completed*
- *$15,000 – demand grows more slowly than income as services dominate economic growth and basic household energy needs are met*
- *$25,000 – economic growth requires little additional energy*[12]

While this is the typical model for economic development, there are other examples. For example, India has a developed, relatively energy efficient services sector which might provide a model for bypassing the development of energy intensive industries to achieve economic growth.

Threats to National Security and Economic Stability

While the world's endowment of energy resources is sufficient to meet anticipated demand through the century, lagging production and distribution infrastructure have strained global energy supplies. In addition, 75 percent of the world's proven oil reserves, and currently half and growing share of oil production, are currently from OPEC nations. The increasing geographic concentration of oil and natural gas supplies, particularly from politically unstable or antagonistic nations, places international energy supplies at risk from regional conflict and from deliberate strategic decisions.

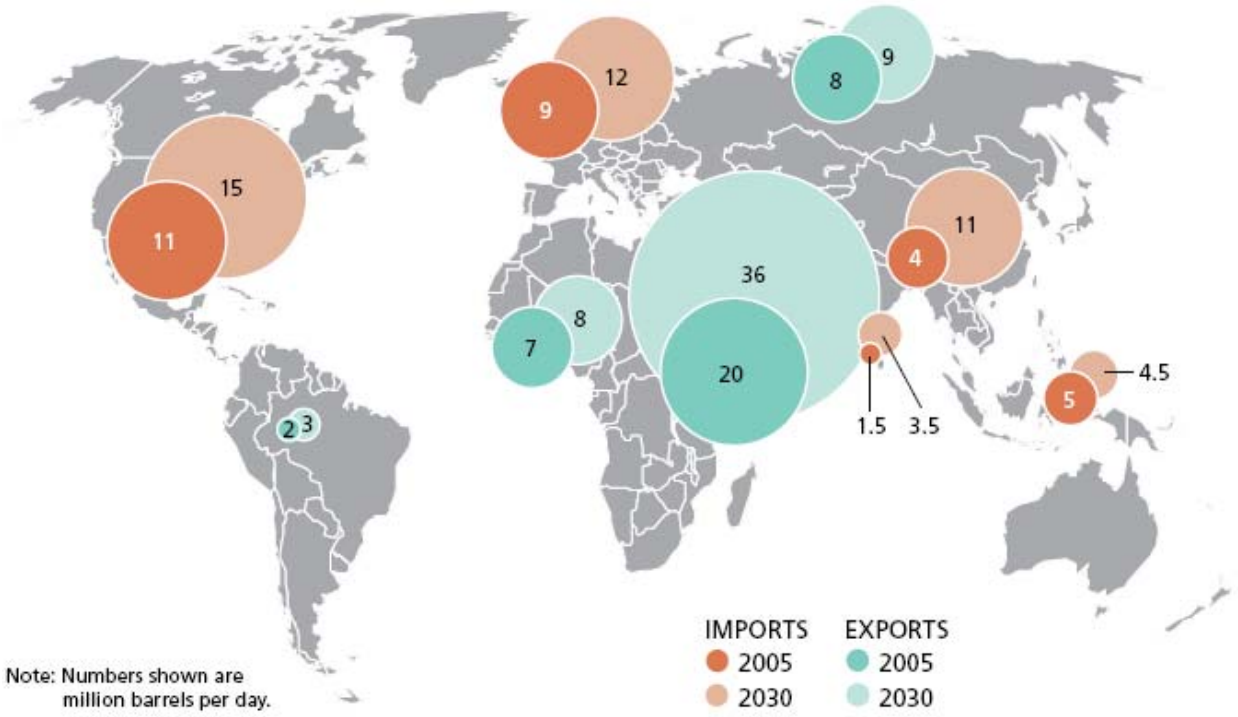
As the world's largest oil consumer, the United States is vulnerable to supply disruptions and price spikes. In the last 10 years, the price of oil has increased 6-fold, to more than $100 a barrel in January of 2008. Increased energy prices have impacts across the economy, particularly affecting households with low discretionary income. Price volatility is also increasing in frequency and scale, impacting the business and investment climate. While 2009 prices have reduced from the peaks of 2008, long-term forecasts, based upon increased global demand, project increasing supply/demand pressures that will inflate prices over the long term.

Dependence on fossil fuels and imported oil poses a growing risk to the national economy, the environment, and national security:

- 75% of the world's proven oil reserves and 50% of the oil production are from OPEC nations, many of whom are controlled by unstable, if not antagonistic, regimes. Six of the seven largest holders of proven reserves are OPEC members, and five of them are in the Middle East.
- Of the 2006 U.S. fossil fuel use of 84.8 quadrillion BTUs, 35.3% was imported[13] (0% of the 26.7% from coal, 4.5% of the 26.4% from NG, and 72.6% of the 46.9% from oil). By comparison, of the 1996 U.S. fossil fuel use of 79.9 quadrillion BTUs, 29.2% was imported.
- The U.S. is the world's largest oil consumer, using a little over 20 million barrels per day in total petroleum products. The five nations currently exporting the most crude oil to the U.S. are Canada, Saudi Arabia, Mexico, Nigeria, and Venezuela. Of these nations, Saudi Arabia, Nigeria, and Venezuela are members of OPEC. In total, OPEC countries supply almost half of U.S. imports.[14]

Supply disruptions and price spikes have direct costs to U.S. consumers and social costs to the U.S. economy. In addition, some fraction of U.S. military expenditures might be attributed to maintaining a stable global energy supply. Finally, global competition for oil has geopolitical and security implications.

Figure 2.5.B-7 shows net regional oil exports and imports from each region of the world in 2005 and projected for 2030.[15]



Source: IEA, *World Energy Outlook 2006*, Reference Case.

Figure 2.5.B-7. Net Regional Oil Exports and Imports.

Overall, under the business-as-usual scenario for Figure 2.5.B-7, the U.S., EU, China and India will increase their net imports over the next 25 years, with the majority of new supply coming from the Middle East. Since the U.S. is the world's largest oil consumer, using four times more oil in absolute terms than any other nation,[16] an increase in U.S. demand can raise world oil prices, although it is unclear if a reduction in U.S. demand will have an elastic effect on world oil prices, particularly with global demand growing strongly. Additionally, increasing U.S. demand for oil can effectively increase the U.S. trade deficit as a result of the higher price paid to foreign suppliers. Moreover, there are environmental costs of oil and other fossil fuel usage that are not priced to the consumer.

The U.S. uses various instruments and organizations to maintain a stable global supply of energy. The United States has incurred substantial military costs in the Persian Gulf and in other oil exporting regions. As a strategic commodity with no viable substitutes in the near future, oil and gas are increasingly being used for geopolitical leverage resulting in global tensions.[17]

Sustainable Energy Technologies and Transition Issues

Reduction in our dependence on fossil fuels, domestically and globally, offers multiple benefits, including energy diversity, increased domestic energy production with concomitant economic

impacts, and reduced environmental impacts. Sustainable energy resources, such as (sustainably harvested) biomass, geothermal, wind, solar, ocean and kinetic energy, hydropower, will require significant increase in scale over the next decades to contribute in a meaningful way to the domestic and global energy economy.[2] Note that no singular definition of "sustainable" is in use globally. For example, large scale hydropower is often not considered "sustainable" given the major environmental impacts associated with damning. Some constituents argue that nuclear power should be reconsidered as sustainable, as it has no carbon emissions; however, it has a separate set of environmental and national security issues.

Adoption rates of these new energy technologies, however, are restrained by the stock of current energy technologies and by energy policy. Capital-intensive energy infrastructure technologies have long lifetimes, ranging from 20 to 100 years. They are also subject to both technological and institutional "lock-in," (i.e., a number of positive feedback loops that favor existing technologies).

Several studies have assessed the potential contribution from various technological options. A number of distinct measures of potential, not necessarily comparable, are often reported. Such measures might reflect the following:

- Theoretical potential (up to the maximum permitted by the laws of physics)
- Technological potential
- Economic potential based on social value
- Economic potential based on private internal value
- Economic potential accounting for market failures, existing infrastructure, and a realistic policy future

The last of these measures is the most applicable and likely the most important in making policy decisions. In addition, it is likely that there is no single technological silver bullet for producing changes at the scale and speed desired. Portfolios of complementary options will need to be evaluated that also take into account the complexities of energy security, the environment, and economic prosperity.

Technology advancements are often thought of in relation to consumer goods and services, such as cell phones and the internet. However, there is a long and deep literature on technology advancement that clearly indicates very large differences in the rate of technology advancement and adoption as a function of the market characteristics, size of individual purchase or investment, and a number of other factors.[18]

Essential findings are as follows:

- Social contagion, social learning, and network effects all influence market penetration of new technology. Exogenous factors such as government mandates and "directional leadership" may play a significant role in shaping those factors.[19]
- Historically, academic R&D precedes product introduction by 4–10 years.[20]
- The average time from invention to "sales takeoff" (accelerating market penetration) is more than 42 years[21]
- Market penetration of consumer technologies (e.g., cell phones, liquid crystal displays

[2] Details on the current status of these technologies is available in the Renewable Energy Data Book, avalailable at http://www1.eere.energy.gov/maps_data/

[LCDs], etc.) is accelerating to less than 10 years

- Consumer-oriented energy technologies show unusually slow adoption rates (e.g., 10 years to more than 50 years for "consumer energy goods" such as low E windows, compact fluorescent light bulbs).[22]
- Capital-intensive, long-lived infrastructure technologies have adoption rates on the order of decades to a century.

Further details are provided in Appendix A.

Many publications address the challenges posed by our current energy economy. Esteemed organizations such as the Council on Foreign Relations and many others have argued that energy (and energy-related climate emissions) threats are now more acute than ever and call for strong leadership.[23] Moreover, the policy environment has evolved to address three principal goals: the environment, economic prosperity, and security.[24]

No one technology approach is envisaged to be capable of addressing the complexities of the energy economy challenges, particularly in the 20–50 year timeframe, with such a large installed capital base. The suite of technologies ranges from conservation to forest management to non-emitting power technologies, to carbon capture and storage (CCS) to biofuels, as discussed further below.

Having benefited from significant long-term expansion of U.S., European and some Asian economies, the developed nations of the world, and particularly the U.S. and Europe continue to create and adapt to help sustain economic development. Moreover, some leaders have been extremely cautious of crafting policies or rules that would inhibit economic development for the benefit of either the environment or security. Others, however, are fully committed to the environmental aspects of energy while crafting policies and strategies to address the economics and security aspects. Going forward, the increasingly strong growth of two new economic powerhouses, China and India, is reshaping the debate.[25] Both with billion-level populations and economies growing 8-10% /yr., their economic power, combined with their foreign investment management, have increasing implications for significant growth in energy requirements. If not approached mindfully, increasing competition among nations for energy supply, conversion technology, and workforce may be anticipated to drive nationalistic interests and decisions. Rapid expansion of the Indian and Chinese energy economies, often with least cost technologies that are not adaptable for tighter environmental requirements, is anticipated to lead to a large installed base where energy lock-in may persist, with significant implications for environment.

There is a growing recognition of the critical role energy plays in the U.S. and global economy and world order. While the U.S. economy has proved (perhaps surprisingly to some) resilient to the recent rise in oil prices, supply disruptions such as the Northeast electricity blackout of 2003 and production and refinery outages due to Hurricane Katrina indicate the relative fragility of the energy system. Similarly, there is a growing recognition of the increasingly complex world energy interdependence and challenges associated with national and global security concerns ranging from supply disruption to geopolitical interventions, from nationalization of assets to price manipulation.[26] A wide range of countries (U.S., China, Japan, India, Brazil and many others) have recently moved to secure long-term supply and to increase development and use of domestic resources in order to lessen their dependencies on imported energy from unstable regions.

Recognizing the characteristics of today's energy economy (including scale, strong lock-in, and the duration of energy infrastructure) indicates that today's decisions have lasting impacts for many decades. For example, the age distribution of the U.S. fleet of coal power plants, as shown in Figure 2.5B-8 below, indicates that such capital-intensive infrastructure may be kept "alive" well past its economic life.[27]

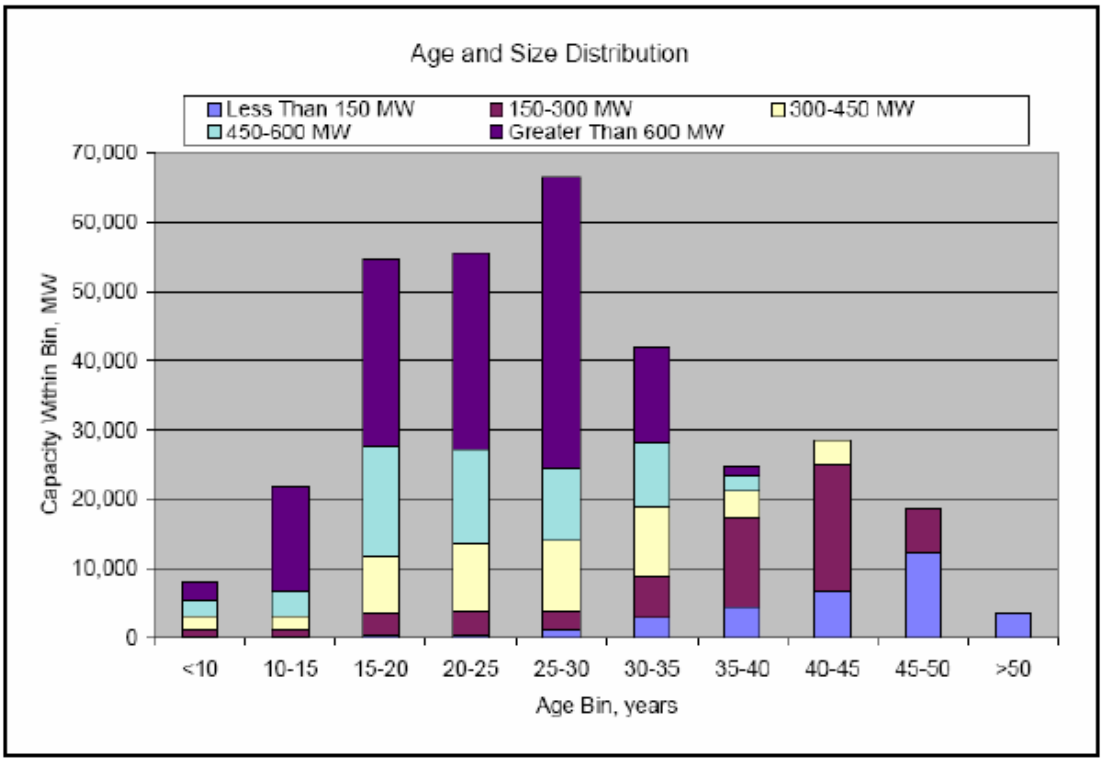


Figure 2.5.B-8. Age Distribution of U.S. Coal Power Generation Fleet. [28]

Given the recent IEA estimate that ~$1-–2 trillion/year for the next 20 years will be spent on energy infrastructure, without significant exogenous input, today's decisions will not only be with us for many decades, but they may have implications lasting well into the next century.

Within the context addressing the complexities of environment, economic prosperity and security, it is increasingly clear that a portfolio approach toward our energy future is required. No single technology can be deployed at scale with sufficient speed to address the urgency of the challenge. Moreover, it is increasingly clear that pursuing a low carbon energy economy is pivotal to reducing global greenhouse gas (GHG) emissions and eventually stabilizing world CO_2 concentrations. What is less clear, however, are the tradeoffs required to "decarbonize" the energy economy. Typically, these tradeoffs are expressed as increased costs for new or low-carbon energy technologies. However, more recent analyses have pointed out additional benefits that include reduced price fluctuation exposure, reduced negative health impacts, increased workforce/job development, and national and global security benefits. Other reports also evaluate the potential damages related to climate change and evaluate the economic cost/benefit tradeoffs. Globally, the Stern Report, conducted for the UK Government, and lead by Nicholas Stern, conducted the most comprehensive cost/benefit analysis. Using a consensus set of financial assumptions such as discount rate, Stern found that immediate action at scale and speed

significant enough to impact the global concentrations would have materially reduced economic impacts when viewed over the next few centuries.[29]

In the U.S., many reports have been completed on the possible impacts of various forms of clean energy legislation, including state-level renewable portfolio standards (RPSs), national RPSs, and other variants.[30,31,32] In all cases, increased use of renewable energy technologies induced positive economic and environmental outcomes. Consumer impacts varied from slightly negative to slightly positive on an annual expenditure basis.

The following are considered to be the essential characteristics of a feasible long-term approach to energy sustainability:

1. A future global sustainable energy infrastructure and economy must simultaneously have the following characteristics:
   - Economically affordable and stable
   - Secure, accessible, and reliable
   - Environmentally sustainable
2. There is a time urgency that demands a serious departure from business as usual.
3. The new approach will surely require a blend of technology innovation, market restructuring, new government policies, and education of our citizens.
4. This long-term solution will likely require investments in technological innovation, development of human talent needed for such innovation, and policy development orders of magnitude larger than have been made to date.

Climate Change and Energy Security Tradeoffs. Finally, in order to highlight the tradeoffs between energy security and climate change, the World Resources Institute produced the following assessment of energy options.

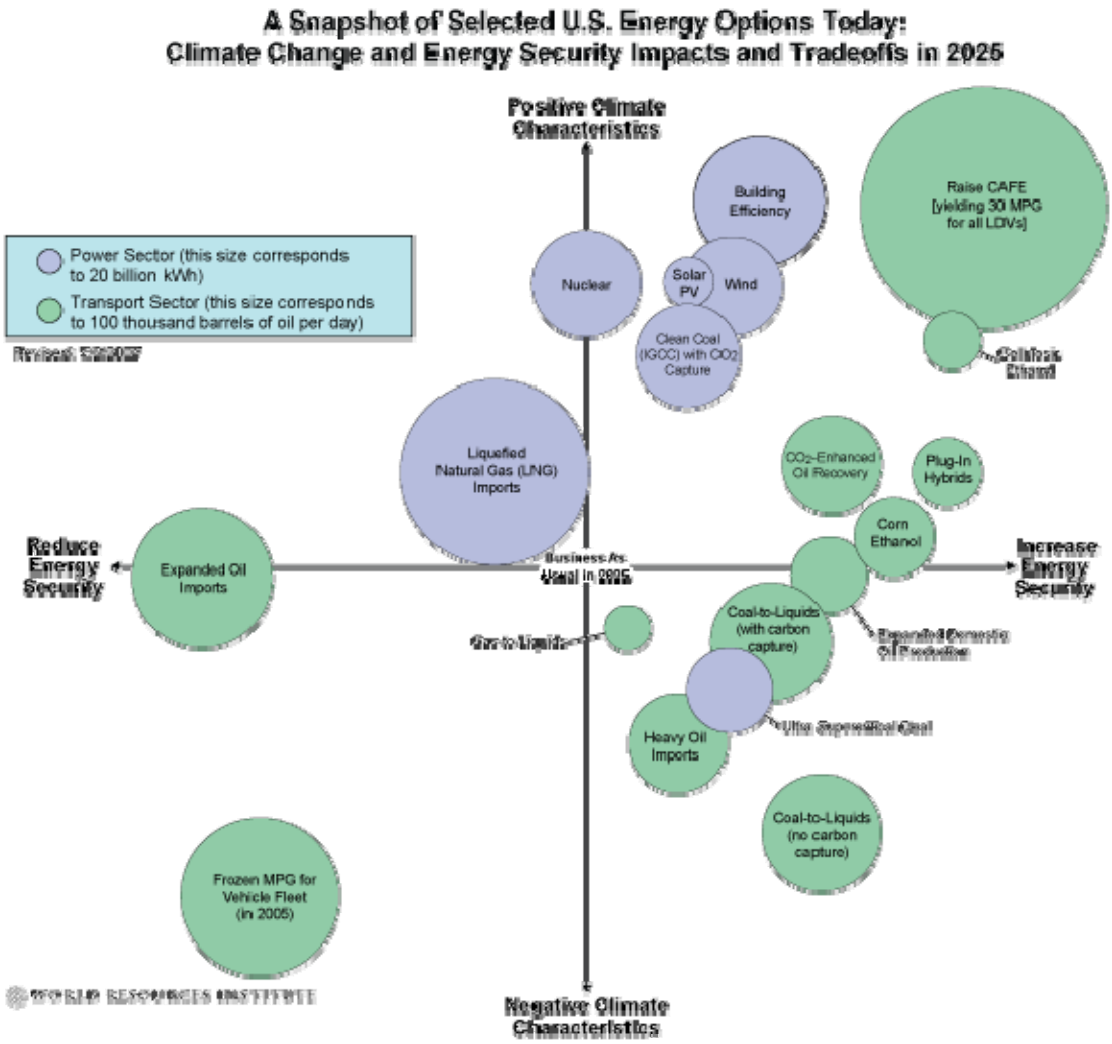


Figure 2.5.B-9. A Snapshot of Selected U.S. Energy Options Today: Climate and Energy Security Impacts and Tradeoffs in 2025.

Figure 2.5.B-9 compares the energy security and climate characteristics of different energy options. The bubble size corresponds to incremental energy provided or avoided in 2025. The reference point is the "business as usual" mix in 2025. The horizontal axis includes sustainability, as well as traditional aspects of sufficiency, reliability, and affordability. The vertical axis illustrates lifecycle greenhouse gas intensity. Bubble placements are based on quantitative analysis and WRI expert judgment. [33]

The Challenge

Many technology options exist that can contribute to an energy sustainable future, including energy efficient techniques, non-emitting power generation, biofuels, forest and land management, and carbon dioxide capture and storage. A challenge of this scale needs to be addressed at the global level, but here in the U.S., an appropriate strategy will also be required across Federal, state, and local governments, as well as across businesses, non-profit organizations, and other stakeholders.

Summary and Conclusions

This paper has attempted to synthesize the very broad and deep literature on energy issues related to security, prosperity and sustainability. The key issues that have been raised in the literature include:

1. Global energy demand is anticipated to grow 2×–-3× over the next century as population and global GDP continue to expand.
2. The world's energy endowment is sufficient to meet the anticipated demand through 2100, but it will become increasingly strained unless other energy pathways are developed. Additionally, continued extraction and use of fossil fuels has material implications for environmental and security goals.
3. U.S. energy use has developed in accordance with the availability and use of coal, petroleum, natural gas and nuclear energy. Recent advances in renewable energy make up approximately 2.5% of the U.S. energy supply.
4. U.S. energy policy has historically dealt relatively independently with petroleum resources and use vs. electricity generation, transmission, and distribution. Policy and regulations have been the primary drivers of changes in the electricity generation portfolio, automobile fuel economy, the built environment, and are increasingly complicated by ties to agricultural policy via biofuels and international policy via energy and national security and trade.
5. Capital-intensive energy infrastructure technologies change slowly, with lifetimes from 20–100 years, and are subject to strong lock-in.
6. Anthropogenic emissions of greenhouse gases, predominantly from our energy economy, are the major contributors to increased concentrations of GHGs.
7. Many technologies exist today that can contribute to energy for a sustainable future, including energy efficiency approaches, non-emitting power generation, biofuels, forest and land management, and CO_2 capture and storage.

Taken as a whole, the assessments summarized indicate that a transformational change is required for the global energy economy. The need for this change is urgent, and the scale of the change is enormous.

Appendix A: Details of Technology Change and Adoption

A. Technology Introduction and Gestation Time

Studies have approached the question of how long it takes a technology to be introduced into the market from a number of perspectives. First, research by Mansfield[34] on the impact of academic research on technology introduction found that average lag time between the most recent published academic research finding and its first commercial introduction over seven industries was 7 years. Mansfield's findings are consistent with those of Gellman, who studied innovations from 1953 to 1973, finding an average lag time of 7.2 years.[35] Moreover, Mansfield reported a range of lag time from 4.2 years (instrument sector) to 9.8 years (metals) and found that industry believed substantial delays of 9 years (or more) would have occurred in technology introduction without academic research in which such research played a significant role in technology development.

In addition to the work of Mansfield and Gellman, analysis of 30 product-development and market sales cycles over 150 years was performed by Agarwal and Bayus[36] in which they conclude that average time from invention to "sales takeoff" (accelerating market penetration) is

more than 42 years. That is, the length of time from invention to measurable market penetration—without accounting for any fundamental R&D that preceded the specific product invention—is 42 years across a broad range of consumer products. They based their estimate of the year of invention on several published sources[37] and analyses[38] which, in general, defined *invention* as the year of *the most recent patent*, not the first patent or publication of fundamental technology. Their analysis is summarized in Figure 2.5.B.A-1.

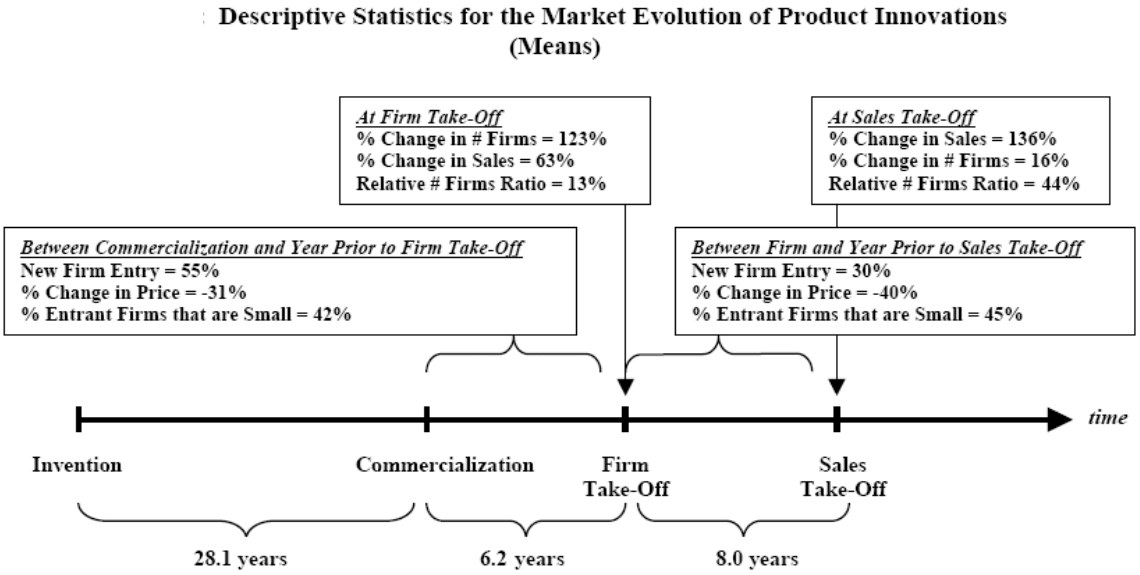


Figure 2.5.B.A-1. Descriptive Statistics for Market Evolution of Product Innovations Derived from Analysis of 20 Products by Agarwal and Bayus. [39]

The further measure of the gestation time for fundamental R&D is also significant relative to any measure of advancement of technology into the market by public R&D of fundamental technology. Several metrics exist to measure the effectiveness of R&D as a function of time-from-research, including citation trees and patent citation temporal studies, but they do not quantify impact on temporal change of product introduction. In the energy field, previous studies indicate that the research-market pipeline may be particularly long compared to other sectors.[40, 41]

The research of Agarwal and Bayus also suggests that the number of firms commercializing a technology impacts the time to sales takeoff. They postulate that sales takeoff is caused by outward shifting supply and demand curves, and firm entry not only affects supply, but also affects demand for the product. This is because product improvements, expanded distribution, and increased customer awareness result from increased marketing activities.

B. Energy Sector Infrastructure Adoption Rates

Long-lived, and highly capital intensive infrastructure displays markedly different technology adoption characteristics than consumer products. As shown in Figure 2.5.B.A-2, the shares of total primary energy in the United States supplied by different fuels (top panel) and estimates for the consumption of major primary fuels worldwide (bottom panel) change on a much slower rate than consumer goods. Here, major changes in energy fuel supply occur on the order of a century.

As with the evolution of particular technologies and infrastructures, the diffusion of energy sources follows a similar pattern. The market share of a fuel expands initially slowly as gradually expanding niche markets are filled; a more lengthy process of pervasive diffusion follows, eventually saturating and declining as a superior competitor enters and diffuses. Despite the simultaneous competition and interaction among many different sources of energy, each historical period is characterized by a clear dominance of a single energy source, which corresponds with the main technological clusters: fuelwood (and feed), followed by coal, and later by oil and natural gas. The rates for diffusion are also similar for all of the major fuels (wood, coal, oil, gas), which required a century to achieve their ultimate saturation. The diffusion rate for nuclear is not yet evident because of its relatively low market share. [42]

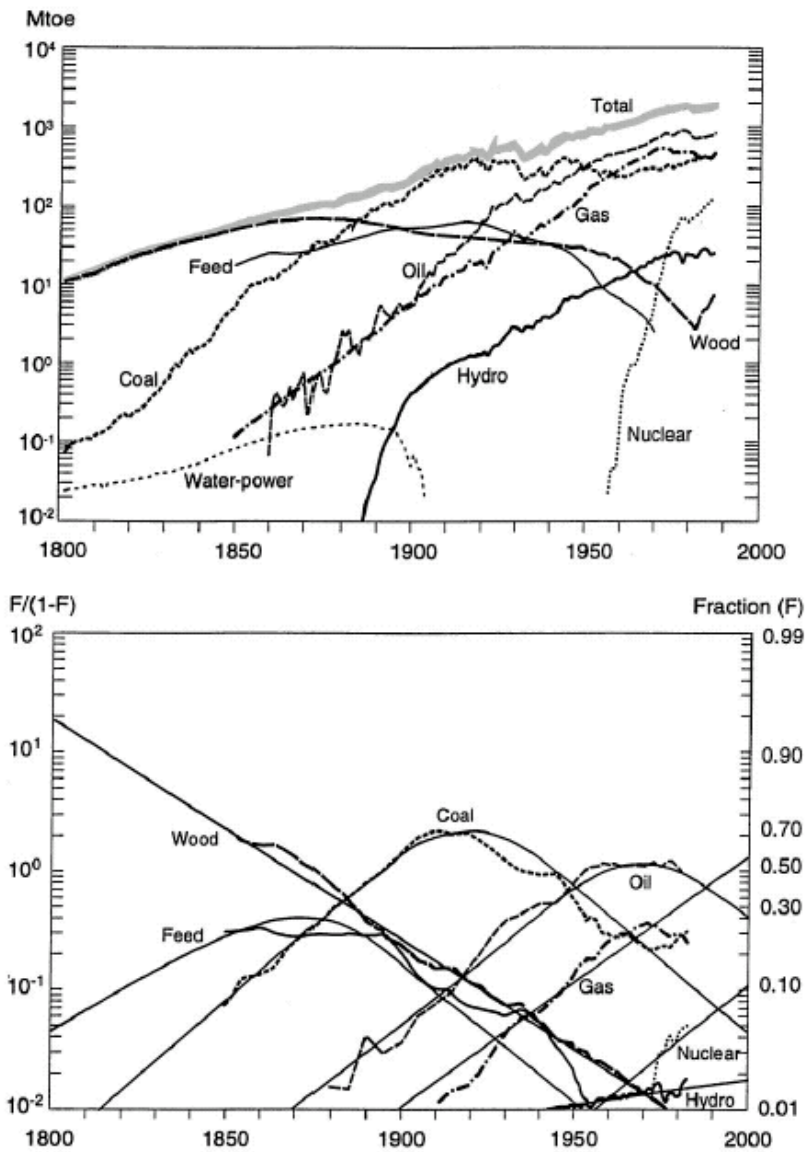


Figure 2.5.B.A-2. Selection of Energy Source Supply in the United States from 1800 to 2000. [43]

Using average values from the literature—excluding preceding academic R&D—the time from invention to market penetration of 90% is about 80 years, with a range as low as 20 years and as high as two centuries for long-lived infrastructure technologies.

C. Energy "Technology Lock-in" and Learning

Technology Lock-in. Particularly for capital-intensive, long-lived infrastructure, it has been argued that the development of technologies both influences and is influenced by the social, economic and cultural setting in which they develop.[44] Technology pathways have been shown to depend on the particular characteristics of initial markets, the institutional and regulatory factors governing its introduction and the expectations of consumers. This "technology lock-in" has been shown to depend on four major classes of technology and social influences: *scale economies, learning effects, adaptive expectations* and *network economies*, which contribute to this positive feedback in favor of existing technologies.[45] For capital intensive energy technologies, each factor is particularly relevant: (1) the economic scale is on the order of trillions of dollars per annum, (2) learning drives competitive costs (see below), (3) consumer expectations have been formed over multiple generations, and (4) economics of networks such as electricity transmission and distribution or oil/gas collection/transportation/processing/distribution constitute a significant investment and influence on possible alternatives.

Additionally, as modern technological systems are deeply embedded in institutional structures, institutional lock-in can interact with and reinforce the drivers of technological lock-in (e.g., slow technology change). Here, examples such as the regulatory and oversight institutions, as well as the corporate and political institutions, cover a broad range of interests in the energy sector. It has been suggested that modern technological systems, such as the carbon-based energy system, have been strongly influenced by both technological and institutional factors, resulting in exceptionally strong lock-in.

Technology Learning. Technology learning is classically described as the decrease in costs as a function of cumulative production. (Initially, learning was intended to capture both economies of scale and R&D advancement in a single descriptor. More recent advances in the theory have shown that multiple factors can influence overall learning rates and one can separate the various influences on overall technology cost & performance.[46]) Numerous researchers have investigated the nature of learning curves for energy technologies. [47] The International Energy Agency has also published data and trends, as shown in Figure 2.5.B.A-3. From the IEA data for electric generation technologies, photovoltaics have the fastest learning (drop in price with cumulative production), followed by wind power and then biomass power. Coal and natural gas combined cycle technologies have shown limited cost reductions after the initial TW of installations.

↑ Back to the top

---

[1] International Energy Agency, World Energy Outlook, 2008.

[2] Grubler, Arnulf, "Energy transitions," *Encyclopedia of Earth*, (2007), Cutler J. Cleveland (ed.). Available online at: http://www.eoearth.org/article/Energy_transitions.

[3] "Table 11.1 World Primary Energy by Source, 1970-2004," Energy Information Administration, International Energy Database. Available online at: http://www.eia.doe.gov/emeu/aer/txt/stb1101.xls.

[4] M. I. Hoffert et. al., "Energy Implications of Future Atmospheric Stabilization of CO2 Content," *Nature* 395 (1998): 881-884. Available online at: www.nature.com/nature/journal/v395/n6705/pdf/395881a0.pdf

[5] M. I. Hoffert et. al., "Energy Implications of Future Atmospheric Stabilization of CO2 Content," *Nature* 395 (1998): 881-884. Available online at: www.nature.com/nature/journal/v395/n6705/pdf/395881a0.pdf

[6] See for example scenarios at the Pew Center for Climate Change, Union of Concerned Scientists, Electric Power Group, and the Electric Power Research Institute for additional scenarios.

[7] Greenpeace International and European Renewable Energy Council, *Energy Revolution: A Sustainable Energy Outlook* (January 2007), p 25.

[8] *Ibid.*

[9] *Ibid.*

[10] Greenpeace International and European Renewable Energy Council, *Energy Revolution: A Sustainable Energy Outlook* (January 2007), p 44.

[11] M. I. Hoffert et. al., "Energy Implications of Future Atmospheric Stabilization of CO2 Content," *Nature* 395 (1998): 881-884. Available online at: www.nature.com/nature/journal/v395/n6705/pdf/395881a0.pdf

[12] *Ibid.*

[13] "Table 1.2 Energy Production by Primary Energy Source, 1949-2006," U.S. Department of Energy. Available online at: www.eia.doe.gov/emeu/aer/pdf/pages/sec1_7.pdf.

[14] Energy Information Administration (EIA). *International Energy Outlook 2007*. Available online at: http://www.eia.doe.gov/oiaf/ieo/index.html

[15] International Energy Agency (IEA). *World Energy Outlook 2007: China and India Insights*.

[16] Parry, Ian and Joel Dramstadter, "The Cost of US Oil Dependency," National Commission on Energy Policy: Washington, DC (November 17, 2004).

[17] Recent examples include China's relationship with oil rich countries like Sudan and Iran and Russia's leverage of its natural gas supply to Europe.

[18] For example, see a summary publication, H. Peyton Young, "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," (working paper, Santa Fe Institute, 2007). Available online at: http://www.santafe.edu/research/publications/workingpapers/07-09-038.pdf

[19] See, Paul A. David , "Zvi Griliches on Diffusion, Lags and Productivity Growth …Connecting the Dots," (working paper, Stanford University, 2003). Available online at http://www-econ.stanford.edu/faculty/workp/swp04002.pdf.

[20] Mansfield, E., "Academic research and industrial innovation," *Research Policy* 20 (1991): 1-12; and E. Mansfield, "Academic research and industrial innovation: An update of Empirical Findings," *Research Policy* 26 (1998): 773-776.

[21] Agarwal, Rajshree and Barry L. Bayus, *Management Science*, 48:8 (August 2002): 1024-1041.

[22] See for example, D.B. Elliott, et al, "Methodological Framework for Analysis of Buildings-Related Programs: The GPRA Metrics Effort," *PNNL -14697*, Pacific Northwest National Laboratory, Richland, WA (2004).

[23] Bennet, C., **"**Risk and Opportunity in the Gathering Climate Change Storm", The Conference Board (2007); and J. Busby, "Climate Change and National Security" Council on Foreign Relations (November 2007).

[24] See for example, the "The Hard Truths"; National Petroleum Council, 2007.

[25] The IEA 2007 forecast highlights the growth of India and China economically and in relation to the energy economy. See http://www.worldenergyoutlook.org/2007.asp.

[26] Verrastro F., (Center for Strategic and International Studies), Testimony before the Committee on Energy and Natural Resources, United States Senate, "Comments and Observations on the Topic of U.S. Energy Independence" (March 7, 2006).

[27] Coal Power database, U.S. Department of Energy, National Energy Technology Laboratory. Available online at: http://www.netl.doe.gov/energy-analyses/technology.html.

[28] Coal Power database, U.S. Department of Energy, National Energy Technology Laboratory. Available online at: http://www.netl.doe.gov/energy-analyses/technology.html.

[29] The Stern Review reports that effective global mitigation, focused on strong actions in the next 10-20 years will cost 5x-20x less in terms of global GDP impacts than latter actions of mitigation and adaptation. See, Nicholas Stern, *Stern Review: The economics of climate change*, HM Treasury, Cambridge University Press: Cambridge. Available online at: http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/stern_review_report.cfm.

[30] Cory, K.S. and B.G. Swezey, "Renewable Portfolio Standards in the States: Balancing Goals and Rules," *Electricity Journal* 40:4 (May 2007): 21-32.

[31] Union of Concerned Scientists, "Analysis of a 20% National Renewable Portfolio Standard," (2004). Updated October 2007 for a 15% National Standard.

[32] EIA has also analyzed proposed national RPS legislation. See http://www.eia.doe.gov/oiaf/servicerpt/rps2/methodology.html.

[33] Logan, Jeffrey and John Venezia, "Weighing U.S. Energy Options: The WRI Bubble Chart," World Resources Institute, *WRI Policy Note* (July 2007), pg. 3. Available online at: http://pdf.wri.org/weighing_energy_options.pdf.

[34] Mansfield, E., "Academic research and industrial innovation," *Research Policy* 20 (1991): 1-12; and E. Mansfield, "Academic research and industrial innovation: An update of Empirical Findings," *Research Policy* 26 (1998) 773-776.

[35] Gellman Associates "Indicators of International Trends in Technology Innovation" report to the National Science Foundation, 1976.

[36] Agarwal R. and B. Bayus, "The Market Evolution and Sales Take-Off of Product Innovations," (working paper, University of Illinois at Urbana, 2002).

[37] Giscard d'Estaing, V., *The Second World Almanac Book of Inventions*, Pharos Books: New York (1986).

[38] Kohli, R., D. Lehmann, and J. Pae, "Extend and Impact of Incubation Time in New Product Diffusion," *J. Product Innovation Mgmt,* 16:2 (1999): 123-144; and references therein.

[39] Agarwal R. and B. Bayus, "The Market Evolution and Sales Take-Off of Product Innovations," (working paper, University of Illinois at Urbana, 2002).

[40] Margolis, R. and D.M. Kammen, "Energy R&D and Innovation: Challenges and Opportunities" in Schneider, S.A Rosencranz, and J. Nices, (ed.), *A Reader in Climate Change Policy*, Island Press: Washington D.C. (2001).

[41] Margolis R. and D.M. Kammen, "Underinvestment: the energy technology and R&D policy challenge," *Science* 285 (1999): 690-692.

[42] Grubler, A., N. Nakicenovic, and D. Victor, "Dynamics of energy technologies and global change," *Energy Policy* 27 (1999) 247-280; and Grubler, A., "Time for a Change: On the Patterns of Diffusion and Innovation," in *Technological Trajectories and The Human Environment*, National Academies Press, (1997).

[43] Grubler, A., N. Nakicenovic, and D. Victor, "Dynamics of energy technologies and global change," *Energy Policy* 27 (1999) 247-280; and Grubler, A., "Time for a Change: On the Patterns of Diffusion and Innovation," in *Technological Trajectories and The Human Environment*, National Academies Press, (1997).

[44] See, Timothy J. Foxon, "Technological lock-in and the role of innovation," Imperial College, London, (2005). Avaialable online at: http://www.hm-treasury.gov.uk/media/C/0/climatechange_imp_3.pdf.

[45] Arthur, W. Brian, "Competing Technologies, Increasing Returns, and Lock-In By Historical Events," *The Economic Journal* 99 (1989): 116-131.

[46] Jamasb, Tooraj, "Technical Change Theory and Learning Curves: Patterns of Progress in Energy Technologies," University of Cambridge (March 2006). Available online at: http://www.dspace.cam.ac.uk/bitstream/1810/131682/1/eprg0608.pdf.

[47] See, for example, G.F. Nemet, "Policy and innovation in low-carbon energy technologies," in *Energy and Resources Group*, University of California: Berkeley, CA (2007).

2.5. Food

2.5.A. World Food Availability and the Natural Land Resources Base (Jeffrey Steiner, Timothy Griffin)

Authors and Organizations: Jeffrey Steiner, Organization: Agricultural Research Service, United States Department of Agriculture; Timothy Griffin, Tufts University
Contact Information: jeffrey.steiner@osec.usda.gov

As long as people have lived, utilized the Earth's natural resources, and created industry, social, economic and cultural systems have evolved and will continue to change as the world becomes more populated, urban, and interconnected and as people continue to derive their livelihood from the planet's resources. Therefore, as we discuss how future world population projections are related to food needs, we need also need to consider how changing population and changing wealth patterns will not only impact food availability and consumption patterns, but also our inter-related needs for energy and water. We also must recognize that the Earth's land area is finite, so increasing population and migration to cities will result in (1) more rural land area to be taken out of agricultural production around population centers and (2) the continued loss of vital ecosystem services that include the provision of clean water, air, wildlife habitat, aesthetic landscapes, and maintenance of biodiversity.

The dependable availability food and natural resources will most certainly be affected by the uncertainties of climate change. Decreased agricultural output resulting from shifts in precipitation and increased temperatures has the potential to negatively impact not only food availability in developing countries whose populations already live close to subsistence levels and depend on a greater portion of their national income from agriculture and whose citizens,[1] but also industrialized countries who are planning on replacing non-renewable energy and other consumer products with bio-based analogs produced from the land.

In addition to meeting nutritive needs, food is also a significant trade good for most nations, so the profitable production, processing, and marketing of food will depend on the ready availability and quality of natural resources at all scales, from farm to global. The significance of this is that agriculture is not only the basis of income to producers and rural communities, but from the farm through the rest of the food chain, 17% of the jobs in the U.S. workforce are provided by industries supported by agricultural production.

It is estimated that 83% of the present earth's surface already has marks of human influence,[2] and we can expect a continual shrinking of the land resources base.

This paper discusses (1) the likely trends in food consumption, food accessibility, and food prices, (2) the potential destabilizing effects of these trends on people in developed and developing economies, and (3) a discussion of how the present and future capacity of the land and the natural resources base can be used to meet the needs of people around the world for abundant, affordable, and safe foods.

Projected Consumption Trends

The predicted trends for food consumption are based on the compounded effects of continued increasing population and changes in dietary patterns due to economic growth, particularly in the

developing world.[3] The United Nations projections indicate that the population of the developed world will likely remain about constant at 1.25 billion through 2050.[4] This is in stark contrast to the predicted developing and transitioning world population, which is expected to grow from 5.6 to 7.9 billion people. Much of this growth will take place in already populous and food insecure areas, especially in Asia and Africa, which together could contain 75–80% of world population. Food insecurity is defined as the gap in distribution of food so that the average consumption by people of a population is less than 2,100 calories per person per day. It is estimated that in 2009 there are 833-million food-insecure people worldwide.[5]

There will also be a significant shift in demographics during the 2010-2050 period due to projected declines in birth rate in the developing world, with the most rapidly growing age cohort being those over 60 years. There will be a continued movement of people from rural to urban areas, with emigration from the developing to the developed world that contribute to a significant global population flux.

Income growth and the resulting changes in food consumption patterns are significant elements of shifts in global food demand and agricultural trade. Real average per capita income grew by nearly 100 percent for most countries from 1960 to 2000. For low-income countries, per capita income in 1998 was just over US$500, compared to US$28,000 for high-income nations. Income growth from 1961 to 1998 for developing countries was 221%, and it surpassed 173% growth for income in developed countries. Such great gains for developing countries have resulted in significant changes in the kinds of foods consumed, particularly for higher income developing countries.[6] Substantial economic growth in developing countries has translated into a more rapid increase in the rate of global energy use than relative population growth.[7] Before consumers gain the ability to purchase cars and transportation fuels, they achieve the ability to afford more food and also different kinds of food. Therefore, increased world demand for cereals is projected to be 40% by 2020, and meat demand is projected to double by 2050.As populations become increasingly urbanized and as per capita income grows, particularly in countries like China and India, dietary patterns are expected to change markedly in the coming decades, with concomitant changes not only in the types of food consumed, but also in the supply chain infrastructure needed to meet demand. The implications for these changes in Asia are that diets will continue to become more similar to Western diets, meaning the use of more processed food products that have greater energy density.[8] Of particular note is the projected increase in the consumption of meat and dairy products, and fish,[9] as this will represent both a diversion of grains from human consumption to their use as animal feeds, and the associated changes in land use patterns in many regions to accommodate that change in kind of agricultural products produced. The globalization of the food production and distribution system will result in increasing similarity, or convergence, of developed and developing countries' food system characteristics.[10] As a result, the food marketing system will need to adjust to greater concentrations of people in urban areas, along transportation routes, and to coastal regions,[11] and there will be need to develop strategies that adjust to the variability in the size and growth rate of different populations.[12]

In the United States, 50% of the population lives in coastal areas, and for every person added to the population, 1.7 acres of rural land are developed. As pointed out earlier, with continued population growth, demands from land and natural resources are expected to increase. Such demand will grow increasingly greater as there is a shift away from the use of non-renewable resources for the production of fuels and other consumer products, as well as other bio-based consumer products that originate from agricultural production. Just as significant, our rural lands

also are the source of clean air, water, and other ecosystem services upon which all of our society depends. Because of rising production costs, some sectors of the agricultural economy have experienced significant demographic shifts. Production of many crops that require human labor for cultivation and harvest has moved to developing countries where less expensive labor is abundant and where transportation fuel costs are low so that distant markets can be accessed affordably.

Hunger and Food Access

The first Millennium Development Goal is to "eradicate extreme hunger and poverty" by 2015.[13] However, food security is projected to deteriorate over the next decade.[14] There has been significant progress in a number of regions, but advances have been slower than anticipated, and the current economic crisis has in some cases reversed previous gains. As discussed by FAO, gains have been made in Asia, the Pacific, and Latin America, while hunger has *increased* in Sub-Saharan Africa.[15] In many areas, particularly in rural ones, hunger and poverty coincide. In these situations, agricultural productivity is a key driver for poverty alleviation, and rural and agricultural developments are critical for long-term stability. However, solutions to enhance agricultural productivity must be combined concurrently with solutions to address broader public health, education, and infrastructure development challenges.[16]

Rapid increase in food prices from 2006-2008 were larger in magnitude, but not unlike major increases in 1971 to 1974 and 1994 to 1996. These price surges were the results of a combination of factors that include the depreciation of the U.S. dollar, strong worldwide demand for agricultural products (including most recently for biofuels), supply shocks, and policy responses by major trading partners.[17] Such "price shocks" not only reduced the ability of food-insecure people to buy food, but they also made international provision of food aid more expensive.[18] The recent rapid escalation in food prices led to protests, some of which were not limited to low-income countries. A decrease in export earnings and capital receipts could increase the number of food-insecure people worldwide by 100 million.[19]

Biophysical Limitations to Increased Food Availability

There is general agreement that by 2050, the world will have nine billion people. The Earth's arable land area is finite, so as we move away from using non-renewable resources, more will be expected from our working lands not only to produce food, feed, and fiber, but also to provide the raw materials needed for the production of a wide array of biobased consumer products, including fuels and energy.

Eleven percent of the globe's land surface (1.5 billion hectares [ha]) is used in crop production. This area represents 36 percent of the land estimated to be to some degree suitable for crop production. The fact that there remains some 2.7 billion ha with crop production potential suggests that there is still land available for further expansion of agriculture.[20] In the developed world, there are few constraints to agricultural production on existing land areas, but large expansion of currently cultivated areas is unlikely due to environmental concerns and limitations to extending use of water resources for irrigation to supplement natural net primary productivity. In developing countries, the United Nations estimates that enough food could be produced if 79 percent of the need were met through increased yields by use of intensified cropping practices, with the remaining 21 percent of need met through expansion of the land area under cultivation. This would represent a net increase from 848 million ha in 2000 to 1076 million ha in 2030, not accounting for land losses due to soil degradation, urbanization, and land conversion to industrial

use.[21]

In developed countries, overall agricultural production has increased, but for commodity crops like maize and soybean, intensified production over the last 60 years has only resulted in essentially linear crop yield increases. For example, U.S. maize yields increased 3.3% per year, and soybean yields increased 1.8% per year[22]. The trends for agricultural productivity in developing countries have not been as encouraging as developed countries. Global cereal production per capita has slowed since 1984 due to long-term productivity declines in sub-Saharan Africa, as well as more recent declines in North America, Europe, Oceania, and the former Soviet Union, even with the steady increase in yields that have been achieved in Asia.[23] In parts of the developing world, the greatest increases in crop productivity took place 25 to 40 years ago as the Green Revolution was implemented.[24] Since that time, yield increases have stagnated due to (1) reduced use of inputs and other technologies that could boost yields, (2) low cereal prices, (3) market and infrastructure constraints, and (4) low levels of investment in agricultural research and technology.[25]

The capacity of land to produce food, fiber, and fuels varies by place due to natural biophysical limitations to the amounts of natural carbon that can be converted into plant material by photosynthesis.[26] Production can be intensified through practices that boost productivity, but at the cost of import of resources from outside the natural production system, and also at the expense of additional cost and energy. Production potential can also be decreased by poor soil resource management such as cultivation that leads to erosion, improper irrigation that leads to salt accumulation, or removal of harvested material without replacement of nutrients. Such lost potential is typically localized. For example, average soil degradation does not threaten food security at the global level,[27] but it is a serious limiter of productivity in regions such as sub-Saharan Africa.

Global food production has grown faster than population in recent decades.[28] Per capita global availability of meat, fruits, and vegetables increased by more than 60% between 1961 and 1998, while the supply of roots and tubers decreased by 21%. World cereal supply also increased by almost 17% during the same period.[29] For example, the FAO aggregate crop production index grew at an annual rate of 2.3% from 1961 to 1999, and crop production per capita has increased 0.6% per year in all regions but Africa.

All facets of food production will be impacted by climate change. Although climate change represents a significant source of uncertainty in this regard, there is general consensus that the negative impacts will be borne disproportionally in those regions where food insecurity and hunger are most problematic. This can be attributed not only to the geographic distribution of direct climate change impacts brought about by changes in temperature and precipitation, but also on the lack of resources that farmers in these areas can use to adapt to climate change.[30] The increased variability in climate and the increased incidence and severity of extreme weather events impacts not only production of food, but also food availability, food safety, and overall food security.[31]

Given that there is a finite land base to work with to increase total world food output to meet global foods demands by the middle of this century, total dependence on bringing large tracks of the natural land base into agricultural production, as well as urban development, is not an entirely viable option. Adequate provision of ecosystem services, including sourcing of clean water and air, the sequestration of soil carbon, and the maintenance of natural habitats to

preserve critical wildlife habitats and biodiversity are critical and must be managed along with agricultural production. Therefore, the preferred strategy is to intensify the amounts of agricultural products produced per unit of area.

Viable strategies for intensifying production efficiency include the use of improved management practices such as the incorporation of new technologies, the use of soil amendments to increase soil water and nutrient availability for crop growth, and utilization of improved or adapted varieties. There is significant interest in using agro-ecological methods to meet these needs[32] and to reduce the need for purchased inputs. However, any technologies that help producers and nations adjust to increased risks due to extreme weather associated with the effects of climate will be needed. A systems approach will be required because the reliance of any one technology to overcome limitations to productivity will likely not produce resilient results and analyses of past successes do not demonstrate a single responsible factor.

Analyses of the sources of corn yield gains in one American state over the past 50 years showed that about half of the increase was due to the use of improved genetic materials, particularly hybrid varieties, while the other half were due to the use of intensified cultural practices, including the use of fertilizers, pesticides, and irrigation.[33] With more recent advances in the introduction of biotechnology varieties (specifically corn, soybean, and cotton), the greatest gains have been in terms of increased production efficiency as a result of reduced energy use, where conservation tillage methods have been employed, along with lowered annual soil erosion rates. To date, some genetically improved crop varieties that utilize biotechnology have been developed in conjunction with the need for the use of another purchased input such as herbicides. Also, new challenges have emerged due to over-reliance on single technologies such as crop herbicide resistance, with the development of herbicide tolerant weed populations that require additional management inputs or greater management skill to control.

The utilization of biotechnology can positively impact both productivity[34] and economic return,[35] but their adoption continues to face significant cultural and trade policy obstacles in many countries. Results demonstrating the potential for biotechnology to moderate plant response to abiotic stresses like drought[36] and salinity[37] are promising in concept, but they may be more easily achieved through conventional breeding methods. As indicated above, crop improvement and the response to a variety of inputs have contributed to past yield increases. Farmers who do not have access to these inputs are unlikely to experience potential significant increases in productivity by the use of advanced genetic materials unless soil productivity is improved and agricultural inputs are available. Biotechnologies implemented in agriculture have primarily been to incorporate herbicide and insect resistance into corn, soybean, and cotton. Many believe that the potential yields of crops have not been increased, but rather a greater portion of the existing potential yield has been exploited.[38] Therefore, the use of a multiple input management approach has been cited as a strategy that could achieve significant improvement in food production.[39]

Presently, the greatest net primary productivity potential for agriculture is in the tropics, where agricultural development has been the least effective. The adoption of improved crop varieties in sub-Saharan Africa has resulted in far less yield increases than other developing regions.[40] However, when agricultural development across this region expands, there will be significantly greater demand for inputs such as improved varieties and soil amendments such as fertilizers. Greater such demand for fertilizers could present a significant challenge to agricultural production in developed countries that presently utilize great amounts of imported fertilizers. For example, more than 52% of nitrogen and 85% of potassium used in the U.S. was imported in

2007. U.S. domestic production capacity of fertilizers is limited, so any increases in demand will need to be met by greater imports.[41] From 2000-2006, U.S. ammonia production declined 44%, while imports increased 115%. U.S.-produced ammonia dropped in the domestic aggregate supply from 80 to 55%, with imports increasing from 15 to 42%. Natural gas is used as the energy input to turn atmospheric dinitrogen gas (N_2) gas into ammonia, so as a result of increased natural gas costs between 2000 and 2006, the price farmers paid for ammonia increased from $227 to $521 per ton, a 130% increase.[42]

Strategic Considerations

- Agricultural productivity must be a central component of strategies to relieve food insecurity. Continual progress is needed in identifying regions and populations that experience chronic food insecurity, and implementation strategies and practices that ameliorate insecurity that impacts regional populations as described in the Millenium Development Goals and other initiatives,
- Foresight is needed to prepare for and make adjustments to the shifting demographic structure that is expected in many regions, and that are complicated by concomitant shifts in food consumption patterns and demands brought on by population growth and migration.
- Degradation of natural resources, most significantly soils (loss as impacted by water and wind erosion and loss of soil carbon that reduces productivity) and water (quality, quantity, and access) limits the capacity for agricultural productivity improvement responses to natural biophysical conditions that are otherwise favorable and to improved management and use of inputs. In resource-poor and food insecure regions, this can contribute to cultural and political instability.
- Investment in research focused on practices and systems that increase output while mitigating negative environmental impacts is needed in both industrialized and developed countries. This should include an assessment of the appropriateness of emerging technologies, along with application of agro ecological principles for all systems. A whole-systems approach is needed that not only focuses on the short-term productivity of systems, but how all aspects of commerce are affected throughout the entire supply chain, now and in the future.
- The potential impacts of global climate change on regional food systems threatens to be a tremendous destabilizing force globally, particularly because of the disproportionate impact on populations that are ill-equipped to adapt to these changes.

↑ Back to the top

---

[1] National Intelligence Council (2008). Global Trends 2025: "A transformed world." Superintendent of Documents, US Government Printing Office. November.

[2] Sanderson, E.W., M.J. Jaiteh, M.A. Levy, K.H. Redford, A.V. Wannebo, and G. Woolmer (2002). "The human footprint and the last of the wild." BioScience 52:891-904.

[3] von Braun, J. (2007). "The world food situation: new driving forces and required actions." *International Food Policy Research Institute.* DOI: 10.2499/0896295303

[4] "World Population Prospects: The 2008 Revision, Highlights." United Nations, Department of Economic and Social Affairs, Population Division (2009). Working Paper No. ESA/P/WP.210.

[5] Shapouri, S., S. Rosen, B. Meade, F. Gale (2009). "Food security assessment, 2008-09." GFA-20. Economic Research Service, USDA.

[6] Regmi, A., and M. Gehlhar. 2001. "Forces shaping global food demand and agricultural trade." *Agricultural Outlook*, May. Economic Research Service, USDA.

[7] "World energy and economic outlook" (2007) Energy Information Administration. DOE/EIA-0480. Office of Integrated Analysis and Forecasting, U.S. Department of Energy, Washington, DC.

[8] Pingali, P. (2006). "Westernization of Asian diets and the transformation of food systems: Implications for research and policy." *Food Policy* 32:281-298.

[9] "The state of food and agriculture in Asia and the Pacific Region" (2008). RAP Publication 2008/03. United Nations, Food and Agriculture Organization (FAO).

[10] Regmi, A., H. Takeshima, and L. Unnevahr. (2008). "Convergence in global food demand and delivery," USDA-ERS , *Economic Research Report* #56.

[11] National Intelligence Council (2008). "Global Trends 2025: A transformed world." Superintendent of Documents, U.S. Government Printing Office. November.

[12] Coyle, W., B. Gilmour, and W.J. Armbruster. 2004. "Where will demographics take the Asia-Pacific food system?" *Amber Waves*, June. Economic Research Service, USDA.

[13] For a summary of the MDGs and associated targets related to hunger, see http://www.un.org/millenniumgoals/poverty.shtml verified: 12 August 2009.

[14] Shapouri, S., S. Rosen, B. Meade, F. Gale (2009). "Food security assessment, 2008-09." GFA-20. Economic Research Service, USDA.

[15] United Nations, Food and Agriculture Organization. (2007). The state of food insecurity in the world 2006: Eradicating world hunger - taking stock ten years after the World Food Summit. Sub-Saharan Africa contains 13% of the global population, but 25% of the undernourished and 33% of the chronically hungry. The situation is particularly recalcitrant in areas were per capita food consumption is less than 2200 kCal/day.

[16] Sanchez, P., C. Palm, J. Sachs, G. Denning, R. Flor, R. Harawa, B. Jama, T. Kifflemariam, B. Konecky, R. Kozar, E. Leleral, A. Malik, V. Modl, P. Mutuo, Am Nlang, H. Okoth, F. Place, S.E. Sachs, A. Said, D. Siriri, A. Teklehalmanot, K. Wand, J. Wangila, and C. Zamba (2007). "The African Millennium villages." *Proceeding National Academy of Sciences* 104:16775-16780.

[17] Peters, M., S. Langley, P. Westcott (2009). "Agricultural commodity price spikes in the 1970s and 1990s, Valuable lessons for today." *Amber Waves* 7: 16-23. Economic Research Service, USDA.

[18] Shapouri, S., and S. Rosen. (2007). "Energy price implications for food security in developing countries." *USDA-Economic Research Service Report* GFA-18.

[19] Shapouri, S., S. Rosen, B. Meade, F. Gale (2009). "Food security assessment, 2008-09." GFA-20. Economic Research Service, USDA.

[20] FAO (2003). "World agriculture; Towards 2015/2030, an FAO perspective."

[21] Agriculture and environmental challenges of the Twenty-First Century: A strategic approach for FAO. (2009). Committee on Agriculture COAG/2009/3. Rome, 22-25 April.

[22] Egli, D.B. (2008). "Comparison of corn and soybean yields in the United States: historical trends and future prospects." *Agronomy Journal.* 100:S79-88.

[23] Wiebe, K (2006). "Global resources and productivity." In K. Wiebe and N. Gollehon (ed.), *Agricultural resources environmental indicators, 2006 edition.* Economic Information Bulletin 16. Economic Research Service, USDA.

[24] See for example: Conway, G., and G. Toenniessen. (1999). "Feeding the world in the 21st century." *Nature*. 402:C55-C58.

[25] "World agriculture: towards 2030/2050 interim report – prospects for food, nutrition, agriculture, and major commodity groups." United Nations Food and Agriculture Organization (FAO). (2006). (See Figure 3.4 for example.)

[26] Grieser, J., R. Gommes, S. Cofield, and M. Bernardi (2006). "World maps of climatological net primary production of biomass, NPP." Food and Agriculture Organization of the United Nations, Climpag. http://www.fao.org/nr/climpag/globgrids/NPP_en.asp verified: 31 August 2009.

[27] Wiebe, K (2006). "Global resources and productivity." In K. Wiebe and N. Gollehon (ed.), *Agricultural resources environmental indicators*, 2006 edition. Economic Information Bulletin 16. Economic Research Service, USDA.

[28] Wiebe, K. (2006). "Global resources and productivity." In K. Wiebe and N. Gollehon (eds.), *Agricultural Resources and Environmental Indicators, 2006 Edition.* Economic Information Bulletin 16, Economic Research Service, USDA.

[29] Regmi, A., and M. Gehlhar. 2001. "Forces shaping global food demand & agricultural trade." *Agricultural Outlook,* May. Economic Research Service, USDA.

[30] Parry, M., C. Rosenzweig, and M. Livermore. (2005). "Climate change, global food supply and risk of hunger." *Philosophical Transactions of the Royal Society*, B. 360:2125-2138. Projections for increased yield potential in developed countries, but yield reductions of 10% are possible in developing countries.

[31] Schmidhuber, J., and F.N. Tubiello. (2007). "Global food security under climate change." *Proceedings of the National Academies of Science.* 104:19103-19108.

[32] The United Nations Food and Agriculture Organization (FAO) convened a conference in 2007 that addressed organic agriculture and its ability to increase food availability and security. For example, see: Scialabba, N. (2007). "Organic agriculture and food security." UN-FAO OFS/2007/05. See also: http://www.fao.org/organicag.

[33] An analysis of Minnesota, USA corn yield increases between 1930 and 1979 estimated that genetic improvements contributed to 58% of the half-century yield increase, while the combined effects of improved management from commercial nitrogen fertilizer and herbicide use was 70%. From: Cardwell, V.B. (1982). "Fifty years of Minnesota corn production: sources of yield increase," *Agronomy Journal* 74: 984-990.

[34] Qaim, M., and D. Zilberman. (2003). "Yield effects of genetically modified crops in developing countries." *Science*. 299:900-902.

[35] Raney, T. (2006). "Economic impact of transgenic crops in developing countries." *Current Opinion in Biotechnology.* 17:174–178.

[36] Umezawa, T., M. Fujita , Y. Fujita, K. Yamaguchi-Shinozaki and K. Shinozaki. (2006). "Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future." *Current Opinion in Biotechnology.* 17:113–122.

[37] Yamaguchi, T., and E. Blumwald. (2005). "Developing salt-tolerant crop plants: challenges and opportunities." *Trends in Plant Science.* 10: 615-620.

[38] Tannura, M., S. Irwin, and D. Good (2008). "Are corn trend yields increasing at a faster rate?" *Marketing and Outlook Brief 08-02*, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Feb. 20.

[39] Shapouri, S., S. Rosen, B. Meade, F. Gale (2009). "Food security assessment," 2008-09. GFA-20. Economic Research Service, USDA.

[40] Special Panel on Impact Assessment, Contribution Made by the CGIAR and Its Partners to Agricultural Development in sub-Saharan Africa (Consultative Group on International Agricultural Research, Washington, DC, 2001.

[41] Huang, W. 2009. "U.S. fertilizer imports/exports: Summary of the data findings." Economic Research Service, USDA. http://www.ers.usda.gov/Data/FertilizerTrade/Summary.htm verified: 12 August 2009.

[42] Huang, W. 2007. "Impact of rising natural gas prices on U.S. ammonia supply." USDA, WRS_0702, August.

2.5.B. Food Production in Arid Regions as Related to Salinity (Donald Suarez)

Author and Organizatin: Donald Suarez, Salinity Laboratory, Agricultural Research Service, US Department of Agriculture
Contact Information: Donald.Suarez@ars.usda.gov

Abstract

Arid and semiarid regions of the world are generally associated with high population density and lower–than-average per capita incomes and living standards. These regions are vulnerable to food shortages due to current, unsustainable use of fresh water for irrigation and increasing soil salinization. In many countries, current crop production cannot be sustained without adoption of new management practices and alternative water supplies. This paper discusses the issue of water supply, the impact of salinity, the potential for water reuse (both of irrigation drainage water and municipal waste water), the utilization of saline waters for crop production and the potential problems associated with the use of these alternative waters for irrigation.

Introduction

There has been a dramatic increase in total global food production, as well as an increase in food production on a per-acre basis, over the last 50 years. This increase is generally attributed to improved crop varieties and management practices (green revolution), however at least part of this increase is related to an increase in the amount of irrigated acreage. Irrigated lands have much higher productivity and economic return per acre as compared to non-irrigated lands. It is estimated that globally, irrigated lands represent 15% of the cultivated land, yet they produce over 30%[1] to 40%[2] of the world's food. The 35% increase in irrigated land from 1970 to the late 1980s thus provided a significant part of the increase in world food production, and was a major factor in avoiding large scale famine.

Since the 1980s, there has been a significant decline in the rate of growth in the world's irrigated land, with a leveling off of the total irrigated acreage in the first decade of this century. The reason for the leveling is due partially to the limitations of suitable land, but more importantly it is due to the lack of new developable water supplies in most of the arid and semiarid regions that can most benefit from irrigation. Globally, irrigated agriculture uses approximately ⅔ of the total water used, with industrial and municipal use making up the balance. Supplemental irrigation in more humid regions has increased, masking the decline in irrigation in arid regions. California, as an example, has experienced significant declines in irrigated acreage, some of which are considered temporary as a result of drought, while others are more permanent due to declining ground water supplies as well as increasing urban and environmental water demands.

Unfortunately, most arid regions do not have new developable surface waters and the large fresh water extractions of ground water required for irrigated agriculture cannot be increased. Of greater concern is the consideration that irrigation is not sustainable at current fresh water utilization rates. This over utilization of fresh water is particularly severe in drier regions of the world, where population density, poverty, and food demands are greatest. Over drafting of groundwater has resulted in declining water tables, loss of shallow fresh water for municipal use, and sea water intrusion in coastal regions. In early 1990s, approximately one fifth of the U.S.'s

irrigated lands were extracting groundwater water in excess of the natural recharge.[3] The data are not completely known, but the situation appears more severe in many less developed nations.

Increasing population results in increasing total demand for fresh water for municipal and industrial use as well as for increased food production. Increased fresh water needs are also related to increased per capita water usage associated with improved economic conditions in a region. Increases in living standards are not only related to increased domestic per capita water consumption, but they also result in increased in water consumption related to food production on a per capita basis. This increased demand with improved living standards is related to the increased water requirement for meat production versus grain production (expressed as gallons of water per kcal). It will be a major challenge just to maintain the existing level of irrigation and associated food production in the arid and semiarid regions of the world. An increase in living standards will require yet more water.

Salinity

In addition to the unsustainable extraction of fresh water, there is a related decline in water quality of existing supplies. A substantial part of the increasing salinity of water supplies in arid regions is unavoidable when irrigating, unless the drainage water from irrigation is isolated from fresh water supplies. Salinity increases in drainage water relative to irrigation water are inevitable due to the fact that plants extract water and leave most of the salts behind, concentrating these salts in the remaining water. Typically, plants extract only 5–10% of the salt associated with the water that they extract. Hence, more efficient irrigation (generally resulting in less water applied more uniformly), while desirable, results in smaller volumes of drainage water, but of greater salinity. The salinity increase is approximately inversely proportional to the change in volume (inverse to volume of irrigation water/volume of drainage water).

It is estimated that over 20% of irrigated land in the U.S. is salt-affected.[4] Globally, it is estimated that ⅓ of all irrigated land (76 million hectares [Mha]) is salt affected as a result of human activity,[5] meaning that crop yields are reduced. Salinity is a major threat to current irrigation projects and to the remaining near-surface fresh water supplies in arid regions. The extent of the salinity problem has not stabilized; instead, it is estimated that as much as 2 Mha of irrigated land, representing 1% of the total, is lost from production due to salinity each year.[6]

Most of the salt affected lands are in Asia and Africa. In contrast to salinization of water supplies, soil salinization is more readily controlled. Most soil salinization has occurred as a result of over irrigation, resulting in drainage volumes to the subsurface in excess of natural drainage capabilities and subsequent water logging, evaporation of water, and deposition of salts at or near the soil surface. Ensuring that soils are not over irrigated and maximizing food production per unit of water applied requires changes in irrigation systems. For example, conversion of surface flooding or furrow to drip irrigation allows for more uniform application of water, reduced need for irrigation water and reduced drainage volumes. These system changes and associated changes in management practices require capital investments and education programs for irrigators.

Increased salinization in arid and semiarid regions is also caused by leaching of existing salts from the soil during irrigation in regions with high salt containing strata, as well as by application of waters of low quality without proper management. In the instance of soils high in native salts, regional salinization of ground and surface waters is aggravated by excessive water applications. This impact is particularly important when implementing a new irrigation project,

but the impacts of leaching salts present before irrigation may be observed for in excess of 120 years after initiation of the irrigation project (Grand Valley, Colorado), depending on the hydrology of the system and the type of salts present. In this instance, improved water management is essential for salinity control and for maintenance of existing fresh water supplies. Improvements in irrigation system infrastructure and management in Grand valley Colorado, including concrete lining of canals and laterals, installation of closed pipe delivery systems, and irrigation scheduling has reduced the salt load to the Colorado River by approximately 500,000 tons per year.

Since saline water is not generally usable by agriculture with present technology, salinization of water resources represents a loss of useable water and can be a source of conflict among nations. Development of new irrigation projects upstream in a river basin inevitably results in adverse consequences to downstream users, either with reduced waters flows, increased salinity, or a combination of both.

Water Reuse

Municipal water demands throughout the world are high and increasing over time. These increases are greatest in the arid and semiarid regions, where the population growth is also greatest. However, municipal use cannot be considered a completely competitive use for fresh water. Only a fraction (typically 25%) of the water diverted for municipal use is actually consumed (by evaporation or plant transpiration). The majority of the municipal water is either degraded or lost due to leakage in the delivery system. Water currently lost via leakage is generally not available, as it may mix with saline or contaminated ground waters and thus be unusable, depending on local hydrologic conditions. Improvements in the municipal water delivery system can reduce the current all too common leakage of up to 50% of extracted water. These improvements, coupled with collection of wastewaters in a sewage system (with minimal losses in the system) can result in recovery of up to 70% of the allocated municipal water. It can thus be considered that municipal users primarily degrade rather than consume water. Increasing municipal demands for fresh water will result in corresponding increases in municipal wastewater.

In order to sustain agricultural productivity and irrigated acreage, especially in the arid and semiarid regions of the world, there will need to be a dramatic shift from fresh water to either treated wastewaters or alternative saline waters. Use of desalination for irrigation water does not yet appear economically viable, despite continuing advances in the technology and current increases in its use for drinking water. At present, costs are typically in the $1,000-per-acre-foot range when subsidies are fully considered,[7] and a large portion of the cost is associated with energy consumption. These high capital and production costs may be viable for a few very high value specialty crops in developed nations if there is no other water source, but it is not an option for production of basic food needs in less developed nations. Nonetheless, desalination is increasingly being used for municipal drinking supplies as it becomes cost effective with alternative supplies in arid coastal regions (such as in southern California and in Gulf states in the Middle East).

Marginal waters for irrigation contain (1) elevated concentrations of salts, typically above 2 dS/m in electrical conductivity, (2) elevated values of sodium adsorption ratio (SAR), which is defined as $Na/(Ca + Mg)^{0.5}$, where concentrations are expressed in terms of mmol/L (3) sometimes elevated pH, and (4) boron (B) concentrations that are potentially toxic to plants.

The adverse effects of SAR and elevated pH on hydraulic conductivity (soil property related to infiltration) when combined with low salinity are well documented.[8,9] Increased sodium levels relative to calcium (Ca) and magnesium (Mg) are associated with a loss of soil structure, clay migration, and reduction in water infiltration rates. These losses can be sufficient to cause severe and at times even total loss of soil productivity. Once the lands are degraded in this manner restoration of productivity is difficult and expensive, hence it is important to avoid degradation.

Current guidelines for irrigation[10] based on laboratory studies and field observation indicate that waters of SAR 10 or below can be used with no reduction in rate of infiltration when the salinity of the irrigation water is at or above 2 dS/m. On this basis, it has been assumed that most treated waste waters can be used without loss in infiltration. However, more recent research[11,12] indicates that over an irrigation season, losses in infiltration are significant even at lower SAR (SAR 4-6) and that for systems using both rainfall and irrigation (the general situation for irrigated regions), any increase in SAR reduces the infiltration rate.

A reduction in the soil infiltration rate is undesirable for several reasons. Reduced infiltration rates result in greater time to infiltrate irrigation water, causing water ponding on the surface and possibly insufficient water infiltration to meet crop water needs. It may also result in generation of oxygen deficient soil conditions, all with adverse crop effect on crop production. Perhaps most importantly, a reduction in infiltration rate will result in decreased infiltration of rainfall, increased runoff, and increased soil erosion. In many irrigated regions rainfall is limited but especially winter rains are an important component of the irrigation system, enabling leaching of salts from the soil during and enabling a non saline environment for seedling development in the spring.

In addition to a decrease in the water infiltration rate, loss of infiltration will result in a loss in recharge from rain events. This is very adverse in irrigated regions that are using low quality waters since the rain is often a major component of leaching. A good example is the leaching of salts from the soil profile in the Mediterranean region due to winter rains. Rains during the pre-plant time are especially valuable since for most crops the growth stage most sensitive to salts is during the early seedling stages. Winter rain provides a less saline environment at this critical stage, enabling irrigation with more saline water during the later stages of plant growth. It is this factor that minimizes yield loss in the Mediterranean basin when irrigating with moderately saline waters. Loss of this winter rain would adversely affect crop production. Regions lacking winter rain may not be as successful when using saline waters for irrigation.

Most regions that are critically short of water are already starting to utilize waste waters for irrigation out of necessity. In many instances, these waters are not properly treated, and health risks due to food contamination with pathogens are of concern. Additional problems with reuse of all waste waters, even those that are tertiary treated, are associated with the elevated salinity SAR, pH and often boron present at concentrations adverse to plant growth (species dependent). Municipal and industrial users of water discharge added salts in their wastewater. Current waste water treatment processes utilize chlorine products to reduce or eliminate harmful microorganisms; this process inevitably results in additional increase in salinity of the treated water. Municipal treated wastewaters are always more saline than the water delivered to the municipal users. Perhaps more important than elevated salinity levels of treated wastewater is the increased SAR and pH levels relative to fresh water.

Sustainable use of treated waste water requires that we consider the impact of these waters on soil properties. The detrimental effects of elevated SAR and pH are cumulative and may take years to become evident. Use of amendments such as gypsum may be required for sustainable use of these waters, depending on site-specific conditions such as water composition, soil properties, and the extent and timing of rain events. In the absence of proper management, use of these waters may aversely affect crop production and future soil productivity. This threat is greatest in less developed nations where agricultural extension services are minimal and where the farmers are least able to afford the needed amendments to maintain productivity. This is a concern because municipal wastewater supplies are increasing with increasing municipal water demands, and the utilization of the wastewater is increasing even more rapidly. The one positive factor is that wastewater is a more reliable irrigation water supply than many current surface or groundwater sources.

Saline waters are generally abundant in arid regions. Unfortunately, most currently utilized crops of high value are salt sensitive, and most salt tolerant crops are of lower economic value. Plants also often adapt to elevated salinity by reduced transpiration and reduced growth. Current recommendations[13] are to select increasingly salt tolerant crops with increasing irrigation water salinity, thereby avoiding yield loss. These recommendations need to be reformulated in terms of absolute production (not relative) and economic return. Salt-tolerant forages such as wheat grass are available; however, optimal biomass production is lower than with some salt-sensitive forages. Therefore, total biomass production of wheatgrass under moderate salinity with no yield loss may still result in lower biomass production than would alfalfa, with its less-than-optimal yield under this salinity level.[14]

Waters currently deemed too saline for irrigation, may nonetheless be utilized in less developed nations in arid regions. Reduced yields due to salinity, while not economic in developed nations in humid regions may still be economically acceptable in more arid regions if there are no alternative water supplies. Use of saline waters will require changes in irrigation systems, crops, and management practices. For example, sprinkler systems are not recommended for application of saline waters as foliar salt damage occurs. Saline waters might be best utilized in a cyclic cropping pattern, applying saline water for more salt-tolerant species and less saline water for more salt sensitive species. Thus use of saline waters for irrigation may require significant changes in management systems and extensive farmer education to implement them.

Reuse of irrigation drainage water is another component of improved water utilization. In this instance, there is collection of the drainage water and irrigation with a more salt-tolerant crop. This can be repeated with several cycles of reuse, ending in an evaporation pond for collection of the remaining salts. A major benefit of this system is that it can avoid degradation of the existing fresh water. Such a system would enable full utilization of a water source. The feasibility of this has been demonstrated in various field studies,[15] but it is currently not widely adopted and will require infrastructure development. By implementing these reuse practices on a large scale and by keeping the drainage water separate from other water supplies, may be the only ways to avoid degradation of large supplies of moderately saline water underlying irrigated lands. These marginal waters are critical, as they will need to be utilized to sustain agriculture in arid regions.

The potential to develop salt-tolerant varieties of major food crops has not yet been realized, but increased understanding of the genetic traits of salt-tolerant plants holds the promise for breeding new suitable varieties, either via genetic engineering or by conventional breeding methods.

Conclusions

Current agricultural production in arid regions is dependent on irrigation. These arid regions are generally characterized by small irrigated parcels, which are critical to providing food to the farming family. Extraction of water in excess of what is replenished will result in reductions in available fresh water in arid regions. This reduction, along with continuing salinization of water and soils, threatens the ability of these regions to even maintain their current levels of food production and make it unlikely that they can grow sufficient food to feed their populations without technological advancements, infrastructure investment and an extensive extension system to transfer existing technology to the farmers. In the short and intermediate time frame, brackish and moderately saline ground waters are relatively abundant and can be potentially utilized to grow many, but not all crops. These practices can extend current high utilization of irrigation for decades in many countries, but will require improved and often costly management practices. The major hazard to using these waters is their elevated SAR and potential for loss of soil productivity, especially if there is significant rain. Improvements in irrigation practices, investment in new technologies, and development of salt-tolerant plant varieties may enable these regions to utilize more abundant saline waters for irrigation and slow the rate of degradation of fresh water supplies and thus extend their critical use for drinking supplies.

↑ Back to the top

---

[1] Ghassemi, I., A.J. Jakeman, and H.A.Nix. 1995. Salinisation of Land and Water Resources. University of New South Wales Press LLTD. Sydney. 526pp.
[2] Postel, S. 1999. Pilar of Sand: Can the Irrigation Miracle Last? W.W. Norton & Co. London.
[3] Postel, S. 1997. Last Oasis: Facing Water Scarcity. W.W. Norton & Co. New York.
[4] Postel, S. 1999. Pilar of Sand: Can the Irrigation Miracle Last? W.W. Norton & Co. London.
[5] Ghassemi, I., A.J. Jakeman, and H.A.Nix. 1995. Salinisation of Land and Water Resources. University of New South Wales Press LLTD. Sydney. 526pp.
[6] Postel, S. 1997. Last Oasis: Facing Water Scarcity. W.W. Norton & Co. New York.
[7] National Research Council. 2008. Desalination: A Nation Perspective. Washington D.C.
[8] McNeal, B.L., and N.T. Coleman. 1966. Effect of solution composition on soil hydraulic conductivity. *Soil Sci. Soc. Am. Proc*. 30:308-312.
[9] Suarez, D.L., J.D. Rhoades, R. Lavado, and C.M. Grieve. 1984. Effect of pH on saturated hydraulic conductivity and soil dispersion. *Soil Sci. Soc. Am. J.* 48:50-55.
[10] Ayers, R.S., and D.W. Westcot. 1985. Water quality for agriculture. FAO Irrigation and Drainage Paper 29 Rev.1. FAO, Rome.
[11] Suarez, D.L., J.D. Wood, and S.M. Lesch. 2006. Effect of SAR on water infiltration under a sequential rain-irrigation management system. *Agric. Water Mgmt*. 86:150-164.
[12] Suarez, D.L., J.D. Wood, and S.M. Lesch. 2008. Impact of sequential applications of rain and irrigation water on infiltration into a cropped soil. *J. Environ. Qual.* 37:S169-S179.
[13] Ayers, R.S., and D.W. Westcot. 1985. Water quality for agriculture. FAO Irrigation and Drainage Paper 29 Rev.1. FAO, Rome.
[14] Grattan, S.R., C.M. Grieve, J.A. Poss, P.H. Robinson, D.L. Suarez, and S.E. Benes. Evaluation of salt tolerant forages for sequential water reuse systems: I Biomass production. *Agric. Water Mgmt*. 70:109-120.
[15] Rhoades, J.D, F.T. Bingham, J. Letey, G.J. Hoffman, A.R. Dedrick, P.J. Pinter, and J.A. Replogle. Use of saline drainage water for irrigation: Imperial Valley study. *Agric. Water Mgmt.* 16:25-36.

2.6. Epidemics

2.6.A. Epidemics: A Thumbnail Sketch of the Past, Present and Future (Debarati Guha-Sapir)

Author and Organization: Debarati Guha-Sapir, Professor, Department of Epidemiology and Public Health; Director, WHO Collaborating Centre for Research on Epidemiology of Disasters and Conflicts (CRED), and University of Louvain School of Medicine, Brussels
Contact Information: Debarati.Guha@uclouvain.be

Introduction

Epidemics are generally defined as the occurrence of more cases of a disease than would be expected in a community and that spreads sustainably during a given time period. The definition as such includes chronic diseases, but it commonly refers only to infectious or communicable diseases.

A related concept is a pandemic which spreads over many countries and sometimes the world and frequently involves the emergence of a new disease. It is not a pandemic not only because it is widespread or kills many people; it must also be infectious.[1]

In 2008, WHO reported 26 outbreaks of which, one was a new virus from the *arenaviridae* family and one was the avian flu pandemic[2]. Other outbreaks may go unreported, as surveillance systems are weak or dysfunctional in many part of the world where infectious diseases mostly occur. Recent pandemics include Sudden Acute Respiratory Syndrome (SARS) in 2003, West Nile Fever, and the ongoing H1N1 (swine) flu. Other pandemics which mobilized the global community were bovine spongiform encephalopathy (BSE, or mad cow disease), and the more sustainably active epidemics such as drug-resistant tuberculosis and HIV/AIDS. Outbreaks of disease such as cholera in Zimbabwe in 2008–2009, the plague in India in 1994, Ebola and hepatitis outbreaks in conflict affected regions of the Democratic Republic of Congo (DRC), Angola and Uganda are nationally contained, but nonetheless they have wide regional and sociopolitical relevance.

Epidemics generate global fear that is essentially linked to people's uncertainty of the size, spread, and individual outcomes. The basic questions underlying these fears are "How bad is it going to get?" Or "am I likely to get sick with this disease?" The next section briefly describes the techniques epidemiology has to offer that provide some answers to these questions. The following section discusses current issues under debate and their global impact. Section 3.7.A-3 presents selected global initiatives of interest, and Section 3.7.A-4 sets out priority areas for policy and programmes.

This paper is not intended to be complete or unequivocal. Not only are many interesting and promising items excluded here, the positions taken are eminently debatable.

Epidemics and Pandemics –Measuring Spread and Impact

Epidemics are fast-moving disease events which are usually limited in time. Exceptions to this are the rise and spread of HIV/AIDS, or the re-emergence of tuberculosis on the global level. Highly virulent pathogens tend to limit the spread and duration of an epidemic simply because the infected person dies shortly after contracting the disease and has little time to actually infect many others. The pathogen therefore needs to mutate to less virulent forms over time to survive in nature.

This characteristic is only one example of the immense biological and sociopolitical complexities of an epidemic. At very fundamental levels, the dynamics of any epidemic are dependent on the rate of transmission from infectious to susceptible hosts. In many disease models, this rate is captured in a compound probability of transmission. This parameter can be broken down into a number of variables, including the rate at which the infected person comes in contact with others, whether those contacts actually result in disease transmission, and variations in severity of illness or periods of infectivity. [3] Understanding of each of these sub-parameters allows health professionals to decide what action to take to address one of these sub-parameters, and it allows policy makers to establish the most effective containment policy. Each of these sub-parameters is more or less actionable according to the context and available resources. Finally, an indisputable advantage for effective epidemic control is early detection – the recognition of an epidemic soon after it has started. Like many population-based phenomena, early detection provides the best chance of success in containment.

Today there are many infectious disease models. Most are based on the basic Susceptible – Infectives – Recovery - SIR approach. In its more complex forms, this approach provides practical insights into how fast an epidemic will progress, who are those at high risk, and what actions can be taken at what stage to contain it. They involve complex mathematical algorithms that were first introduced by David Bernoulli in 1760s following a small pox epidemic.[4] His approach is still considered one of the soundest approaches to modeling disease spread. Modern approaches often underestimate herd immunity (the proportion of individuals who are not susceptible to the disease), which determines the vaccination threshold that has to be exceeded to eliminate an infection and the basic reproduction number (the inverse of the proportion of susceptible individuals).[5] Today's wide-ranging disease models are exceptionally intricate, but they are all variants of the original.

Another important concept is the "Basic Reproduction Number" (R0), which represents the number of new infected persons that can be generated by a single infected individual in a susceptible population. This depends on the level of the disease's infectiousness, the time of infection, the clinical onset of the disease (latency period), and of course, the density of susceptibles near outbreak.[6]

This above framework illustrates that epidemics are both a social process (determinants include contact or mobility) and a biological process (transmission and virulence). The fear of an epidemic and its response is also a perception process.

With regard to public health, security and stability, these factors have been closely associated over centuries. Disease control has its roots in military medicine, which was and is uniquely mass medicine of a certain type. In the 17th and 18th centuries, only the military had the logistic and the operational structure for dealing with large health threats such as epidemics. Public health was an issue of public security, and the main public health measures concerned control of people movements and mandatory health check for special categories (e.g., mobile

populations, migrants, prostitutes). The cholera pandemics of 1830 which invaded Europe from Asia, served as a watershed for military medicine applied to civil society and the origin of public medicine.[7] The "cordon sanitaire"[8] so effectively imposed by the military in Naples against bubonic plague in the early 1700s proved to be totally ineffective against pandemic cholera transmission.[9] Military intervention for epidemic disease control was called "folly" by the British medical establishment and fell out of favor as the limitations of population-based isolation for different diseases were better understood.

Current Issues

In the last decades, epidemic disease combined with bioterrorism is again a matter for security and military. Preparedness for bioterrorism or germ warfare touches upon issues such as the politisation of microbial spread, concepts of civil liberties, or the role of biomedical and surveillance technologies in public health. These are delicate areas of public policy and require wide discussion and open debate.

But the need for better epidemic control should not only be a priority for bioterrorism or germ warfare preparedness. It is also (and maybe more concretely) linked to legitimate global transactions such as trade and tourism. Massive increases in world tourism and increasingly dynamic world trade are highly desirable developments for all citizens, but they are accompanied by risks and hazards, including transmission of disease.

The global perceptions of cholera outbreaks lead to restricting travel to and from countries where a cholera outbreak is occurring, or imposing import restrictions on certain foods. For example, the cholera outbreak in Peru in 1991 cost the country US$770 million due to food trade embargoes and adverse effects on tourism.[10] Announcement of cholera epidemics in Bangladesh has a negative effect on its export market for its seafood, plunging many of its small-scale fishermen into rapid destitution.

Transparency becomes a victim to the legitimate needs of a poor nation to maintain the economic viability of its people. The current cholera outbreak in Ethiopia is referred to as "acute watery diarrhoea," although cholera is likely.[11] The safety of the global community and the survival of the local community are in play, and policies to protect both must be devised.

Whether from a global security platform or from the standpoint of disease or trade and tourism, the military continues to have the best human and material resources, as well as the best technology, for effective, rapid action against epidemics and pandemics. Laboratory facilities, medical personnel, and engineers are more readily available in certain countries over and above what almost any health ministry could offer in a crisis. However, use of military resources for civil action, even in democracies, is deeply controversial. Nonetheless, the use of military resources merits sound examination, especially for volatile, high-stake events such as epidemics.

Given the available evidence and scientific studies, it is difficult to argue convincingly that disease outbreaks, however massive, have directly led to sociopolitical upheaval. However, their impacts on factors that are critical for national and international stability are significant. The HIV/AIDS epidemic in many countries in sub-Saharan Africa illustrates how infectious disease can destroy the viability or the ability of a State to function. The 1996 Ebola outbreak in Kikwit and some of the subsequent outbreaks in Angola, Uganda, and Sudan, killed a large

proportion of desperately scarce health personnel, further sinking the country/region into under-development status.

From sociopolitical perspectives, there are at least two potentially problematic issues. First, pandemics or perceptions of pandemics can fuel the growth of insularity and suspicion of foreign goods or people as potential transmitters of diseases, in itself a driver of social instability. Second, the introduction of the notion of a group at risk (encountered for example, in the control of HIV/AIDS), is a key concept in epidemiology, but it is a delicate notion in public policy arenas.

Balancing all of the above—military and civilian policy for global disease control, trade, tourism, trade, epidemic disease and survival, discrimination of high risk groups—will be key challenges for global disease control.

Selected Global Initiatives

The importance of global monitoring of epidemic disease has not gone unrecognized in the recent past. Lead public health institutions such as the World Health Organization (WHO), the US Centers for Disease Control and Prevention (CDC), and now the more recently created European Centre for Disease Prevention and Control (ECDC) have launched successful initiatives for global disease surveillance and monitoring, as have the non-organized health and medical community.

Among the global effort, one notable initiative is the Global Public Health Intelligence Network (GPHIN), an Internet-based multilingual early-warning tool that continuously searches global media sources for events of potential international public health concern. These events include disease outbreaks, infectious diseases, contaminated food and water, bio-terrorism, exposure to chemical and radio-nuclear agents, and natural disasters. Another highly successful and useful field-based monitoring system is ProMed-mail, the Program for Monitoring Emerging Diseases, which is also an Internet-based reporting system dedicated to rapid global dissemination of information on outbreaks of infectious diseases and acute exposures to toxins that affect human health. In contrast to GPHIN, it is commonly used by practitioners, including those in remote and insecure areas of the world, and it has about 20,000 members worldwide. The WHO has also started up Global Outbreak Alert and Response Network (GOARN), which networks initiatives such those described above.

Some Directions for the Future

How significant is the threat of epidemics for global sociopolitical stability? In a period of competing demands for scarce resources, is this a defensible basis for resource reallocations? Fortunately, actions that address the global control of epidemics as drivers of social instability and the disease control goals of the international public health community are not very different, a serendipitous coincidence that kills two birds with one stone. First, systems for controlling diseases with epidemic and pandemic potential require more creative thinking. These systems must also account for the priorities and perceptions of the local population. Incentives to think outside the classical disease surveillance and monitoring systems paradigms are needed to increase innovation in this area.

The current surveillance models are designed for wealthy, stable countries with reasonable sized health budgets and educated population bases. Evidence over the last decades has illustrated that these models cannot be reliably operational in poor countries or in insecure

regions. Initiatives that are purely based on private voluntary organizations (PVOs) are not sustainable, institutional solutions. While they are admittedly often of high quality, they are nonetheless ad hoc and transient. New paradigms for disease surveillance that account for weak infrastructure and limited human resources need to be developed. One such example is the monitoring of infectious disease outbreaks using "syndromic surveillance" of pre-diagnostic data. These data include medication sales, patients' symptoms, and absenteeism from work.[12]

Cooperation for laboratory facilities and technology exchange between Europe, North America, and other economically comparable countries (e.g., Japan, Korea, Mexico, and Singapore) is essential. The recent experiences with SARS, Avian flu, and H1N1 flu provide an excellent scare opportunity to push a cooperation agenda forward, as most countries are open to such initiatives and aware of the potential political costs at home. Some of this is being addressed and should be supported. At a further stage, the cooperation should also include medium-development countries such as Turkey, Algeria or Thailand. China and India should be encouraged to take on more effective national and regional responsibilities for epidemic and pandemic monitoring. Effective cooperation in laboratory networking and technology transfers will vastly improve monitoring of trends in drug resistance and the appearance of new disease strains, aspects that are central to effective disease control and epidemic response and in the interests of all.

International aid programmers for populations affected by conflicts, including preventive diplomacy efforts, should give higher priority to the public health components of disease control. Today nearly 600 million people are included in populations affected by conflicts, and most of them are in areas where epidemics are most likely to break out and remain undetected for some time.[13]

Global epidemic control is commonly perceived to be an emerging disease issue. This involves early childhood diseases such as measles, diarrhea and acute respiratory infection (ARI), still the top three killers of children worldwide. Weakened and unvaccinated children are susceptible hosts to new and virulent strains. Efforts to stamp out or substantially reduce the incidence of these diseases must absolutely remain a priority, as these contribute indisputably to the bulk of infectious disease deaths.

↑ Back to the top

---

[1] For instance, cancer is responsible for many deaths but is not considered a pandemic because the disease is not infectious or contagious.

[2] WHO Global Alert and Response Disease Outbreak News http://www.who.int/csr/don/archive/year/2008/en/index.html

[3] L. A. Real and R. Biek., Infectious disease modeling and the dynamics of transmission *Current Topics in Microbiology and Immunology*; 315:33-49, Feb. 2007

[4] Daniel Bernoulli introduced this by his analysis of smallpox epidemic in 1760s. He expressed the proportion of susceptible individuals of an endemic infection in terms of the force of infection (virulence) and life expectancy (survival).

[5] Klaus Dietz[1] & J. A. P. Heesterbeek, Bernoulli was ahead of modern epidemiology Correspondence, *Nature* 408, 513-514 (30 November 2000) | doi:10.1038/35046270

[6] Preventing an epidemic thus becomes equivalent to keeping *R0*<1

[7] The key concepts of epidemic control of "quarantine – from the French quarantaine or 40 days as well as *cordon sanitaire* or isolation originated at this time.

[8] The Austrian cordon to prevent the disease from entering from east to west was 10–20 miles wide and 1200 miles long. It consisted of a permanent line of soldiers and sentry posts to prevent anyone from passing and to shoot on sight a unauthorized person. who had not been isolated for the quarantine – (French for 40 days)

[9] Snowden.Frank M. *Naples in the Time of Cholera, 1884-1911*, Cambridge University Press,Cambridge, UK,1995

[10] Heymann,DL., Rodier,GR., Global surveillance of communicable diseases, *Emerging Infectious Diseases*, Vol 4, No 3, July – Sept. 1998

[11] Personnel communication with international medical NGO

[12] Chretien J-P, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, et al. (2008) Syndromic surveillance: Adapting innovations to developing settings. *PLoS Med* 5(3): e72.

[13] Guha-Sapir,D., and D'Aoust O.„The Uncounted Millions*, CRED Working Paper, Brussels, 2009*

2.6.B. Infectious Disease and Social Instability: Prevent, Respond, Repair (Daniel Strickman)

Author and Organization: Daniel Strickman, Agricultural Research Service, USDA
Contact Information: Daniel.strickman@ars.usda.gov

Introduction

The range of effects of infectious diseases on human populations is nothing short of astounding. Historically we read of the great plagues of history and how they appeared to be as important to the outcome of nations as military preparedness, government stability, and sophistication of populations. One only needs to mention black plague (*Yersinia pestis*), yellow fever (a flavivirus), and epidemic typhus (*Rickettsia prowazekii*) to conjure images of societal chaos and devastation. Although these diseases still occur in the modern world, the current immediate threats are different: human immunodeficiency virus (HIV), tuberculosis (*Mycobacterium tuberculosis*), and malaria (*Plasmodium falciparum, vivax, ovale,* and *malariae*) among them. Of course, we could add many pathogens that are feared but which have not yet impacted populations in a way that alters demographics.

Infectious Disease Threat Classification

The kinds of infectious disease threats can be classified in a number of different ways. Academically, they are usually divided taxonomically into viruses, bacteria, protists, nematodes, etc., with a great deal of effort expended on classification according to accurate phylogenetic criteria. These efforts often seem pedantic to laymen, though phylogenetically accurate classification can lead to many practical results and enables focused, more effective research. However, in considering the range of pathogens' effects on social instability, such taxonomic classification is not going to be very useful.

Diseases and Social Stability

The aspects of diseases that are likely to be important to those equally concerned with social stability include severity to the individual, available medical interventions for prevention and treatment, and the means of transmission. Some diseases have a very high fatality rate once the infection occurs, such as HIV and pulmonary anthrax. Others cause their share of deaths in adults but are mainly characterized by their debilitating effects (e.g., dengue, malaria). Sometimes the same infection is much more dangerous in children than adults (e.g., dengue, malaria) or vice versa (e.g., polio). Preventive measures vary from large-scale infrastructure (e.g., sewers, safe food handling) to promotion of good habits (e.g., hand washing following defecation) to vaccines. Treatment ranges from very effective antibiotics to supportive therapy that allows the patient's own immune system to overcome the infection. Transmission is a complicated subject because it represents the ecology of the disease from the standpoint of the pathogen itself. Thankfully, few pathogens are carried in aerosolized droplets or picked up directly from surfaces (e.g., influenza), but close physical contact offers a wider range of pathogens a chance to infect a person. The range of species infected also varies a great deal so that some pathogens only circulate between people (e.g., polio virus) while many others have a range of hosts (e.g., influenza virus) or even live in the environment (e.g., cholera [*Vibrio cholerae*] and tetanus [*Clostridium tetani*]). Although it sounds specialized, one particular form

of transmission has been very problematic for humans. This is transmission by arthropods, especially when the pathogen is specifically adapted to a cycle that involves biological amplification in the arthropod.

Table 2.7.B-1. Examples of Diseases that Could Affect Social Stability.

Disease	Pathogen	Severity	Prevention	Treatment	Transmission
Influenza	virus	low mortality, short illness	vaccine, sanitation	antivirals	aerosol, contact
Toxigenic *E. coli*	*Escherischia coli* O17H57	low mortality, short illness	sanitation	antibiotics	fecal-oral, contamination oral
Plague	*Yersinia pestis*	high mortality, short illness	avoidance, vector control, rodent control	antibiotics	aerosol, fleas from rodents
Malaria	*Plasmodium* spp.	low mortality, long illness	drugs, vector control	drugs	mosquitoes from other people
Epidemic typhus	*Rickettsia prowazekii*	high mortality, short illness	vector control	drugs	body lice from other people
AIDS	virus	high mortality, long illness	sexual discipline	antivirals	sexual contact
tuberculosis	*Mycobacterium tuberculosis*	moderate mortality, long illness	sanitation, detection and treatment, vaccine (not in US)	antibiotics	aerosol

The examples of diseases in Table 2.7.B-1 represent some familiar examples that would seem to have societal implications. Influenza has a relatively low rate of mortality and is accepted as an inconvenience by many, but it is so easily transmitted that it sweeps through the world population and affects normal activities. The fear of a form of the virus that causes more mortality is well justified by the experience with the 1918 epidemic.[1] Toxigenic *E. coli*, especially the O157:H7 serotype, has affected thousands of people in individual outbreaks caused by food contamination.[2] It is rarely fatal, but its sudden appearance among those who shared the same source of food causes disruptions. Plague is largely tamed at the present time thanks to efficient methods to deal with incipient outbreaks; however, reservoirs of plague occur worldwide, and historically, it caused the death of large percentages of some populations. Malaria has insidious effects on a population because it contributes to shocking childhood death rates and tends to lower the vigor of the adult population in a very persistent manner (World Health Organization 2008).[3] Malaria has been difficult to control in large areas, though it has been brought under control in most temperate climates. Epidemic typhus was a great killer through WWII,[4] usually favored by economic disruptions and subsequent increases in body lice. With the advent of tetracycline antibiotics and efficient insecticides in the 1940s and 1950s, epidemics of this disease can be brought under control[5]. The sudden appearance of

HIV in the early 1980s is likely the most recent salient example of an emerging disease, having apparently originated from a non-human retrovirus within modern times. Fortunately, transmission of HIV is not robust and generally requires venereal contact, which provides important strategies for prevention. Finally, tuberculosis was at one time responsible for about 25% of deaths in some regions. Vigorous sanitation measures and development of effective drugs made tuberculosis a persistent but manageable problem. The appearance of antibiotic-resistant forms of the pathogen raises the specter of the return of epidemics of this terrible disease. There are, of course, other current and historical examples of pathogens with societal impacts, but this short list illustrates the variety of issues involved in evaluating and managing them.

Emerging Infectious Diseases

Before moving on to some general strategies for preventing societal impacts of infectious diseases, it is important to discuss the concept of emerging infectious diseases. The appearance of HIV was a startling development at the time: an apparently new pathogen that was 100% fatal and unaffected by any medical intervention. Since then, drug therapy has reduced the impact of HIV infection for the individual, but at a great societal cost. More important has been the tedious task of educating the public about a problem that involves moral taboos, subjects little discussed, and unwelcome alterations of behavior. Once convinced that the problem is preventable and deadly serious, many populations have complied with the necessity to exercise sexual discipline. The problem is not over, but society in general deals with it at a cost that is largely accepted.

The experience with HIV raises the question as to whether studying the past really is the key to preparing for the future. While extensive study is conducted of the historical diseases that changed national histories—plague, epidemic typhus, tuberculosis—will nature provide a new or modified pathogen that surprises us? The haunting possibility that what seems like a small problem now (e.g., rodent-borne hantavirus) could become a large-scale epidemic if conditions changed stimulates interest in studying what appear to be obscure infections. A recent example is chikungunya disease. Caused by a virus that has appeared in severe epidemics in the Indian Ocean basin and Africa, the disease is seldom fatal, but it causes painful symptoms for about 30 days. Historically it has appeared in large outbreaks and then disappeared for many years at a time. It was principally transmitted by a single mosquito species, the yellow fever mosquito (*Aedes aegypti*), which itself has been transported by humans throughout the tropical world.[6] In 2007, during a very large epidemic, the virus underwent a single mutation that suddenly made it susceptible to transmission by the Asian tiger mosquito (*Aedes albopictus*), a native of southeastern Asia.[7] The addition of a form of the virus transmissable by the Asian tiger mosquito was particularly bad because this mosquito had only recently (since 1986) been transported to many more locations in the used tire trade.[8,9] Unlike the yellow fever mosquito, the Asian tiger mosquito can live through cold winters, so that its expanding distribution includes Europe and the US, as well as a wide distribution in the tropics. As a direct result of the human-induced expansion of the distribution of Asian tiger mosquito and the natural mutation of the chikungunya virus, the disease broke out for the first time outside of its normal range in 2007 in Italy.[10] Fortunately, this occurred in an area with a vigorous mosquito abatement function, and vector control limited the number of Italian cases to about 300. At this point, it seems like just a matter of time before many more people are affected in other parts of Europe and the United States.

How can society prepare itself for potential epidemics? We have elaborate and largely effective systems in many parts of the world that protect populations from epidemics. These systems depend on reasonable access by individuals to medical treatment, public health infrastructure that gathers statistics and responds to outbreaks, and a population that is aware of the importance of infectious disease and therefore willing to support public health measures. In a current context, there is much frustration with the cost of this system and the gradual erosion of public health services in countries like the U.S. To the extent that infectious disease control needs to be strengthened, the key improvement is probably the same as for so many other systems that support modern human life: integration. Integrated Disease Management (IDM) would further organize preparedness and response by performing two important functions. First, it would educate responsible parties at all levels, including the public, on the functions of IDM. Those functions are risk assessment, surveillance, prevention and control, and sustainable support.

Functions of Integrated Disease Management (IDM)

Risk assessment is the research and data gathering required to know which problems are likely to be important and which can be safely downplayed. Academic research, especially which is innovative and motivated by curiosity, is an important layer of risk assessment that is most likely to identify those problems that the medical community at large does not even suspect. Data gathering agencies like the Centers for Disease Control and Prevention also perform an important function in risk assessment when they gather disease data and provide analysis. Improvements in the efficiency and effectiveness of these processes will go a long way toward protecting society. They depend heavily on small-scale attention to health by individual physicians and county public health departments. Above that level, the contribution is largely collation of results, though excellent work is often accomplished by specialized data gathering teams at a state or national level in response to particular outbreaks.

Surveillance can contribute to risk assessment, as discussed above, but it should also inform those who must take action to prevent or limit an outbreak. One example that worked well recently was the evaluation of dead crows for West Nile virus in the U.S. Using a newly invented dipstick diagnostic device that cost about $2.50 per test, technicians in the field were able to determine infection in dead birds (particularly crows that have very high viremia) within 15 minutes. As a result, they could call for concentration of mosquito abatement resources in those locations where the virus was most active. Had there been a delay of even a few days between the test and the result, the consequences would be much less relevant to those responsible for controlling the vector mosquitoes.

Prevention and control of disease includes the myriad activities in education, sanitation, vaccination, and treatment that stop infection of individuals and groups. Some diseases are easily controlled once detected, and others are not. Vaccines are only available for a handful of pathogens. Some sanitary measures have become entrenched in society, which is a good development, but the public at large has forgotten why those measures were started in the first place, which weakens economic support. Treatment is the last resort to spare the individual and is a very expensive activity for society.

Sustainable support for the application of IDM requires constant input of data to the medical system. In the best scenarios, the medical system continues to pay attention to and be prepared for the particular infection. In the worst scenarios, the skills and materials for prevention and

treatment are lost until an outbreak motivates to rebuild the capability. The current U.S. system for sustainability depends on professional advocates like CDC and county public health officers.

Integration of Contributors

The second function of successful IDM would stress integration of all the contributors to effective programs. In many situations, there is a large divide between those activities usually considered clinical (vaccination, treatment, case detection, and recording) and those usually considered environmental (pollution control, mosquito abatement, wholesale food safety). In actual practice, a large number of people, all with very different backgrounds, are involved in keeping the public healthy, and it is difficult to bring these parties together in meaningful ways. Consider a county-level meeting that includes the county's parks department, public health, agriculture commissioner, mosquito abatement manager, water district manager, urban planners, public works manager, and educators. These individuals do not typically share the same professional languages and may not even recognize their common purpose in protecting health. There are many outstanding examples of these various groups coordinating their efforts; however, it seems clear that we have a long way to go in this area.

Epidemics Affecting Societal Functions

How likely are we to have epidemics that affect societal functions? In many parts of the world affecting the majority of the human population, those epidemics take place already. Malaria and tuberculosis in Africa, dengue in Southeast Asia, and diarrheal disease in South Asia cause great costs to the societies involved. Sometimes those costs are accepted with the tragic consequence of bad population level health and the consequent decrease in economic development and quality of life. On the hopeful side, science continues to provide new insights into disease management, most recently fueled by dramatic advances in genomics. Less hopefully, populations continue to strain resources available, and there are clear signs that agriculture and industry are approaching firm limits. In a particularly cruel example of a positive feedback loop, human use of fossil fuels has increased uncertainty by initiating a chain of physical effects on climate. One of the consequences of those strains can be greater susceptibility to epidemics from old and new diseases as the science, medical, and environmental health infrastructure suffers along with the rest of the economy. The possibility of an actual decrease in those infrastructures would make epidemics all the more likely as the effective means for controlling epidemics are lost.

The steps necessary to maintain and improve the ability to prevent, manage, and recover from epidemics are clear enough. First, continued support of institutions that have effectively protected much of the world population from epidemics is needed. Second, continued support of scientific efforts driven by curiosity about infectious disease is needed. Third, resources must be concentrated on public and environmental health, possibly at the cost of funding for advancements in individual medical treatment. Finally, creativity must be focused on solving the problems of accomplishing the best IDM that is humanly possible.

↑ Back to the top

[1] Reid, A.H., J.K. Taubenberger, and T.G. Fanning. 2004. "Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus." *Nature Reviews Microbiology* 2: 909-914.

[2] Duffy, G., E. Cummins, P. Nally, S. O'Brien, and F. Butler. 2006. "A review of quantitative microbial risk assessment in the management of *Escherichia coli* O157:H7 on beef." *Meat Science* 74: 76-88.

[3] World Health Organization. 2008. *World Malaria Report 2008*. WHO/HTM/GMP/2008.1, xvii+190 pp.

[4] Moe, J.B., and C.E. Pedersen, Jr. 1980. "The impact of rickettsial diseases on military operations." *Military Medicine* 145: 780-785.

[5] Hechemy, K.E., and A.F. Azad. 2001. Epidemic typhus. In: Service, M.W. (ed.), *Encyclopedia of Arthropod-transmitted Infections of Man and Domestic Animals.* CABI Publishing, NY. Pp. 170-174.

[6] Woodall, J. 2001. Chikungunya virus. In: Service, M.W. (ed.), *Encyclopedia of Arthropod-transmitted Infections of Man and Domestic Animals.* CABI Publishing, NY. Pp. 15-19.

[7] Pagès F, Peyrefitte CN, Mve MT, Jarjaval F, Brisse S, et al. 2009. "*Aedes albopictus* Mosquito: The Main Vector of the 2007 Chikungunya Outbreak in Gabon." *PLoS ONE* 4(3): e4691. doi:10.1371/journal.pone.0004691.

[8] Hawley, W.A. 1988. "The biology of *Aedes albopictus." J. Am. Mosquito Control Assoc.* (Supplement) 4: 1-40.

[9] Schaffner, F., B. Bouletreau, B. Guillet, J. Guilloteau, and S. Karch. 2001. *Aedes albopictus* (Skuse, 1894) established in metropolitan France. *The European Mosquito Bulletin* Issue 9, pp. 1-3.

[10] Rezza, G., L. Nicoletti, R. Angelini, R. Romi, A.C. Finarelli, M. Panning, P.Cordioli, C. Fortuna, S. Boros, F. Magurano, G. Silvi, P. Angelini, M. Dotton, M.G. Ciufolini, G.C. Majon, A. Cassone, for the CHIKV study group. 2007. "Infection with chikungunya virus in Italy: an outbreak in a temperate region." *The Lancet* 370: 1840-1846.

3. Social Science Modeling and Analysis Approaches

3.1. Analyzing Political and Social Regional Stability with Statistics: Challenges and Opportunities (Victor Asal, Stephen Shellman)

Authors and Organizations: Victor Asal, University at Albany; NYSU; Stephen Shellman, College of William and Mary
Contact Information: vasal@email.albany.edu; steveshellman@gmail.com

The use of statistics in the understanding and forecasting of political and social regional stability offers some important advantages and shows important promise as a policy aid in evaluating possible interventions, but it faces important challenges that need to be overcome in terms of (1) the data that are necessary to perform accurate assessments, and (2) the best means to choose the most effective methods for performing assessments.[1] (By *statistics*, we mean the use of mathematical analysis to identify, quantify, and summarize relations between numerical values that represent real world observations.) Most importantly, statistical analysis allows the application of a well-used approach to the analysis of stability. This permits the analyst to put confidence intervals around the possible relationships between the possible causes analyzed and the outcome of instability, as well as the impact of possible interventions. This also allows the analyst to appropriately caveat the resulting analysis. On the other hand, a statistical approach requires that the information to be studied be entered into a spreadsheet format, as well as a numerical format. This may be relatively easy with certain phenomena such as inches of rain a year or the gross domestic product of a country, but it is much more challenging to translate such phenomena as discrimination or violence into numbers. Furthermore, certain kinds of phenomena of interest are difficult to study because of their rare occurrence or an almost complete lack of relevant data.

If we understand the task of analysis to be one of explaining the phenomena and of forecasting the phenomena, statistical work on political and social instability has progressed far in terms of explaining what factors are related to instability. However, despite years of pioneering work by a few researchers, there has been very little accomplished in the area of forecasting related to political and social instability. This paper provides a brief overview of some of the research on the causes of stability and instability conducted using statistical analysis. The paper then discusses efforts that have been made in the area of forecasting. Finally, the paper examines the challenges related to statistical analysis when it comes to issues of data and method.

Explaining

As mentioned above, unlike narratives or comparative case studies or formal models, statistical analyses offer the advantage of being able to more formally control a variety of competing factors.[2] Statistical analysis provides a check on opinion by offering a transparent methodology for identifying causal links and by providing confidence intervals around the assessments given. This is true because, while subject matter experts bring a deep reservoir of knowledge about their subject matter, they "… are not mere objective experts. They develop belief systems, operational codes, theories, and agendas; they are subject to the same cognitive and social psychological and group dynamics that affect decision makers." In addition, "Evidence from surveys suggests that forecasts of decisions in conflicts are typically based on experts'

unaided judgments."[3] Studies have found that expert predictions are not better than those made by non-experts and that they often ignore data that disagree with their views.[4,5] Analysts use statistical methods belief systems about their subjects as well, but (1) the statistical analysis can act to reject relations that are assumed to be true, (2) the transparency of the approach allows for others to check their analysis and to incorporate other measures to identify different dynamics at work in a clear fashion.

The growing sources of cross-national data have led to a greater use of statistics, for if there are no numbers, relevant data need to be collected, or statistical analysis cannot be used. Not only has there been enormous growth in the area of hand-coded data sets, but developments in automated coding have allowed for a much more intense focus on event analysis.[6,7,8] In addition to the greater availability of data, there has been a growing sophistication in the methods used that have been facilitated by technological developments, especially in the area of dealing with the challenges created by time-series cross-sectional data.[9]

Recent research has found important correlations between political violence and other forms of instability. For example, there is evidence that certain kinds of geographic attributes (mountainous terrain) make violence more likely,[10] that institutional arrangements crucial for economic development are more likely (or less likely),[11] and that whether civilians become displaced in the first place and whether they end up as internally displaced or become international refugees (a key source of instability) is impacted by the kind of violence and the actor (state or non-state violence) perpetrating the violence.[12,13] Indeed there is an enormous literature using statistics to analyze different aspects of political and social instability, particularly in the area of the likelihood of political violence. For example, in addition to mountainous terrain, [14] natural disasters,[15] economic conditions[16] (especially low income levels[17]), regional conditions of conflict or stability,[18] types of natural resources,[19] and discrimination[20] have all been found to have an impact on the likelihood of political violence.

As the list above suggests, a great deal of work has been done identifying country-level factors that lead to political and social instability. It is important to note that much of this analysis focuses on the state level of analysis. This primary focus on the state level constrains the analysis for two reasons. First, as Gleditsch argues, instability is often a regional affair and can spread from one country to another. [21] For example, the risk of civil war is a result not only of the character of the country but the nature of the links that country has with other countries.[22] Salehyan, offering support for the importance of the regional context, identifies a series of international factors, and in particular the presence of refugees, impacts domestic conflict. Amongst other transnational factors, Salehyan finds that refugee diasporas increase the likelihood of a rebellion.[23] Second, instability is often does not "happen" at the state level but is tied to a particular geographic area or population within a country and thus the country level of analysis may obscure what is going on.[24] (The data section of this paper expands upon this point.)

Forecasting

Almost all of the analysis discussed up until now has focused on identifying potential causal relations between different factors and different types of instability. This kind of analysis allows for the identification of "danger signs" and provides a better understanding of the regional and state-level playing field if conditions should change. What is missing in much of the literature that is important particularly from a policy perspective is an effort to go beyond

what in-sample analysis will allow. To attempt to forecast likely outcomes there is a need to go beyond in-sample analysis and do out-of-sample analysis by partitioning the data.[25] What this means is that most people who have studied political and social instability have used all the data they have in one analysis, producing a model that fits all their data. In-sample and out-of-sample analysis means that the analyst partitions the data and builds a model with one part of the data and then uses the other part of the data to see how well the model constructed on different data fits different data. If the model works "well" on data that were not used to construct it, then there is a strong possibility that it will work well on other data, and most importantly, it may be able to forecast into the future with some degree of accuracy.

We should note that advances in this kind of forecasting research have often been hampered by a lack of collaboration across disciplines, with computational and social scientists rarely working together.[26] This kind of collaboration between social scientists, who are experts in the phenomena and the theories related to political and social instability, with computational social scientists, who focus on advancing sophisticated analysis methods, is beginning to happen more frequently and hopefully will result in advances of this kind of analysis in the near future. Despite the lack of focus on this kind of research amongst most social scientists, the efforts of researchers that have attempted to go beyond in-sample analysis have produced very interesting and substantive results.[27,28,29,30] For example, O'Brien examines which factors lead to conflict and instability. O'Brien focuses on structural factors using a pattern classification approach, as well as an in-sample out-of-sample-approach to identify how well his model can forecast the likelihood of instability. Using factors like GDP per capita, infant mortality, democracy, political and civil rights, and the presence of a youth bulge, his model is about 80% accurate in its predictions. [31]Bennett and Stam's analysis of the likely time for the conventional war in Iraq in 2003 illustrates some of the advantages and challenges of this kind of work. The invasion of Iraq is a classic example of both an important intervention in an unstable region, as well as a key example of how interventions can have unintended consequences. Using a model created for explaining war duration, in March 2003, Bennett and Stam took "…the coefficient estimates from it and make a prediction about the duration of the ongoing war with Iraq."[32] Presenting a variety of scenarios based on different statistical analyses, the Bennett and Stam prediction was between 27 days and 2.5 months. As the authors recognize, the analysis is dependent on making contingent estimates about the likely terrain where key battles will be fought and the strategies used by the opponents. Nonetheless, the performance of the model, when focused on one particular case projecting into the future, is not bad, especially since the original model was not designed for forecasting. The more difficult challenge though, as the authors recognize, is to do the same kind of analysis for the insurgency phase of the conflict. In this circumstance, Bennett and Stam predicted in 2005 that this phase of the conflict would last seven years (83 months). (Bennett and Stam also go on to predict the likely length of possible wars between the United States Syria or between the United States and North Korea.)[33] Like all such predictions (and this is one of the reasons that forecasting is a daunting task within the academic context with certain types of rewards and punishments), this can only be validated in hindsight.

Data

Whether explaining or trying to forecast political or social instability, all statistical approaches face particular challenges. One challenge is that existing data are not often shared. For example, it is only within the last several years that data on domestic terrorist behavior have

become publically available (see http://www.start.umd.edu/gtd/ and http://wits.nctc.gov/) such that researchers can use and report on their results, including methodologies for how they coded and transformed the data.[34] Key advantages of statistics are that (1) the results are transparent and easily replicable, and (2) results are open to a modification of assumptions. Therefore, the free availability of data is crucial to maximize the utility of this kind of analysis. When data are used that cannot be shared, a secondary check on the analysis is not possible. The increasing public availability of all kinds of data, as well as the increasing aggregation of useful data, is greatly benefitting the study of instability.

As mentioned above, another key challenge has been a lack of numerical data, particularly at the sub-state level regarding non-state actors. In order for statistical analysis to be used at the sub-state level, there is a need for a great deal more geographic data, as well as coding at the non-state actor level. While the collection of event data involving violence has enhanced our knowledge of this violence, we cannot get traction on understanding the data without more data about the groups themselves, in addition to data only on their behaviors.

For example, when it comes to the destabilizing phenomena of terrorism, most of the statistical literature examines what factors account for the amount of terrorism in a country or how much terrorism a country "exports."[35] While this kind of analysis is useful for some policy issues, it obscures the important issue of how or why the organizations or groups who are actually carrying out the destabilizing behavior act in certain ways.[36] To draw an analogy, we know that the structure of a state impacts how it uses violence. Democracies, for example, are much less likely to target other democracies for violence.[37] When it comes to understanding violent non-state actors and how their attributes impact their behavior the data that does exist is very sparse. Using this data Asal and Rethemeyer have found that organizations that are large, very networked and religious are more likely to be very lethal.(The data outlining terrorist organizational attributes is available upon request by emailing vasal@email.albany.edu.)[38,39] Unfortunately, this analysis must be caveated because the data that do exist are only for one observation per group and are limited to the time frame of 1998–2005. With this kind of limited data, forecasting is simply impossible. More robust data over a longer time period would allow for a much broader analysis and would also allow for attempts at forecasting. In order for statistics to contribute more fruitfully to an analysis of social and political instability, the lack of sub-state data needs to be addressed.

Another issue with data is how a concept is operationalized (i.e., how is a concept like discrimination is turned into a number that can then be analyzed). There is a preponderance of literature indicating that when groups are the subject of discrimination from governments, they are more likely to use violence in the pursuit of their political goals.[40] The problem is how the statistician operationalizes discrimination. Several researchers have created proxies for discrimination using ethnic fractionalization as a measure, and they have found that it does not lead to more violence in a country.[41,42] The finding is suspect because, while ethnic fractionalization may lead to a group being more likely to suffer discrimination, it is not the same as discrimination. Regan and Norton's analysis at the group level found that indeed discrimination is a significant contributor to ethnic violence[43] as has Wimmer and Min at the state level.[44] On the other hand, the data that both of these papers used are limited in that they only consider discrimination of certain kinds of groups, which may create issues of selection bias. This finding of selection bias has led to a large effort to code a broader set of ethnic

groups.[45] This examination of operationalization underlines both the utility and the importance of the transparency that is possible using a statistical approach.

Method

As mentioned above, advances in time series analysis, advances in controlling for different kinds of data, and automated coding have allowed analysis to move from aggregate country-level analysis to weekly forecasting models that can focus on one particular country.[46] These multi-equation modeling approaches using time series data open the possibility for the construction of early warning models that can draw on the broad literature and new sources of data and new methods to construct useful policy tools.

In the same way that real benefits are likely to arise from the collaboration of computational scientists and social scientists, it is clear that multi-methodological approaches that marry statistical analysis to other kinds of analysis make it possible to create hybrid analyses that allow for formal theory, narrative evidence, and quantitative analysis to identify the complex causes of instability. One area that was addressed briefly above but should be highlighted here is the need to integrate the geospatial component much more strongly in into statistical analysis. Drawing on this kind of analysis, for example, Sambanis finds that inter-regional inequality is a much better predictor of civil war than national measures of inequality.[47] In addition, there is a need to integrate social network analysis much more fully into the study of instability. While networks have often been used as a theoretical construct in studying issues related to instability,[48] the actual method has not been integrated as much, although that has started to change.[49,50,51 52,53] Maoz was able to identify network alliance polarization as a key factor in the increase of international conflict.[54]

Conclusion

This short overview cannot fully address many of the efforts underway and the complexity of some of the methodological challenges faced in trying to analyze and forecast. Statistical analysis may benefit from increasing the focus on the need for more and better open source and shared sub-national disaggregated data collection across space and time. The current paucity is beginning to be filled by more aggressive hand coding and more sophisticated automated coding technology. These efforts are enabling disaggregation to the daily and city levels. If these data are publicly shared and there is increased collaboration across the sciences, there is much that statistical analysis will have to offer to policymakers at the national and multinational levels.

↑ Back to the top

---

[1] Mazarr, Michael J. 2007. "The Iraq War and Agenda Setting." *Foreign Policy Analysis* 3(1): 1-23.

[2] Hechter, M., and D. Okamoto. 2001. "Political Consequences of Minority Group Formation." *Annual Review of Political Science* 4 (1):189-215.

[3] Green, K. C. and J. S. Armstrong (2007). "The Ombudsman: Value of Expertise for Forecasting Decisions in Conflicts." *Interfaces* 37(3): 287.

[4] Tetlock, P. E. (1999). "Theory-Driven Reasoning About Plausible Pasts and Probable Futures in World Politics: Are We Prisoners of Our Preconceptions?" *American Journal of Political Science* 43(2): 335-366.

[5] Green, K. C. and J. S. Armstrong (2007). "The Ombudsman: Value of Expertise for Forecasting Decisions in Conflicts." *Interfaces* 37(3): 287.
[6] Gerner, D. J., P. A. Schrodt, et al. (1994). "The analysis of political events using machine coded data." *International Studies Quarterly* 38(1): 91-119.
[7] Schrodt, P. A. and D. J. Gerner (1994). "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982-92." *American Journal of Political Science* 38(3): 825-854.
[8] Schrodt, P. A. and D. J. Gerner (2004). "An Event Data Analysis of Third-Party Mediation in the Middle East and Balkans." *Journal of Conflict Resolution* 48(3): 310.
[9] Beck, N. and J. N. Katz (1995). "What to do (and not to do) with time-series cross-section data." *American Political Science Review* 89(3): 634-647.
[10] Fearon, J. and D. Laitin. (2003). "Ethnicity, Insurgency, and Civil War Replication data." Retrieved April 3 2004, 2004, from http://www.stanford.edu/~jfearon/.
[11] Rodrik, D., A. Subramanian, et al. (2004). "Institutions rule: the primacy of institutions over geography and integration in economic development." *Journal of Economic Growth* 9(2): 131-165.
[12] Moore, W. H. and S. M. Shellman (2004). "Fear of Persecution: Forced Migration, 1952-1995." *Journal of Conflict Resolution* 48(5): 723-745.
[13] Moore, W. and S. Shellman (2006). "Refugee or Internally Displaced Person?: To Where Should One Flee?" *Comparative Political Studies* 39(5): 599.
[14] Fearon, J. and D. Laitin. (2003). "Ethnicity, Insurgency, and Civil War Replication Data." Retrieved April 3 2004, 2004, from http://www.stanford.edu/~jfearon/.
[15] Nel, P. and M. Righarts (2008). "Natural Disasters and the Risk of Violent Civil Conflict." *International Studies Quarterly* 52(1): 159-185.
[16] Blomberg, S. B., G. D. Hess, et al. (2004). "Economic conditions and terrorism." *European Journal of Political Economy* 20(2): 463-478.
[17] Hegre, H. and N. Sambanis (2006). "Sensitivity analysis of empirical results on civil war onset." Journal of Conflict Resolution 50(4): 508.
[18] Gleditsch, K. S. (2002). *All International Politics Is Local: The Diffusion of Conflict, Integration, and Democratization.* Ann Arbor: University of Michigan Press.
[19] Humphreys, M. (2005). "Natural resources, conflict, and conflict resolution: Uncovering the mechanisms." *Journal of Conflict Resolution* 49(4): 508.
[20] Regan, P. M. and D. Norton (2005). "Greed, Grievance, and Mobilization in Civil Wars." *Journal of Conflict Resolution* 49(3): 319-336.
[21] Gleditsch, K. S. (2002). *All International Politics Is Local: The Diffusion of Conflict, Integration, and Democratization.* Ann Arbor: University of Michigan Press.
[22] Gleditsch, K. and M. Ward (2006). "Diffusion and the international context of democratization." International Organization 60(04): 911-933.
[23] Salehyan, I. (2007). "Transnational Rebels: Neighboring States as Sanctuary for Rebel Groups." World Politics 59(2): 217-242.
[24] Weidmann, N. (2007). Group Concentration and Ethnic Conflict: An Empirical Study.
[25] King, G. and L. Zeng (2001). "Improving Forecasts of State Failure." World Politics 53(4): 623-658.
[26] Bankes, S., R. Lempert, et al. (2002). "Making Computational Social Science Effective: Epistemology, Methodology, and Technology." Social Science Computer Review 20(4): 377.
[27] Schrodt, P. A. (2000). Pattern Recognition of International Crises Using Hidden Markov Models. Political Complexity: Nonlinear Models of Politics.
[28] King, G. and L. Zeng (2001). "Improving Forecasts of State Failure." World Politics 53(4): 623-658.
[29] O'Brien, S. P. (2002). "Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis." Journal of Conflict Resolution 46(6): 791-811.
[30] Bennett, D. S. and A. C. Stam (2006). "Predicting the Length of the 2003 US–Iraq War." Foreign Policy Analysis 2(2): 101-116.

[31] O'Brien, S. P. (2002). "Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis." Journal of Conflict Resolution 46(6): 791-811.
[32] Bennett, D. S. and A. C. Stam (2006). "Predicting the Length of the 2003 US–Iraq War." Foreign Policy Analysis 2(2): 101-116. (The original analysis can be found at http://www.personal.psu.edu/dsb10/warduration/Report-Iraq_War_Duration_2003.pdf.)
[33] Bennett, D. S. and A. C. Stam (2006). "Predicting the Length of the 2003 US–Iraq War." Foreign Policy Analysis 2(2): 101-116. (The original analysis can be found at http://www.personal.psu.edu/dsb10/warduration/Report-Iraq_War_Duration_2003.pdf.)
[34] LaFree, G. and L. Dugan (2007). "Introducing the Global Terrorism Database." Terrorism and Political Violence 19(2): 181 - 204.
[35] Enders, W. and T. Sandler (2006). The Political Economy of Terrorism, Cambridge University Press.
[36] Asal, V. and R. K. Rethemeyer (2008). "Dilettantes, Ideologues, and the Weak: Terrorists Who Don't Kill." Conflict Management and Peace Science 25(3): 244-263.
[37] Huth, P. K. and T. L. Allee (2003). The Democratic Peace and Territorial Conflict in the Twentieth Century. Cambridge, Cambridge University Press.
[38] Asal, V. and R. K. Rethemeyer (2008). "Dilettantes, Ideologues, and the Weak: Terrorists Who Don't Kill." Conflict Management and Peace Science 25(3): 244-263.
[39] Asal, V. and R. K. Rethemeyer (2008). "The Nature of the Beast: Terrorist Organizational Characteristics and Organizational Lethality." Journal of Politics 70(2): 437-449.
[40] Gurr, T. R. (2000). People vs. States. Washington D.C, United States Institute of Peace.
[41] Collier, P. (2002). Greed and Grievance in Civil Wars. Working Paper Series In 2002-01. Oxford, Centre for the Study of African Economies.
[42] Fearon, J. and D. Laitin. (2003). "Ethnicity, Insurgency, and Civil War Replication data." Retrieved April 3 2004, 2004, from http://www.stanford.edu/~jfearon/.
[43] Regan, P. M. and D. Norton (2005). "Greed, Grievance, and Mobilization in Civil Wars." Journal of Conflict Resolution 49(3): 319-336.
[44] Wimmer, A. and B. Min (2006). "From Empire to Nation-State: Explaining Wars in the Modern World, 1816–2001." American Sociological Review 71(6): 867-897.
[45] Birnir, J. and J. Wilkenfeld (2007). Minorities at Risk: Addressing Selection Bias Issues and Group Inclusion Criteria for Ethnopolitical Research. University of Maryland, National Science Foundation.
[46] Shellman, S. and B. Stewart (2007). "Predicting Risk Factors Associated with Forced Migration: An Early Warning Model of Haitian Flight." Civil Wars 9(2): 174-199.
[47] Sambanis, N. (2004). "Poverty and the Organization of Political Violence." Brookings Trade Forum 2004: 165-211.
[48] Keck, M. E. and K. Sikkink (1999). Transnational advocacy networks in international and regional politics. International Social Science Journal, Blackwell Publishing Limited. 51: 89.
[49] Maoz, Z. (2001). "Democratic Networks: Connecting National, Dyadic, and Systemic Levels of Analysis in the Study of Democracy and War." War in a Changing World: 143–182.
[50] Maoz, Z., R. D. Kuperman, et al. (2006). "Structural Equivalence and International Conflict: A Social Networks Analysis." Journal of Conflict Resolution 50(5): 664.
[51] Maoz, Z. (2006). "Network Polarization, Network Interdependence, and International Conflict, 1816-2002." Journal of Peace Research 43(4): 391.
[52] Asal, V. and R. K. Rethemeyer (2008). "Dilettantes, Ideologues, and the Weak: Terrorists Who Don't Kill." Conflict Management and Peace Science 25(3): 244-263.
[53] Asal, V. and R. K. Rethemeyer (2008). "The Nature of the Beast: Terrorist Organizational Characteristics and Organizational Lethality." Journal of Politics 70(2): 437-449.
[54] Maoz, Z. (2006). "Network Polarization, Network Interdependence, and International Conflict, 1816-2002." Journal of Peace Research 43(4): 391.

3.2. The Social Stability of Societies: An Anthropological View (Lawrence Kuznar)

Author and Organization: Lawrence A. Kuznar, National Security Innovations, Inc.
Contact Information: lkuznar@nsiteam.com

The features of a society that promote its stability become apparent when they fail and lead to instability and collapse. Evidence of social collapse, on large or small scales, is commonplace in the news today. Somalia is the world's most failed state.[1] Indigenous Peruvians riot and attack police over oil drilling leases.[2] The global economy is in a tailspin. Certainly, insight into the drivers of social stability or instability can be gleaned from these contemporary examples. However, instability does not emerge overnight, and longer term historical perspectives are necessary to understand the latent factors that often slowly accumulate and then result in dramatic social collapse. Anthropology is a field that provides this deep, historical perspective on social collapse and instability, and this chapter reviews insights that anthropologists have gained concerning why some societies fail and others prosper.

The most essential findings of anthropological analyses implicate (1) a combination of changes in the material support for a society (especially environmental support), and (2) the ways that natural patterns of human decision making subvert effective social responses to change.

This chapter begins with a review of basic anthropological theory concerning the support of social systems and a consideration of the role environment and natural resources often play in sustaining them. Then, a brief overview of theories of social collapse is presented, outlining the key factors that are involved in social collapse. The role of human decision making is reviewed to highlight how even obvious and solvable social problems often go unresolved. The factors identified by anthropologists are also compared to social stability factors identified elsewhere in this volume (see Hewitt, Rieger). Finally, these factors are illustrated with a brief example of social unrest in Iran in June and July 2009.

A Social Ontology

What constitutes a society or culture and what supports it? There is no clear consensus in the anthropological community, although repeated analyses consistently identify similar factors. The clearest and most consistent frameworks have elements of a cultural materialist paradigm, in which the most basic factors undergirding a society are material, upon which people construct social structures and systems of values.[3,4,5] From this point of view, societies come into being as people employ technologies and social organizations to harvest food and shelter from the environment, and to provide access to mates and security for their families (Figure 3.2.A-1). People in turn construct rituals and ideological systems that reinforce the social relationships and values that provide for their material, biological and security needs. Within this framework, there is room for feedback from the ideological and ritual levels back upon the social and technological, although the primary causal influence is from material factors to social to ideological factors. Despite the fact that many anthropologists prefer to begin their analyses of culture at an ideological level, the cultural materialist paradigm has proven to be an extremely robust analytical framework for which there are no major contradictory analyses.[6]

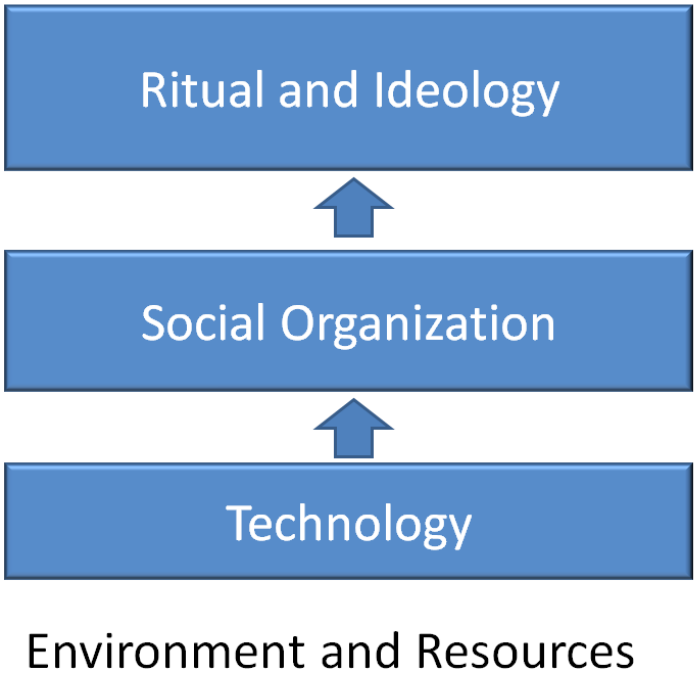


Figure 3.2.A-1. Simplified Cultural Materialist Ontology.

A key component of a cultural materialist paradigm is recognition that environmental factors often play a pivotal role in social stability; stable and productive environmental resources are a necessary component for a sustainably stable social system. A social system can experience instability in a permissive environment if social relations that promote trade and/or political stability are undermined (see Hewitt, Rieger, this volume). However, whenever there are major environmental changes that undermine the flow of goods and wealth in a society, social instability is guaranteed.

The cultural materialist paradigm is useful not only as a scientifically defensible view of society, but also because this paradigm explicitly identifies the factors that support a society; implicitly, the absence or failure of these factors are what theoretically cause social instability and collapse.

Theories of Social Collapse

Anthropological theories of social collapse can be found in the ancient writings of the Greek historian Herodotus, the medieval writings of the Arab scholar Ibn Khaldun, and in Edward Gibbon's 18th century *Rise and Fall of the Roman Empire*. Anthropologists have long held a fascination with studies on the rise and fall of civilization studies, and one landmark analysis is Tainter's comparative analysis of the collapse of Roman civilization, the Anasazi of the American southwest, and the ancient Maya.[7] Tainter's thesis is that as societies confront problems, such as diminishing agricultural yields, people invent higher levels of social complexity to solve the problems. These higher forms of complexity include social and economic differentiation, development of increasingly complex forms of symbolic communication, and the invention of non-productive bureaucrats to address problems. While these measures may help to solve social problems in the short run, their marginal costs become too great for a society to bear, resulting in the collapse of the social system as it existed.

The most recent comprehensive analysis of social collapse is Jared Diamond's *Collapse: How Societies Choose to Fail or Succeed*.[8] Diamond departs from Tainter's focus on marginal returns and focuses instead on environmental factors and failures in human decision making. Their theories are not necessarily mutually exclusive. Diamond's analysis is also comparative and focuses on the collapse of agricultural systems in Montana, the collapse of Easter Island society, the collapse of Mayan civilization, the failure of the Viking settlement of Greenland, and modern examples such as genocide in Rwanda and social instability in Haiti. Diamond proposes a five-point framework of factors that contribute to social collapse. These include the following:

1. Environmental damage
2. Climate change
3. Hostile neighbors who may attack a social system or sap its resources
4. Friendly trade partners who may turn hostile or indifferent, denying a social system the inputs required for its functioning (e.g., OPEC Oil crisis in the 1970s)
5. Failures in a society's response to these challenges

These factors are similar to those identified elsewhere in this volume. For instance, Hewitt's Political Instability Task Force (PITF) identifies high infant mortality (as a barometer of material well-being) and lack of integration in the world economy as key factors leading to social instability. Hewitt also finds that instability of neighboring states is a key factor leading to instability in a state. Similarly, in Gallup's model of instability (Rieger, this volume), basic material needs, living conditions, safety and security, and perception of economic conditions are key measures of social stability.

Each of Diamond's factors delineates key variables in a cultural materialist paradigm. The fifth factor, "failures in a society's response to these challenges," is always present in social collapse and is therefore a necessary condition. Typical reasons why people fail to respond to system-threatening challenges include the following:

1. Failure to anticipate a problem
2. Failure to perceive a problem
3. Failure to try to solve the problem
4. Failure of solutions when tried

Problems can be unanticipated simply due to their novelty. One hope for modern computational modeling methods is that they may help analysts anticipate the range of possible futures a society may face.[9,10,11] Absent these methods, it can be very difficult if not impossible for decision makers in a society to anticipate what problems they may face.

Problems may not be perceived. Many changes are gradual, creating noticeable changes only once one can compare conditions across many years. For instance, such "landscape amnesia" is apparent when people fail to notice receding snowlines due to global warming trends.[12]

If a problem is perceived, decision makers in a society can fail to solve the problem for a wide range of reasons. One common reason is a clash of interests among decision makers. Decision makers' goals may conflict, and each decision maker may very well make well-informed, rational decisions to attain their own selfish goals while at the same time pursuing courses of action that harm the social system upon which they depend. Well-known examples include the ancient Mayan elite's reservation of prime agricultural land for prestige crops instead of food-producing crops,[13] and the common practice of producing large families to support individual farms, creating a population boom that threatens the entire agrarian basis of an economy.[14] The

"Tragedy of the Commons," in which individual users overexploit open common resources, is a classic example.[15,16] Another common problem in decision making includes conflicts between short-term gains and long-term discounted profits; if the expected value of exploiting a resource today is greater than its long-term expected value, then it is often overexploited. This tendency is not only seen in modern industries such as logging, as Diamond details, but is a pervasive feature of non-Western people in non-monetary economies, such as hunter-gatherers in Madagascar.[17]

Hewitt's (this volume) stability variable, political institution consistency, may tap into both Tainter and Diamond's focus on failed organizational decision making. A country with a mix of authoritarian and democratic institutions may very well experience difficulties coordinating decision making to address problems, or the country may even experience difficulties with perceiving the same problems across its political spectrum.

Another factor of importance in social stability is competition for prestige goods, which often creates highly unsustainable economic practices.[18] The distribution of status-conferring goods (income, house size, herd size, exotic items) is always uneven and is characterized by sharp differences between social classes.[19,20] These sharp differences create greed and envy as individuals struggle and take chances to attain higher social status.[21] This risk-taking behavior predicts political risk taking and social unrest,[22] including contemporary terrorism.[23] Ironically, it is often more affluent, upwardly mobile sectors of a society, like educated middle and upper middle classes), that more strongly feel these wealth differences.

In addition to these classic rational forms of decision failure, there are also forms of decision failure based on more irrational tendencies in human decision making (see also Rieger, this volume). For instance, people typically ignore "sunk costs"; people tend to stay with a failing investment because they have already sunk resources into it. This not only applies to investment behavior on Wall Street, but was also a contributing factor to the collapse of Anasazi civilization as populations remained near large religious structures long after the surrounding environment could no longer support their populations.[24]

These features of social collapse and faulty decision-making are not exhaustive. However, they exemplify key elements researchers have identified in the failure of social systems. Examination of the current social unrest in Iran serves as an example of how these features may manifest in social instability, and possibly collapse.

Iran 2009: A Case Study of Social Instability

Iran appears to be currently in a state of political flux, with dissention among the ranks of the influential Shi'ite clergy, apparent tension between the supreme leader Ayatollah Khamenei and the president Mahmood Ahmadinejad,[25] and a vocal opposition, led by Mir-Houssein Mousavi that is currently protesting the June 13, 2009 election results. It would be premature to declare Iran collapsed, but there is clearly social instability that threatens state power holders at this time. (As of this writing, 9 July 2009, the situation is extremely fluid, alternating between government crackdowns and repeated public protests. If Iran possesses the state capacity to suppress dissent, then it can weather the current crisis. What is clear is that Iran is currently unstable.) Headlines highlight power struggles between clergy and secular leaders, between incumbents and challengers, and between street demonstrations that are occasionally met with violent repression. However, closer examination of the underpinnings of this crisis provides a test of Tainter's and Diamond's propositions, and it exposes instabilities in Iran's social

infrastructure. Not only does this analysis provide a deeper understanding of the Iranian crisis, but it also suggests leading indicators of social stability that can help forecast similar social breakdowns elsewhere.

Iran's Environmental Infrastructure

Iran possesses an abundance of natural resources in fossil fuels, and oil and gas revenues constitute 80% of export earnings and 50–70% of government revenues.[26] This provides the Iranian government with a major source of material support, although that support is concentrated and therefore vulnerable to world market fluctuations. Environmental changes in Iran have a small effect on the overall economic picture, but they have a dramatic effect on small but key segments of the Iranian population. In particular, overgrazing and climate change are causing water and pasture scarcity that has been undermining the traditional economy of many of Iran's traditional people, including minorities that are often hostile to the central state government. These environmental changes have led to migration to urban centers and the growth of impoverished segments in Iranian society.[27] Therefore, Iran's environmental and resource base, while in some ways strong, has clear vulnerabilities.

Iran's Political Economy

The Iranian economy and its political system are inextricably linked. In order to maintain Iran's social services and energy subsidies to the Iranian populace, the government controls much of the Iranian economy. Approximately 60% of the economy is controlled by a complex central state bureaucracy, with perhaps as much as another 20% of the economy controlled by Bonyads, religious foundations connected to the politically powerful Shi'ite religious hierarchy.[28] Both Iran's political system and its Shi'ite religious hierarchy are characterized by extremely complex, interlocking associations of politicians, elected officials, appointees, and clerics with often conflicting goals. The mix of authoritarian and democratic institutions, along with the overall complexity of the Iranian political system, may represent the sort of inconsistency Hewitt associates with instability (this volume). (Note that Hewitt codes Iran's political consistency as high; the suggestion is that this coding may be reconsidered.) Despite the government's domination of the economy and recent historically high fuel prices on the world market, it still lacks the resources necessary to provide services and maintain its military spending. Recently new taxes have been levied on the population in order to boost revenues. These new taxes have even met with public resistance and rioting.[29]

Iranian Demography: A Ticking Time Bomb

Iran has a very young population, with two thirds under the age of 30. The population is, overall and especially in urban areas, well educated; the literacy rate is 77%.[30] University attendance is relatively high, and women now represent more than half of all university students. A large, urbanized young sector is a source of potential disruption in any society.[31] When that age imbalance is matched with lack of opportunity, the combination can be explosive. Iran's official unemployment rate is currently 12.5%,[32] although Iran typically ignores both young and old (above age 55) and women who are seeking employment in its unemployment figures. Re-estimates of unemployment from the early 2000s indicate that the unemployment rate was 34% among 15–24 year olds and projected to reach over 50%.[33]

Anthropological Theories of Collapse and Iranian Social Instability

As Diamond notes, some form of environmental hazard often lies at the root of social instability. In the case of Iran, it is a deceptively strong economic base in fossil fuels that is actually highly vulnerable and that provides little employment to the population. On a smaller scale, environmental desiccation is displacing potentially disruptive minority populations, creating a potential pool of dissent in Iranian society. Demographically, Iran has a very young, well-educated middle class population whose members have high aspirations for social advancement, but those aspirations are thwarted by unemployment and lack of opportunity. Despite rhetoric concerning clerical divisions and secular power-grabs, the demography of Iranian dissent suggests more material and mundane sources of support for challenging the current government.

Mousavi and his clerical supporters are known as reformist candidates and are popular with the large segment of Iran's youth, especially students. This is precisely the segment of Iranian society most failed by the current administration and the Supreme leader's religious following. In addition, preliminary indicators suggest that middle class supporters of Mousavi were a major source of protest,[12] once again reinforcing the fact that those whose aspirations have been most thwarted are the ones taking to the streets. This is consistent with the type of relative deprivation that emerges from wealth differentials and thwarted aspirations. Many Iranians clearly have lost confidence in the current regime, satisfying one of Gallup's key indicators of instability. Furthermore, this instability is manifest in violence in the aftermath of a clear "matchstick event," an election (see Rieger, this volume). It is also apparent that Mousavi's political support in this crisis comes from groups advocating democratic reform, whereas Ahmadinejad's political support comes more from the autocratic elements, reinforcing Hewitt's research on regime consistency.

The Iranian government's response to their vulnerable economy has been to develop an extremely complex and invasive system of government control. This is entirely consistent with Tainter's proposition that societies use social complexity to solve problems. Furthermore, the failure of these government efforts to sustain themselves perfectly fits Tainter's prediction that such efforts provide diminishing marginal returns, further sapping the material basis for that society.

As Diamond stresses, underlying factors of instability alone do not destabilize a society. It is failures in the social response to these challenges that create instability, and in extreme cases, collapse. Previous crackdowns and censorship of social networking sites and the Internet indicate that the Iranian government at least perceived, and perhaps anticipated the growing dissent in their society. They also clearly attempted to solve the growing problem of dissent through censorship and the maintenance of low domestic fuel prices as a palliative to the masses. However, the Iranian government's solutions have clearly failed to address the material underpinnings of instability; they have not only continued their inefficient control of the economy, but they have exacerbated the problem by levying more taxes on the population. This fifth failure of decision making is the key ingredient in any social collapse.

How the current protest and instability in Iran will play out is difficult to predict. So far, the underlying factors and subsequent failures in decision making follow the hallmarks of social instability noted in anthropological theories of social collapse. However, collapse can only come about if dissenting elements can martial the power necessary to transform or destroy the system. The supreme leader and president's control of the Revolutionary Guard, al Quds forces, and the

Basij militias, will make any overthrow of the Iranian government difficult unless these centers of force defect to the dissenting parties. However, should the Iranian government continue its singular dependence on one volatile commodity and its ineffective control of the overall economy, the underlying conditions that lead to dissent and instability will only grow.

Summary Analysis: Indicators of Social Instability

The Iranian crisis is a test-case in process, but it illustrates several key points drawn from anthropological theories of social collapse. These key variables also serve as indicators of a society that is moving toward greater instability and therefore may aid in analysis and planning efforts designed to anticipate where U.S. resources may need to be applied. These key factors are detailed below.

Weakening economy

Economies can become weakened due to environmental changes, especially in Third World countries where the agrarian economy is important. Recent attention to social unrest in Sudan and global climate change,[34] and Department of Defense projections of the impact of global climate change on U.S. forces[35] are good examples of these concerns. Economies can also be threatened by a lack of diversification and a focus on primary export commodities, such as oil.[36] Finally, social investments in bureaucracy and social control designed to address problems, but which are inefficient and provide diminishing marginal returns, sap the material support of a society and hasten its demise. Iran exhibits these challenges.

High unemployment among young and middle class

Young people require new resources to sustain themselves, attract mates and get married, and subsequently support their families. Whenever a society has a large proportion of young people with such demands, the social system is pressed to acquire the new resources necessary to keep the young satisfied. When the young sector of society has been provided with good education, training and skills development, but without the possibility of realizing their aspirations to use these skills, then a large, passionate, and vocal pool of dissent is created in the society. The casebook example of this is the growth of the Italian red Brigades in the 1970s as increased educational opportunities plus economic stagnation combined to create a large pool of disaffected and eventually radicalized young people in Italy.[37] The demographic of Iranian dissent today indicates that similar challenges exist for the Iranian system.

Unprecedented taxes needed to support system

Both Tainter's proposition about diminishing returns and Diamond's focus on faulty decision making are manifest when states levy taxes on dissenters in order to try to support a failing system. A classic example of such folly would be the British stamp acts in Colonial America that, while not burdensome, fueled revolutionary fervor. The fact that increased taxes were met with riots in Iran indicates that these measures to increase state revenues to address social problems will likely fail.

Poor decision making

The advantage of Diamond's framework is that it identifies the specific ways that decision making can fail. Many of these are evident in this brief analysis of Iranian social instability. Whenever decision makers appear ignorant of obvious problems, when they persist in failed policies that drain resources from the state and the population, and when their policies

exacerbate rather than resolve problems, the social system as a whole is clearly at risk. At the current time, the hard-line taken by the Iranian supreme leader and president, along with the continued inefficiencies of the state-controlled economy, indicate that their solutions are failing.

Policy Implications of Anthropological Theories of Collapse

Anthropological theories of social collapse can help analysts to understand social change and to anticipate it by monitoring leading indicators. The practical policy implications of this research concern which indicators analysts should monitor, on which time scales, and which can be effectively impacted.

The first class of indicators includes environmental variables. These are often overlooked in policy analysis, but they can fundamentally alter the physical and consequent political environment. Political factions that can be blamed for failing to respond to disruptive environmental changes will be weakened to the benefit of their rivals. Key environmental variables impact economic resources, and include desertification, salinization of soils, flooding, major storm activity, and any other factor that impacts a society's economy. The time scale on which these variables operate is over years if not decades, and therefore requires a long-term commitment to data gathering and analysis. The best use of these data is forecasting of environmental trends and their impacts; there is little to nothing that can be done about environmental change or the occurrence of a devastating hurricane.

The second class of indicators includes economic variables. Simple changes in aggregate measures like GDP and poverty rates are of limited use. More detailed information on shifts in the distribution of wealth and productivity in a society, however, can reveal eventual political rifts in a society, before these rifts become politicized. Key economic variables would include income, land ownership, investment, employment, occupational specialties and primary exports (for example, oil). Economic variables operate over years, and the effects of dramatic changes can have political impacts in months. One policy implication is that economic variables need to be monitored at the nation level, province level, and faction level (who controls what resources, what relative wealth differences exist between factions), on a long-term basis. Impacting economic trends is difficult and unpredictable, but at least these human-created factors have the potential of being altered, and analysts must stay current on human social cultural and behavioral theories that indicate what effects could be anticipated from altering economic variables. Training on current social theory must be better incorporated into government business practices.

A third class of indicators includes how decisions are made in a society. Data relevant to decision making includes structural data on social organization and decision flow (how governments, businesses and other organizations are organized, gather information and act upon it). These kinds of data are best gathered through ethnographic field research, although email records and other communications can reveal social structure through social network analytic methods. It is important to recognize that, although electronic communication has become widespread in industrialized societies, face-to-face and word-of-mouth communication continues to be fundamentally important in most corners of the world. Ethnographic methods require months of personal observation. One policy implication is that more support must be made for ethnographic research in key contexts that may impact national security. This can be done in an open manner, consistent with American Anthropological Association ethics codes, provided that subjects of the research understand that their participation in a study is designed to

reveal how their organizations operate, and research findings can be made to the subjects for their own use. Social organizations tend to change slowly, however, crises in decision making can precipitate rapid change in organizations; consider the rapid failure of large corporations due to scandals such as ENRON. Therefore, studies of social organization and decision making must be periodically updated, and also conducted on an ongoing basis. Behavioral economics and evolutionary psychology provides powerful theoretical tools for understanding how decision making will change in different social contexts. Training in these fields should be made available to analysts in order to bring these insights to bear on real situations.

Conclusions

Anthropologists' long-term view of social processes and unique databases (sometimes spanning thousands of years of social evolution), provide them with a unique viewpoint from which to understand social collapse. Societies collapse (their fundamental institutions cease to function to hold the social fabric together, basic human needs are not met, violence erupts) due to long-term erosions of, or effective challenges to, their economic base (often due to environmental change). When change comes, it is often rapid and dramatic. The key to anticipating these rapid changes, and possibly for mitigating them, is to identify the indicators of social erosion long before rapid changes erupt. These indicators include environmental, economic, and decision making variables. Much of this data is openly accessible, and powerful theories are under constant development to explain the drivers of collapse. However, adequately monitoring these variables requires a commitment to long-term research by governments, NGOs, and the academic community, backed by training in relevant areas of research.

↑ Back to the top

---

[1] Rice, Susan E. and Stewart Patrick, 2008. *Index of State Weakness in the Developing World.* Brookings Institute, Washington D.C. http://www.globalpolicy.org/images/pdfs/02weakness.pdf

[2] "Many Missing" after Peru Riots. http://news.bbc.co.uk/2/hi/americas/8088350.stm June 8, 2009.

[3] Harris, M. (1979). *Cultural Materialism: The Struggle for a Science of Culture*. New York: Vintage Books.

[4] Steward, J. (1955). *Theory of Culture Change: The Methodology of Multilinear Evolution.* Urbana, Illinois: University of Illinois Press.

[5] Wilson, D. J. (1999). *Indigenous South Americans of the Past and Present: An Ecological Perspective.* Boulder, Colorado: Westview Press.

[6] Kuznar, L. A., & Sanderson, S. K. (Eds.). (2006). *Studying Societies and Cultures: Marvin Harris's Cultural Materialism and Its Legacy*. Boulder, Colorado: Pergamon Press.

[7] Tainter, J. (1988). *The Collapse of Complex Societies.* Cambridge: Cambridge University Press.

[8] Diamond, J. (2005). *Collapse: How Societies Choose to Fail or Succeed.* New Yok: Penguin Books.

[9] Bankes, S. (1993). "Exploratory Modeling for Policy Analysis." *Operations Research*, 41(3), 435-449.

[10] Kuznar, L. A. (2006). "Exploratory Modeling of Ensembles for Testing Decision Theory Paradigms." In D. L. Sallach, C. M. Macal & M. J. North (Eds*.), Proceedings of the AGENT 2006 Conference: Social Agents: Results and Prospects* (pp. 157-166). Chicago, Illinois: Argonne National Laboratory.

[11] Zacharias, G. L., MacMillan, J., & Hemel, S. B. V. (Eds.). (2008). *Behavioral Modeling and Simulation: From Individuals to Societies.* Washington, D.C.: National Academies Press.

[12] Diamond, J. (2005). *Collapse: How Societies Choose to Fail or Succeed.* New York: Penguin Books.

[13] Webster, D., Freter, A., & Golin, N. (2000). *Copán: The Rise and Fall of an Ancient Maya Kingdom.* Fort Worth, Texas: Harcourt Brace.

[14] Handwerker, W. P. (1986). "The Modern Demographic Transition: An Analysis of Subsistence Choices and Reproductive Consequences." *American Anthropologist*, 88, 400-417.
[15] Alvard, M. S. (1998). "Evolutionary Ecology and Resource Conservation." *Evolutionary Anthropology*, 7, 62-74.
[16] Hardin, G. (1993). *Living within Limits: Ecology, Economics, and Population Taboos.* Oxford: Oxford University Press.
[17] Tucker, B. (2006). "A Future Discounting Explanation for the Persistence of a Mixed Foraging-Horticulture Strategy among the Mikea of Madagascar." In D. J. Kennett & B. Winterhalder (Eds.), *Behavioral Ecology and the Transition to Agriculture* (pp. 22-40). Berkeley: University of California Press.
[18] Diamond, J. (2005). *Collapse: How Societies Choose to Fail or Succeed.* New York: Penguin Books.
[19] Kuznar, L. A. (2001). "Risk Sensitivity and Value among Andean Pastoralists: Measures, Models and Empirical Tests." *Current Anthropology*, 42(3), 432-440.
[20] Kuznar, L. A. (2002). "Evolutionary Applications of Risk Sensitivity Models to Socially Stratified Species: Comparison of Sigmoid, Concave and Linear Functions." *Evolution and Human Behavior,* 23(4), 265-280.
[21] Friedman, M., & Savage, L. J. (1948). "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 4, 279-304.
[22] Kuznar, L. A., & Frederick, W. G. (2003). "Environmental Constraints and Sigmoid Utility: Implications for Value, Risk Sensitivity, and Social Status." *Ecological Economics*, 46, 293-306.
[23] Kuznar, L. A. (2007). "Rationality Wars and the War on Terror: Explaining Terrorism and Social Unrest." *American Anthropologist*, 109(2), 318-329.
[24] Janssen, M. A., Kohler, T. A., & Scheffer, M. (2003). "Sunk-Cost Effects and Vulnerability to Collapse in Ancient Societies." *Current Anthropology*, 44(5), 722-728.
[25] Fathi, Nazila (2008). "Ahmadinejad Loses Favor with Khamenei, Iran's Top Leader." *NY Times Online*, January 7, 2009 http://www.nytimes.com/2008/01/07/world/africa/07iht-tehran.4.9063096.html
[26] *CIA World Factbook*, Iran https://www.cia.gov/library/publications/the-world-factbook/geos/IR.html
[27] Beck, L. (1998). "Use of Land by Nomadic Pastoralists in Iran, 1970-1998." *Bulletin of the Yale School of Forestry and Environmental Science*, 103, 58-80.
[28] http://en.wikipedia.org/wiki/Iran_economy
[29] *CIA World Factbook*, Iran https://www.cia.gov/library/publications/the-world-factbook/geos/IR.html
[30] *CIA World Factbook*, Iran https://www.cia.gov/library/publications/the-world-factbook/geos/IR.html
[31] Hudson, V. M., & Den Boer, A. (2002). "A Surplus of Men, a Deficit of Peace." *International Security,* 26(4), 5-38.
[32] *CIA World Factbook,* Iran https://www.cia.gov/library/publications/the-world-factbook/geos/IR.html
[33] Amuzegar, Jahangir (2004). "Iran's Unemployment Crisis." *Middle East Economic Survey* 47(41) http://www.mees.com/postedarticles/oped/a47n41d01.htm
[34] *United Nations Environment Programme*, 2007
[35] Goodman, S. (2007). "National Security and the Threat of Climate Change." Unpublished manuscript, Center for Naval Analyses (CNA) Corporation.
[36] Collier, P., & Hoeffler, A. (2004). "Greed and Grievance in Civil War." *Oxford Economic Papers*, 56, 563-595.
[37] Gurr, T. R. (1990). "Terrorism in Democracies: Its Social and Political Bases." In W. Reich (Ed.), *Origins of Terrorism: Psychologies, Ideologies, Theologies, States of Mind* (pp. 86-102). Baltimore: Johns Hopkins University Press.

3.3. Quantitative Content Analysis (Laura Leets)

Author and Organization: Laura Leets, NSI and Georgetown University
Contact Information: leets@nsiteam.com

Abstract

This white paper provides a readable overview of content analysis, an analytical technique that examines message content. Specifically, it focuses on the central concepts underlying content analysis and how to conduct effective content analysis research.

Introduction

Content analysis is one of a variety of social science methods employed by researchers to find answers to questions. Simply, it is a technique for gathering and analyzing the content of text. The content refers to anything written, visual, or spoken that serves as a medium for communication (e.g., words, pictures, symbols). It is a technique with a long history and one that is used in many fields. In particular, there are two general forms of content analysis: *quantitative* and *qualitative*. The quantitative approach is normally referred to as basic content analysis, and the qualitative approach is usually called discourse analysis or textual analysis.[1] The emphasis of this white paper is on the quantitative approach.

A classic quantitative definition of content analysis is "a research technique for the *objective, systematic, and quantitative* description of the manifest content of communication."[2] It is important to understand the words objective, systematic and quantitative. As Stempel notes, in this approach the content is (a) selected according to explicit and consistently applied rules (systematic), (b) not influenced by the researcher's personal bias (objective), and is (c) quantified, that is, numerical values are recorded and analyzed (quantitative).[3] Historically, content analysis focused on the manifest (surface) meaning of a message, as opposed to the latent (underlying) meaning. Today, most researchers recognize the value of both analyses and often advocate the examination of both manifest and latent meanings if possible.

There are numerous reasons for conducting a content analysis; ten common reasons are provided in the table below.

Table 3.3-1. Reasons to Conduct a Content Analysis[4]

1. Describe trends in content over time	6. Trace conceptual development in intellectual history
2. Describe the relative focus of attention for a set of topics	7. Compare *actual* content with *intended* content
3. Compare international differences in content	8. Expose use of biased terms in propaganda research
4. Compare group differences in content	9. Test hypotheses about cultural and symbolic use of terms
5. Compare individual differences in communication style	10. Code open-ended survey items

The following section presents a methodology to conduct a basic content analysis.

How to Conduct Content Analysis Research

Specifically, content analysis involves (1) selecting and defining a set of content categories, (2) defining and then sampling the elements of the text that are described by the categories, (3) quantifying the categories, such as by counting their frequency of occurrence, and (4) drawing conclusions from the findings. These steps are clearly captured in Neuendorf's (2002)[5] flowchart of a typical content analysis process, which is adapted below. In the remainder of the section, the flowchart is discussed step by step.

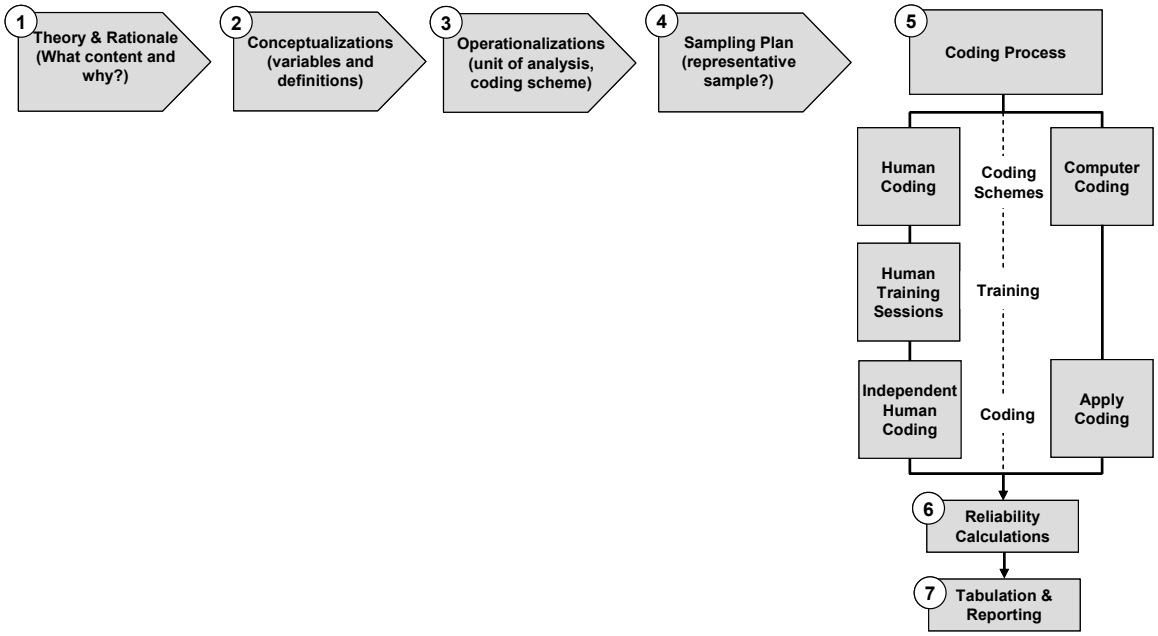


Figure 3.3-1. Content Analysis Flowchart.[6]

Step 1 Theory and Rationale

Content analysis, like all research, begins with a research problem: "what content will be examined and why?"[7] A sound content analysis states explicitly a rationale that supports why certain messages are worth examining. The research question should be based on existing theory, prior research, or practical problems.

For example, decades of communication research on television violence demonstrated that exposure to violence in mass media led to (1) learning of aggressive attitudes and behavior, (2) fear, and (3) desensitization to violence. As a result, the National Cable Television Association funded a national television violence study in which academics from four universities conducted a three-year content analysis to describe the violence on broadcast television. This allowed the industry to monitor the issue before Congress passed legislation that would force them to do so.

Step 2 Conceptualizations

It is crucial that the researcher defines the concepts or variables of interest (conceptualization) and develops specific methods for observing the concepts in relation to that framework (operationalization).

There are two main approaches to defining the meaning of communication content. One approach is to define the content by its surface or visible meaning (manifest content). For example, a researcher may count the number of times a weapon appears in a television show. The meaning of this symbol is on the surface, and what you see is what you get. Alternatively, the researcher can define the content by examining the underlying, implicit meaning of the content of a text (latent content). This requires reading between the lines and requires the coders to understand the relationship between language and social meaning. It is less accurate and harder than manifest coding. For example, assessing the political orientation of a newspaper editorial would constitute latent content.

In the national television violence study, the researchers defined their main topic of interest, violence, as the following: "Violence is defined as any overt depiction of a credible threat of physical force of the actual use of such force intended to physical harm an animate being or a group of animate beings. Violence also includes certain depictions of physically harmful consequences against an animate being (or group of beings) that occur as a result of unseen violent means. Thus, there are three primary types of violent depictions: credible threats, behavioral acts and harmful consequences."[8] It is clear that the coding of violence is a complex task and is usually regarded as latent content, requiring a long list of coding rules.

Step 3 Operationalizations

Variables in content analysis are operationalized using a coding system, which is a set of instructions or rules on how to systematically observe and record content from text. In creating a coding system, the researcher first selects and defines a *unit of analysis*. Units of analysis are what the coders actually count. Depending on the aim of the project, the unit of analysis may be newspapers, speeches, photographs, television violence, words, sentences, or something else entirely. The unit of analysis must be related to the purpose of the study. It is important that the unit of analysis is clear and precise, and the rules for inclusion are apparent and easily observable. For example, the national television violence study used a unit of analysis they termed "violent incident." Violent incidents were often indicated by the acronym "PAT," as they included the interaction between a *perpetrator* (P) and a *target* (T) involved in a violent *act* (A).

There are two approaches to coding data: *emergent* and a *priori* coding. With emergent coding, categories are established following some preliminary examination of the data. The researcher reviews the material and comes up with a set of features that form a coding scheme. If multiple researchers are involved, they compare notes and reconcile any differences. The second approach is *a priori* coding, where the categories are established prior to the analysis based upon some theory. The researcher(s) decide on the categories, and the coding is applied to the data.

Regardless of which approach is used, Holsti provides five general guidelines to follow in constructing a set of content analysis categories or a *category scheme*.[9] The first is that the categories should reflect the purposes of the research. The second rule for constructing a set of categories is that the categories be exhaustive or that there must be a category into which each relevant item can be placed. The third guideline states the content should only be capable of being placed in one category, which is known as mutually exclusive categories. The fourth rule is known as independence of categories. The assignment of one item to a given category should not affect the assignment of other items. The fifth category deals with the unit of analysis. One cannot mix different levels of analysis in the categories. Failing to follow these rules (especially

mutually exclusive and exhaustive categories) can result in a faulty content analysis because it compromises the study's reliability and validity.

The national television violence study provided an *a priori* category scheme for their unit of analysis, violent incident. One of the several variables included in this scheme was titled "reason for violence." For any given violent incident addressed by the coders, a determination was made of which reason for violence applied most clearly, choosing between a set of ten potential coding values. Other variables related to the unit of analysis were similarly coded. The reason-for-violence potential values are included in the table below; note that they adhere well to Holsti's guidelines.

Table 3.3-2. Reason for Violence.	
1. Self Defense	Violence intended to protect the self against harm
2. Protect Others	Violence intended to protect other characters
3. Ideology/Belief	Violence accomplished to promote or adhere to an ideology
4. Anger	Violence triggered by feelings of hostility, rage, resentment
5. Retaliation	Violence triggered by feelings of retribution
6. Personal Gain	Violence for purposes of selfish gain (material goods, power)
7. Amusement	Violence for the sake of ones own enjoyment
8. Mental Incompetence	Violence from a person who lacks mental competence
9. Forced Compliance	Violence as a result of orders from more powerful others
10. Accident	Violence occurred even though perpetrator didn't intent for it to

Step 4 Sampling Plan

Often there is a large amount of relevant content, and with limited time and funding constraints, it is rare to analyze all the potential content. As a result, the researcher almost always selects an appropriate sample from the population. To provide useful descriptions of the content, the sample content must be selected so as to represent the entire population, which requires using precise statistical techniques. There are a number of science-based probability sampling techniques that the researcher can use,[10] which are no different from the sampling procedures used in survey research.

The national television violence study examined the largest and most representative sample of television content ever evaluated in a single study. For three years, programs on 23 television channels were randomly selected over a nine-month period to create a composite week of programming.[11]

Step 5 Coding Process

As Babbie notes, "content analysis is essentially a coding operation. Coding is the process of transforming raw data into a standardized form."[12] This is usually the most time-consuming part of a content analysis.

Researchers can choose to use either manual coding (humans) or automatic coding (computers). If the study employs humans, the people who do the coding are called coders. The number of coders involved in a content analysis is usually small (two to three). Their job is to place a unit of analysis into a content category. The researcher needs to provide definitions and coding rules for the coders to use in assigning units to categories. This is provided in a codebook, in which all the variable measures are fully explained. Coders participate in training sessions to increase their accuracy and to become calibrated with each other. The coders as a group code sample material and discuss the results afterward. When the coders become consistent and comfortable with the coding task, they code the data independent of one another.

Some researchers choose to save time and select automated coding for the repetitive tasks (e.g., searching, listing, and counting). There are many different software packages available that help facilitate coding. The researcher must provide a codebook or data dictionary that defines all the variables and the method for applying them. Although computers can aid in identifying content, there is often a problem with contextual bias or computers misidentifying irrelevant content. As Krippendorf points out, it is important to determine what humans and computers do best. *Computers do not perform the analysis, but they aid it*.[13]

Regardless of whether human or computer coding is employed, the overall value of the content analysis depends on coding accuracy or reliability, which is discussed below.

Step 6 Reliability Calculations

When the coding is finished, the researcher establishes the reliability of the coding. Reliability helps the researcher explain how the data were collected in order to justify the authenticity and accuracy of the coding. The reliability of a content analysis is discussed with regard to two dimensions. The first is called stability or *intra-rater reliability*. This refers to the coder's ability to consistently re-code the same data in the same way over a period of time. The second is titled reproducibility or *inter-rater reliability*. This refers to whether the coding-schemes enable the same text to be coded in the same category by different people. As Weber (1990) points out, "reliability problems usually grow out of the ambiguity of word meanings, category definitions, or other coding rules."[14]

Researchers establish reliability among coders through statistical analysis and a statistical coefficient that determines the degree of consistency among coders (e.g., Scott's pi or Cohen's Kappa). A general rule of thumb is that reliability estimates above .80 demonstrate good reliability among coders, estimates between .70–.80 illustrate moderate reliability, estimates between .60–.70 show fair reliability, and coefficients below .60 are unacceptable. For example, in the national television violence study, the overall reliability estimates were very high (e.g., .90).

Step 7 Tabulation & Reporting

It is important to remember that content analysis *describes* what is in the text. It cannot demonstrate effects or draw causal inferences (x causes y). Validation of the inferences made from content analysis usually demands the use of multiple sources of information and may require the use of benchmark data to know if a value is high or low. Content analysis itself is only valid and meaningful to the extent that the results are related to other measures.[15] Hence, analysts should be cautious about the statements that are drawn from a content analysis and should verify that they are appropriate for the analysis.

For example, the national television violence study provided descriptive data about violent content on broadcast television. The first year provided a benchmark from which to compare the following years. Moreover, the researchers were aware of the appropriate use of content analysis, and they knew that a discussion of the potential effects of the violent content on viewers would require a different method—an experiment.

Conclusion

The applications of content analysis are extensive. For example, with regard to counterinsurgency operations, there are many interesting questions that can be answered by conducting a quantitative content analysis. Successful operations depend on thoroughly understanding the society and culture within which they are being conducted.[16] One way to examine the society is to content analyze messages (television, radio, paper, speeches) in order to describe (1) relationships and tensions among groups, (2) ideologies and narratives that resonate with groups, (3) values of groups, interests and motivations, and (4) how a particular culture defines effective and legitimate governance. Moreover, intelligence and counter-terrorism agencies absorb massive amounts of information, and content analysis provides a mechanism for converting it to actionable data.

This white paper was written to help readers determine when to use a content analysis and to understand what steps are involved in designing an effective study.

↑ Back to the top

---

[1] Krippendorf, K. (2004). *Content Analysis: An Introduction to its Methodology.* 2nd ed. Thousand Oaks, CA: Sage Publications.
[2] Berelson, B. (1952). *Content Analysis in Communication Research.* New York, NY: Free Press.
[3] Stempel, G. (1989). "Content Analysis." In G. Stempel & B. Westley (Eds.) *Research Methods in Mass Communication.* Prentice Hall.
[4] Berelson, B. (1952). *Content Analysis in Communication Research*. New York, NY: Free Press.
[5] Neuendorf, K. (2002). *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.
[6] Neuendorf, K. (2002). *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.
[7] Neuendorf, K. (2002). *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.
[8] Seawell, M. (Ed.) (1998). *National Television Violence Study* (Vol. 3). Thousand Oaks, CA: Sage Publications.
[9] Holsti, O. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.
[10] Babbie, E. (2004). *The Practice of Social Research.* 10th ed. Belmont, CA: Thomson Wadsworth.
[11] Seawell, M. (Ed.) (1998). *National Television Violence Study* (Vol. 3). Thousand Oaks, CA: Sage Publications.
[12] Babbie, E. (2004). The Practice of Social Research. 10th ed. Belmont, CA: Thomson Wadsworth.
[13] Krippendorf, K. (2004). *Content Analysis: An Introduction to its Methodology.* 2nd ed. Thousand Oaks, CA: Sage Publications.
[14] Weber, R. (1990). *Basic Content Analysis* 2nd ed. Thousand Oaks, CA: Sage Publications.
[15] Shapiro, G. & Markoff, J. (1997). "A Matter of Definition." In C. Roberts (Ed.). *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Text and Transcripts.* Mahwah, NJ: Lawrence Erlbaum Associates.
[16] *The Counterinsurgency Field Manual*, 2007.

3.4. The Peace and Conflict Instability Ledger Ranking States on Future Risks (J. Joseph Hewitt)

Author and Organization: J. Joseph Hewitt, Center for International Development and Conflict Management, University of Maryland
Contact Information: jhewitt@cidcm.umd.edu

Introduction

Over the past two years, the risks of instability and conflict have increased significantly in the regions of the world where these dangers were already very high. This is one of the most important conclusions to be drawn from the most recent analyses that produce the 2010 *Peace and Conflict Instability Ledger*. The heightened risks are not the result of worsening government effectiveness in delivering services to the population or deteriorating economies. The heightened estimated risks are associated with a development that is normally welcomed—the initial steps toward democratic governance. Additionally, since the last publication of rankings, new armed conflicts have outnumbered those that have terminated, driving up estimated risks of instability in many regions of the world where neighborhood security has now worsened. This chapter presents country rankings based on newly calculated risk estimates and discusses some of the key results from the analysis, including the pivotal relationship between democratization and risk of instability.

The *Peace and Conflict Instability Ledger* ("the ledger," for short) is a ranking of 162 countries based on their estimated risk of experiencing major bouts of political instability or armed conflict in the three-year period of 2008–2010. The estimates are obtained from a statistical forecasting model that uses 2007 data (the most current data available) for several variables that correlate strongly with the onset of political instability or armed conflict. The ledger represents a synthesis of some of the leading research on explaining and forecasting state instability. As such, the selection of factors accounted for in the ledger's underlying forecasting models was based on identifying variables for which agreement was strong among researchers about their relative importance. An abridged version of the full rankings appears at the end of the chapter. (The full listing is included with Hewitt [forthcoming][1]). Readers are encouraged to consult this information regularly while proceeding through this overview.

The Peace and Conflict Instability Ledger

Figure 3.4-1 shows how the countries in the analysis were classified according to their estimated risk scores. A quick review of the map offers a broad overview of what the geographic landscape looks like from the perspective of the risks of instability. Undoubtedly, Africa remains the most serious concern. More than half the countries on the continent qualify for the high or highest risk categories. Of all the countries worldwide that qualify in those categories, African countries make up more than 75 percent of the states (28 of the 36 total). A similar concentration of states qualifying at high or highest risk exists in South Asia, a grouping that contains crucial states like Pakistan (newly classified as high-risk) and Afghanistan, which are pivotal because their fates have direct repercussions for global trends in terrorism.

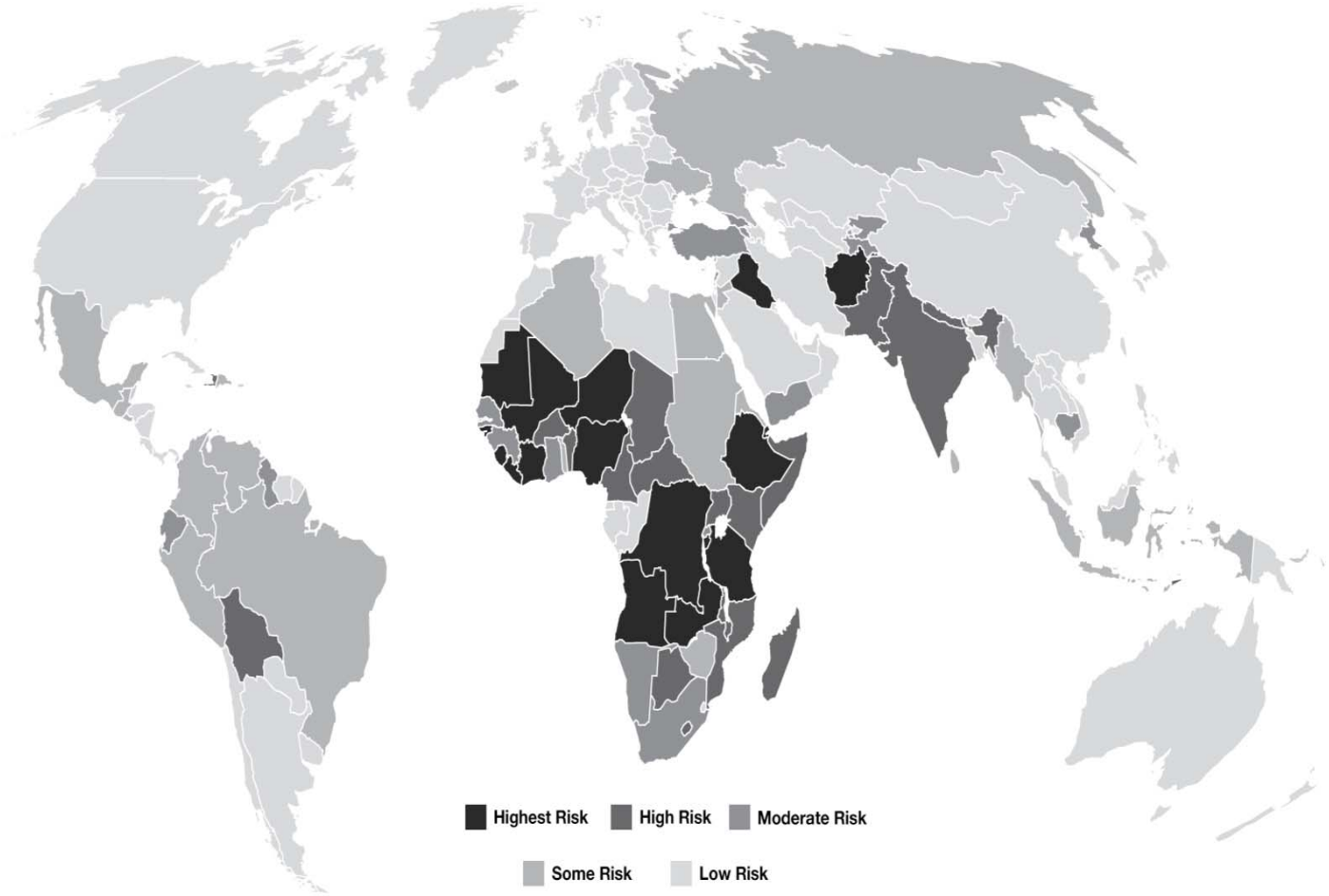


Figure 3.4-1. Risk of Future Instability, 2008-2010.

The ledger's conceptualization of political instability relies on the definition developed by the Political Instability Task Force (PITF). (The initial compilation of state failure events for the Task Force was conducted at CIDCM in 1994–1995 under the direction of Ted Robert Gurr. The roster of genocides and politicides was provided by Barbara Harff. The PITF presents full definitions for revolutionary wars, ethnic wars, adverse regime changes, genocide, and politicide in Esty et al. [1999]).[2] This definition, which has guided the task force's comprehensive compilation of state failure events covering the period 1955–2006, encompasses a wide variety of event types. These include revolutionary wars, ethnic wars, adverse regime changes, and genocides or politicides. The onset of any of these types of episodes for a state marks the beginning of an instability event. While this set of events is quite heterogeneous, they all share a fundamental similarity—the onset of any one of these events signals the arrival of a period in which government's capacity to deliver core services and to exercise meaningful authority has been disrupted, threatening its overall stability.

Empirical studies using 60 years of historical data show that instability can emerge from a combination of five factors in four domains of government and society. (Readers interested in some of the more significant recent contributions to this literature should consult Collier et al. [2003],[3] Collier and Hoeffler [2004],[4] Esty et al. [1999][5], Fearon and Laitin [2003],[6] Goldstone

et al. [2005],[7] Hegre and Sambanis [2006],[8] Hegre et al. [2001],[9] King and Zeng [2001],[10] Sambanis [2002, 2004],[11,12] and the United States Agency for International Development [2005][13]).The key factor in the political domain is the institutional consistency of a country's governmental institutions. In the economic domain, it is openness to international trade: the more interdependent a country's economy with others, the less likely a country will experience instability in the near future. In the societal domain, the infant mortality rate is a crucial indicator of socioeconomic well-being. And in the security domain there are two factors: one is the extent to which a country is militarized, and the other is whether neighboring countries have armed conflict. Table 4.4-1 provides a brief overview of the theoretical relationship between each of these factors and risks of instability. A fuller discussion is given in *Peace and Conflict* 2008.[14]

Leveraging the strong historical relationships that exist among the five factors and the risk of future instability, the ledger uses a statistical model to obtain risk scores for all countries having a population of at least 500,000 in 2007 (162 countries total). The data collection that serves as the foundation for this analysis contains an annual observation for each country for every year that data exist for the five factors. Each annual observation in the data collection records whether the country experienced an onset of a new instability event in any of the three years following the year of the observation. In this fashion, the data can be analyzed to assess the empirical relationship that the five factors have with the risk of future instability. To maintain comparability with the results presented in the previous volume of *Peace and Conflict*, we continue to estimate the model using data from 1950–2003. The logistic regression procedure for estimating the model on this data (sometimes called "training data") produces weights for each factor that reflect the relative influence that each has on explaining future instability. The previous ledger results,[15] which used 2004 data to produce forecasts for the period 2005–2007, were based on the same 1950–2003 training data. For the updated ledger, we now use 2007 data (the last year for which complete data are available for all five factors) to produce a three-year forecast indicating the risk of instability at any time during the period 2008–2010. It should be noted that in the absence of significant change to any of the five factors, risks change only gradually from year to year. Therefore, a high-risk country that experiences no major structural change to its regime, socioeconomic status, or security situation in the period 2008–2010 will likely remain at high risk beyond this forecast period.

Table 3.4-1. Factors Influencing the Risk of Instability.

Factor	***Domain***	***Description***
Institutional Consistency	Political	The ledger accounts for the impact of institutional consistency. This refers to the extent to which the institutions comprising a country's political system are uniformly and consistently autocratic or democratic. Political institutions with a mix of democratic and autocratic features are inconsistent, a common attribute of polities in the midst of a democratic transition. Based on a series of findings reported in the academic literature, we expect regimes with inconsistent institutions to be more likely to experience political instability.[16,17,18]
Economic Openness	Economic	The ledger accounts for the impact of economic openness, which is the extent to which a country's economy is integrated with the global economy. Countries that are more tightly connected to global markets have been found to experience less instability.[19,20]

Infant Mortality Rates	Economic, Social	The ledger examines the impact of infant mortality rates, an indicator that serves as a proxy for a country's overall economic development, its level of advancement in social welfare policy, and its capacity to deliver core services to the population. In this respect, this indicator taps into both the economic and social domains of a country. Research findings reported by the PITF have been especially notable for the strong relationship found between high infant mortality rates and the likelihood of future instability.[21,22]
Militarization	Security	To account for the security domain, the ledger focuses on a country's level of militarization. Instability is most likely in countries where the opportunities for armed conflict are greatest. In societies where the infrastructure and capital for organized armed conflict are more plentiful and accessible, the likelihood for civil conflict increases.[23] Extensive militarization in a country typically implies that a large portion of the society's population has military skill and training, weapons stocks are more widely available, and other pieces of military equipment are more diffused throughout the country. The likelihood of instability is greater in this setting because increased access to and availability of these resources multiplies the opportunities for organizing and mobilizing.
Neighborhood Security	Security	The likelihood of political instability in a state increases substantially when a neighboring state is currently experiencing armed conflict. This risk is especially acute when ethnic or other communal groups span across borders. A number of studies have shown that neighborhood conflict is a significant predictor of political instability.[24,25,26]

The full listing of all 162 countries is presented at the end of the chapter. The table includes an indication of how each country is performing on each of the risk factors, which enables a quick assessment of how the ultimate risk estimate relates to each indicator. In this fashion, the full ledger table serves as a diagnostic tool, offering comprehensive information about all countries so that comparisons can be drawn about how the levels of each factor influence risk.

To ease interpretation of the results, the ledger presents each country's likelihood of future instability as a risk ratio. The risk ratio gives the relative risk of instability in a country compared to the average estimated likelihood of instability for 28 members of the Organization for Economic Cooperation and Development (OECD). The OECD serves as a useful baseline because its membership is widely viewed to contain the most stable countries in the world. The estimated probability of the average OECD country's experiencing an instability event in the period 2008–2010 is 0.007. To illustrate, Nicaragua's estimated probability of experiencing instability in the next three years is 0.029, which yields a risk ratio of approximately 4.1. Presented in this way, the analysis indicates that Nicaragua's risk of instability is about four times greater than an average OECD country—a more useful characterization of its risk than the simple probability 0.029 by itself.

The risk ratios appearing in the ledger are statistical estimates and, accordingly, are accompanied by varying levels of confidence, depending on the particular attributes of a given country. An under-appreciated characteristic of statistical inferences is that they are always associated with some level of uncertainty. For instance, in the model used to create the ledger, infant mortality rates were found to be positively related to the onset of instability. The level of uncertainty for that estimate was sufficiently small to rule out the possibility that the model was pointing erroneously to a positive relationship when the "true" relationship was actually negative (or nonexistent). However, uncertainty around the estimate remains. The uncertainty exists because

many countries with high infant mortality rates have not experienced instability (e.g., Malawi, Saudi Arabia, or Bolivia) and some with a low rate do (e.g., Israel). These outlier states create "noise" in the estimated relationship between instability and infant mortality rates. Each of the variables in the model is accompanied by this kind of uncertainty or noise.

Information extracted from the statistical model for instability can be used to compute the total amount of uncertainty surrounding an individual country's estimate for instability risk. The ledger reports this level of uncertainty. For each country, the ledger reports a single best estimate of the overall risk of instability. Additionally, the ledger reports a range of values within which the best estimate lies. Statistically speaking, the "true" risk of instability lies within this range with a confidence probability of 95 percent. The graphical display of the confidence range shows how it extends across risk categories. For some countries, the confidence range is confined largely within one category. For others, large segments may extend across multiple categories, suggesting that assessments about the country's status should be drawn with more caution.

The updated ledger presented in this volume features an innovation that leverages information from the confidence intervals to inform the classification of countries into separate risk categories. In the previous ledger, the three-category classification scheme (high risk, moderate risk, and low risk) was based simply on the percentile ranking of states. This method has been improved by incorporating greater gradation into the set of risk categories and informing the classification of states based on statistical confidence. A brief comment is warranted about the new classification procedure to better appreciate its advantages.

With the list of 162 countries sorted from highest score to lowest, the country with the highest risk score (Afghanistan) establishes a new grouping of states. Moving down the list, any country whose upper bound on its confidence range is higher than Afghanistan's risk score is assigned to the same group as Afghanistan (Niger, Burundi, DRC, and Djibouti). For each of these countries, the extensive overlap across their respective confidence ranges indicates little qualitative difference in their respective risk scores. The first country with an upper bound less than Afghanistan's score begins a new grouping. In this case, that would be Ethiopia. Moving further down the list, any country with an upper bound value greater than Ethiopia's risk score of 26.8 would be grouped with Ethiopia. These two groups are combined to form the "highest risk" category.

Moving down the list in a similar manner, the third grouping of states (beginning with Kenya) will be assigned to the "high risk" category. Proceeding in the same fashion, the next grouping makes up the "moderate risk" category, followed in turn by the next grouping that makes up the "some risk" category. All remaining groups below the "some risk" grouping are combined into the "low risk" category.

The new procedure establishes risk categories that have qualitative meaning. Within each grouping, a solid empirical basis exists that indicates that the identified states are quite comparable in terms of risk. Moreover, we can be more confident that states assigned to different groupings are qualitatively distinct. Since the groupings were created by using information from the confidence ranges, it is unlikely that states in lower groupings have a true estimated risk that is higher than countries assigned to higher risk categories. In sum, by moving away from arbitrary categories of risk, the new classification procedure provides better guidance for making more accurate country assessments.

Overview of Results

Table 3.4-2 lists the 25 countries with the highest risk scores. With the exception of three states (Afghanistan, Iraq, and Nepal), all of them are African. Indeed, the concentration of African states among the grouping of states with the highest estimated risk is even higher than the previous findings reported in 2008, when 19 of the 25 states were from Africa. Higher estimated risk scores in the Democratic Republic of the Congo, Guinea-Bissau, and Mauritania have led these countries into the top 25, supplanting Bangladesh, Lebanon, and Haiti. No doubt, a quick review of the top 25 list reveals that problems in Africa remain acute, and, as the discussion below suggests, they are likely getting worse.

Overall, the mean instability scores across countries formerly classified at moderate or high risk have increased significantly since the last report. Among countries classified at high risk in the previous analyses, the average risk score was 14.1. For the same set of countries, the average is now 17.3, a difference that is statistically significant. For the countries previously classified at moderate risk, the previous average was 5.3. It is now 7.4 for those countries, a difference that is also statistically significant. Among countries classified previously at low risk, the average score did increase somewhat (from 1.5 to 1.6), but that difference is not statistically meaningful. What factors explain the upward shift in risk estimates among states that are already vulnerable to instability and conflict?

Table 3.4-2. Highest Estimated Risk for Instability, 2008-2010.

Rank	***Country***	***Risk Score***
1	Afghanistan	38.9
2	Niger	33.1
3	Burundi	30.3
4	Congo, Democratic Republic*	29.1
5	Djibouti	28.2
6	Ethiopia	26.8
7	Mali	25.9
8	Nigeria	25.6
9	Tanzania	24.5
10	Zambia	24.2
11	Sierra Leone	23.3
12	Liberia	22.7
13	Mauritania*	21.4
14	Guinea-Bissau*	20.2
15	Angola	20.0
16	Iraq	19.7
17	Côte d'Ivoire	19.5
18	Kenya	18.0
19	Central African Republic	17.6
20	Somalia	16.9
21	Chad	16.6
22	Benin	16.0
23	Mozambique	15.8
24	Malawi	15.5
25	Nepal	15.2

* New to the top 25 in the most recent rankings

The modest push of democratization and the inherent inconsistency of regime that follows is one of the key factors contributing to increased risks of instability. In the most recent data available on regime characteristics, seven countries formally classified at moderate risk or high risk transitioned to partial democracies from more autocratic regime types. No countries in either of those categories of risk transitioned to consolidated democracies. (In the low-risk category, Chile and the Slovak Republic transitioned to more coherent democracies, which led to significantly lower risk scores for both countries.) One country, Bangladesh, experienced a setback in its democratic transition and was reclassified as a more autocratic state in the most recent data, which contributed to a lower estimated risk. On the whole, though, the net effect of regime changes in the class of states with at least moderate risk was an exertion of more upward pressure on risk scores.

A slight increase in the number of active armed conflicts around the world has also contributed to the overall increase in the risk of instability. With more neighborhoods experiencing the volatile externalities generated by violent conflict—refugee flows, arms trafficking, threats of intervention—the risk of instability for the residents has increased. (Since the last publication of the rankings, a slight improvement has been implemented to the definition of a neighborhood conflict. As a result, several states that were originally coded as involved in active conflict are no longer coded as such. Previously, a neighbor was classified at war when it was involved in any armed conflict, even a conflict not being fought on its territory. But that definition was inconsistent with the theoretical expectation about how a neighborhood conflict relates to the risks of instability. In almost all imaginable circumstances, the spillover effects of neighborhood conflict operate only when the conflict is being fought directly on the neighbor's territory. As a result of this coding change, the only significant coding change implemented in the new analyses, the risk scores for some states declined. Examples of such states include any state that bordered a state that had contributed troops to the multilateral efforts in Iraq or Afghanistan. This set of countries tended to have relative low risks of instability in the first place, so the impact of the definitional change was slight.) All told, risk estimates for 13 countries increased because a new conflict erupted in a neighboring state (or an old conflict resurged). Only one state—Papua New Guinea—has a lower risk estimate because a conflict subsided in a neighboring state (the conflict in the Aceh territory in Indonesia, which subsided in 2006).

Changes Since 2007

To see more clearly how democratization and neighborhood conflict can have an immediate impact on increasing the risks of instability, let us take a more detailed look at the circumstances in some countries that experienced significant change from the previous rankings.

For purposes of identifying cases in which significant change has occurred, we adopt a clear standard that utilizes information from each country's confidence interval. To identify cases in which the risk of instability has significantly increased, we require that the lower end of the new confidence range be greater than the risk estimate from the previous analysis. This standard allows us to conclude that the change in underlying risk has increased or has remained largely unchanged, but we can be nearly certain that it has not decreased. Cases of significant improvement require that the upper bound of the new confidence range be less than the risk estimate from the previous analysis. In this case, the standard permits an interpretation that the risk in a given country has likely declined or remained unchanged, but it has not increased.

Table 3.4-3 lists the 17 countries with the largest increase in risk scores. Of these, the change in 13 of the countries satisfies the standard for a significant worsening. For each of the countries, data are presented corresponding to the previous ledger rankings (based on the forecast period 2005–2007) alongside data used for the current forecast period (2008–2010). The net change in risk score is listed with an indication (*) of whether the change satisfies the threshold defined above. In most of these cases, the increase in risk can be traced to two separate factors: transitions to democratic governance and armed conflict in neighboring countries.

Undoubtedly, the process of democratization is a welcome development because it brings desirable qualities to governance (e.g., greater citizen participation, broader competition for leadership positions, more expansive civil liberties, etc.). For many observers, though, the heightened dangers of instability during this period are often under-appreciated.

Table 3.4-3. Largest Increases in Risk of Instability.

Forecast Period	Country	Risk Score	Net Change	Confidence Range	Regime Consistency	Partial Democracy	Infant Mortality	Economic Openness	Militarization	Neighborhood War
2005-07	Congo, Dem.	6.9		3.7 – 11.8	0	○	129	70%	115	●
2008–10	Rep.	29.1	22.2*	18.0 – 42.5	25	●	128	65%	115	●
2005–07	Burundi	11.1		6.5 – 18.0	0	○	114	40%	1112	○
2008–10		30.3	19.2*	19.0 – 45.2	36	●	109	59%	112	●
2005–07	Mauritania	5.1		3.1–7.5	36	○	78	25%	671	●
2008–10		21.4	16.3*	12.6–33.0	16	○	77	123%	690	●
2005–07	Nigeria	13.4		7.6–21.5	16	●	101	92%	124	○
2008–10		25.6	12.2*	15.9–37.4	16	●	98	69%	112	●
2005–07	Djibouti	17.1		8.4–31.3	4	●	101	133%	1412	~
2008–10		28.2	11.1	15.7–45.3	4	●	86	137%	1412	●
2005–07	Guinea–	9.3		4.8–16.6	1	○	125	87%	585	○
2008–10	Bissau	20.2	10.9*	11.3–32.7	36	●	129	89%	547	○
2005–07	Angola	10.5		4.7–20.6	4	○	154	125%	762	○
2008–10		20.0	9.5	9.2–35.8	4	○	154	108%	664	●
2005–07	Zambia	14.8		9.1–23.1	25	●	101	47%	139	○
2008–10		24.2	9.4*	14.9–36.8	25	●	101	78%	137	●
2005–07	Pakistan	5.2		3.3–7.9	25	○	80	31%	606	●
2008–10		14.6	9.4*	9.2–22.0	4	○	77	36%	581	●
2005–07	Nepal	6.4		3.8–10.0	36	○	59	48%	493	●
2008–10		15.2	8.8*	10.6–21.1	36	●	46	41%	474	●
2005–07	Uganda	4.9		2.8–8.1	16	○	80	41%	198	●
2008–10		12.9	8.0*	8.0–19.8	1	○	77	45%	198	●
2005–07	Burkina Faso	8.3		5.0–12.9	0	○	97	32%	80	○
2008–10		14.1	5.8*	8.8–21.4	0	○	122	38%	80	○
2005–07	Cameroon	6.8		4.2–10.5	16	○	87	39%	143	○
2008–10		12.5	5.7*	7.3–19.7	16	○	86	43%	143	●
2005–07	Tanzania	18.9	5.6	12.3–27.9	4	●	78	46%	74	●
2008–10		24.5		16.2–35.5	1	●	74	50%	74	●
2005–07	Chad	11.2	5.4	5.4–20.7	4	○	78	46%	74	●
2008–10		16.6		8.8–27.7	4	○	124	83%	360	●

Table 3.4-3. Largest Increases in Risk of Instability.

Forecast Period	Country	Risk Score	Net Change	Confidence Range	Regime Consistency	Partial Democracy	Infant Mortality	Economic Openness	Militarization	Neighborhood War
2005–07	Kyrgyz	3.5	5.3*	1.7–6.2	9	○	59	96%	333	●
2008–10	Republic	8.8		4.7–14.6	9	●	36	116%	404	○
2005–07	Bolivia	7.6	5.2*	4.5–12.1	64	●	54	58%	759	○
2008–10		12.8		7.8–19.3	64	●	49	66%	887	●

NOTE: An asterisk (*) indicates a net change that qualifies as significant according to the definition offered above. The numbers in the infant mortality column are the total infant deaths per 1,000 live births. The percentage in the economic openness column refers to the percentage of a country's GDP accounted for by the value of its imports plus exports. The number in the militarization column refers to the number of active military personnel per 100,000 people. Finally, the symbol ● means "yes" and the symbol ○ means "no."

Partial democracies are at greater risk for instability than autocracies or full democracies. Repressive tactics adopted by autocratic governments often smother the potential for political instability. Coherent and mature democracies possess the capacity to address group grievances and manage the competition between groups that vie for political power and other resources, thereby reducing the risks of instability. Partial democracies typically possess neither of the qualities of full autocracies nor those of democracies, leaving them more vulnerable to the drivers of instability and conflict.[27] Indeed, the historical data over the past half-century shows a strong empirical relationship between partial democracy and the future onset of instability or conflict.

The bout of instability that raged in Kenya after its December 2007 presidential election serves as a recent illustration. In the ledger published in *Peace and Conflict 2008*, data from 2004 were used to produce Kenya's risk estimate for the period 2005–2007. At the time, Kenya was classified as a partial democracy, according to the Polity project, because the competitiveness of political participation was still seen as in transition from competition that had been largely suppressed by the regime. Its risk score, 12.9, reflected the potential for instability associated with these regime characteristics, placing it squarely among other high-risk states in Africa, despite a record of relative stability compared to neighbors like Ethiopia, Somalia, and Sudan. Those institutions would prove to lack the resilience to withstand forces that developed in the months preceding the close presidential election, as well as the pressures that were released in its aftermath. After months of campaigning with appeals to rectifying injustices based on advantages accorded to ethnic Kikuyus, opposition candidate Raila Odinga appeared to be positioned to win the election based on pre-election polling. When reports indicated that he had lost by a slim margin, amid widespread reports of election fraud, the forces driving the potential for instability were catalyzed. In the weeks following the election, approximately 1,000 people were killed in ethnically based violence throughout the country.

The experience of Kenya in late 2007 and early 2008 illustrates the vulnerabilities of partial democracies to some of the forces that can catalyze major episodes of instability. Kenya continues to be classified as a partial democracy, receiving an updated risk score of 18.0, which reflects some of the worsening conditions in the country in the aftermath of recent instability. With Kenya's experience in mind, it is worthwhile to briefly note some other countries that have

been newly classified as partial democracies in the updated rankings, a change that has significantly increased their respective risk for future instability.

The Democratic Republic of Congo, Burundi, Mauritania, Guinea-Bissau, Nepal, and the Kyrgyz Republic all transitioned in varying extents to partial democratic rule between 2005–2007. For any of these countries, much could be written about how the transition to more democratic governing arrangements influences the estimated risk for instability. Space constraints permit a focus on just one of these states. For the purposes of illustrating the impact that democratic transition can have on the estimated risk of instability, the case of the Kyrgyz Republic will be most suitable.

In the previous ledger rankings, the Kyrgyz Republic estimate for risk of instability from 2005 to 2007 was based on data from 2004. In that year, governing arrangements in the country tended toward autocracy, although constitutional provisions did allow for some competitive elections and fewer restrictions on political participation. Still, Kyrgyzstan did not qualify as a partial democracy, which contributed to an estimate of only moderate risk for instability (3.5). By late 2006, a new constitution was in place that gave more political authority to the parliament. The changes in regime characteristics were sufficient to reclassify Kyrgyzstan as a partial democracy according to the Polity project's coding rules. In subsequent months, that authority would shift back to the presidency, but Kyrgyzstan continues to be coded as partially democratic. Predictably, the tenuous step toward democratization in Kyrgyzstan led to an increase in the estimated risk of instability (8.8). What makes the case of Kyrgyzstan notable is that the estimated risk of instability increased despite significant improvement in other areas. From 2004 to 2007, the country's infant mortality rate declined from roughly 58 deaths per 1,000 live births to 36. Moreover, a low-intensity armed conflict in neighboring Uzbekistan in 2004 had subsided by 2007. Despite these changes, which exert modest downward pressure on risk estimates, the movement toward democracy had a more powerful impact on pushing the estimated risk upward.

For other countries listed in Table 3.4-3, the heightened risk of instability is due to the onset (or recurrence) of armed conflict in a neighboring state. In Burundi, for example, the risk of instability increased substantially, due to renewed fighting in the neighboring Democratic Republic of the Congo (DRC). That same recurrence of conflict in the DRC is responsible for the heightened risk estimates in Angola. Nigeria's risk of instability increased with the recurrence of conflict in Chad in 2005 and less intense violence in Niger in 2007.

Table 3.4-4 presents a list of ten countries that showed the largest improvement in risk scores. Of these countries, seven made improvements that satisfied the requirements for significance outlined above. Glancing down the "Net Change" column of the table, it can be seen that the absolute level of reductions in risk is much lower than the absolute level of increases observed in Table 3.4-3. Given the overall global trend toward greater levels of instability, these differences should be expected.

Just as the risk of instability can increase substantially in a short period of time, it can decrease just as suddenly. For instance, the termination of a neighborhood conflict can reduce the risk of instability for a country abruptly. As mentioned previously, the lower risk score for Papua New Guinea is the result of the cessation of serious armed violence in neighboring Indonesia in 2006—the only conflict termination since the previous publication of the ledger.

For some countries in Table 3.4-4, the estimated risk of instability decreased because the country experienced a setback in its transition to democracy. Both Fiji and Bangladesh were coded as

partial democracies in the previous rankings. In the most recent data, governing arrangements in both countries have shifted toward greater autocracy, which produces lower estimated risk scores. In Bangladesh, the lower risk estimate may be short-lived. Parliamentary elections were postponed in January 2007 due to serious concerns of potential electoral fraud and corruption. President Iajuddin Ahmed resigned from office, and a series of measures were implemented as part of a general state of emergency. The changes led to the Polity project's recoding Bangladesh's regime as essentially autocratic in 2007. Since the parliamentary elections were held in December 2008, it is conceivable that Bangladesh may qualify as a partial democracy in a future release of Polity data, which would cause risk estimates to return to higher levels.

Table 3.4-4. Largest Reduction in Risk of Instability.

Forecast Period	Country	Risk Score	Net Change	Confidence Range	Regime Consistency	Partial Democracy	Infant Mortality	Economic Openness	Militarization	Neighborhood War
2005–07	Iraq	29.9	–10.2*	20.0 – 43.2	0	●	101	22%	668	●
2008–10		19.7		12.0 – 28.8	0	●	50	62%	668	●
2005–07	Bangladesh	13.1	–10.1*	9.1 – 18.7	36	●	56	36%	180	●
2008–10		3.0		19.0 – 5.2	36	○	51	51%	137	●
2005–07	Serbia	4.5	–2.8*	2.4–8.0	36	●	13	74%	1350	●
2008–10		1.7		0.7–3.5	64	●	7	76%	324	○
2005–07	Fiji	3.6	–2.8*	1.9–6.0	36	●	16	40%	357	○
2008–10		0.8		0.3–1.7	16	○	16	128%	480	○
2005–07	Papua New	5.1	–2.7	2.5–9.3	100	○	69	32%	52	●
2008–10	Guinea	2.4		1.1–4.8	100	○	55	147%	48	○
2005–07	Honduras	6.6	–2.4	3.9–9.3	49	●	32	85%	284	●
2008–10		4.2		2.2–6.8	49	●	23	130%	287	○
2005–07	Albania	4.5	–1.9*	2.6–7.3	49	●	16	65%	691	●
2008–10		2.6		1.4–4.3	81	●	15	79%	363	○
2005–07	Nicaragua	5.9	–1.8	3.4–95	64	●	31	81%	260	●
2008–10		4.1		2.2–6.7	81	●	29	120%	253	○
2005–07	Guatemala	7.3	–1.7	4.8–11.0	64	●	33	48%	390	●
2008–10		5.6		3.5–8.7	64	●	31	66%	269	○
2005–07	Vietnam	2.3	–1.7*	0.6–5.8	49	○	17	140%	6772	○
2008–10		0.6		0.2–1.1	49	○	15	159%	589	○

NOTE: An asterisk (*) indicates a net change that qualifies as significant according to the definition offered above. The numbers in the infant mortality column are the total infant deaths per 1,000 live births. The percentage in the economic openness column refers to the percentage of a country's GDP accounted for by the value of its imports plus exports. The number in the militarization column refers to the number of active military personnel per 100,000 people. Finally, the symbol ●means "yes" and the symbol ○ means "no."

For many of the other countries in Table 3.4-4, improvements in the economic and social domains were most responsible for reductions in the underlying risk of instability. Iraq posted the largest improvement from the previous rankings for the 2005–2007 period. Large improvements in infant mortality rates and in economic openness are the main sources of the reduced risk estimate. The figures from 2004 were particularly bad, an unsurprising artifact of the vast disruption caused by the war that began in 2003. Iraq's infant mortality rate (as estimated in the *CIA Factbook*) was approximately 100 per 1,000 live births, the twentieth highest figure of all countries worldwide. The proportion of its GDP accounted for by trade was just 22 percent, ninth worst in the world. By 2007, those indicators had improved. Now, Iraq's infant mortality (about 50 deaths per 1,000 live births) is more comparable to that of Bangladesh.

In addition, the indicator for economic openness has improved substantially to 62 percent. While the improvements are notable, Iraq's overall risk score (19.7) continues to place it in the highest risk category, a solemn reminder of how grim the circumstances remain there.

In Serbia, improvements on several dimensions contributed to substantial reductions in its risk score. Although coded as a partial democracy, Serbia's regime consistency score improved from 36 in 2004 to 64, reflecting steady movement toward greater democratic consolidation. Infant mortality rates decreased from about 12 per 1,000 births to about 7. Finally, the militarization indictor dropped significantly, reflecting a sizeable reduction in the number of active armed personnel in Serbia from 2004 to 2007. In Albania and Vietnam, combined improvements in infant mortality, economic openness, and militarization all contributed to modest, yet significant, reductions in estimated risk. The achievement in Albania is notable because it qualifies as a partial democracy, illustrating how countries in the midst of a democratic transition can mitigate the profound risks inherent in that transformation through effective governance.

Conclusion

The sudden and significant increase in the estimated risk of instability across countries that were already vulnerable should be a cause for concern for policy-makers involved with managing conflict and addressing the larger challenges to state stability. There are two distinct causes for heightened levels of risk, which means that policy-makers should be attentive to how policy responses relate to each.

As regimes transform from autocracies to partial democracies, the estimated risks of major instability events increases dramatically. Policy responses that address the specific vulnerabilities of such regimes have the potential to mitigate instability risks. For example, any government policies that reduce the extent of factional-based political competition can increase the prospects that multiple sub-national groups (ethnic or non-ethnic) see themselves as stakeholders in the current set of institutional arrangements. A greater sensitivity to the importance of transparency in electoral procedures can reduce the catalytic potential for tightly contested elections to trigger instability. And, of course, while the volatile transition to consolidated democracy occurs, it is crucial that attention be paid to policies that enhance government's ability to deliver core services to the population (as illustrated by Albania's recent experience). Doing so will enhance the likelihood that it is viewed as legitimate, mitigating the risks faced by typical partial democracies.

At the same time, estimated risks may suddenly become elevated because of the onset or recurrence of a neighborhood conflict. In these cases, appropriate policy responses should address some of the contagion effects of conflicts. For a country with a neighbor involved in civil conflict, attention should be paid to the relationship that ethnic groups located near the border may have with warring parties in the country at war. Where there is potential for cross-border activity, appropriate responses may include heightened border monitoring and control to prevent the transfer of arms or the movement of soldiers into the warring country.

Ultimately, the key to effective policy responses to heightened risks of instability depends heavily on an ability to trace back from the estimate to the particular factors that exert the most influence on it. The *Peace and Conflict Instability Ledger* places an emphasis on making information about the risk estimates as accessible and interpretable as possible, so that diagnosing the foundations of these risks can be more effective. Moreover, by explicitly reporting confidence ranges associated with each country estimate, the ledger offers policy-

makers enhanced leverage for making more confident assertions about the substantive importance of any year-to-year change observed in a particular country—a crucial necessity for making precise assessments about progress in at-risk countries. This chapter has offered several brief discussions of cases to be suggestive of how information from the ledger can be used to help clarify risk trends in a particular country. Employed alongside the detailed information (both qualitative and quantitative) available to country experts, the ledger can be a powerful diagnostic tool in any policy-makers' toolkit for assessing risk levels across countries.

The *Peace and Conflict Instability Ledger* ranks states according to the forecasted risk of future instability. See notes following these listings for a description of the codes for each indicator and also a detailed explanation of the confidence range (note 10).

Table 3.4-5 - The Peace and Conflict Instability Ledger (abridged).

Region	*Recent Instability*	*Country*	*Regime Consistency*	*Infant Mortality*	*Economic Openness*	*Militarization*	*Neighborhood War*	*Risk Category*	*Risk Score*	*Confidence Range*	
										Lower	*Upper*
Africa (Top 15)		Niger	⊙	●	●	○	●	●	38.9	21.3	47.6
		Burundi	⊙	●	⊙	●	●	●	30.3	19.0	45.2
	■	Dem. Rep. of Congo	●	●	⊙	●	●	●	29.1	18.0	42.5
		Djibouti	●	●	○	●	●	●	28.2	15.7	38.7
		Ethiopia	●	⊙	●	⊙	●	●	26.8	18.2	38.0
		Mali	⊙	●	⊙	○	●	●	25.9	15.7	38.7
		Nigeria	●	●	⊙	○	●	●	25.6	15.9	37.4
		Tanzania	●	⊙	⊙	○	●	●	24.5	16.2	35.5
		Zambia	●	●	⊙	○	●	●	24.2	14.9	36.8
		Sierra Leone	⊙	●	⊙	○	○	●	23.3	12.4	37.7
		Liberia	⊙	●	⊙	○	○	●	22.7	12.3	38.1
		Mauritania	●	⊙	⊙	○	○	●	21.4	12.6	33.0
		Guinea-Bissau	⊙	●	⊙	⊙	○	●	20.2	11.3	32.7
		Angola	⊙	●	⊙	⊙	●	●	20.0	9.2	35.8
	■	Côte d'Ivoire	●	●	⊙	⊙	○	●	19.5	11.7	30.1
Asia (Top 10)	■	Afghanistan	●	●	⊙	○	●	●	38.9	23.7	58.2
		Nepal	⊙	⊙	●	⊙	●	⊙	15.2	10.6	21.1
	■	Pakistan	⊙	⊙	●	⊙	●	⊙	14.6	9.2	22.0
		Timor-Leste	⊙	⊙	●	○	○	⊙	14.3	14.3	20.3
	■	India	⊙	⊙	●	⊙	●	⊙	12.0	7.1	18.2
		Kyrgyz Republic	●	⊙	○	⊙	○	⊙	8.8	4.7	14.6
		Korea, Dem. Rep.	○	⊙	●	●	●	⊙	8.7	3.2	17.9
		Tajikistan	⊙	⊙	⊙	⊙	●	⊙	8.3	4.7	14.0
		Cambodia	⊙	⊙	○	●	○	⊙	7.8	4.1	13.6
	■	Myanmar	○	⊙	●	●	●	○	6.0	3.9	8.0
Eastern Europe (Top 10)		Armenia	●	⊙	⊙	●	●	⊙	10.8	6.0	17.2
		Georgia	⊙	⊙	⊙	⊙	●	⊙	9.0	5.5	13.6
		Russia	●	⊙	⊙	●	●	○	6.7	3.4	11.4
		Ukraine	⊙	⊙	⊙	⊙	●	○	5.8	3.4	9.6
		Bosnia	●	⊙	○	⊙	○	○	4.2	1.6	8.7
		Azerbaijan	○	⊙	⊙	●	●	○	3.9	2.3	6.4
		Romania	⊙	⊙	⊙	⊙	○	○	3.1	1.7	5.4
		Montenegro	⊙	○	○	●	○	○	2.9	1.3	5.5
		Bulgaria	⊙	⊙	○	●	●	○	2.8	1.4	4.9
		Moldova	⊙	⊙	○	⊙	○	○	2.7	1.5	4.9
& the C		Haiti	●	⊙	⊙	○	○	⊙	13.1	8.3	19.1
		Bolivia	⊙	⊙	⊙	●	●	⊙	12.8	7.8	19.3

Table 3.4-5 - The Peace and Conflict Instability Ledger (abridged).

		Ecuador	●	⭘	⭘	⭘	●	⭘	7.8	4.5	12.7
		Guyana	⭘	⭘	○	⭘	○	⭘	7.7	4.0	13.6
		Mexico	⭘	⭘	⭘	⭘	●	○	7.3	4.6	11.1
		Venezuela	●	⭘	●	⭘	●	○	7.2	4.0	11.7
	■	Columbia	⭘	⭘	●	●	●	○	6.6	4.1	10.8
		Brazil	⭘	⭘	●	⭘	●	○	5.9	4.0	10.3
		Peru	⭘	⭘	●	⭘	●	○	5.9	3.6	9.0
		Guatemala	⭘	⭘	⭘	⭘	○	○	5.6	3.5	8.7

Region	*Recent Instability*	*Country*	*Regime Consistency*	*Infant Mortality*	*Economic Openness*	*Militarization*	*Neighborhood War*	*Risk Category*	*Risk Score*	*Confidence Range*	
										Lower	*Upper*
Middle East and North Africa (Top 10)	■	Iraq	●	⭘	⭘	⭘	●	●	19.7	12.0	28.8
	■	Yemen	⭘	⭘	●	⭘	○	⭘	11.6	7.7	16.5
		Lebanon	⭘	⭘	⭘	●	●	⭘	10.5	8.3	6.0
	■	Turkey	⭘	⭘	●	●	●	⭘	8.3	5.3	12.3
		Egypt	⭘	⭘	⭘	●	●	○	6.0	3.4	9.9
		Algeria	⭘	⭘	⭘	●	○	○	5.0	2.8	8.1
		Jordan	⭘	⭘	○	●	●	○	4.7	2.3	8.1
		Tunisia	⭘	⭘	○	⭘	●	○	3.2	1.6	5.6
		Morocco	○	⭘	⭘	●	●	○	2.2	1.2	3.9
		Iran	○	⭘	⭘	●	●	○	2.1	1.1	3.7
North Atlantic (Top 5)		Macedonia	⭘	⭘	○	●	○	○	2.5	1.3	4.3
		United States	○	○	●	⭘	●	○	0.9	0.4	1.9
		Belgium	⭘	○	○	⭘	○	○	0.8	0.2	1.8
		Greece	○	○	●	●	●	○	0.6	0.2	1.4
		Cyprus	○	○	●	●	●	○	0.6	0.2	1.3

Explanations for the Ledger

The ledger is based on a model that estimates the statistical relationship between the future likelihood of instability and each of the five factors in the chapter. The model is estimated based on data for the period 1950–2003, which indicated that each of the five factors was strongly related to the future risk of instability. Using the model estimates for the causal weight assigned to each factor, data from 2007 (the last year for which complete data are available for all five of our factors) were used to produce a three-year forecast indicating the risk of instability in the period 2008–2010. The color codes used in the ledger to present a country's standing on each of the five factors are based on the values in 2007. The notes below explain the various codings.

(1) Recent Instability: indicates (with a ■) whether the country has been coded by the Political Instability Task Force (PITF) as being involved in an instability event as of the end of 2006. The country's risk score (see column 10) provides an assessment of the likelihood of the country's experiencing future instability. One might interpret the risk score for countries currently experiencing instability as the risk of continued instability, but readers are cautioned that the causal factors that drive the continuation of instability are likely not the same as the factors that drive the onset of instability.

(2) Country: the ledger examines only those countries with populations greater than 500,000 in 2007.

(3) Regime Consistency: the risk of future instability is strongly related to the extent to which the institutions comprising a country's political system are uniformly and consistently autocratic or democratic. Political institutions with a mix of democratic and autocratic features are deemed inconsistent, a common attribute of polities in the midst of a democratic transition (or a reversal from democratic rule to more autocratic governance). Regimes with inconsistent institutions are expected to be more likely to experience political instability. In the ledger, highly consistent democracies (polity score ≥6) and autocracies (polity score ≤ -6) receive a green marker. The symbol ● has been assigned to regimes with inconsistent characteristics that also qualify as partial democracies according to PITF. Regimes with these characteristics have been found to have the highest risk for instability. The symbol ⭘ has been assigned to partial autocracies because the propensity for instability in these regimes is somewhat less than in partial democracies.

(4) Infant Mortality: infant mortality rates serve as a proxy for overall governmental effectiveness in executing policies and delivering services that improve social welfare in a country. High infant mortality rates are associated with an increased likelihood of future instability. The states with the best records are indicated with the symbol ○ (scoring in the bottom 25th percentile of global infant mortality rates). States with the worst record (scoring in the highest 25th percentile) are indicated with the symbol ●. States in the middle 50th percentile are indicated with the symbol ⭘.

(5) Economic Openness Closer integration with global markets reduces the likelihood of armed civil conflict and political instability. Policies that integrate global and domestic markets can produce higher growth rates and sometimes reduce inequality. To that extent, economic openness can remove or weaken common drivers for civil unrest related to economic grievances. We focus on the proportion of a country's GDP accounted for by the value of all trade (exports plus imports) as a measure for economic openness. The countries with the lowest score for economic openness are considered to be at the highest risk for instability. These states are designated with the symbol ●. The states in the highest 25th percentile receive the symbol ○ in the ledger. The middle 50th percentile receives the symbol ⭘.

(6) Militarization Instability is most likely in countries where the opportunities for armed conflict are greatest. In societies where the infrastructure and capital for organized armed conflict are more plentiful and accessible, the likelihood for civil conflict increases. The ledger measures militarization as the number of individuals in a country's active armed forces as a percentage of the country's total population. Countries with militarization scores in the bottom 25th percentile are indicated with the symbol ○. Countries in the top 25th percentile are presented with the symbol ●. The middle 50th percentile is indicated with the symbol ⭘.

(7) Neighborhood War The presence of an armed conflict in a neighboring state (internal or interstate) increases the risk of state instability. The contagion effects of regional armed conflict can heighten the risk of state instability, especially when ethnic or other communal groups span across borders. The most recent data release from the Uppsala Conflict Data Project at the International Peace Research Institute is used to determine the conflict status of states in 2007.[28] For a neighbor to be considered involved in armed conflict, the conflict must produce 25 or more battle-related fatalities per year. The symbol indicates when at least one neighbor is involved in

armed conflict. The symbol ○ indicates the absence of armed conflict in all neighboring states.

(8) Risk Category States have been placed in one of five categories corresponding to their risk score. The chapter text discusses the procedure for assigning states to the highest risk category (●), the high risk category (◉), the moderate risk category (◉), the some risk category (○), or the low risk category (○).

(9) Risk Score The risk score gives a three-year forecast of the relative risk (compared to an average member of the OECD) of experiencing instability. The score is computed based on the results of estimating a statistical model using global data from the period 1950–2003. Then, using the model estimates, data from 2007 were used to obtain the three-year forecasts for each country for the period 2008–2010.

(10) Confidence Range The confidence range provides information about the degree of uncertainty corresponding to a country's estimated risk score. Statistically speaking, the "true" risk of instability lies within this range with a confidence probability of 95 percent.

↑ Back to the top

---

[1] Hewitt, J. Joseph. forthcoming. "The Peace and Conflict Instability Ledger: Ranking States Based on Future Risks," in Peace and Conflict 2010 (J. Joseph Hewitt, Jonathan Wilkenfeld, Ted Gurr). Boulder: Paradigm Publishers.

[2] Esty, Daniel C., Jack Goldstone, Ted Robert Gurr, Barbara Harff, Marc Levy, Geoffrey D. Dabelko, Pamela T. Surko, and Alan N. Unger. 1999. The State Failure Report: Phase II Findings.

[3] Collier, Paul, V. L. Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis. 2003. Breaking the Conflict Trap: Civil War and Development Policy. Washington, DC: The World Bank.

[4] Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War," Oxford Economic Papers 56:563-95.

[5] Esty, Daniel C., Jack Goldstone, Ted Robert Gurr, Barbara Harff, Marc Levy, Geoffrey D. Dabelko, Pamela T. Surko, and Alan N. Unger. 1999. The State Failure Report: Phase II Findings.

[6] Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." American Political Science Review 97:75-90.

[7] Goldstone, Jack A. Robert H. Bates, Ted Robert Gurr, Michael Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2005. "A Global Forecasting Model of Political Instability." Presented at the Annual Meeting of the American Political Science Association, Washington, DC. September 1-4.

[8] Hegre, Håvard, and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." Journal of Conflict Resolution 50:508-35.

[9] Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001. "Toward a Democratci Civil Peace? Democracy, Political Change, and Civil War, 1816-1992." American Political Science Review 95 (1):33-48.

[10] King, Gary, and Langche Zeng. 2001. "Improving Forecasts of State Failure." World Politics 53:623-58.

[11] Sambanis, Nicholas. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil Wars." Defense and Peace Economics 13:215-43.

[12] Sambanis, Nicholas. 2004. "What Is a Civil War? Conceptual and Empirical Complexities of an Operational Definition." Journal of Conflict Resolution 48:814-58.

[13] United States Agency for International Development (2005) Measuring Fragility: Indicators and Methods for Rating State Performance. Washington DC.

[14] Hewitt, J. Joseph. 2008. "The Peace and Conflict Instability Ledger: Ranking States Based on Future Risks," in Peace and Conflict 2008 (J. Joseph Hewitt, Jonathan Wilkenfeld, Ted Gurr). Boulder: Paradigm Publishers.

[15] Hewitt, J. Joseph. 2008. "The Peace and Conflict Instability Ledger: Ranking States Based on Future Risks," in Peace and Conflict 2008 (J. Joseph Hewitt, Jonathan Wilkenfeld, Ted Gurr). Boulder: Paradigm Publishers.

[16] Gurr, Ted Robert. 1974. "Persistence and Change in Political Systems, 1800□1971." American Political Science Review 68(4):1482□1504.

[17] Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2006. "Institutional Inconsistency and Political Instabiliity: Polity Duration, 1800-2000." *American Journal of Political Science* 50(4):893-908.

[18] Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001. "Toward a Democratci Civil Peace? Democracy, Political Change, and Civil War, 1816-1992." American Political Science Review 95 (1):33-48.

[19] Hegre, Håvard, Ranveig Gissinger, and Nils Petter Gleditsch. 2003. "Globalization and Internal Conflict," in *Globalization and Armed Conflict*, by Gerald Schneider, Katherine Barbieri, and Nils Petter Gleditsch. Lanham, Maryland: Rowman & Littlefield Publishers.

[20] Goldstone, Jack A., Ted Robert Gurr, Barbara Harff, Marc A. Levy, Monty G. Marshall, Robert H. Bates, David L. Epstein, Colin H. Kahl, Pamela T. Surkom, John Ulfelder, and Alan N. Unger. 2000. "State Failure Task Force Report: Phase III Findings." McLean, VA: Science Applications International Corporation (SAIC).

[21] Esty, Daniel C., Jack Goldstone, Ted Robert Gurr, Barbara Harff, Marc Levy, Geoffrey D. Dabelko, Pamela T. Surko, and Alan N. Unger. 1999. The State Failure Report: Phase II Findings.

[22] Goldstone, Jack A. Robert H. Bates, Ted Robert Gurr, Michael Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2005. "A Global Forecasting Model of Political Instability." Presented at the Annual Meeting of the American Political Science Association, Washington, DC. September 1-4.

[23] Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War," Oxford Economic Papers 56:563-95.

[24] Hegre, Håvard, and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." Journal of Conflict Resolution 50:508-35.

[25] Goldstone, Jack A. Robert H. Bates, Ted Robert Gurr, Michael Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2005. "A Global Forecasting Model of Political Instability." Presented at the Annual Meeting of the American Political Science Association, Washington, DC. September 1-4.

[26] Sambanis, Nicholas. 2001. "Do Ethnic and Non-Ethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution* 45:259-282.

[27] Pate, Amy. 2008. "Trends in Democratization: A Focus on Instability in Anocracies." in *Peace and Conflict 2008* by J. Joseph Hewitt, Jonathan Wilkenfeld, and Ted R. Gurr. Boulder, Colorado: Paradigm Publishers.

[28] Gleditsch, Nils Petter, 2002. "The Future of Armed Conflict." Ramat Gan, Israel, The Begin-Sadat Center for Strategic Studies, Bar-Ilan University.

3.5. Perception is Reality: Stability through the Eyes of the Populace (Tom Rieger)

Author and Organization: Tom Rieger, Gallup
Contact Information: tom_rieger@gallup.com

Introduction

It has been said that the only constant is change. Countries that are stable today may quickly deteriorate under certain conditions. Catastrophic events may be weathered better by some countries than by others. Spillover from instability elsewhere may threaten neighbors, or it may destabilize an entire region. Having a robust model of the level of stability for a given population can be of tremendous value in being able to plan for, and possible help to avoid, significant human suffering as a result of instability. Many different models of state stability are currently in use or are regularly published from a variety of sources. While providing valuable insight, these models also offer some limitations.

Sources for Stability Models

To date, development of stability models has been limited by the sources available. The first and most obvious source available to the modeler is official or government statistics. However, that means trying to determine if a country is stable based on metrics compiled by a potentially unstable government. As an example, it has been argued that there are numerous problems in the statistics available from countries like Nepal for monitoring progress toward the Millennium Development Goals, ranging from gaps in data, lack of benchmarks, inconsistent methodologies, and controversy over basic definitions.[1] Others have argued that official statistics are often either not available, not comparable across countries, not comparable across time, and lack sub-national breakouts.[2] In other words, even when statistics are available and reliable, there are no global uniform standards regarding methodology and definition of those statistics, meaning that the data points used in creating the model will likely not be entirely comparable.

Another readily available source of data is articles and news content. While content analysis is highly useful for a number of applications, it is subject to self-selection bias, as well as the bias of the news source running the content. News media do not run every single story that crosses their desks. The editors choose which stories are worth reporting and which ones will likely increase readership or sell more advertising. Therefore, news media should not be mistaken to represent a comprehensive and representative overview of everything that is happening in a particular geography. In addition, many countries have state-controlled media, which only report what they want the populace and the outside world to know. Even within more open societies, including the United States, different news sources are characterized by and subject to their own biases.[3]

All of these limitations mean that the media content as an input to a model will by its nature be incomplete, biased, and self-selected. Since a model is only as good as the data that feeds it, content-based models will have these limitations, as well. But even beyond these limitations, there are two other serious problems with state-level estimates of stability based on metrics representing actual outcomes. These issues are discussed in the following sections.

Perceptions of Outcomes

The first is that actual outcomes do not necessarily reflect satisfaction with those outcomes. For example, according to the *Gallup World Poll*, in the fourth quarter of 2008, the percent of people saying the Afghanistan economy was poor was 35%. The percent of people saying the United States economy was poor was 62%, nearly twice as bad as the Afghanistan level,[4] and yet per-capita GDP in the United States is over $43,000, compared to only $270 in Afghanistan.[5]

One possible reason for this apparent disconnect can be found in Kahneman and Tversky's *Prospect Theory*. They found that the value of an outcome is determined not by the outcome itself, but rather by a perceptual evaluation of that outcome relative to a reference point, and that there is much greater pain in being below that reference point than joy in being above it.[6]

While Kahneman and Tversky's original work was primarily focused on decisions under risk, these principles have since been related to perceptual evaluations of progress toward goals and outcomes.[7,8]

In other words, if a person were to live in Chicago and that person was used to having 24 hours of electricity per day, facing a day with only 18 hours of electricity would be viewed as a catastrophe. If a person were to live in Fallujah in 2007, that same 18 hours of electricity per day would have been viewed as a godsend.

The United Nations Development Group acknowledged this limitation, in that statistics about poverty are limited because they "do not reflect people's feeling about relative deprivation" or even risks to their health.[9]

Lack of Homogeneity

The second problem with reliance on state-level outcome-based stability estimates is that levels of stability are rarely, if ever, completely homogenous across an entire nation, especially when conditions are first starting to destabilize.

Chechnya experienced violent unrest; Moscow did not. The Northwest Frontier Provinces of Pakistan are currently in a state of open conflict, while Punjab is relatively calm. During the worst of "The Troubles," Northern Ireland was much more unstable than Ireland or any other part of the United Kingdom. The problems in Sudan are much more severe in Darfur than they are in Khartoum. Even here in the United States, South Central Los Angeles is rife with crime and economic instability, while just a few miles away, South Orange County remains safe and relatively prosperous.

If a stability estimate is an amalgamation of some very stable areas and some very unstable areas, the estimate would be somewhere in the middle between stability and instability and would not be an accurate portrayal of either.

An Ideal Model

In order to avoid all of these limitations, a model should ideally (1) avoid official statistics or content analysis, (2) be based on perceptions of satisfaction with outcomes rather than the level of the outcome itself, and (3) have the ability to be easily applied to any subgroup of the population at any time.

One such model is Gallup Leading Assessment of State Stability (GLASS), which was developed entirely based on representative surveys of the populations of different countries using

the Gallup World Poll. The World Poll methodology is quite robust, including a mix of urban and rural coverage, a randomized choice of household members, weighting to non-responders to account for non-response bias, and several other quality control procedures.[10]

The model was developed under the assumption that people's expectations and satisfaction with their conditions will be a better measure of the likelihood of civil unrest than the outcome itself, for the reasons stated above.

To test this assumption, data from 69 countries from the 2007 Gallup World Poll were utilized, representing approximately 98,000 respondents. An initial set of independent and non-overlapping factors were statistically derived based on these data. The factors that were distilled through the analysis included the following:

- Basic needs, such as food and water
- Living conditions, covering "quality of life" issues
- Perceptions of economic conditions
- Government confidence, including local and national leadership, as well as the judicial system, electoral process, military, etc.
- Safety and security

Weighting the Factors

The next step was to determine the relative weight for each of these factors.

Each of the 69 countries included in the analysis was classified into one of three groups based on their current level of civil unrest. Those that were completely at peace with strong economies and relatively low crime were classified as a 2, indicating "stable." Those that were currently experiencing some form of significant civil unrest were classified as a 0, indicating "unstable." Countries that were somewhere in the middle were classified as a 1, indicating "moderate."

A linear regression was then run to determine the degree to which each factor could predict the current level of stability. Since the factors were derived in a manner that ensured statistical independence (orthogonality), a straight linear regression could be run without concerns over multicolinearity. The results of the regression were quite compelling. Overall, the model explained 71% of the variance (r-squared=.71).[11]

95% confidence intervals were calculated to then determine the appropriate ranges for stable, moderate, and unstable GLASS levels.

Testing the Model's Replicability

In order to test the replicability of the model, data sets were analyzed from an additional 30 countries that were not included in the original analysis, including some countries that were highly stable (such as Norway and Australia), some that were more moderate (such as Jamaica and Indonesia), and some that were clearly unstable (such as Chad, Afghanistan, and Haiti). In each case, the GLASS model accurately predicted the level of instability or civil unrest at the time of measurement.

In looking across the original set of 69 countries, however, there were a few countries that the model indicated as unstable, even though they were at the time relatively calm. Those countries included, among others, Georgia, Kenya, Peru, Mauritania, and Moldova. In each of these countries, within a year *after* measurement took place, stability markedly declined.

Other countries that were estimated to be unstable became significantly worse off, such as Nigeria, Zimbabwe, Madagascar, Haiti, as well as many of the countries experiencing foot riots in 2008.[12]

As further validation, a list of countries was compiled that had an election scheduled within a year following the interviews. Each of these countries was then classified, once the election occurred, into two groups based on the level of post-election violence (no violence, violence).[13]

The average level of stability as determined by GLASS for the countries that experienced post-election violence was firmly in the unstable zone, while those that did not experience violence were in the stable zone:

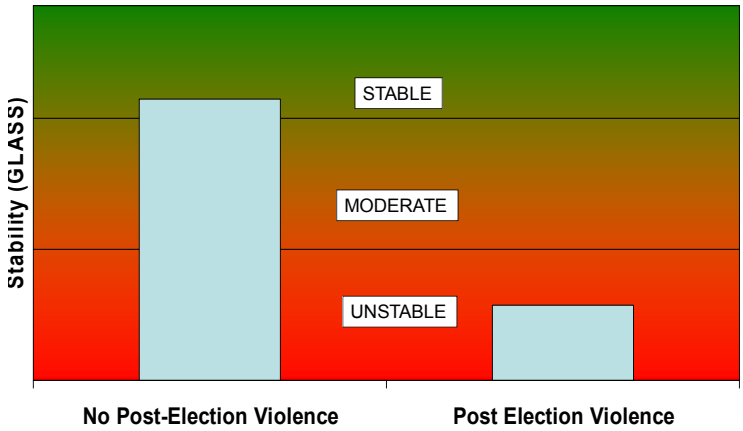


Figure 3.5-1. Post-measurement Election Outcome.

It is important to emphasize that these elections occurred after the time that interviews were collected, indicating that the GLASS estimate was a *leading* indicator of potential instability, and not just a reflection of the current state of affairs.

Measuring Potential for Instability

This lag effect apparent in the data presents an interesting implication for the model. Based on these results, as well as the countries that the model predicted will worsen, it appears that what the model is measuring is *potential for instability*. In other words, if an area can be thought of as a big pile of sticks, the model can help indicate their "flammability": whether they are soaked in water, soaked in gasoline, or dry wood. Before unrest occurs, there may be the need for the occurrence of a "matchstick event" such as an election, a spike in food prices, or some other inflammatory occurrence.

While the model has proven useful for a number of applications and initiatives, there are some important caveats to keep in mind.

Limitations

First, since the model is survey based, it can only be applied where it is possible to conduct surveys. While there are very few denied areas for research and polling, some do exist (such as North Korea, Darfur, etc.). GLASS estimates cannot be gathered for these populations.

Secondly, where sub-group analysis has been possible (due to large enough sample size), levels of stability, as well as the primary causes of instability, differ dramatically within countries. One area of a country may be highly stable, while another may be highly unstable. Even within

unstable areas, the most critical needs tend to differ. Perhaps in one village the water supply is not adequate, while in another, there may be no jobs. In a third, there may be extremely high crime. The model provides the ability to easily create these estimates for different subgroups (virtually any combination of respondents), but doing so requires adequate sample size, as well as the necessary classification variables included in the fieldwork.

Finally, cultural factors must be considered. While great care has been taken in the design of the World Poll to find questions that are generalizable across populations, there are some instances where, for the purposes of the model, questions may not apply. For example, in Thailand, it is not sufficient to simply ask about confidence in government leadership since it is against the law to say anything critical of the king (who enjoys extremely high support from the population regardless), but there is quite a bit of dissatisfaction with other leaders. When such issues exist, it is critical to ensure that these distinctions are taken into account in data collection.

Conclusion

They key point to consider is that ultimately, unrest occurs among the population. Determining the potential for unrest must therefore take into account how that population feels about their conditions. When used appropriately, taking the approach described herein can be valuable for determining where the next hot spot may occur, perhaps with enough advance notice to prevent instability from occurring, or lessen its impact.

↑ Back to the top

---

[1] Kayastha, R.P. (2007). "Statistical Problems for Monitoring of the Millennium Development Goals in Nepal." *International Conference on the Millennium Development Goals Statistics (ICMDGS)*, Manila, Philippines, October 1-3, 2007.

[2] Maligalig, D.S. (2003). "Measuring the Millennium Development Goals Indicators." *Concluding Workshop, RETA 6007: Enhancing Social and Gender Statistics.* Bangkok, Thailand, June 24-27, 2003.

[3] Roseclose, TG and Milyo, J. (2005) "A Measure of Media Bias." *Quarterly Journal of Economics*. VOL CXX, Issue 4, Nov 2005.

[4] "Gallup World Poll 2008 Results" (unpublished). Gallup, Inc., Washington D.C.

[5] Brown, S., et. al. (2009). *The Economist: Pocket World in Figures*. London, UK: Profile Books, Ltd. p. 28.

[6] Kahneman, D. and Tversky, A. (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*. Vol. 47 No. 2: pp. 263-91.

[7] Heath, C., Larrick, R. and Wu, G. (1999). "Goals and Reference Points." *Cognitive Psychology,* 38: pp. 79-109.

[8] Kahneman, D. Personal interviews, August 10, 2006 and August 30, 2007, Berkeley, CA.

[9] United Nations Development Group. (2003). "Indicators for Monitoring the Millennium Development Goals: Definitions, Rationale, Concepts and Sources," p. 7.

[10] Gallup, Inc. (2009) "Gallup World Poll: World Poll Methodology." Gallup, Inc., Washington D.C.

[11] Rieger, T. (2008). White Paper: "GLASS: Gallup Leading Assessment of State Stability, Methodological Overview Paper" (unpublished).

[12] "2007–2008 world food price crisis," retrieved June 2, 2009, from Wikipedia: http://en.wikipedia.org/wiki/2007%E2%80%932008_world_food_price_crisis.

[13] *IFES Election Guide*, retrieved August 2008, from: http://www.electionguide.org/index.php.

3.6. Assessing the Stability of Interstate Relationships Using Game Theory (Frank Zagare)

Author and Organization: Frank C. Zagare, Department of Political Science, University at Buffalo, SUNY
Contact Information: FCZagare@buffalo.edu

Abstract

Game models have long been used to analyze the robustness of contentious interstate relationships. Early attempts to analyze deterrence game-theoretically, however, were limited by certain deficiencies of the methodology. But, many of these limitations have now been overcome. The most recent generation of game models is dynamic, more realistic and easily adaptable to a wide range of strategic relationships. This paper explains the sense in which game models can be used to establish the stability, or lack thereof, of typical deterrence situations and to understand the context of their policy recommendations.

Early Literature

There is perhaps no tool more appropriate for analyzing the stability of contentious interstate competitions, nuclear or otherwise, than the mathematical theory of games. When it was originally developed by John von Neumann and Oskar Morgenstern in 1944 in their magisterial *Theory of Games and Economic Behavior*, game theory seemed to offer a systematic framework for analyzing the acrimonious relationship of the United States and the Soviet Union. However, it was not until the late 1950s and the early 1960s that this potential began to be realized. Among the early contributors to the ensuing debate about the strategic implications of nuclear weapons were such giants as Morgenstern himself,[1,2] Herman Kahn,[3,4,5] Albert Wohlstetter,[6] Glenn Snyder (1961),[7] Daniel Ellsberg,[8,9] and, of course, Thomas Schelling.[10,11]

Much of the early literature of nuclear deterrence focused on the well-known game of "Chicken" (see Figure 3.7-1). Chicken is a simple game that many strategic analysts initially believed offered a powerful theoretical context in which to think about interstate conflict. Eventually, however, it came to be recognized that the chicken analogy was inadequate. For one, it was a static game that took common knowledge about preferences (i.e., payoffs) as axiomatic. Additionally, it assumed that the payoffs to the players (i.e., states) were constant across time and issue areas. As the literature of deterrence evolved and matured over the next twenty years, game-theoretic models became more sophisticated. The first to be dismissed was the static framework and the assumption of complete information.[12] Shortly thereafter, the simplifying assumption of uniform preferences was abandoned.[13] Consequently, recent game-theoretic attempts to penetrate strategic issues have become more relevant to policy makers.

		State B	
		Cooperate (C)	Defect (D)
State A	Cooperate (C)	*Status Quo* (3,3)	*B Wins* (2,4)*
	Defect (D)	*A Wins* (4,2)*	*Conflict* (1,1)

Key: (x,y) = payoff to State A, payoff to State B
4 = best; 3 = next-best; 2 = next-worst; 1 = worst
* = Nash equilibrium

Figure 3.6-1. Chicken.

Assessing Strategic Stability

There are many reasons why game-theory has proven to be the methodology of choice among academic researchers who study deterrence, both early on and currently. For one, game models provide a formal framework that enhances the probability of logical consistency. Additionally, the assumption that state actors make purposeful choices is more than compatible with standard approaches to security studies. But perhaps the most important reason for the application of this methodology to strategic issues is the conceptual foundation of almost all game-theoretic solution concepts, that is, the concept of an equilibrium outcome. The equilibria of a game model are uniquely appropriate for assessing the stability of competitive interstate relationships.

Equilibrium can be thought of as a state of decision-making stasis. An equilibrium is said to exist when no one player in a game has an incentive to switch to another course of action (i.e., to another strategy or policy). Thus, an equilibrium outcome is a natural measure of strategic stability. Conversely, when at least one actor has an incentive to abandon its strategy, either by attacking the other, demanding an alteration of the status quo, or more generally, by unilaterally adjusting its relationship with another actor, equilibrium cannot be said to exist; that is, the strategic relationship is unstable.

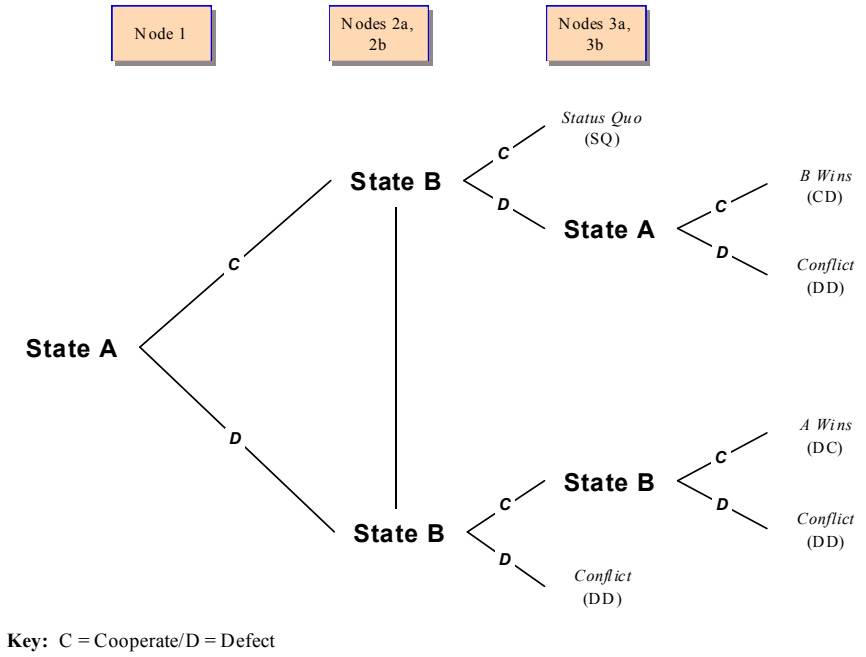
As it turns out, there are more than a few distinct definitions of a game-theoretic equilibrium, and some of these definitions are quite elaborate mathematically. The most well-known and commonly employed definitions are theoretically consistent, differing mainly in the circumstances to which they pertain. For example, the well-know definition of Nash equilibrium is the standard measure of rational behavior in static games of complete information like "chicken." By contrast, in dynamic games, rationality is gauged by subgame perfect equilibria in

games of complete information and by perfect Bayesian equilibrium in games of incomplete information. Nonetheless, leaving technicalities aside, the underlying notions are quite similar. Equilibria, however they are defined, indicate that no incentive exists for an actor to do anything other than what it is doing at the present moment and that it will persist in that behavior until something material changes. In other words, a stable strategic relationship can quite naturally be associated with a status quo that is in equilibrium, and conversely.

Knowing that a strategic relationship is stable—that is, that it is in a peaceful equilibrium—can be quite useful, not to mention reassuring, but knowing the conditions that give rise to that stability, and the conditions that might upset it, is even more important to decision-makers since the political and military policies they pursue sometimes are determinative.

For example, consider for now the simple game model, called the "generalized mutual deterrence game," depicted in Figure 3.6-2. At the start of this game, both players (i.e., States A and B) simultaneously choose to cooperate (C) or to defect (D). If both choose either C or D, the game ends, but if one chooses C and the other chooses D, the player choosing C is provided with another opportunity to defect, that is, to retaliate. If, at the end of the game, both players have chosen C, the *status quo* prevails; if both have chosen D, *conflict* (or war) results. If one player chooses C and the other D, the defecting player gains an advantage—either *A wins* or *B wins*.

The generalized mutual deterrence game is austere, but not quite as austere as chicken because it is dynamic. It has also been examined systematically under both complete and incomplete information and under a variety of preference configurations. It is, therefore, a much more useful model since it is readily adaptable to whatever empirical circumstances or analytic purposes are at hand. It is not the only model that can be used to study contentious interstate relationships. Other more elaborate, and perhaps more pertinent models exist. The generalized mutual deterrence game is simply an example of one possible strategic relationship, albeit one that is both important and fundamental.



Key: C = Cooperate/D = Defect

Figure 3.6-2. Generalized Mutual Deterrence Game.

In Figure 3.6-2, the vertical line connecting State B's moves at node 2 is called an *information set*. Information sets are simply graphical devices used to indicate what a player knows, or does not know, about prior choices at a particular decision point in a game. As drawn here, the information set joining nodes 2a and 2b indicates that when State B makes a choice at decision node 2 it does not know exactly what strategy State A has selected.

In analyzing this game model, one might be interested in the conditions that give rise to stable deterrence. These conditions are associated with the existence of an equilibrium outcome or strategy pair. It might not be surprising to learn that decision-makers may choose to cooperate with one another in this game rationally as long as each actor's deterrent threat is sufficiently credible (or believable). This is a straightforward conclusion that is consistent with intuition. What might also be surprising, however, is that these same conditions generally give rise to another equilibrium under which neither player cooperates. When this equilibrium is selected by the actors, the status quo is upset, and conflict (or war) is implied. While this outcome can be said to be stable in a game-theoretic sense, it is not the sense in which most strategic analysts think of stability.

The co-existence of two rational outcomes, one associated with mutual cooperation and the other with conflict, strongly suggests that dyadic strategic relationships may not be as stable as some strategic thinkers have concluded.[14] It also suggests that the policies that decision-makers adopt, and not just the so-called strategic balance, may make the difference between war and peace. As it turns out, in this illustrative example, conditionally cooperative (i.e., tit-for-tat) strategies, when they are implemented by both parties, offer the greatest prospects for mutual cooperation. This is a fairly robust theoretical finding that persists across a number of distinct game models.

It is interesting to observe that stable deterrence is much more likely to occur in one-sided or unilateral deterrence relationships than in those situations in which all the players have an incentive to overturn the status quo. This straightforward conclusion has some important policy implications: under certain conditions, ongoing conflicts are best settled unilaterally by a dominating actor rather than being negotiated bilaterally. There is systematic empirical support for this prescription: imposed settlements generally last longer than negotiated settlements, regardless of the circumstances under which the conflict take place.[15]

Conclusions

Game-theoretic models are a natural and intuitively satisfying framework in which to assess the stability of contentious interstate relationships. Current modeling efforts are much more realistic, and hence, potentiality much more useful for policy makers than the first wave of academic studies.[16] They are also easily modified to take account of shifting strategic conditions. The models can be used not only to assess the robustness of strategic relationships, but also to generate and evaluate competing policy prescriptions.[17,18]

↑ Back to the top

---

[1] Morgenstern, Oskar (1959). *The Question of National Defense*. New York: Random House.
[2] Morgenstern, Oskar (1961). "The Cold War Is Cold Poker." *New York Times Magazine*, February 5.
[3] Kahn, Herman (1960). *On Thermonuclear War*. Princeton: Princeton University Press.

[4] Kahn, Herman (1962). *Thinking About the Unthinkable*. New York: Horizon Press.
[5] Kahn, Herman (1965). *On Escalation*. New York: Praeger.
[6] Wohlstetter, Albert (1959). "The Delicate Balance of Terror." *Foreign Affairs*, 37: 211 – 234.
[7] Snyder, Glenn H. (1961). *Deterrence and Defense: Toward a Theory of National Security*. Princeton: Princeton University Press.
[8] Ellsberg, Daniel (1959). "The Theory and Practice of Blackmail." Lecture at the Lowell Institute, Boston, MA, March 10. Reprinted in Oran R. Young, [ed.], *Bargaining: Formal Theories of Negotiation*. Urbana, IL: University of Illinois Press, (1975).
[9] Ellsberg, Daniel (1961). "The Crude Analysis of Strategic Choice." *American Economic Review*, 51: 472 478.
[10] Schelling, Thomas C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
[11] Schelling, Thomas C. (1966). *Arms and Influence*. New Haven: Yale University Press.
[12] Powell, Robert (1990). *Nuclear Deterrence Theory: The Search for Credibility*. New York: Cambridge University Press.
[13] Zagare, Frank C. and D. Marc Kilgour (2000). *Perfect Deterrence*. Cambridge: Cambridge University Press.
[14] Waltz, Kenneth N. (1964). "The Stability of the Bipolar World." *Daedalus*, 93: 881 – 909.
[15] Senese, Paul D. and Stephen L. Quackenbush (2003). "Sowing the Seeds of Conflict: The Effect of Dispute Settlements on Durations of Peace." *Journal of Politics*, 65: 696 – 717.
[16] Zagare, Frank C. and Branislav L. Slantchev (2009). "Game Theory and Other Modeling Approaches." In Robert A. Denemark et al., [eds.], The *International Studies Compendium Project*. Oxford: Wiley-Blackwell, in press.
[17] Powell Robert (2003.) "Nuclear Deterrence Theory, Nuclear Proliferation, and National Missile Defense." *International Security*, 27: 86 – 118.
[18] Quackenbush, Stephen L. (2006). "National Missile Defense and Deterrence." *Political Research Quarterly*, 59: 533 – 41.

3.7. Theoretical Foundations of Partnerships for Economic Development (Robert Axelrod)

Author and Organization: Robert Axelrod, Ford School of Public Policy, University of Michigan
Contact Information: axe@umich.edu

Introduction

In recent years, the international development community has come to recognize that "despite the potential benefits of globalization and technological change, world poverty has increased and growth prospects have dimmed for developing countries during the 1980s and 90s."[1] This realization has led to a change in focus from donations from the rich to the poor—often with the imposition of strict conditions—to a new emphasis on partnership. (The theme of partnership can be traced back to 1969, when the Pearson Commission called for "a new partnership based on an informal understanding expressing the reciprocal rights and obligations of donors and recipients."[2,3]) For example, at the World Bank, the intention is

> *. . . to treat partnerships as an organic process, in which trust is built over time, in which steps are taken to weave a "fabric of sustainability"; and to consider how mutual accountability may be built, perhaps in the form of a contractual arrangement*[4].

The purpose of this essay is to help clarify the theoretical foundations of partnerships in a development context. The essay proceeds by explicitly taking into account that a donor and recipient may have divergent as well as complementary interests. Thus, a partnership between a donor and recipient can be analyzed in terms of a mixed-motive game between two players. In particular, the iterated "prisoner's dilemma" game is helpful for analyzing the theoretical foundations of partnerships. Analyzing partnerships from this theoretical perspective suggests implications for selecting a partner, setting up a partnership, choosing a *modus operandi*, building trust, achieving selectivity, and performing monitoring and evaluation.

The theoretical framework offered here is deliberately kept very simple. The simplicity of the framework helps to highlight some important principles. The goal is to promote insight about some fundamental issues rather than to provide a fully accurate rendition of all the complexities of partnerships for economic development.

How Partnerships Differ from Other Kinds of Relationships

A working definition of *partnership* is "a collaborative relationship between entities to work toward shared objectives through a mutually agreed division of labor."[5] While this working definition is not very precise, it does help distinguish partnerships from other forms of aid relationship.

1. A partnership is not a gift. A partnership aims at taking advantage of what the recipient, as well as the donor, can bring to the relationship. This can include local expertise, on-site workers, and a better understanding of priorities, needs, and constraints. Even more important, a partnership seeks joint "ownership" of the relationship and tries to build the capacity of the recipient government to undertake sustainable development.

2. A partnership is not a relationship based on conditionality as a one-off event. Such conditionality presumes that donors can impose conditions to coerce poor countries to do things they do not want to do in order to obtain resources they need. A partnership is consistent, however, with conditionality seen as a process over time with mutual accountability.[6] Some researchers even question whether strict conditionality has been effective at promoting development.[7] Partnership recognizes that both sides must be involved in defining the terms of the relationship.
3. Similarly, a partnership is not a principal-agent relationship between a donor and a recipient. In a partnership, the donor cannot prescribe the terms of the relationship in the way that an employer can specify terms of employment when hiring a worker. (The principal-agent framework is very useful for the examination of issues that arise from information asymmetries. For example, in partnerships for economic development the recipient often knows more than the donor about what is happening at the local level.)
4. A partnership is not a simply "team" activity. In an ideal sports team, everyone has exactly the same interest in winning, and everyone on the team either wins or everyone on the team loses. While the members of a partnership for economic development certainly have strong interests in common, they are also likely to have some divergent interests, as well.

Finally, although the formal terms of a partnership may be expressed in a valid contract under international law, the donor and recipient usually have no intention to use courts to resolve their conflicts. Instead, like nations bound by treaties, partners rely mainly on each other's need to maintain a good reputation to secure future agreements.

A Partnership as a Prisoner's Dilemma

In order to understand the theoretical foundations of a partnership, it is helpful to use game theory. In game theory, a relationship is defined by (1) identifying the players, (2) specifying the choices they have, and (3) spelling out the consequences of these choices in terms of the payoffs to the players. Analyzing the strategies available to the players reveals what is needed to foster mutually beneficial outcomes.

In the case of a partnership for economic development, the game can be specified as shown below. (Game theory has previously been used to model other aspects of the development aid. For example, conditionality has been modeled as a principal-agent game,[8] and the enforcement of debt service has been modeled as a game with a choice to default.[9]

1. The players are the donor and the recipient. The donor might be an international financial institution such as the World Bank, a non-governmental organization, a foundation, a group of countries such as the OECD, or a single country. The donor might even be a coalition of diverse contributors. The recipient in a partnership is often the government of a single country, although it might be a group of countries, a ministry in charge of some sector of a single country, or even a local organization.
2. The available choices can be quite complex. To highlight the theoretical aspects of a partnership, the choices can be reduced to just two: cooperate or defect. Choosing to cooperate at a given point in time means that the player performs the agreed upon terms of the partnership promptly and fully. Choosing to defect means that the player fulfills its obligations less than completely, or slower than agreed upon.

3. The consequences depend upon what each player has chosen. The basis of a partnership is that there is mutual gain to be had by working together. Thus, if both cooperate, both do well. However, there is typically a temptation to defect. Some reasons why a donor might be tempted to avoid some of its commitments include a desire to increase the amount of tied aid, a desire to reduce costs once it got credit for "flag raising," and a reluctance in practice to support leadership by the recipient government. All three concerns were voiced in the workshop *Making Partnerships Work on the Ground.*[10] On the other hand, a recipient might be tempted to use some of its limited human and financial resources that were supposed to be devoted to the partnership for other pressing needs. In general, each side would typically prefer that the other fulfill its commitments, but have freedom to choose which of its own commitments will be fulfilled. Finally, if both sides defect by ignoring their commitments, neither does well because neither side gets the advantages of the partnership. These consequences are represented in the payoff matrix of Table 3.7-1.

Table 3.7-1. A Partnership as a Prisoner's Dilemma.

		Donor	
		Corporate	***Defect***
Recipient	***Corporate***	R = 3, R = 3	S = 0, T = 5
	Defect	T = 5, S = 0	P = 1, P = 1

R = Reward for mutual cooperation, T = Temptation to defect, S = Sucker's payoff, P = Punishment for mutual defection. The first number in each cell is the payoff to the Recipient, and the second number is the payoff to the Donor.

The payoffs need not be symmetric, or even comparable. What matters is the rank order of each player's payoffs. The payoff matrix of Table 3.7-1 is a fundamental tool for understanding the theoretical foundations of partnerships. The numbers are meant to be suggestive only. What matters is their rank order. Since the payoffs reflect the incentives of the players, each side wants to get as high a payoff as possible. For the reasons just described, the rank order of the payoffs in a partnership relationship (from best to worst) are: (1) the Temptation payoff for defecting when the other cooperates, (2) the Reward for mutual cooperation, (3) the Punishment for mutual defection, and (4) the Sucker's payoff for cooperating when the other defects. (Other games arise with other payoff structures. For example, if mutual cooperation is preferred to unilateral defection, there is no temptation to defect. The only issue in the resulting Assurance Game is for the players to assure each other that they are sensible. This essay deliberately focuses on the harder case of the prisoner's dilemma, in which the players need to overcome their short run temptation to exploit the situation at the other's expense. With payoffs in this order, the game is a prisoner's dilemma. Inspection of Table 3.7-1 reveals that in the prisoner's dilemma, each side has an incentive to defect, no matter what the other side does. However, if both defect, both do worse than they could have done had they both cooperated. That is the dilemma. (The original story involves two prisoners who are being interrogated by the police. Each is given an incentive to confess to help convict the other, but if both confess, the police do not need either's testimony. Technically, the game is a "prisoner's dilemma" if the payoffs for both players satisfy not only $T > R > P > S$, but also $R > [T + S]/2$.)

Note that the payoff matrix presents the game as if the relationship is completely symmetric. Actually, this is unnecessary. The payoffs need not even be comparable between players. All that

is necessary is that the payoffs for each player separately fulfill the conditions of the prisoner's dilemma, and this is likely to be true in almost any development partnership. Game theory provides methods for analyzing a very wide range of strategic problems. Since the prisoner's dilemma occurs in many contexts, it is one of most thoroughly studied games. (A good introduction to game theory is Dixit and Skeath [1999].[11] A more advanced game theory text is Fudenberg and Tirole [1992].[12] For reviews on the theoretical and empirical literature on the prisoner's dilemma, see Axelrod and Dion [1992][13] and Hoffman [2000][14]).

If the Prisoner's Dilemma is played only once, there is no getting around the problem that each side has an incentive to defect, no matter what the other side does. This would result in a failure to achieve the benefits of mutual cooperation. However, if the game is played over and over, the strategic possibilities are very much richer. Fortunately, this is exactly the situation in a partnership where the relationship is not based on a single choice, but an ongoing interaction between the donor and recipient.[15]

In the iterated prisoner's dilemma, many strategies are possible since each player can make its current choice contingent on what has previously happened in the relationship. Cooperation theory is the study of what strategy a sophisticated player should use and what conditions will sustain mutual cooperation. To study these questions, Axelrod[16] organized a series of round robin computer tournaments in which both experts and amateurs were invited to submit a computer program that embodied their favorite strategy. In both rounds of these tournaments, there was one strategy that did best of all. This was the strategy called "tit for tat." Tit for tat cooperates on the first move and then does whatever the other side did on the previous move.

By cooperating on the first move and then doing whatever the other player did on the previous move, tit for tat managed to do well when interacting with a wide variety of more or less sophisticated strategies. Tit for tat's robust success was due to its being nice, provocable, and forgiving. Its niceness means that it was never the first to defect. Not being the first to defect prevented it from getting into unnecessary trouble. Its provocability meant that it retaliated when the other side defected. Provocability discouraged the other side from persisting in defection. By cooperating again as soon as the other player did, the forgiveness of tit for tat helped restore mutual cooperation.

Tit for tat is not the only strategy that can succeed by being nice, able to be provoked, and forgiving. The important thing is reciprocity. For example, the strategy of tit for two tats needs two defections to be provoked. One advantage of tit for tat over other forms of reciprocity is that it is easy for the other player to recognize. Once recognized, it is clear that the best way to respond to a tit for tat player is to cooperate.

An especially interesting aspect of tit for tat's success is that it never did better than any of the other strategies it played with. In fact, it cannot do better since it never defects more often than the other player. This curious fact highlights the difference between a non-zero sum game like the iterated prisoner's dilemma and a zero-sum game like chess or basketball. The single most important theoretical lesson of game theory in general, and the iterated prisoner's dilemma in particular, is that most interactions are not zero-sum games in which one side's loss is the other side's gain.

Perhaps the biggest surprise from the tournaments is the complementary point that to maintain stable cooperation in an iterated prisoner's dilemma, it is necessary to be provocable. In fact, the

sooner one responds to a defection by the other side, the easier it is to restore mutual cooperation by establishing that defection does not pay.

A basic theoretical result is that the potential for mutual cooperation depends on "the shadow of the future," which is measured by *w*, the rate at which the payoff of the next move is discounted relative to the present. Cooperation based on tit for tat is stable if and only if *w* is at least as great as the larger of (T-R)/(T-P) and (T-R)/(R-S).[17] Thus, as long as the shadow of the future is large enough, when one player is using tit for tat, the other player can do no better than using tit for tat as well, and this establishes mutual cooperation on a stable basis.

An unavoidable feature of partnerships is the possibility of misimplemenation or misunderstanding. With misimplementation, a player intends to do one thing but actually does another. For example, a leader may decide to cooperate, but the rest of the organization may not properly implement the leader's choice. Misunderstanding occurs when a player does one thing, but the other player perceives it as something else. In either case, a single error can cause an echo of unending trouble between two tit for tat players.[18] To deal with either form of error, a player should be slightly more generous than simple tit for tat. For example, a player might be generous at the 10% level, meaning that 10% of the time that the player would otherwise defect, it cooperates instead. If a player realizes that it was responsible for the trouble due to its own misimplementation, it could even avoid responding to the other player's defection after its own unintended defection. This is called "contrition." Deductive theory and computer simulation demonstrate that a generous tit for tat strategy is effective at dealing with either type of error. Adding contrition further minimizes the disruptive effects of misimplementation.[19]

Implications for Partnerships

1. Implications for Selecting a Partner

The principle of "selectivity" is usually taken to mean that program aid should be reserved for countries that have adopted efficient pro-development policies for themselves.[20] Cooperation theory based on the iterated prisoner's dilemma suggests a related but somewhat different set of considerations in choosing a partner.

1. The more favorable the payoff structure, the easier it will be to achieve and sustain mutual cooperation. This implies that a potential donor or recipient consider more than just the potential gains from mutual cooperation. In selecting a partner, one should also consider the magnitude of the temptation to defect that the other will face.
2. Similarly, the stronger the shadow of the future, the easier it will be to achieve and sustain mutual cooperation. A good partner is one who has a long enough time horizon that it will not want to exploit the relationship for immediate gain.
3. A partner needs the organizational and managerial capacity to fulfill its responsibilities. It should not be overextended with other commitments.
4. The potential for follow-on projects, especially at a larger scale, makes a partner a good prospect. This prospect of a follow-on project, especially one at a larger scale, makes success in the current project an investment in the continuing relationship.
5. Selecting a partner who has a good reputation for cooperation has two benefits: it rewards good behavior in the past, and it increases the chances that the current partnership will succeed.

2. Implications for Setting up the Partnership

Once a donor and recipient have agreed to set up a partnership, the terms of the agreement need to be worked out. Theoretical and empirical research suggests that there are three aspects of an ongoing relationship that facilitate cooperation based upon reciprocity: clarity of obligations, promptness of feedback, and the institutionalization of reciprocity. In setting up a partnership, the participants should take care that these three conditions will be met throughout the partnership.

1. Clarity of obligations is needed so that a cooperative choice can be distinguished from a defection. If the partners are not able to agree on whether a given action was consistent with their obligations, they are likely to get into mutual recriminations. Unfortunately, clarity of obligations competes with the flexibility of implementation. Until a deep bond of trust develops between the two sides, it may be better to err on the side of clarity of obligations rather than flexibility in setting up a partnership.
2. Promptness of feedback means that each side can monitor the behavior of the other so that any problems can be addressed before they become major grievances. Promptness of feedback strengthens the shadow of the future by allowing a timely response to a perceived problem. Thus, in setting up the partnership, the obligations of the two sides should be chosen so that their performance (or nonperformance) is observable as quickly as possible.
3. The institutionalization of reciprocity can help by specifying in advance just what form of review and redress is open if either side has a complaint about the other. Maxwell and Riddell[21] even suggest that what is needed is some kind of WTO agreement for aid administration. In the absence of such agreements, even informal norms can help prevent problems from echoing out of control.[22]

3. Implications for Modus Operandi

Once a partnership is underway, the question of modus operandi deals with how the partners should behave in their relationship. Obviously, mutual cooperation is the ideal. But what should one partner do if the other does not cooperate by fulfilling its commitments in a timely manner? Theory and simulations show that the sooner one is provoked, the smaller the response needs to be in order to be effective.[23] Not responding promptly to an uncalled for defection risks sending the wrong signal. The longer defections go unchallenged, the more likely it is that the other player will draw the conclusion that defection can pay. Furthermore, the more strongly this pattern is established, the more difficult it will be to break.

The importance of provocability is increased when the relationship is being observed by others who are or may become partners with either of the participants. The reason is that reputation matters. This gives a recipient some leverage over a donor. For example, if a recipient is provoked, it cannot only withhold part of its contribution to the partnership, but it can also threaten to criticize the donor in forums that are important to the donor's reputation and self-image.[24]

Provocability is also important for the donor. If, for example, a lending organization gets a reputation of allowing its partners to ignore their commitments without consequences, it may be difficult to hold other partners accountable in the future. This is the reason why the Grameen Banks, for example, are so insistent on prompt repayment of their microcredit loans (Bornstein,

1996).[25] The Grameen Bank is willing, however, to allow late payment when there is a large external shock beyond the foresight or control of the recipient, such as an unusually devastating monsoon. (Another reason for adjusting the terms of the partnership would be if the premises of the project were revealed to be inaccurate. Suppose, for example, the partnership was designed to supply water from wells. If the wells do not give the expected yield, the partners should renegotiate their commitments, perhaps switching to laying pipes. The new agreement would redefine what counts as cooperation or defection in the partnership.)

In general, reciprocity is effective. In addition to being provocable, it is important to restore mutual cooperation as soon as possible by being forgiving when the other side returns to cooperation.

In a partnership for economic development, a recipient government may not always have the capacity to implement the decisions it reaches. In fact, even donors sometimes make promises that they fail to keep despite their best intentions. Moreover, there may be cultural differences between the donor and recipient that can contribute to misunderstandings. Thus, an uncalled-for defection might be due to misimplementation or misperception. In either case, it pays to be somewhat generous to prevent echoing of conflict. This can be done by not responding to some small proportion of the other side's defections, or by responding in a somewhat milder manner than usual. In either case, however, too much generosity invites exploitation. Also, if the other side's defection was simply a response to one's own unintended defection, it pays to be contrite by not letting the conflict continue.

4. Implications for Building Trust

Trust is crucial for establishing cooperation based on mutual obligation. (This point is confirmed by a study of 17 joint ventures in the United Kingdom and Malaysia.)[26]

Since building trust takes time, it is likely to take place mostly between projects rather than within projects. In theoretical terms, this suggests the importance of distinguishing two different time scales of a partnership: a particular project and the long-term relationship. A project is typically based on a single agreement lasting several years. Within this time scale there may be many "moves" as the partners fulfill their obligations. On a longer time scale, each project can be regarded as a single "move" in the relationship between the partners. An example is the relationship between Conservation International and Starbucks.[27] In the first partnership project, Starbucks committed $150,000 over three years to promote shade-grown coffee. At the end of this period, Conservation International publicly praised Starbucks for helping farmers and for improving the natural environment. With this endorsement and with consumer acceptance of shade grown coffee, Starbucks signed a new three-year partnership quadrupling its commitment.

Theoretically, playing a game on a longer time scale has the advantage of providing more reliable feedback, but the disadvantage of having slower feedback. For the sake of building trust over time, the partners need a long time horizon as embodied in a strong shadow of the future. Theoretical work confirms that starting with small stakes and raising them as the relationship develops helps the partners avoid having to take "leaps of faith."[28]

5. Implications for Achieving Selectivity

Donors need to be selective about the partnerships they enter, lest they spread their managerial and financial resources too thinly. Cooperation theory suggests some guidelines for deciding just

how to be selective. We have already discussed the characteristics of a good partner. Here we add some strategic considerations based on how reciprocity can work at the project level.

1. If the recipient has not fulfilled the obligations of the current project, the donor can decide not to expand the project in the next phase or not to enter into new projects with this recipient. This has the advantage of helping to establish that cooperation needs to be mutual. (Numerous simulation studies show that the ability and willingness to leave an unsatisfactory relationship actually fosters cooperation in the long run. For a review see Axelrod [2000, p. 146f]).[29] Put another way, to overcome a recipient's temptation to exploit the relationship, the donor needs to be provocable at the long time scale as well as the short time scale.
2. To avoid retreating from a country in need, a donor can be selective about the sector rather than the country. Thus if a project in one sector has not worked well, the donor may switch to a project in another sector of the same country. Assuming that the decision makers are at least somewhat different in the two sectors, this can be an effective strategy. (To the degree that aid is fungible across sectors, the effectiveness of this approach is reduced.)[30]

6. Implications for Monitoring and Evaluation

Monitoring is the continuous assessment of project implementation, while *evaluation* is the periodic assessment of the project. In the aid community, the evaluation process is intended to serve two functions: institutional credibility and organizational learning. For institutional credibility, the acid test is performance on the ground. For this reason, accountability for result-based assistance has become a fundamental imperative.[31] For organizational learning, the goal is not only to improve individual programs, but to make the results available to the global evaluation community.

From the point of view of cooperation theory, monitoring and evaluation can play an active role in sustaining the conditions needed for mutual cooperation. As discussed earlier, to achieve stable cooperation in an iterated prisoner's dilemma, each player needs to know in a timely fashion whether the other is fulfilling its commitments. In particular, knowing when the other side has defected is needed for provocability and forgiveness. Moreover, knowing when one's own side has defected (perhaps unintentionally) is needed for contrition.

Determining whether the partners are doing what they promised is a different task than the traditional role of evaluation of how well the partnership is achieving its goals. Periodic evaluation can be joined with continuous monitoring to assure the partners when mutual cooperation is being achieved, and to warn them when it is not.

Theoretically, the stability of cooperation relies on *both* partners having timely information about the actions of the other. In practice, the donor is more likely to monitor and evaluate the recipient than the other way around.

> *Whereas the behaviour and performance of low-income developing countries is measured and assessed in ever-increasing detail within the international community, the behavior and performance of the donor "partners" receives only cursory attention, except at the aggregate level which is of little operational usefulness to individual recipients.*[32]

A step in the right direction is the World Bank's recently established Inspection Panel which investigates claims by local stakeholders that the Bank may have failed to adhere to its own

operational policies, procedures in design, appraisal, or implementation of ongoing or new operations.[33]

The unique position of the World Bank allows it to play an important role in economic development activities even when it is not a direct participant. The World Bank's multinational character, size and experience allow it to play an increasingly active role as catalyst in identifying and promoting promising activities.[34] Cooperation theory suggests that even after a partnership is set up, there are valuable services that could be performed by a credible outside organization such as the World Bank. Potential services include the following:

1. Promoting accounting standards to allow timely feedback on the performance of partners.
2. Certifying that partners are being responsible in fulfilling their obligations.
3. Sponsoring institutional mechanisms for redress if one side of a partnership has a complaint about the other.

Conclusion

A partnership for economic development includes divergent interests, as well as shared objectives. Analyzing a partnership as an iterated prisoner's dilemma reveals the theoretical conditions needed to sustain mutual cooperation. These conditions lead to important implications for how to promote the effectiveness of such partnerships. The implications concern selecting a partner, setting up a partnership, choosing a modus operandi, building trust, achieving selectivity, and performing monitoring and evaluation.

↑ Back to the top

---

[1] Maxwell, Simon and Tim Conway (2000). "Perspectives on Partnership," World Bank OED Working Paper No. 6. Available at www.worldbank.org/html/oed.

[2] Pearson, Lester (1969). *Partners in Development: Report of the Commission on International Development* (NY: Praeger).

[3] Helleiner, Gerry (2000). "Towards Balance in Aid Relationships: Donor Performance Monitoring in Low-Income Developing Countries," *Cooperation South* (UNDP), no. 2, p. 21-35. Available at www.ecdc.net.cn/events/co_south200002/4.htm

[4] Maxwell, Simon and Tim Conway (2000). "Perspectives on Partnership," World Bank OED Working Paper No. 6. Available at www.worldbank.org/html/oed.

[5] World Bank (1998). "Partnership for Development: Proposed Actions for the World Bank," SecM98-421, May 28, 1998.

[6] Branson, William and Nagy Hanna (2000). "Ownership and Conditionality," World Bank OED Working Paper No. 8. Available at www.worldbank.org/html/oed.

[7] Gilbert, Christopher L., Andrew Powell, and David Vines (2000). "Positioning the World Bank," in Christopher L. Gilbert and David Vines (eds.), *The World Bank: Structure and Policies* (Cambridge, UK: Cambridge University Press).

[8] Squire, Lyn (2000). "Why the World Bank Should be Involved in Development Research," in Christopher L. Gilbert and David Vines (eds.), *The World Bank: Structure and Policies* (Cambridge, UK: Cambridge University Press).

[9] Anderson, R. W., C. L. Gilbert, and A. Powell (1991). "Securitising Development Finance: The Role of Partial Guarantees and Commodity Contingency," in T. Privolos and R. C. Duncan (eds.), *Commodity Risk Management and Finance* (Oxford: Oxford University Press).

[10] Making Partnerships Work on the Ground (1999). *Making Partnerships Work on the Ground* (Stockholm, Sweden: Ministry of Foreign Affairs), pp. 5 and 14.
[11] Dixit, Avinash and Susan Skeath (1999). *Games of Strategy* (NY: Norton).
[12] Fudenberg, Drew and Jean Tirole (1992). *Game Theory* (Cambridge, MA: MIT Press).
[13] Robert Axelrod and Douglas Dion, "The Further Evolution of Cooperation," *Science*, 242 (9 December 1988), pp. 1385-1390.
[14] Hoffmann, Robert (2000). "Twenty Years on: The Evolution of Cooperation Revisited," *Journal of Artificial Societies and Social Simulation*, vol. 3, no. 2. Available at www.soc.surrey.ac.uk/JASSS/3/2/forum/1.html
[15] Branson, William and Nagy Hanna (2000). "Ownership and Conditionality," *World Bank OED.* Working Paper No. 8. Available at www.worldbank.org/html/oed.
[16] Axelrod, Robert (1984). *The Evolution of Cooperation* (NY: Basic Books).
[17] Axelrod, Robert (1984). *The Evolution of Cooperation* (NY: Basic Books).
[18] Molander, Per (1985). "The Optimal Level of Generosity in a Selfish, Uncertain Environment," *Journal of Conflict Resolution*, vol. 29. pp.611-618.
[19] Wu, Jianzhong and Robert Axelrod (1995). "How to Cope with Noise in the Iterated Prisoner's Dilemma," *Journal of Conflict Resolution*, vol. 39, 183-189.
[20] Hopkins, Raul, Andrew Powell, Amlan Roy, and Christopher L. Gilbert (2000). "The World Bank, Conditionality and the Comprehensive Development Framework," in Christopher L. Gilbert and David Vines (eds.), *The World Bank: Structure and Policies* (Cambridge, UK: Cambridge University Press).
[21] Maxwell, Simon and Roger Riddell (1998). "Conditionality or Contract: Perspectives on Partnership for Development," *Journal of International Development*, vol. 10, 257-268.
[22] Axelrod, Robert and Robert Keohane (1985). "Achieving Cooperation Under Anarchy," *World Politics*, vol. 38, 226-54. Reprinted in Kenneth A. Oye (ed.), *Cooperation Under Anarchy* (Princeton, NJ: Princeton University Press, 1985).
[23] Axelrod, Robert (1984). *The Evolution of Cooperation* (NY: Basic Books), pp 119 and 184f.
[24] Fox, Jonathan A. (2000), "The World Bank Inspection Panel: Lessons from the First Five Years," *Global Governance*, vol. 3, pp. 279-318.
[25] Bornstein, David (1996). *The Price of a Dream* (New York: Simon and Schuster).
[26] Butler, Richard and Jas Gill (1999). "The Dynamics of Trust in Partnership Formation and Operation: Project Summary," Mimeo. University of Bradford Management Centre.
[27] Purcell, Randall (2000). "Good Practices And Regional Programs" (Washington, DC: World Bank), 66f.
[28] Roberts, Gilbert and Thomas N. Sherratt (1998), "Development of Cooperative Relationships through Increasing Investment," *Nature*, vol. 394, p. 175-179.
[29] Axelrod, Robert (2000). "On Six Advances in Cooperation Theory," *Analyse & Kritik*, vol. 22, 130-151, 146f.
[30] Devarajan, Shantayanan and Vinya Swaroop (2000). "The Implications of Foreign Aid Fungibility for Development Assistance," in Christopher L. Gilbert and David Vines (eds.), *The World Bank: Structure and Policies* (Cambridge, UK: Cambridge University Press).
[31] Picciotto, Robert (forthcoming). "Evaluation in the World Bank: Antecedents, Methods, and Instruments," in Jan-Eric Furubo, Ray C. Rist, and Rolf Sandahl (eds.) *International Atlas of Evaluation* (Piscataway, NJ: Transaction).
[32] Helleiner, Gerry (2000). "Towards Balance in Aid Relationships: Donor Performance Monitoring in Low-Income Developing Countries," *Cooperation South* (UNDP), no. 2, p. 21-35. Available at www.ecdc.net.cn/events/co_south200002/4.htm
[33] Fox, Jonathan A. (2000), "The World Bank Inspection Panel: Lessons from the First Five Years," *Global Governance*, vol. 3, pp. 279-318.

[34] Gilbert, Christopher L., Andrew Powell, and David Vines (2000). "Positioning the World Bank," in Christopher L. Gilbert and David Vines (eds.), *The World Bank: Structure and Policies* (Cambridge, UK: Cambridge University Press)

3.8. Process Query System as a Framework for Modeling and Analysis of Regional Stability (George Cybenko, Douglas Madory)

Authors and Organizations: George Cybenko, Douglas Madory; Dartmouth College
Contact Information: gvc@dartmouth.edu; douglas.c.madory@dartmouth.edu

Abstract

During the past five years, a study has been conducted of situational-awareness problems that have arisen in multiple application domains including computer security, autonomic computing, sensor networks, video tracking, and social network analysis. The variety of applications investigated suggests that a common analytic foundation underlies many such problems. As a result, a general purpose framework has been developed for modeling situational-awareness across multiple sensor types and environments. This framework is called Process Query System (PQS). In addition, the Human Behavioral Modeling Language (HBML) has been developed as a schema to support behavioral modeling and analysis by large-scale computational systems across a variety of domains. The PQS framework, which is based on process detection, together with HBML, can enable modeling and analysis of regional stability by facilitating the comparison of competing theories or aggregating the conclusions of multiple domain specific-analytical approaches.

Introduction

Regional stability can be viewed as an attribute of the social, political and economic dynamics of a geographic area. As such, regional stability is measured and forecasted in terms of macro-level behavioral properties of human systems, which in turn may be dependent on and derived from micro-level human behaviors.

Behavioral modeling of systems, both natural and engineered, is an area of growing interest.[1,2,3,4] In particular, quantitative behavioral modeling of humans is of interest at all scales and for a multitude of reasons.[5,6] Nonetheless, in at least some disciplines, the notion of what a behavior is and how it is represented has not been broadly formalized.[7] To avoid confusion or misunderstanding, this paper briefly reviews an approach to behavioral modeling and its uses.

Atomic activities are the basic elements of behaviors in this approach, in which behaviors are composed of activities. The granularity of an activity is defined by the modeler, not by the underlying application so that an activity in one context can be a behavior in another. For instance, "voting" can be viewed as an activity atomic element in the context of modeling the democratic process. On the other hand, "voting" can be view as a behavior comprised of atomic activities (such as "transit to a polling station," "voter identification/authentication," "casting confidential votes," or "tallying votes").

In the framework, *behaviors* are collections of atomic activities. Research efforts have led to three distinct categories of behaviors.[8]

- **A 0th-order behavior** is merely a list of activities. For example, people commonly list their hobbies on their resumes: cycling, skiing, chess, etc. Such a list of activities, with no

indication of frequencies or interdependencies, constitutes a 0th order model of behavior in this terminology.

- **A 1st order behavior** is an enumeration of activities together with their relative frequencies, such as a person saying that he or she goes to the bank once a week, watches television for 2 hours a day, or spends three hours preparing for a one hour-lecture.
- **A 2nd order behavioral** model consists of causal dependencies between activities, described in either deterministic or statistical terms. For example, if a person says that he or she has coffee at the end of every meal, this relates two activities to each other in a causal manner. That is the probability of having coffee, given that a meal was eaten, is 1.0. Statistical assertions, such as the probability of coffee given a meal, is some number less than 1 and is also within the scope of a 2nd order behavior.

Roughly speaking, 0th, 1st and 2nd order behavioral models can be regarded, respectively, as collections of basic activities, relative frequencies of basic activities, and dependencies between activities.

Furthermore, this review of behavioral model across various domains has suggested that there are three major uses of behavior models.

- **Enable prediction.** The first and most ambitious use is to enable prediction. Specifically, given a behavioral model and evidence that a system occupies an estimated state according to the behavioral model, prediction strives to forecast the next state or to assign probabilities to the next state.
- **Characterize a system under observation.** A second use of behavioral models seeks to characterize, or to assign to a clustered class, a system under observation. For example, investment records may indicate that an individual is "conservative," as opposed to "risky," in his or her investment decisions. Such a characterization may be possible, while an accurate prediction or forecast may not be. For this reason, that prediction or forecasting can be more challenging than characterization.
- **Detect change or anomaly.** The last major use of behavioral modeling is change or anomaly detection. In this use, a system's behavior is first empirically modeled and characterized, and then it is monitored to detect when the behavior deviates sufficiently from its original characterization.

A large collection, perhaps all, regional stability models can be formulated in terms of these different orders of behaviors and behavioral uses. This research has been focused on (1) generic technologies, theories, and software systems for modeling behaviors in these ways, (2) representing behaviors uniformly so that they can be shared and made operational, and (3) detecting known or learning new behaviors from empirical observations.

Below a description is provided of a software system for detecting behavioral patterns from a stream of observations, namely process query systems (PQSs),[9] as well as an XML framework for representing various orders of behaviors in general systems, be they human, engineered, standalone or networked.[10] Operational, computational approaches to modeling regional stability require at least these two technical capabilities. Specifically, models of regional stability should be and could be codified in standardized schema (e.g., HBML) to allow them to be used simultaneously in simulations and analyses. Computational engines (e.g., PQS) should be created

for associating evidence, events, observations and indicators with hypothesized models to estimate which behavioral models are instantiated and to determine the status of those models.

Process Query Systems

The ability to instrument and collect data from numerous disparate environments has increased dramatically in recent years. Computer systems and networks now routinely include various performance monitors, firewalls, intrusion-detection systems, and application-logging agents. In addition, sensor networks have been deployed in physical environments to record acoustic, seismic, infrared, video, electromagnetic, and other types of measurements.[11, 12] These networks also can monitor and extensively archive communications, as well as financial and social transactions among large communities of people and organizations.[13, 14]

While the technology to instrument such environments has matured significantly, the ability to obtain situational awareness from the resulting enormous data streams has not kept pace. Sensed-event data are often displayed in relatively raw formats, leaving analysis and interpretation up to human operators, which ultimately is not scalable.

Neither traditional database technology nor rule-based expert systems have proved to be up to the task of closing the gap between low-level sensor events and high-level situational awareness. Database technologies, including extensions to data-stream processing, effectively store, index, and retrieve sensor reports, but they do not provide analysis beyond rudimentary report generation. On the other hand, decision trees and logical-rule processing inherit the well-known brittleness and scalability problems associated with expert systems. A new approach for extracting situational awareness from sensed data is therefore needed.

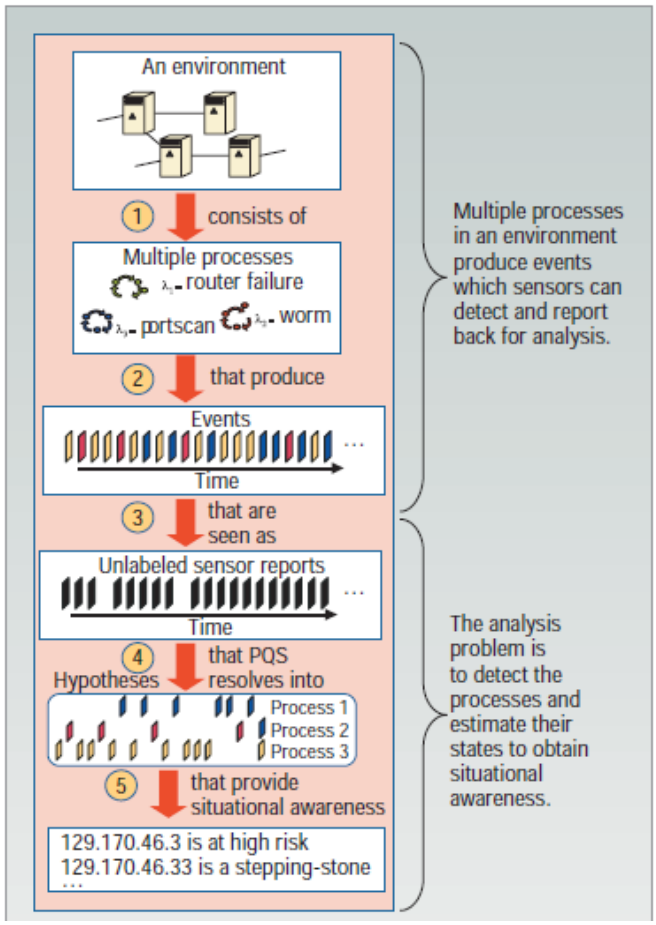


Figure 3.8-1. PQS Framework for Network Security.

During the past five years, situational-awareness problems have been studied that have arisen in multiple application domains including computer security, autonomic computing, sensor networks, video tracking, and social network analysis. The variety of applications suggests that a common analytic foundation underlies many such problems. A general-purpose framework has been developed for modeling situational awareness across multiple sensor types and environments. Using this framework, which is based on process detection, a PQS has been implemented and tested in several applications. The PQS modeling framework consists of several steps, as illustrated in Figure 3.8-1.

Step 1. A sensed environment consists of processes with distinct states, dynamics, and observables. In the case of computer security, computer attacks are the processes. Possible states of these attack processes include reconnaissance, intrusion, exploitation, and data exfiltration. Attack processes change their states over time as determined by the attack model's dynamics.

Step 2. The states and dynamics of processes in an environment cannot be observed directly, but they do produce observable events and artifacts.

Step 3. The sensing infrastructure detects events and communicates them back to an analysis center. The events provide evidence of the processes' states but are not identical to them. For example, it is usually impossible to know precisely about a computer attack's abstract internal state—only sensing and detecting observable artifacts of attacks such as network packets, file system changes, and system behaviors, including processor utilization, can be expected.

Step 4. Sensor observations about the hidden, internal processes' states often consist of inconclusive and noisy evidence, which makes gaining situational awareness of the environment a challenge.

Step 5. Knowledge of the processes and their states within an environment provides the desired situational awareness. Whether in computer security, sensor networks, social network analysis, or modeling regional stability, a key complicating factor in these application domains is that many active processes are possible, and observations of the processes are interwoven, ambiguous, and unlabeled. Some observations can be missed, while others can enter the data stream as noise.

The table below summarizes how the PQS framework applies to various application domains.

Table 3.8-1. The PQS Framework as it Applies to Various Application Domains.

Application environment	***Processes***	***States***	***Observables***
Computer attacks	Host and network behaviors	Normal, scanned, infected, failed, trusted, hostile	Tripwire, applications logs, Snort alerts, host-based logs, file access, user access
National border and physical perimeter defense	Moving objects (people, animals, vehicles)	Position and velocity	Video, infrared images, acoustic data, seismic data, electromagnetic data
Geographic region	Airborne agent diffusion and drift	Releases at times T, locations L	Sensor detection of an airborne agent
Identity theft and management	Consumer, bank, ID thief's activities	Normal, phished, exploited	Credit reports, Web postings, database breaches, pretexts
Social networks	Business and social activity	Stages of a business or social process	Communications and transactions

PQS is software that uses various algorithms to generate and manage hypotheses about an observed event sequence, given a set of process models. It consists of the following components:

Track and hypothesis extension. Given a new event, this component updates the tracks in the current set of hypotheses. Specifically, a track can be extended if its current hypothesized state can transition to a next state that can generate the currently observed event. The new tracks can

be instantiated to accommodate the newly observed event, which creates a new set of hypotheses.

Hypothesis scoring. Because the number of hypotheses can increase exponentially, this component ranks them according to a score that measures the merit of tracks within the hypothesis or some other metric. For example, Occam-type rules can assign high scores to the simplest hypotheses, or, if the models are hidden Markov models (HMMs), a Viterbi-type decoder can score hypotheses by the likelihood of their tracks.

Hypothesis management. This component keeps hypotheses with the best scores and scores above a select threshold. Possible approaches for implementing this component include (1) applying heuristics that simply keep the highest scoring hypotheses, (2) clustering hypotheses and keeping exemplars from each cluster of hypotheses, and (3) maintaining a probability distribution over all possible hypotheses by updating and sampling the distribution using Markov chain Monte Carlo (MCMC) techniques.[15,16]

Situation evaluation. This component uses risk-assessment or variance-estimation computations to address ambiguous situations in which multiple hypotheses can be consistent with the observed events at any given time. For example, given multiple possible explanations for a sequence of events, it might assign reporting priority to those hypotheses that present a higher risk within the environment.

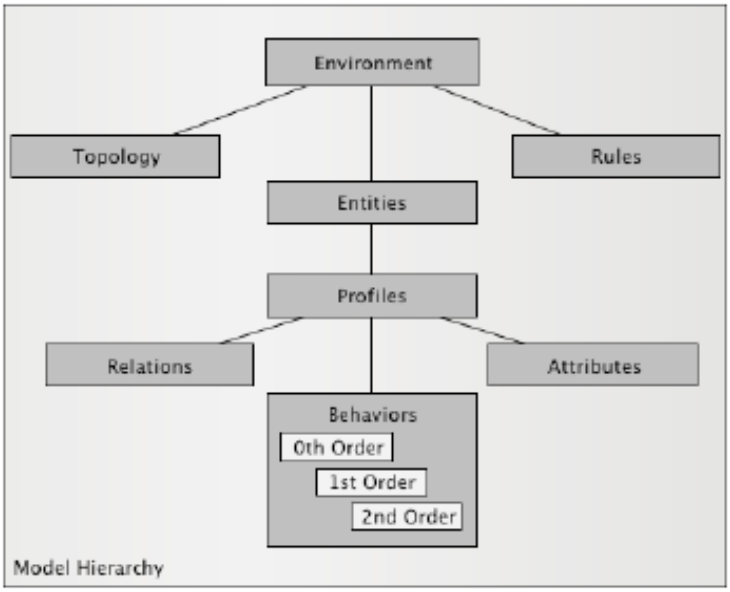


Figure 3.8-2. HBML Object Hierarchy.

The power of the PQS approach derives from the separation of the models from the detection logic. Constructing a decision tree or a rule set to derive the possible hypotheses for a series of observations and a small set of models, for example, is not difficult. However, how the decision tree or rule set would have to be changed if other models were added to the environment must be considered. In the PQS framework, if new model descriptions are added, the algorithmic steps would operate as usual; in contrast, without PQS, the corresponding rule sets and decision trees would have to be extensively and carefully revised to maintain the ability to disambiguate between the different possible observation sequences and resulting hypotheses. The strength of PQS is that it automates the logic of the detection and disambiguation computation, while decision trees and expert system-type rule sets must encode both the models and the detection logic simultaneously.

Human Behavioral Modeling Language (HBML)

Human and machine behavioral modeling and analysis are active areas of study in a variety of domains of current interest including analyzing political and social regional stability. The basic concept of a "behavior" is intuitively easy to grasp, but the precise, quantitative notions of what behaviors are and how they can be represented, compared, shared, integrated, and analyzed are

not well established. They typically vary from domain to domain, even from researcher to researcher. The HBML proposes definitions and a framework which can accelerate progress on computational and algorithmic behavioral modeling and analyses.

The HBML is distinguished from other behavioral modeling languages such as UML by being primarily a language for instantiating process models for profiling and tracking dynamic objects in the human terrain. Therefore, the analysis of the represented objects and their dynamics is data driven rather than simulation oriented. Representations are designed to efficiently capture relationships and behaviors derived from analysis of dense data streams such as network traffic flows, physical traffic flows, SIGINT and HUMINT report streams, etc.

In developing the representational framework, the intent was to capture the essence of the dynamical objects which exist in the human terrain in as generic a framework as possible. The generic nature of the representation language is based on insights drawn from the successful use of graphical models in probability theory. It incorporates generic process modeling and tracking techniques such as those forming the basis of the PQS described in the previous section. Experience in process tracking applications in diverse domains indicates that elements of a useful modeling system tend to fall into four broad generic categories. These four categories are as follows:

- *Environment:* the rule system governing interactions among entities
- *Entities:* the objects of interest with their functional descriptions captured in one or more profiles
- *Behaviors:* the space of potential processes in which an entity may be engaged
- *Observations:* the collected and possibly filtered data which support the process tracking application

In the following section, a generic representation language is presented for describing entities in the human terrain in terms of their range of potential behaviors. The language is organized along two dimensions. The first dimension is scalar, with functional objects organized in multiple layers of aggregation from low level “atomic objects” to progressively larger scales of “aggregate objects.” With properly defined dimension reductions at scale, aggregation may theoretically extend from individuals (atomic entities) through multiple scales of functional groupings, possibly to aggregate entities operating at the global scale. For example, a multi-scalar aggregation in the business domain might be defined in terms of a progression of discrete scales, such as: worker to workgroup to division to subsidiary to multi-national conglomerate.

In addition, each entity is defined in terms of (1) relevant stationary attributes, (2) the entity's locus in the human terrain, and (3) a functional description of potential behaviors in which the entity may be engaged. Process descriptions are captured in the generic Shannon stochastic process framework in terms of their 0th, 1st, and 2nd order descriptions of their process states and interactions.

These entities interact in a specified environment which encodes the topology and rule system determining the interactions of the objects. Examples of environmental systems include (1) a communication network infrastructure, associated protocols, and global behaviors, (2) an electronic auction system and its associated rules of order matching and interaction, (3) transportation and utility systems that support the interactions of entities in a city, (4) a political system, which defines rules for campaign and election conduct, and (5) a social system which encodes guidelines for interaction among social sects.

The table below describes examples of HBML applications in a variety of domains, along with 0th, 1st and 2nd order behaviors associated with these domains.

Table 3.8-2. Examples of HBML Applications.

Domain	***Application***	***0th Order Behavior***	***1st Order Behavior***	***2nd Order Behavior***
Cyber	Network Use Profiling	Sending e-mail, visiting web sites, accessing fileservers, etc.	Probability of accessing certain websites at certain times of day	Probability of visiting web sites conditioned on browsing history
Social	Corporate Interactions	Meeting, e-mail, collaborating, calling, etc.	Probability of contact with others	Chance of meeting someone from Group A after Group B
Spatial	Tracking	Frequented locations	Probability of being at a location conditioned on time or context	Probability of transiting to and from locations
Logistics	Stocks	Bids, Asks, Position	Frequency of buying and selling, position for a stock	Tendency to move up and down in position

Applicability to Modeling and Analysis of Regional Stability

We believe that regional stability can be modeled as a function of behaviors of multiple actors at various levels of granularity, interacting within and across the domains listed in Table 3.8-2. The behaviors are responses to situations, stimuli and global events or trends, such as created by climate change, economics, politics and cultural dynamics for example. The levels of behavioral granularity include individual, group, demographic, social, regional and international. Moreover, the behavior of actors at any level of granularity is influenced and conditioned by the context and composition of the actors in question. For example, a religious organization's behavior is influenced by the behaviors of the organization's membership and leadership, as well as the ambient cultural context in which that group exists.

In order to encode behaviors, it is therefore necessary to express actions depending on preconditions and conditioned on contexts. Human Behavioral Modeling Language (HBML) is a representational scheme for achieving this encoding and we propose it as a mechanism for representing, sharing and computing with behaviors across domains and granularity levels.

Once we have a mechanism for encoding behaviors at various levels and of various types, it becomes possible to characterize actors based on their behaviors, cluster actors with similar behaviors, detect deviations from baseline behaviors and simulate systems on the basis of estimated behaviors. Given a collection of behavioral events, the Process Query Systems technology makes it possible to associate events with specific hypothesized behaviors. The most likely behaviors then become the basis for forecasting the dynamics of a group of actors in the context of regional stability analysis.

Summary

The two proposed two technologies, namely PQS and HBML, could form the foundation for a standard, common computational capability that can be used for representing economic, health, political and environmental models related to regional stability and reasoning about those models in the context of data, observations and other evidence. The role of PQS is to associate events, observations, and other evidence with competing or complimentary behavioral models of regional stability. HBML is a proposed starting point for standard representations of behaviors at different scales and of different types.

↑ Back to the top

---

[1] V. H. Berk and G. Cybenko. "Process query systems." *IEEE Computer*, pages 62-70, January 2007. http://www.dartmouth.edu/~humanterrain/papers/CYBENKO-IEEE-Computer-PQS.pdf

[2] V.H. Berk and N. Fox, "Process Query Systems for Network Security Monitoring," *Proc. SPIE—Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, 2005, pp. 520-530; http://people.ists.dartmouth.edu/~vberk/papers/berkfox5778-75.pdf

[3] R. Savell and G. Cybenko. "Mining for social processes in intelligence data streams." *Social Computing,* Jan 2008.

[4] V. H. Berk, W. Chung, V. Crespi, G. Cybenko, R. Gray, D. Hernando, G. Jiang, H. Li, and Y. Sheng. "Process Query Systems for surveillance and awareness". In *Proceedings of the Seventh World Multiconference on Systemics, Cybernetics and Informatics* (SCI 2003), Orlando, Florida, July 2003.

[5] R. Savell and G. Cybenko. "Mining for social processes in intelligence data streams." *Social Computing,* Jan 2008.

[6] N. Sandell, R.Savell, D. Twardowski and G. Cybenko. "HBML: A Language for Quantitative Behavioral Modeling in the Human Terrain." *Social Computing and Behavioral Modeling*, pages 180-190, April 2009.

[7] 6. D. A. Levitis, W. Z. Lidicker Jr. and G. Freund, "Behavioural biologists do not agree on what constitutes behavior," *Animal Behaviour*, Volume 78, Issue 1, July 2009, Pages 103-110.

[8] R. Savell and G. Cybenko. "Mining for social processes in intelligence data streams." *Social Computing,* Jan 2008.

[9] V. H. Berk and G. Cybenko. "Process query systems." *IEEE Computer*, pages 62-70, January 2007. http://www.dartmouth.edu/~humanterrain/papers/CYBENKO-IEEE-Computer-PQS.pdf

[10] R. Savell and G. Cybenko. "Mining for social processes in intelligence data streams." *Social Computing,* Jan 2008.

[11] V. H. Berk and G. Cybenko. "Process query systems." *IEEE Computer*, pages 62-70, January 2007. http://www.dartmouth.edu/~humanterrain/papers/CYBENKO-IEEE-Computer-PQS.pdf

[12] V.H. Berk and N. Fox, "Process Query Systems for Network Security Monitoring," *Proc. SPIE—Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, 2005, pp. 520-530; http://people.ists.dartmouth.edu/~vberk/papers/berkfox5778-75.pdf

[13] V. H. Berk, W. Chung, V. Crespi, G. Cybenko, R. Gray, D. Hernando, G. Jiang, H. Li, and Y. Sheng. "Process Query Systems for surveillance and awareness". In *Proceedings of the Seventh World Multiconference on Systemics, Cybernetics and Informatics* (SCI 2003), Orlando, Florida, July 2003.

[14] R. Savell and G. Cybenko. "Mining for social processes in intelligence data streams." *Social Computing,* Jan 2008.

[15] S.S. Blackman, Multiple-Target Tracking with Radar Applications, Artech House, 1986.

[16] S. Oh et al., A Scalable Real-Time Multiple-Target Tracking Algorithm for Sensor Networks, tech. report UCB//ERL M05/9, Univ. of California, Berkeley, 2005.

3.9. The Use of Cross-impact Analysis for Modeling, Simulation, and Forecasting (Stephen Millett)

Author and Organization: Stephen M. Millett, Futuring Associates, LLC

Contact Information: smillett@futuringassociates.com

The use of systems dynamics for modeling, simulation, and forecasting is well known.[1,2,3,4,5] It was made famous by the work of Jay Forrester and his protégés at MIT, the Club of Rome report, and the subsequent updated versions of *Limits to Growth* by Donella Meadows, et al., and the best-selling book *The Fifth Discipline* by Peter Senge. The use of cross-impact analysis, however, for the same functions is not as well known. The purpose of this paper is to assert that cross-impact analysis may be just as effective and arguably quicker, less expensive, and more robust than systems dynamics. It certainly is a complementary approach that provides further foresight for the benefit of both forward-looking analysts and decision-makers.

Cross-impact analysis was developed in the mid-1960s by Olaf Helmer and Ted Gordon at the RAND Corporation. It was designed to fulfill a need to integrate many trends with each other in a nonlinear way (to capture indirect relationships and feedback loops) and to project them as net assessments into the future. Cross-impact analysis represented a step beyond the expert judgment of the Delphi Method and the hypothetical sequences of events in planning scenarios. It also introduced the application of Bayesian probabilities of future trend outcomes so that the method blended qualitative judgments with quantitative mechanics.[6,7,8]

The cross-impact matrix that RAND originated can be viewed as a tabular way to construct a systems dynamics flow diagram. It is intuitively easier to construct and to interpret, especially for complex models, than the approach taken by MIT at about the same time. A matrix is created by listing all the trends, issues, or variables relevant to a subject matter in the future as both columns and headings. Values are inserted in the cells of the matrix as to how the occurrence of one trend outcome would directly impact the likelihood of other trend outcomes also occurring. In the language of mathematics, the *a priori* probabilities are marginal probabilities, and the cell values are conditional probabilities; the net impacts generate *a posteriori* probabilities, or most likely future occurrences. Much of the process was an exercise in expert judgment, but the rigor of the method and the use of simple mathematics forced many assumptions to be made explicitly and logically consistent.

The RAND modeling and forecasting team broke up during the 1960s, and the cross-impact method proliferated. Gordon created his own company, The Futures Group, near Hartford, Connecticut. He further developed a variation of cross-impact analysis called Trend-Impact Analysis.[9,10,11] Meanwhile, Helmer migrated across Los Angeles to the School of Business at the University of California, where he teamed with Selwyn Enzer to create a cross-impact method and mainframe computer software program called INTERAX. Enzer subsequently collaborated in the 1970s with a team of mathematicians, economists, and social scientists of the Battelle Memorial Institute at its laboratories in Geneva (Switzerland), Frankfort (Germany), and Columbus (Ohio). Battelle in turn created its own version of cross-impact analysis with group dynamics and scenario generation by creating its BASICS method and software program. In the

mid-1980s, Battelle introduced one of the earliest personal computer software programs to perform cross-impact analysis and generate scenarios; a decade later it further simplified the method, updated the software for Windows on a lap top, and renamed it Interactive Future Simulations, or IFS™.[12,13,14,15,16]

While there are many variations on cross-impact analysis, this paper focuses on just the Battelle approach. This method is discussed using an illustrative IFS scenario analysis of potential outcomes for the war in Iraq. This example was run on May 21, 2008 and has not been revised or updated with new information since then.[17] While it is a purely hypothetical example, it does demonstrate the simplicity and robustness of generating both forecasting and planning scenarios by cross-impact analysis. It also shows an example of how to model, simulate, and forecast the future stability in Iraq within the context of the stability of the entire Middle East with implications for U.S. security and global stability.

The IFS method, as with many other forecasting methods, including systems dynamics, begins with a topic question. In this example, the topic question read "What will be the most likely outcomes of the present war in Iraq by 2012?" It was assumed that the reader was familiar with "the present war in Iraq." At the time that this question was posed, the potential success of the Surge was still in question, and the presidential and congressional 2008 elections had not yet been held. The date of 2012 is important, because the time frame calibrates expert judgments about what is possible, likely, and virtually impossible to occur just four years into the future (while a longer period of time might open the imagination up to all kinds of future possibilities).

The second step was to select the most important descriptors (trends, issues, factors, elements, or variables) relative to the topic question to be included in the analysis. In general terms, this phase is called "the specification of the model," and it is the same for both IFS and systems dynamics. Model specification is always an exercise in expert judgment; someone has to decide which descriptors to include and which to omit. In effect, the specification of the model creates a practical although somewhat arbitrary closed system or fixed set, which is one condition for predictability.

Who makes the selection of the descriptors is very important. In full-blown projects, it is recommended that two to five expert focus groups are conducted, including about 12 diverse experts per session. The Nominal Group Technique is a recommended technique to guide the experts toward an expansive list; voting is then used to down-select the most important in rank order. Although the IFS software can accommodate up to 35 descriptors, it is best to limit the set to no more than two dozen so that models do not become too complex to understand without providing any marginal benefits. In fact, complex models often contain redundancies that force the convergence of outcomes through some repetitious and self-reinforcing fate. For "quick and dirty" feasibility studies, 6 to 12 descriptors are enough. In this illustrative example, based on previous experience, only six descriptors were selected.

Rather than projecting just one state, or outcome, for each descriptor, the IFS method calls for two to five, typically three, likely future states for each descriptor. The states should be mutually exclusive and exhaustive of all reasonable alternative outcomes for a descriptor by the target year. This again is another design feature that represents a closed system or a fixed set. The states may be defined qualitatively or expressed as quantitative ranges. Along with the specification of alternative states in the example, the *a priori* probabilities of occurrence by 2012 were determined. These *a priori* probabilities were literally "prior" in the sense of reflecting

previous knowledge and present (May 2008) information on the descriptors, and "prior" in the sense of providing a starting point for calculations based on the forthcoming cross-impact matrix. They reflect the judgment of a single expert that should be reviewed and potentially revised by peers.

The list of six descriptors, their alternative outcomes by 2012, and the *a priori* probabilities appear in Table 3.9-1.

Table 3.9-1. The List of Descriptors, Their Alternative States, and *A Priori* Probabilities.

Descriptors	***States***	**A Priori *Probabilities***
1. U.S. Military Presence in Iraq		
	A. >125,000 troops	0.30
	B. 50,000 to 125,000 troops	0.45
	C. <50,000 troops	0.25
2. Stability of the Iraqi Government		
	A. highly stable and self-supporting	0.25
	B. cycles of stability and instability	0.45
	C. unstable and weak	0.30
3. U.S. Middle East Policy		
	A. strongly support Israel and pro-West states	0.50
	B. qualified support for Israel and pro-West states	0.30
	C. limited to oil and economic issues	0.20
4. Saudi Role in Iraqi Affairs		
	A. direct support to the government	0.30
	B. behind the scenes support for Sunni tribes/militias	0.50
	C. little or no involvement	0.20
5. Iranian Role in Iraqi Affairs		
	A. direct support of the government	0.25
	B. direct support of Shiite groups/militias	0.55
	C. little or no direct involvement	0.20
6. Level of Violence in Iraq		
	A. high	0.35
	B. medium	0.40
	C. low	0.25

In a rigorous project, a white paper (or essay) should be written by a knowledgeable person for each descriptor. The white paper should contain (1) a definition of what the descriptor means and covers, (2) why the descriptor is important to the topic question, (3) background and present conditions, (4) consideration of alternative futures (states, or outcomes) for the descriptor, and (5) *a priori* probabilities for the alternative states and reasons. The descriptor analyst is often asked to prepare a qualitative cross-impact guide that can be used to facilitate filling out the cross-impact matrix in the IFS software.

The next step is the most tedious part of the method. It is the completion of a cross-impact matrix. All of the descriptors and their alternative states are arrayed on each of the two axes, as both column and row headings. The procedure is to fill in the cells of the matrix going down columns against rows, left to right. The cells are filed in with values ranging from +3 to -3; +3 means that the column descriptor and state have a large impact on the row descriptor and state. To say it more technically, the cell is filled in with an index value in response to a question: IF descriptor *x*, state *a* were to occur (its initial *a priori* probability set to 1.0), then how would it

increase or decrease the *a priori* probability of descriptor *y*, state *a*, occurring by the target year? A zero means that there is no impact. There is a presumption of symmetry. In arraying the descriptor states and in constructing the matrix, the states are listed in a hierarchical order of "high," "medium," and "low" relative to the definition of the descriptor (*not* relative to the proportionality of the *a priori* probabilities). This design will always align one descriptor's "high" with other descriptors' "high" states. Conversely, the "low" state will always be aligned with other "low" states. This symmetry is not required by the method, but it facilitates the project team in filling out the matrix with the use of the IFS software.

The process of completing a cross-impact matrix is best left up to two or three lead experts from the project team, using cross-impact guides prepared by the descriptor analysis and then conducting a later review by peers for criticism. The completion of a matrix initially may take only a few hours, but typically matrices become iterative with many revisions. It is very much a learning process, as the cross-impacting of some descriptors raises questions about descriptor relationships that have never been thought about or discussed. In performing cross-impact analysis, it quickly becomes clear just how much is known and how much still needs to be found out about the descriptors in the model. This leads to questions for further research and identification of issues that must be monitored more closely.

The completed cross-impact matrix for the illustrative example of the future stability of Iraq appears as Table 3.9-2.

In order to interpret the example cross-impact matrix, starting in the upper left hand corner Descriptor 1 on Descriptor 1 is automatically filled in with zeros. The next is Descriptor 1, U.S. Military Presence in Iraq, impacting Descriptor 2, Stability of the Iraqi Government. Descriptor 1 may be seen as having a significant impact on Descriptor 2 for the period of time in question (2008-2012). Therefore, >125,000 U.S. troops in Iraq (the "high" state of Descriptor 1) has a +3 impact on a highly stable and self-supporting ("high" state) Iraqi government.

The opposite relationship, the impact of Descriptor 2 on Descriptor 1, is judged to be the opposite: a highly stable Iraqi government would lower the *a priori* probability that there would be a high state of U.S. troops in Iraq by 2012. Because every descriptor is impacted by every other descriptor in the matrix, inputs, elements (or "stocks"), outputs, and feedback loops are shown just as they would be in systems dynamics or any other model.

Table 3.9-2. Cross-impact Matrix of the Future stability of Iraq.

	1. U.S. Military Presence in Iraq			*2.Stability of the Iraqi Govt*			*3. U.S. Middle East Policy*			*4. Saudi Role in Iraqi Affairs*			*5. Iranian Role in Iraqi Affairs*			*6. Level of Violence in Iraq*				
Occurrence	*0.30 A. >125,000 troops*	*0.45 B. 50,000 to 125,000 troops*	*0.25 C. <50,000 troops*	*0.25 A. highly stable & self supporting*	*0.45 B. cycles of stability & instability*	*0.30 C. unstable & weak*	*0.50 A. strongly support Israel & pro-West states*	*0.30 B. qualified support for Israel & pro-West states*	*0.20 C. limited to oil and economic issues*	*0.30 A. direct support to the government*	*0.50 B. behind-the-scenes support for Sunny tribes/militias*	*0.20 C. little or no involvement*	*0.25 A. direct support of government*	*0.55 B. direct support of Shiite groups/militias*	*0.20 C. little or no direct involvement*	*0.35 A. high*	*0.40 B. medium*	*0.25 C. low*	***Sum of Values***	***Non-Zero Entries***
1. U.S. Military Presence in Iraq																				
0.30 A. >125,000 troops	*	*	*	-3	-1	3	3	1	-3	-2	0	2	3	1	-3	3	1	-3	2	14
0.45 B. 50,000 to 125,000 troops	*	*	*	-1	1	1	1	1	-1	-1	1	1	1	1	-1	1	1	-1	5	15
0.25 C. <50,000 troops	*	*	*	3	1	-3	-3	-1	3	2	0	-2	-3	-1	3	-3	-1	3	-2	14
2. Stability of the Iraqi Govt.																				
0.25 A. highly stable and self supporting	3	1	-3	*	*	*	-1	0	1	-1	0	1	-2	0	2	-3	-1	3	0	12
0.45 B. cycles of stability and instability	1	1	-1	*	*	*	0	0	0	0	0	0	-1	1	1	-1	1	1	3	9
0.30 C. unstable and weak	-3	-1	3	*	*	*	1	0	-1	1	0	-1	2	0	-2	3	1	-3	0	12
3. U.S. Middle East Policy																				
0.50 A. strongly support Israel and pro-West States	1	0	-1	2	0	-2	*	*	*	-2	0	2	3	1	-3	3	1	-3	2	12
0.30 B. qualified support for Israel and pro-West States	0	0	0	1	1	-1	*	*	*	-1	1	1	1	1	-1	1	1	-1	4	12
0.20 C. limited to oil and economic issues	-1	0	1	-2	0	2	*	*	*	2	0	-2	-3	-1	3	-3	-1	3	-2	12
4. Saudi Role in Iraqi Affairs																				
0.30 A. direct support to the government	-3	-1	3	-1	0	1	-3	-1	3	*	*	*	3	1	-3	3	1	-3	0	14
0.50 B. behind-the-scenes support for Sunni tribes/militias	-1	1	1	0	0	0	-1	1	1	*	*	*	1	1	-1	1	1	-1	4	12
0.20 C. little or no involvement	3	1	-3	1	0	-1	3	1	-3	*	*	*	-3	-1	3	-3	-1	3	0	14
5. Iranian Role in Iraqi Affairs																				
0.25 A. direct support of government	2	0	-2	1	0	-1	3	1	-3	3	1	-3	*	*	*	3	1	-3	3	13
0.55 B. direct support of Shiite groups/militias	1	1	-1	0	0	0	1	1	-1	1	1	-1	*	*	*	1	1	-1	4	12
0.20 C. little or no direct involvement	-2	0	2	-1	0	1	-3	-1	3	-3	-1	3	*	*	*	-3	-1	3	-3	13
6. Level of Violence in Iraq																				
0.35 A. high	1	0	-1	-3	-1	3	1	0	-1	1	0	-1	2	0	-2	*	*	*	-1	11
0.40 B. medium	0	0	0	-1	1	1	0	0	0	0	0	0	1	1	-1	*	*	*	2	6
0.25 C. low	-1	0	1	3	1	-3	-1	0	1	-1	0	1	-2	0	2	*	*	*	1	11
Sum of Values	1	3	-1	-1	3	1	1	3	-1	-1	3	1	3	5	-3	3	5	-3		
Non-Zero Entries	13	7	13	13	7	13	13	9	13	13	5	13	15	11	15	15	15	15		

There is no question that many judgments are made in the IFS method. The judgments should be made by qualified experts, and they should be documented. All judgments are made explicitly,

rather than being buried as so many assumptions and expert judgments are made in econometric and even systems dynamics models. They are subject to peer review and revision by new information. It must be understood that the cross-impact method, especially IFS, is based on what is known about the past and the present; if better knowledge is available in the future, then the model can easily be updated. The current model is not literally depicting the future, but rather the current expectations for the future, subject to new information and model revisions. The model, therefore, can be predictive only to the extent that (1) prescient judgments have been made, (2) reasonable boundaries have been applied to the problem (making a closed system or fixed set), and (3) all known descriptors play out in the future as expected. In the worst case, a good model shows much about the current state, which becomes a foundation for thinking about the future.

The matrix also shows what is considered to be the most important descriptors as "drivers" and "drivens." A "driver" is a descriptor that has many impacts on other descriptors, and a "driven" is a descriptor that is highly impacted by other descriptors. It is possible that a descriptor can be both a driver and a driven, indicating a highly volatile descriptor.

In this example, the single most impactful driver is considered to be Descriptor 6, the Level of Violence in Iraq. The second most impactful descriptor is the Iranian Role in Iraqi Affairs, and the third is overall U.S. policy for stability and security in the Middle East. Conversely, the most driven descriptor, or what might be considered the "dependent variable," is Descriptor 1, U.S. Military Presence in Iraq. This indicates that the number of American troops remaining in Iraq depends upon how many other variations (descriptors) play out from 2008 to 2012. Saudi involvement in the stability of Iraq is highly driven, while Iranian involvement is more proactive (impactful) in Iraqi stability than the Saudi role.

With a completed matrix in the IFS software, the IFS algorithm can be run, which in turn will generate scenarios. Because the digital data and memory requirements are relatively small, the calculations require one or two seconds to run. The algorithm is deterministic, not Monte Carlo, in the sense that it always calculates the numbers the same way, with no randomness. There is nothing deterministic in the method in a metaphysical sense other than the judgments made to complete a cross-impact matrix. The algorithm begins with each descriptor state set to 1.0 and to 0, as the starting points to calculate the *a posteriori* probabilities of descriptor states from the cross-impact matrix and to cluster them into scenarios. In terms of Bayesian probabilities, the cross-impact values are new information that changes the *a priori* probabilities up to 1.0 (will occur) or to 0 (will not occur). The scenario results, which are clusters of identically occurring sets of descriptor states, for our illustrative example appear in Table 3.9-3.

In this example, there were six descriptors and a total of 18 descriptor states (three alternative states for each descriptor). The total number of single simulations, or single scenarios, generated by the IFS software algorithm was 36 (2×18, or one simulation with each descriptor state set to 1.0 and to 0 as the initial point of calculation). The most likely scenario has a frequency of 12 out of 36 single scenarios, each with identical occurring and non-occurring sets of descriptor states. In addition, the *a posteriori* probabilities can be seen for each descriptor state.

Looking at the example in Table 3.9-3, a "1" beside a descriptor state means that it occurred, and a "0" means that alternative descriptor states did not occur. Reading the table down the columns shows scenarios as sets of consistently occurring descriptor states.

It must be restated here that this is only an illustrative example of the IFS method based on more general cross-impact methods. This is a highly simplified model of a very complex question that merits far more attention than what was given to this example. The results should not be over-interpreted or considered a well-considered forecast based on how quickly it was generated by just one expert. Occurrence

Table 3.9-3. Scenario Results from the IFS Software.

Scenario Type	1	2	3	4	5	6	7	*A Priori* Probability	Total Occurrences	Posterior Probability
Frequency	12	4	3	3	2	2	2			
1. U.S. Military Presence in Iraq										
A. >125,000 troops	1	0	0	0	1	0	1	0.30	21	0.58
B. 50,000 to 125,000 troops	0	0	1	1	0	0	0	0.45	9	0.25
C. <50,000 troops	0	1	0	0	0	1	0	0.25	6	0.17
2. Stability of the Iraqi Government										
A. highly stable and self-supporting	0	0	0	0	0	1	0	0.25	4	0.11
B. cycles of stability and instability	0	1	1	1	0	0	1	0.45	15	0.42
C. unstable and weak	1	0	0	0	1	0	0	0.30	17	0.47
3. U.S. Middle East Policy										
A. strongly support Israel and pro-West states	1	0	0	1	1	0	1	0.50	25	0.69
B. qualified support for Israel and pro-West states	0	0	1	0	0	0	0	0.30	5	0.14
C. limited to oil and economic issues	0	1	0	0	0	1	0	0.20	6	0.17
4. Saudi Role in Iraqi Affairs										
A. direct support to the government	0	0	0	0	0	0	0	0.30	1	0.03
B. behind-the-scenes support for Sunni tribes/militias	1	1	1	1	1	1	1	0.50	31	0.86
C. little or no involvement	0	0	0	0	0	0	0	0.20	4	0.11
5. Iranian Role in Iraqi Affairs										
A. direct support of government	1	0	0	0	0	0	1	0.25	18	0.50
B. direct support of Shiite groups/militias	0	0	1	1	1	0	0	0.55	11	0.31
C. little or no direct involvement	0	1	0	0	0	1	0	0.20	7	0.19
6. Level of Violence in Iraq										
A. high	1	0	0	0	1	0	1	0.35	19	0.53
B. medium	0	0	1	1	0	0	0	0.40	10	0.28
C. low	0	1	0	0	0	1	0	0.25	7	0.19

The prevailing scenario in this illustrative, hypothetical case, based on information and expectations existing in May 2008, is that by the year 2012, there will most likely be more than 125,000 American troops in Iraq. In this same scenario, the Iraqi government will be largely unstable, while American policy in the Middle East will emphasize strong support for Israel and pro-Western Arab regimes. One might infer from this scenario a continuation of a policy of projecting strong American values into Iraq and the Middle East. One might also infer a continued Al Qaida resistance, along with other types of violent resistance to American troops in Iraq. In this same scenario, the Saudis would likely support Sunni tribes and militia, while the Iranians would provide direct support to the largely Shiite-dominated government in Bagdad. The level of violence in Iraq would be "high."

The second scenario, which occurred in 4 out of 36 single simulations, was one in which the level of American troops fell to the low state of less than 50,000. The Iraqi government would experience cycles of stability and instability. The level of violence in Iraq would fall to the low

state. American foreign policy would shift to a focus primarily on oil and economic issues. The Saudis would continue to support Sunni tribes and militias, while the Iranian government would have little or no direct influence over the Iraqis.

At least five other scenarios, although less likely, should also be considered.

The scenarios can be interpreted in at least two different ways: anticipatory and normative.

Anticipatory. If these scenarios are taken at face value, the most likely scenario to occur, if all trends were correctly identified and turned out in 2012 as expected in May 2008, is the first scenario, which we might characterize as "Long-Term U.S. Military Presence in Iraq." It is a scenario of instability in Iraq and the entire Middle East. This scenario would be likely to occur if the U.S. did nothing differently than it was doing in May 2008.

Normative. On the other hand, one can also ask which scenario is the preferred scenario. This normative approach results in guidance for future planning and action. If one would like to see Scenario 2 occur, and if the scenario is indeed internally consistent (as is likely because of the cross-impact model), then Scenario 2 indicates what conditions, as identified in the model, will need to occur in order for U.S. troop levels to fall below 50,000 (or even lower).

These scenarios can be interpreted as "baseline" scenarios because they are the results of a closed system and a fixed set of descriptors. Extensive sensitivity analysis can be performed on these scenarios by redefining descriptors and their states and by changing *a priori* probabilities. Cross-impact values could also be changed in the matrix. The most popular type of sensitivity analysis entails changing *a priori* probabilities. Disruptive events can also be introduced to generate completely different scenarios from the baseline.

Addressing now the inadequacies of a closed system or a fixed set, new descriptors can be added as potential disruptive events, also called "wild cards" and "black swans." They are essentially exogenous variables. The disruptive event descriptor is given a title, and one state with 1.00 probability of occurrence. It is then impacted upon the other descriptors in the model, but none of the other descriptors impact it, as it is the most certain and most important descriptor in the simulation. Dozens, even hundreds of such disruptive event simulations can be run, with each one taking only minutes to run.

One major difference between cross-impact analysis (especially IFS) and systems dynamics is that cross-impact analysis can be used for both sensitivity analysis and the generation of substantially different scenarios. The different scenarios give us different stories about the future and show us the conditionality associated with different outcomes, both most likely and most desirable. It can also be argued that the construction of a cross-impact matrix as a model is typically easier, quicker, and less expensive to complete than systems dynamics. Of course, the two approaches are complementary and may be done in parallel, especially in cases where cross-impact analysis could be used for very broad questions, while the systems dynamics analysis could be used for much finer grain forecasts.

Ideally, an organization or agency would draw up plans to meet the challenges of each scenario. Practically, one should develop strategies and plans that are sufficiently robust to be responsive to most if not all scenarios. Certainly, one should look at the most desirable scenario and determine what would be required to make that one happen. In addition to planning, one must also monitor trends and collect new information that might be needed to update and potentially revise the cross-impact matrix and its resulting scenarios.

It is highly recommended that the use of cross-impact analysis be further explored as a supplementary approach to systems dynamics as well as other modeling, simulation, and forecasting methods.

In conclusion, cross impact analysis is an alternative approach to systems dynamics that one may argue is easier to construct and to understand. It provides a matrix organization for modeling and simulation of alternative futures. It is also provides an analytical alternative to intuitive, or strictly qualitative, scenarios. Cross impact analysis relies heavily upon trend analysis and expert judgment, but in ways that such information and judgments must be made explicitly and subject to peer review and revisions over time. In the IFS approach to cross impact analysis and scenario generation, scenarios are structures of alternative sets of different likely outcomes for a fixed set of descriptors. The prescient value of IFS scenarios rests upon the boundary conditions of a designed closed system or fixed set of variable for a defined forecasting topic and the use of Bayesian probabilities. The learning value of the IFS method rests upon the generation of both anticipatory and preferred scenarios, creating a framework for simulating events and strategies that are likely to produce desired results in the future.

↑ Back to the top

---

[1] Jay Forrester, *Industrial Dynamics*. New York: John Wiley & Sons, 1961.

[2] Jay Forrester, *Principles of Systems*. Cambridge, Ma: Pegasus Communications, 1990 (1968).

[3] Donella Meadows, Jorgen Randers, and Dennis Meadows, *Limits to Growth. The 30-Year Update*. White River Junction, VT: Chelsea Green Publishing, 2004.

[4] Donella H. Meadows, *Thinking in Systems*. *A Primer*. Edited by Diana Wright. White River Junction, VT: Chelsea Green Publishing, 2008.

[5] Peter M. Senge, *The Fifth Discipline. The Art & Practice of the Learning Organization*. New York: Currency Doubleday, 2006 (1990).

[6] John G. Stover and Theodore J. Gordon, "Cross-Impact Analysis," in Jib Fowles, ed., *Handbook of Futures Research*. Westport, CT: Greenwood Press, 1978, pp. 301-328.

[7] Harold A. Linstone, *The Delphi Method. Techniques and Applications.* Reading, MA: Addison-Wesley, 1975.

[8] Wendell Bell, *Foundations of Future Studies*. Volume 1, *History, Purpose, and Knowledge*. 2 vols. New Brunswick, NJ: Transaction Publishers, 1997, pp. 265-269.

[9] The Futures Group, "TIA. Trend Impact Analysis." Version 5.0. Report 654-01-64R. Glastonbury, CT: The Futures Group, 1985.

[10] William R. Huss and Edward J. Honton, "Scenario Planning – What Style Should You Use?" *Long Range Planning*, *20*, 1987, pp. 21-29.

[11] Theodore Jay Gordon, "Cross-Impact Method." A publication of the United Nations Development Program's African Futures Project in collaboration with the United Nations University's Millennium Project Feasibility Study-Phase II (1994). Retrieved on August 20, 1998, from www.epa.gov/docs/futures/millenni/methodologies/cross-im.txt.htm.

[12] Huss and Honton, "Scenario Planning."

[13] A. Duval, E. Fontela and A. Gabus, "Cross-Impact Analysis: A Handbook on Concepts and Applications" in Maynard M. Baldwin, ed., *Portraits of Complexity*. *Applications of Systems Methodologies to Societal Problems*. Columbus, OH: Battelle Memorial Institute, 1975, pp. 202-222.

[14] E. J. Honton, G. S. Stacey, and S. M. Millett, "FUTURE SCENARIOS: the BASICS Computational Method." Economics and Policy Analysis Occasional Paper Number 44, October 1984 (Revised July 1985)

[15] "Interactive Future Simulations (IFS) ™ Theory and Computational Method" (Dec. 2008), retrieved on August 3, 2009, from www.futuringassociates.com/thoughts.html.

[16] Stephen M. Millett, "The future of scenarios: challenges and opportunities," Strategy & Leadership, 31 (2003), pp. 16-24.

[17] The IFS method, including the internal formulas used for calculations, was never held proprietary by Battelle. Battelle owned the intellectual property rights to the IFS software program and licensed it to assignees until December 2008, when Battelle sold its intellectual rights to the IFS software program to Futuring Association LLC. This example of the use of IFS was performed by the author under license from Battelle. Battelle did not participate in it and assumes no responsibility or liability for this case study.

3.10. Development of a Framework for Action: Community Resiliency as a Means of Achieving Stability (Ame Stormer, Jessica Wambach, Lt Gen (Ret) John Goodman)

Authors and Organizations: Lt Gen (Ret) John F. Goodman, Ame Stormer, Jessica Wambach, The Center for Excellence in Disaster Management and Humanitarian Assistance
Contact Information: ame.stormer@coe-dmha.org; Jessica.wambach@coe-dmha.org; john.goodman@coe-dmha.org

Introduction

The Center for Excellence in Disaster Management and Humanitarian Assistance supports the United States Geographic Combatant Commands in the achievement of their regional enduring outcomes, such as protecting the homeland. This is accomplished by shaping the security continuum through the fostering of resilient communities and societies in order to enhance stability. Healthy and vibrant communities that are able to resist disruption, to quickly recover from disruption and to more rapidly revitalize following natural and man-made disasters are the most secure and stable. Stable and vibrant communities are better able to resist extreme ideologies and are more likely to contribute meaningfully to the community of nations and to global society. The Department of Defense (DoD) receives significant annual funding for the conduct of humanitarian assistance and civic action projects. However, the DoD lacks a common aim and applied methodology to create synergy and guide the application of efforts. By utilizing the concept of fostering community and society resiliency, the Geographic Combatant Commands will shape the security continuum by directly enhancing community resilience and stability.

With the concept of fostering community and society resiliency in mind, earlier this year The Center identified the requirement to (1) determine what makes an individual household, community, and society resilient before, during, and after disasters, and (2) to identify or design a common framework. This effort would inform the DoD's thoughtful investment in civic action programs around the world as a strategic means of protecting our homeland while achieving regional stability. Such a framework must be general in nature to ensure applicability across the vast spectrum of communities, yet it must provide sufficient application methodology that accommodates the variation in geography, culture, and socio-economic features among regions.

Background

The Center is the principal United States organization that promotes disaster preparedness and community resiliency. The Center's mission is to facilitate education and training, to conduct research, and to assist in international disaster preparedness, disaster mitigation, disaster management, disaster response, health security, humanitarian assistance and societal resiliency.

The Center's study team embarked on the Fostering Resiliency Project by identifying existing frameworks or guidelines that would clearly define how to achieve resiliency across the spectrum of communal and societal functioning. The project team found that there is no single source of information that captures a method of achieving overall community resiliency in the face of disaster. Several resources captured tactical means of achieving individual aspects of

community preparedness, but no document or available framework described mechanisms across the range of community functioning to the level of specificity required by the DoD. The Hyogo Framework for Action, the U.S. State Department's *Post-Conflict Reconstruction Essential Tasks*, and the United Kingdom Department for International Development's *Characteristics of a Disaster-resilient Community* served as the starting points for The Center's efforts to construct the necessary framework beginning at the community level.[1]

Process

Synthesizing the extensive information available on what constitutes resiliency, the project team divided community functioning into five categories to create the pillars of the framework of a resilient community:

> **Enhance Social Capacity** – In order to adequately confront the challenges posed by disasters, a community must have a developed public health system, diverse infrastructure, humanitarian assistance capabilities, and robust disaster preparation.
>
> **Strengthen Knowledge Assets** – An equitable education system, hardened communication and information exchange mechanisms, and an inviting environment for research and innovation are critical to developing the intellectual capital to ensure resilience.
>
> **Enable Resource Independence** – A robust economy and established sustainable development and resource management practices enable a community to withstand and recover from the financial and environmental threats of disasters.
>
> **Create Community Cohesion** – A community that shares some common values and objectives while respecting diversity and striving for social parity can thwart disasters or come together to successfully rebuild in the aftermath.
>
> **Foster Good Governance** – A resilient community requires accountable leadership, just legal and regulatory codes, appropriate security and social protection mechanisms, and the ability to assess its risks.

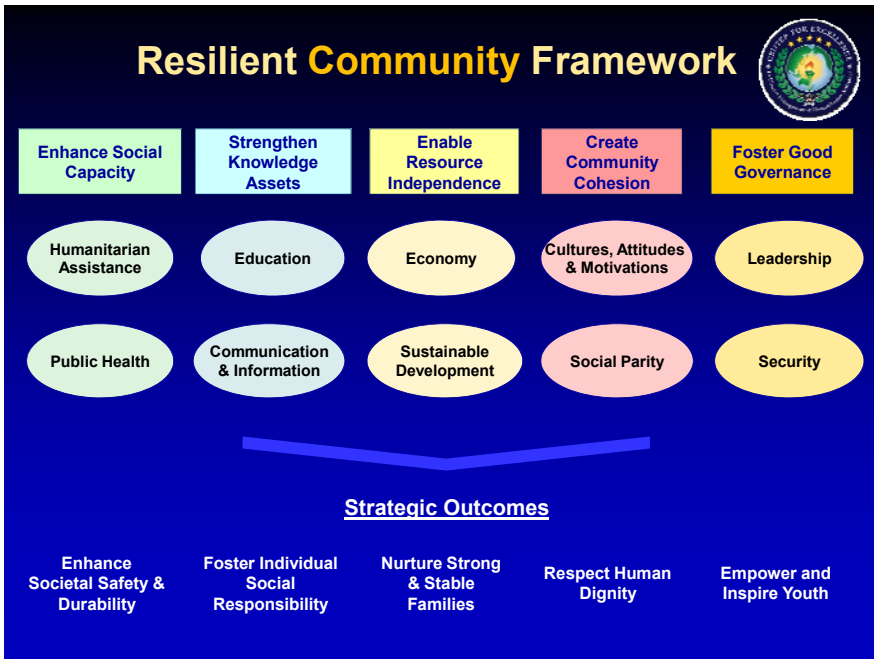


Figure 3.10-1. Resilient Community Framework.

Following this step, the project team listed the components that comprised each pillar and mapped them across three broad phases of Resistance, Recovery, and Revitalization creating a matrix listing the requirements necessary to achieve resilient community functioning within the continuum.

Once the pillars and the components were established, they were plotted against the x-axis that contained the three phases of community and societal resiliency: Resistance, Recovery, and Revitalization. In the first phase of The Center's model, Resistance, the community predicts and anticipates the most likely destructive events and identifies the desired level of community functioning when the destructive event occurs. In the second phase, Recovery, the intent is to allow the community to absorb the shock of impact with the goal of minimizing the disruption in community functioning allowing for rapid response and quick transition to recovery. In Revitalization, the final phase, the community is postured to capitalize on the opportunity to improvise and innovate to enhance community functioning.[2]

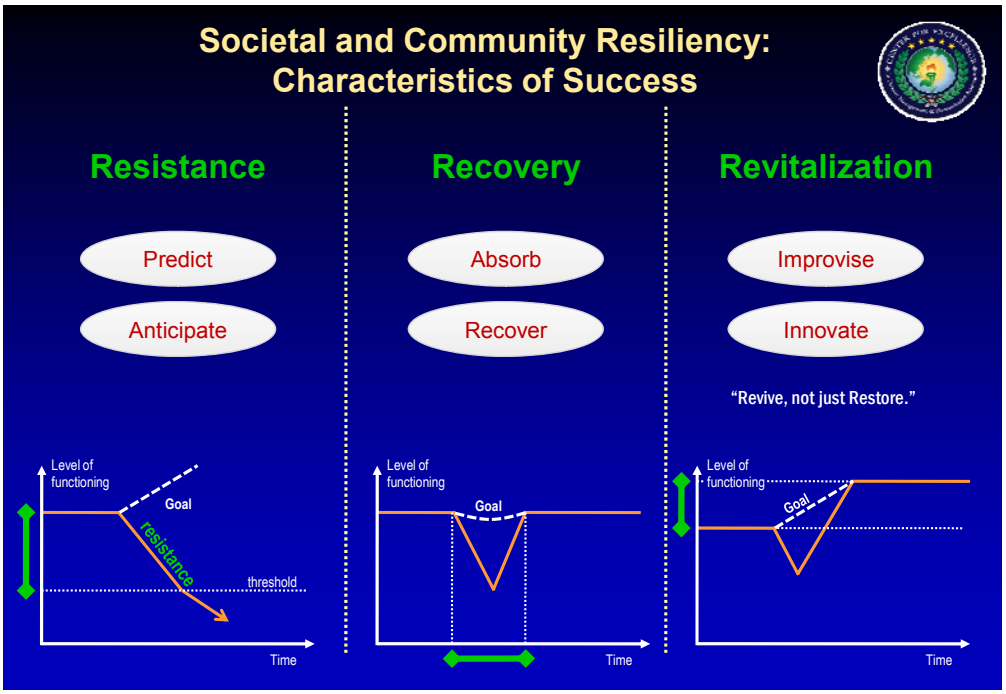


Figure 3.10-2. Societal and Community Resiliency: Characteristics of Success

Outcomes

Populating the requirements within the matrix demanded weeks of extensive research and internal review. The Center's project team scoured existing frameworks and hundreds of the available sources on resilience. Documents such as the *Millennium Villages Handbook,* published by a partnership led by The Earth Institute at Columbia University, were essential in developing the content.[3] The Center's strategic partnerships with the Defense Advanced Research Projects Agency (DARPA), the U.S. Army Corps of Engineers (USACE), Microsoft, and other organizations were consulted. Representatives of those and other organizations attended a workshop in San Diego, California, conducted in the summer of 2009 by The Center. Experts in areas across the five pillars provided feedback and validated the content of the matrix, resulting in enhancement of the framework. The workshop drew the attention of additional United States government agencies and other key stakeholders in the international development community. The feedback The Center has received confirms the assertion that there is a significant need in both the domestic and the international disaster preparedness arenas for a comprehensive resiliency framework to guide disaster recovery, as well as international development efforts.

Next Steps

As a result of the San Diego workshop and through additional review and revision of the original matrix, several follow-on steps have been determined for the Fostering Resiliency Project. The first is the development of a separate response framework to complete the spectrum of phases from disaster prevention through revitalization. Response was not included in the original framework, as the requirements for this phase are in many ways distinct from the long-term prevention and recovery phases from a disaster. The matrix that divides response into several sub-phases is currently under development at The Center.

Second, The Center is also initiating the development a societal resiliency framework that outlines the components and requirements for regional or national resiliency. The linkages between the community and society frameworks will be highlighted to emphasize the interconnections within development at both levels and ensure planners cross-walk community level efforts with societal efforts to ensure policy agreement with desired functioning.

Third, the Resiliency Project will identify sequencing and prioritization of efforts for the planning, implementation, and assessment processes. Sequencing provides the order in which activities could be conducted to optimize recovery and development for governmental and program planners and decision-makers.

These three steps will result in the creation of a "how-to" guide for Geographic Combatant Commands and other USG users which will provide guidance on the activities and the timing of activities for investment in civic action and humanitarian assistance efforts. When combined with a Common Operational Picture, the guide will assist the DoD to better support Humanitarian Assistance and Disaster Management planning, implementation and assessment activities.

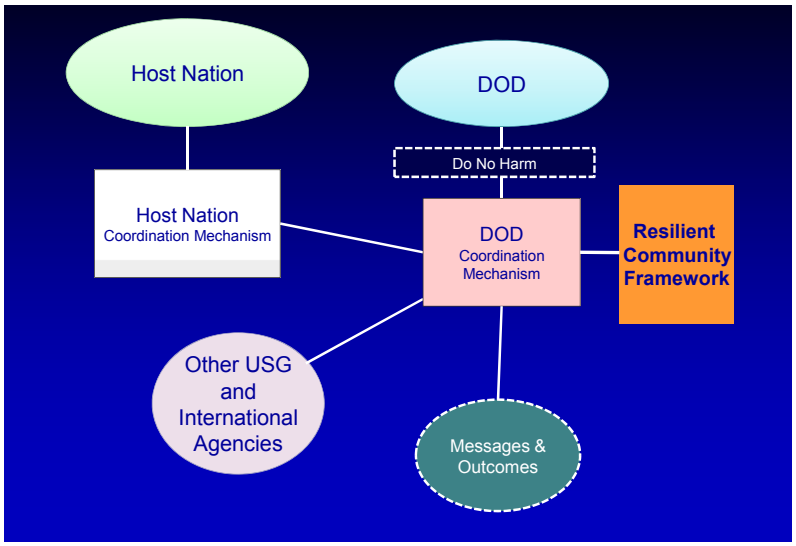


Figure 3.10-3. Coordination Mechanisms.

Conclusions

While a variety of existing frameworks and sources have identified some elements of resiliency, none have comprehensively addressed the community and societal requirements and created an actionable methodology to support functioning, development, recovery, and revitalization. The Center is creating this framework as a means to support DoD initiatives in humanitarian assistance, civic action, and disaster response mission areas. This framework provides the solutions and sequencing of efforts in disaster mitigation and management investments. By fostering resiliency in communities and societies, the application of the framework will enhance regional stability, will promote the United States' long-term interests, and will achieve the desired enduring strategic outcomes to protect our homeland.

↑ Back to the top

[1] World Conference on Disaster Risk Reduction. (2005). *Hyogo Framework for Action 2005-2015*. United States Department of State. (2005). *Post-Conflict Reconstruction Essential Tasks.*

[2] Maguire, B., Hagan, P. (2007) "Disasters and communities: understanding social resilience." Australian Journal of Emergency Management. Vol. 22, No. 2. P.16-20.

[3] The Earth Institute at Columbia University. (2008). *Millennium Villages Handbook: A Practitioner's Guide to the Millennium Villages Approach.*

4. Case Studies

4.1. Sudan Strategic Assessment: A Case Study of Social Science Modeling (Robert Popp)

Author and Organization: Robert Popp, PhD; NSI, Inc.
Contact Information: rpopp@NSIteam.com; 781.864.1347

Background

The Sudan Strategic Assessment Project was a 180-day effort (January – June 2007) conducted for USCENTCOM and funded by the USTRATCOM GISC and OSD DDR&E as part of the Strategic Multilayer Analysis (SMA) project. This was a strategic level proof-of-concept study in which a combination of quantitative and computational social science modeling and analysis approaches were developed and applied to better understand a complex "state" lacking true borders and encompassing many competing interests and complexities (e.g., corruption in the government, desertification, ethnic and religious disputes, border clashes, genocide, starvation, HIV/AIDs, etc). The focus of the effort was to emphasize an evidence-based, empirically-driven approach that would develop recommendation that span the spectrum of the instruments of national power.

Overall Project Objectives

The primary operational objectives for this project, as defined by USCENTCOM, were to develop actionable recommendations to (1) reduce human suffering, (2) promote regional peace and stability, and (3) deny safe haven for international terrorists and criminals, based on applying multiple evidence-based, empirically driven quantitative and computational social science modeling and analysis approaches. The project was also to recommend a set of actionable diplomatic, information, military, and economic (DIME) actions to USCENTCOM in support of their operational and strategic objectives. The project was also to develop a repeatable process for conducting such assessments.

Team

This project assembled a significant team including over 80 different people actively engaged in Sudan efforts across 27 different organizations (see Figure 4.1-1).

Figure 4.1-1. The Multi-organizational, Multidisciplinary Project Team.

Assessment and Analysis Process

Figure 4.1-2 illustrates the overall assessment and analysis process developed for the Sudan Strategic Assessment project. The process started with a set of specific operational objectives that framed the problem, as defined by USCENTCOM. A variety of relevant data were collected, and knowledge and expertise from a group of subject matter experts (SMEs) on Sudanese matters was elicited. A suite of multi-level (micro-to-macro) quantitative/computational social science models was developed based on the three operational objectives (i.e., to reduce human suffering, promote regional peace and stability, and deny safe haven for international terrorists and criminals). The three operational objectives were then decomposed into seven separate hypotheses to which the suite of models was then applied to produce a set of high-level assessments and forecasts.

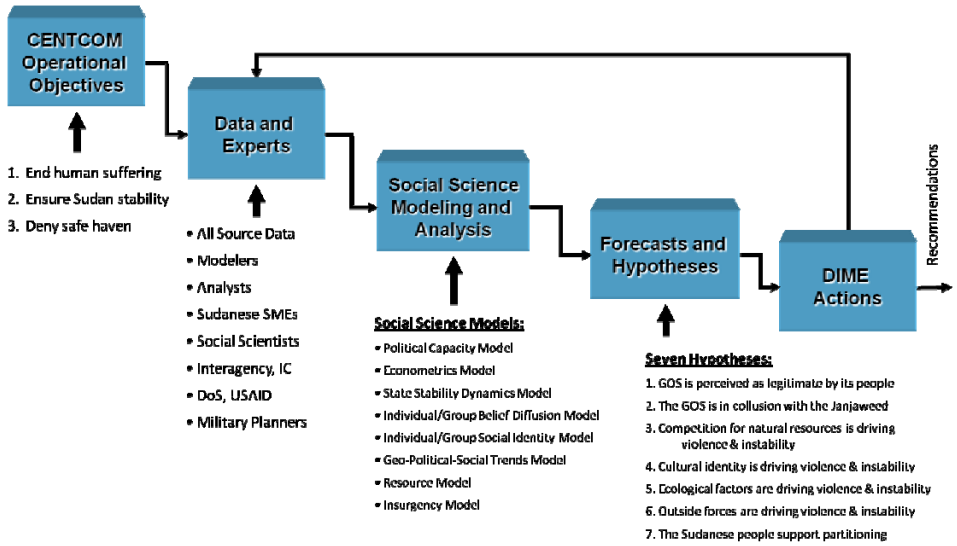


Figure 4.1-2. The Overall Project Assessment and Analysis Process.

These seven hypotheses considered were as follows:

1. The government of Sudan (GoS) is in collusion with the Janjaweed, causing turmoil.
2. Natural resources are not evenly distributed within Sudan, causing turmoil.

3. The GoS is perceived as legitimate by its people.
4. Ecological problems are a root cause of the turmoil in Sudan.
5. Cultural and ethnic intolerance is a root cause of the turmoil in Sudan.
6. The Sudanese people support partitioning of the Sudan.
7. Forces from outside countries are influencing Sudan.

Once the above hypotheses were assessed by the models, forecasts emerged which were then converted into seventeen DIME actions. These DIME actions were integrated and assessed within the overall analytical process (feedback loop), engaging the models and SMEs and consulting additional data and expertise as needed. These DIME actions were as follows:

1. Deploy a UN-led multinational protection force to take over from the under- resourced AMIS.
2. Deploy a hybrid UN-AU multinational protection force to take over from the under-resourced AMIS.
3. Deploy a US-led protection force to take over from the under-resourced AMIS.
4. Influence debt assistance provided to Sudan by international financial institutions to be dependent on the GoS ceasing atrocities.
5. Encourage a major divestment campaign targeting companies that are complicit in the genocide in Darfur.
6. Enforce and enhance a set of targeted economic sanctions.
7. Provide economic assistance to Southern Sudan specifically for infrastructure development and investment in human capital.
8. Pressure the North to more equitably share resources and wealth with the South (i.e., enforcing the CPA).
9. Do nothing.
10. Support a coup whereby the GoS is taken over by African-centric vs. Islamic-centric leaders.
11. Pay the GoS to stop the humanitarian crises in Sudan.
12. Institute programs in the HOA that strengthen and stabilize the area surrounding Sudan.
13. Actively remove key GoS/Janjaweed personnel that play a role in the violence in Sudan.
14. Unite Darfur rebel groups to develop unity militarily, collective bargaining, diplomatically, and economically.
15. Lift US and western sanctions against the GoS, enabling market forces to modify behavior and to provide an influx of Western human rights values to the region.
16. Engage GCC countries to politically influence the GoS to take responsibility for the humanitarian crises within its borders.
17. Enforce the 2005 UN Security Council Resolution demanding that the GoS cease offensive military flights by establishing a "no fly zone."

Social Science Models

Table 4.1-1 provides a brief summary of the various social science models utilized in this project. These models were selected based on their relevance to one or more of the three operational objectives as well as their availability for this project given the aggressive timeline. For more details on each of these models, two compendium documents should be consulted[1].

Table 4.1-1. Summary Social Science Models Developed for This Project.

Model	*Model type*	*Model Class*	*Model Metric*	*Defining Characteristics*
Model #1 (regional/ national)	Keyword-based text filter and categorizer	Qualitative	Political, economic, social indicator trends	Sets indicators for Sudan hypotheses, scores and graphs trends in text over time
Model #2 (national)	Economic (logit regressive) model	Quantitative	Probability civil war	Model focuses on greed (opportunistic) and grievance (rebellion) hypotheses
Model #3 (national/ sub-national)	Regressive dynamic system of equations	Quantitative	Relative political capacity (RPC)	Politico-economic macro-structural model at regional and (sub)national levels
Model #4 (national/ sub-national)	Dynamic system of eqs; Agent-based model	Computational	Likelihood of conflict and conflict intensity	Simulates structural preconditions of interstate conflict residing in notions of power and policy dissatisfaction
Model #5 (sub-national/ group)	State stability system dynamics model	Computational	Violent dissident level and anti-regime acts	Nonlinear dynamics; feedback loops, state capacities and loads, tipping point 5th order system; >140 Eqs, >100 loops
Model #6 (group/ individual)	Socio-cultural agent simulation model	Quantitative; Computational	Individual/group social identities, attitudes	Simulate effect that events have on group opinions, attitudes, identities
Model #7 (group/ individual)	POMDP agent simulation model	Quantitative; Computational	Individual/group actions, believes and behaviors	Models interactions and influence among and between different groups/individuals
Model #8 (group/ individual)	Game and decision theory model	Quantitative; Computational	Individual/group position and support for issues	Simulate stakeholder and decision-maker influence, relationships and negotiations

Key Findings

Figure 4.1-3 provides a summary of the key findings of the various modeling and analyses efforts. While some of the findings may appear to contradict one another, note that each operational objective was modeled and analyzed using a different mix of approaches at different levels of analysis (regional/national, sub-national/group, group/individual). Because of this approach, findings and recommendations for one operational objective may not necessarily be the findings and recommendations to address another objective.

With respect to the seven different hypotheses (listed above), the models and the analyses mostly agreed that the GoS is in collusion with the Janjaweed and that natural resources are not evenly distributed and are therefore a root cause of the problem observed by the international community. The findings are mixed with regard to the GoS being considered legitimate by its peoples. Some models found the government to possess legitimacy but little to no capacity, while others found the government to be illegitimate. There are a number of reasons for these conflicting findings. One is that they may very well depend on the location of the peoples and whether they believe the GoS to be legitimate. Southern and western Sudanese may have a very different view than Sudanese from the North and East. Also, there was some confusion in the study as to what the term “legitimate” meant with regard to a government. While this was addressed at the first Summit, there may have been remnants of ambiguity.

- **Likelihood of conflict increases if resources more evenly distributed**
  - Especially true between North and South
- **Egypt is the dominant regional power in the Horn of Africa (HOA)**
  - Egypt is satisfied with status quo – supports a united Sudan under leadership of North
  - Anticipated massive oil revenues (from South) not expected to change regional distribution of capabilities on HOA – Egypt remain predominant player
- **US and western nations are most important stakeholders in the region**
  - China and Malaysia are large players in oil, but have little security presence currently
- **Sudan is on path towards escalating instability**
  - Existing low levels of regime capacity and high level of dissidence in Sudan makes it difficult to achieve successful outcomes from current mitigation options
- **Enforcement of CPA can lead to consolidation of anti-regime forces**
  - CPA can lead to reduced cleavages among diverse anti-regime factions in South, causing them to organize, mobilize and rebel against GoS
- **Mixed results about the impact of cultural and ethnic intolerance**
  - Grievance model suggests increasing social fractionalization leads to greater conflict
  - Greed model suggests increasing social fractionalization reduces conflict, as it reduces the power of any one group
- **A North-South partition of Sudan has mixed outcomes**
  - Simultaneously reduce the likelihood of conflict within the new "partitioned" states, but increase the possibility of conflict between these states

Figure 4.1-3: A summary of the key findings for the project.

The hypothesis that ecological problems are a cause of unrest in Sudan was generally not supported by the analyses, but rather it appeared that the ecological problems were more a symptom of the instability and humanitarian crises. While there was not much disagreement that cultural and ethnic intolerance is a problem in Sudan, intolerance was not considered to be a cause but rather a symptom. Surprisingly, the hypothesis that the Sudanese people support partition was disconfirmed by most of the models. Based on some of the interviews and SME input, it is perhaps more likely that this may be viewed by many in Sudan as the only way to solve the current problems in the country as opposed to being a desirable option. Lastly, the hypothesis that forces outside the country of Sudan (China, Egypt, and others) are influencing Sudan was not confirmed by the models, SMEs, or analyses. Although it was found that any action that is taken in the region (e.g., a protection force) should consider the dominant actor, Egypt in this case, as a participant.

Conclusions

The Sudan Strategic Assessment "proof of concept" project demonstrated how multiple quantitative and computational social science models, in conjunction with SMEs and other analyses, are an effective, evolutionary step in the analyst toolkit, especially when the need is to provide additional lenses to look at highly complex and ambiguous problems (like the Sudan) to inform the decision-making process. The project also demonstrated the dangers of expecting "pin point predictions" from models for problems such as the Sudan. Instead, these models sought to move the decision maker further from unaided intuition, which is oftentimes the norm. Lastly, this project showed the critical need for (1) an inter-agency assessment group, (2) reach-back capability to a cadre of social science experts and domain SMEs not readily available in the US government, and (3) the need to manage information and experts at multiple classification levels.

↑ Back to the top

[1] U Sudan Strategic Assessment TTPs.doc; U Sudan Strategic Assessment Key Findings.doc

4.2. Analyzing Stability Challenges in Africa: Case Study of an Approach (Tom Mullen)

Author and Organiztion: Tom Mullen, PA Consulting Group
Contact Information: tom.mullen@paconsulting.com

"Afghanistan is not Iraq." This phrase is being used with increasing frequency and urgency in briefing rooms in the Pentagon, in Theater as a greeting to newcomers, and almost every place in between. The phrase is used to counter the classic human tendency to over-generalize, to take lessons learned in one context and rush to apply them in very different circumstances. In some cases, the lessons still hold up very well, while in other cases they can have mixed results, but in others the application of these "lessons" actually may have dangerous unintended consequences when applied in a different setting. Assessing the specific context of a new situation and understanding the reasons why particular lessons worked well is important to avoid unintended consequences, rather than reflexively applying lessons that worked in one setting to every situation. Time pressure naturally limits the amount of analysis that can be done before decisions need to be made and actions taken, but even a rapid, high-level assessment before acting can pay tremendous dividends, especially in the relatively new and uncharted terrain of stability.

With resources spread thin and contingencies likely to continue, better methods are needed to rapidly understand key issues in each new situation and to better understand why particular actions worked (and why others did not). Better methods are also needed to aid in determining where and how those lessons might be more likely to work well. Continual improvement is needed in our ability to assess highly complex situations quickly and to improve our ability to take rapid, effective action. This case study attempts to provide a few lessons drawn from analyzing a number of stability situations on the continent of Africa over the past 3 years. The focus of this paper is on lessons in analyzing complex stability situations rather than on specific actions to apply in a wide range of situations. Broadly applicable lessons for stability remain a work in progress, though certainly there is a growing understanding of the reasons why some actions have helped, which aids in determining where they can best be re-applied.

Background – Four Different Situations and Questions

While various approaches have been used to analyze stability situations in regional hotspots around the world, this paper focuses on four different African stability situations examined in a 3-year period, in which a systems approach was refined for application to stability analysis. The regions covered include the Horn of Africa, Mozambique (retrospective study[1]), and two overlapping regions examined for different confidential purposes. While there are many similar conditions across the four situations (e.g., all had unstable governments, at least a moderate degree of corruption, low GDP per capita, a poorly developed industrial base, and multiple factions vying for power), there are important variations (e.g., degree of religious tension, primary drivers of conflict, amount of high value natural resources). The questions being examined in the regions also had similarities (which potential military and non-military actions would have greatest positive effect on long-term stability) and differences (e.g., some had a greater focus on military action than others, and the US/coalition presence and involvement varied widely).

Approach – Analyzing a Region as a System

The approach for these studies was adapted from an approach (PA's Strategic Issue Analysis methodology) that has been used for decades to analyze corporate strategy issues (e.g., strategies to raise profitability, enter new markets, adapt to regulatory changes), and a few applications to government strategic issues (e.g., transformation of a major government organization, strategic planning, privatization of a European transport system). Using this approach, researchers worked with DoD to examine the challenges of counter-insurgency, and then broadened with these four studies to examine the larger issues of country and regional stability: how to help an unstable region move towards durable stability without the many unintended negative side-effects or resistance to change that so often result from efforts to improve stability. The goal of this approach is not only to identify effective actions, but also to understand why and how these actions aid in learning and decision-making for other situations.

The systems approach developed and used for these studies has seven major steps:

1. ***Background research with a wide range of overlapping sources.***
   Critical to this step is to consult a wide range of sources from *many different perspectives* (e.g., non-military as well as military, indigenous sources, experienced on-the-ground operators, experts in particular domains such as cultural studies and social science, and others who have spent significant time in the area), with *many types of information* (e.g., papers and studies, press reports, structured interviews, unstructured interviews, quantitative data), and *over a broad period of time* (e.g., historical as well as current). When searching this broadly with overlapping sources, many areas of agreement often result, but some key areas of disagreement must be sorted through. During the research, a hypothesis is developed of the key drivers and inter-relationships.

2. ***Identify key players and relevant 'boundary' of system.***
   During the research, key players are identified who are either currently strong actors or relevant parties or could play key roles in the future. These groups are maintained as aggregated as possible (e.g., large groups instead of small groups or individuals). The challenge is to keep the view as simple as possible without oversimplifying to the point of gross distortion. Similarly, in establishing the relevant boundary of the region for analysis, it helps to define as small a view as possible for the issue at hand without omitting critical factors (e.g., not the whole world or continent, but usually more than a single nation or a region in a nation). There is a lot of judgment to such determinations, and the best approach is to (1) conduct thorough, broad research, (2) have several experienced analysts provide input on the tradeoffs, (3) review the logic with others, and (4) be prepared to revise iteratively.

3. ***Map the high-level interconnections.***
   The systems approach used for this research builds on the System Dynamics method. Mapping high-level interconnections involves establishing the key links between different actors and conditions (e.g., influence of insurgent factions on security conditions). Hypotheses about key connections emerge during the first stage of background research and as a result of reviewing and comparing analyses of other regions. This step is particularly expert-intensive, balancing art and science, and also establishing and testing causal hypotheses iteratively to develop a causal map. An example of a resulting high-level causal map is shown below as Figure 4.2-1. In many

cases, the causal map has been a useful tool for facilitating discussion with other analysts and decision-makers.[2]

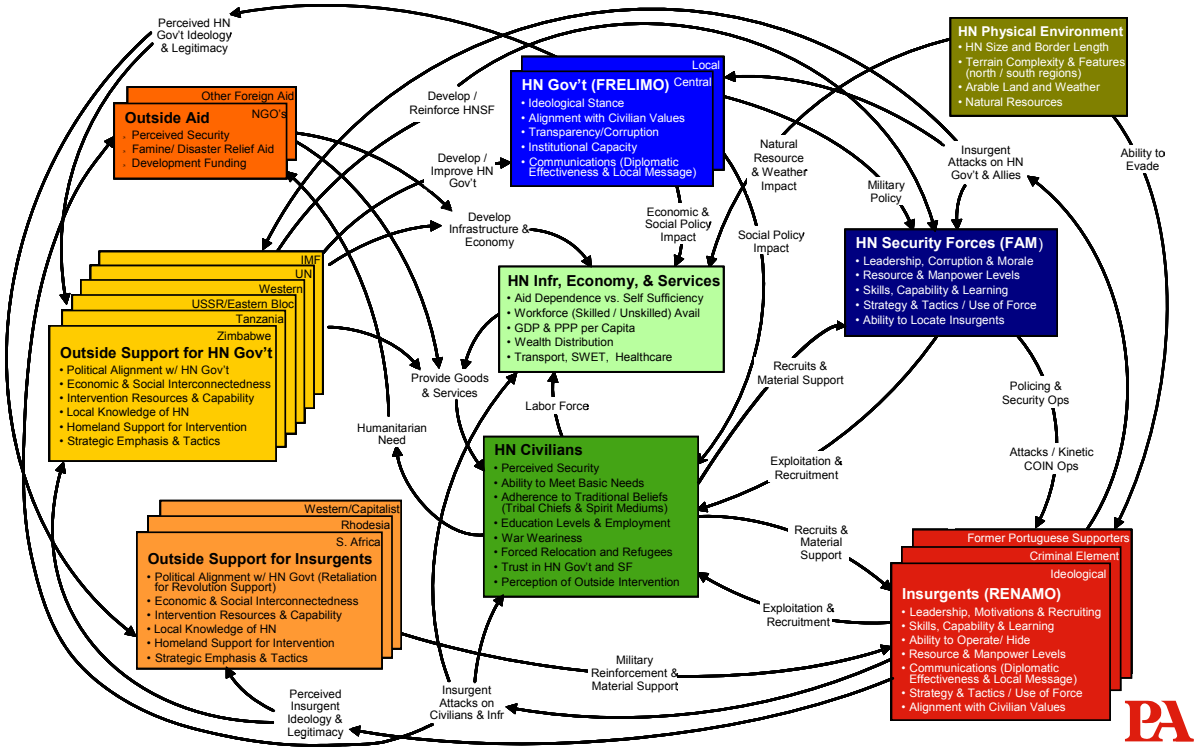


Figure 4.2-1 – An Example High-level Causal Map.[3]

4. ***Consider key scenarios and externalities, and highlight major assumptions.***
While this step often overlaps with the earlier stages, it is important to begin forming a view as to likely scenarios that may be relevant, to help in guiding both analyses and further research. Given real world sources of uncertainty and the limits in our ability to predict behavior, it is helpful to explicitly consider multiple possible scenarios to investigate the robustness of potential actions under a range of plausible futures. Similarly when the boundary of the analysis is established, it is important to consider the impact points of actors outside of the system and potential types of actions they might take. The assumed actions and other key assumptions are often important factors to test in later stages, and it is important to be explicit about them so that assumptions can be refined as more information emerges.[4]

5. ***Assess leverage points under different conditions and key sensitivities.***
The primary purpose for developing the system-level view is to examine the potential impacts (including secondary and tertiary impacts) of contemplated actions, and also to highlight areas where the results of potential actions are particularly uncertain due to their triggering both positive and negative impacts. It is also possible to identify some especially high leverage actions where the direct impacts help produce desired results and also trigger secondary impacts that set in motion further positive benefits. In some

situations, this analysis is performed on a purely qualitative level using the diagram and a structured approach, and at other times, quantitative models with equations are created to simulate the system under different conditions. A key challenge with such models is to establish confidence levels. In some cases the situation is so uncertain and the data so sparse that a quantitative model adds little additional value and is not worth the time and cost, while in other situations, quantitative models can be an important, useful addition. Determining whether to stay qualitative or to progress to quantitative results depends on many factors. If a quantitative approach is used, it is especially important to remind users of the subjectivity and range of uncertainty. There are widely varying views to the use of quantitative models. They may be viewed as completely reliable, they may be rejected out of hand due to their inability to be validated in the same way that physical models are validated, or they may be used in a balanced manner, where the findings are considered and weighted accordingly. It is vital for the analyst to take charge of the use of a model and provide clear guidance on appropriate uses of a model and its findings to avoid misapplication.

6. ***Identify high priority areas for further research.***
   Conducting the leverage and sensitivity analysis of potential actions often will surface issues where further research and/or other analysis methods and approaches could shed important additional light. For example, if the actions of external players are particularly highly leveraged, conducting a classic war game may help to understand the likely actions of external players. Alternately, if the response of one set of actors to others is particularly highly leveraged, an agent-based modeling exercise can help to examine some of the more likely outcomes. A further possibility is that a particular type of data or input from experts is needed to assess and refine certain aspects of the analysis.

7. ***Iterate and refine.***
   Rapid assessment is often carried out through this whole process in three months. Such a quick review can be very useful in providing some early observations and supporting near-term decisions, but in most cases, the rapid review also identifies a number of important ways the analysis should be refined and improved (such as when the rapid review uses only pre-existing data and information and other studies become available, or via review with additional parties, or with the addition of findings from other more detailed studies or newly available data). Even if the results of the initial analyses are well received, if the situation is important, it behooves the analyst to continue refining and challenging all aspects of the analysis. There are undoubtedly many ways that the process can be improved as stability situations are both complex and constantly evolving.

Lessons -- Insights on Analyzing Stability Situations

Through these studies, a number of lessons have been identified on analyzing complex stability situations.

1. ***Determine which stratifications of society are most important.***
   These can be tribal, religious, language, regional, or other dimensions. Determining what is important in this region is one of the most important steps.

2. ***Define the system based on real drivers rather than on geo-political boundaries***
   Other nations with historical connections, shared ethnic or religious connections, trade,

proximity, or other interests, and/or transnational groups (terrorist or otherwise) can have a significant influence and understanding these at the outset is critical.

3. ***Review history as well as current conditions***
   At least a cursory understanding of the region over multiple decades is helpful to understand origins of tensions. In many cases, a longer view also sheds important light.

4. ***Look for drivers, not just conditions.***
   While there is a necessary focus on important conditions in a region—levels of violence, troop levels, balance of power, economic conditions—these are often indicators, and it is important to determine which factors drive these indicators so that the focus can be on the actions that will produce the greatest net positive impact on the key issues of interest.

5. ***Simplify, but stay rigorous and iterate.***
   Simplifying as much as possible (favoring breadth rather than depth at the outset) is important to make rapid headway in an analysis and to be able to provide insight in a timeframe to support the needs of decision-makers. While simplification and speed are necessary, it is also important to maintain rigor and push to keep testing and refining hypotheses against available information, actively challenging conclusions and assumptions, adding complexity where it proves important, and iteratively improving the analysis. Speed cannot be an excuse for sloppiness.

6. ***Acknowledge uncertainty, but do not be paralyzed by it.***
   It is natural for analysts to be enamored of their own analysis and to overlook or minimize aspects not covered. It is important to recognize how complex and uncertain stability situations are and to be explicit about the issues that are not well-covered by an analysis or that are implicit or explicit assumptions. Rather than giving up in the face of the complexity and uncertainty, it is important to persevere to analyze possible implications of actions while acknowledging that specific forecasts are generally not possible.

7. ***Synthesis is challenging and 'soft,' but critical to making the analysis useful.***
   Stability situations tend to be so complex that synthesis of the many factors considered and studies analyzed is important to make the work accessible to decision-makers. Synthesis involves some degree of judgment and must be accomplished by experienced, trusted analysts to be helpful to decision-makers. Done poorly, synthesis can be badly misleading and can destroy the value of a good assessment. Done well, it can make a vast quantity of analysis digestible and useful to a decision-maker so that the analysis is used rather than ignored. However, if a decision-maker does not trust the synthesis, he or she will understandably discount the assessment, even if it is in fact excellent. In the end, analysis of complex human situations is difficult to divorce from the people performing the analysis and synthesis.

8. ***Hard data and local perspective are some of the greatest challenges.***
   A consistent theme in analyzing stability situations is a very limited amount of hard data available to support analyses. Data on population, economic conditions, and violence levels can be difficult to pin down with confidence. These are some of the easier factors to measure, and other information such as popular sentiment, balance of power, influence by different parties, etc., is often speculative at best. Polling and expert elicitation are the best ways of filling such gaps but typically analysts will find data well short of desires in

a rapid assessment. Gaining access to those with a truly indigenous perspective can also be difficult, especially for adversaries whose background varies widely from a typical western perspective and whose motivations are difficult to understand coming from a western viewpoint.

9. ***Quantification has a role, but it is not always better.***
   In some highly uncertain situations, quantified analyses would need to be caveated with such a wide range of uncertainty that they add very limited value. In other situations, even if quantification is more feasible, it may be important first to provide a rapid qualitative framework that helps others have constructive discussions and debates. Quantification often helps enough to be worth performing, but the value of qualitative assessments and frameworks should not be underestimated.

10. ***There are common insights, but there is no universally applicable action plan for stability.***
    Many insights recur in analysis of multiple stability situations: time delays are often frustratingly long (decades when we would prefer months and years); knowing what to do is easier than actually doing it when facing pressure to achieve results or stressful situations in theater; corruption is an insidious condition that changes many leverage points and is extremely difficult to overcome in a durable way; and COIN principles broadly hold up in many situations, though the emphasis varies. However there are undoubtedly some situations in which these insights do not apply and others where they are less important than other factors. The set of actions to improve stability must be tailored to a specific situation.

"Simplify as much as possible, but not further," a suggestion by Albert Einstein, is especially apt for stability situations. Without some simplification, it is easy to be paralyzed by the complexity of rich histories, many players, and issues at every level of abstraction, from individuals on up. Many assessments fall into this trap and never see the light of day. Other assessments simplify too much, leading to potentially dangerous conclusions. The work highlighted in these case studies has attempted to make progress in establishing a workable middle ground, capturing important complexities and drivers and creating an analysis platform that can be used to support decision-making and lesson analysis in a useful timescale. With vital decisions on stability-related actions facing decision-makers for the foreseeable future, it is critical for us to provide rigorous and actionable advice so that these decisions can be as informed as possible.

Acknowledgements: Many thanks to our clients for these studies, for having the wisdom and boldness to try a new approach to analyze these issues, recognizing the importance of the problem and the limitations of more traditional methods. Sponsoring organizations included DARPA, J-8, OUSD(P) FT&R, DDR&E, PA&E, OPNAV, and Northrop Grumman. Thank you also to my colleagues at PA Consulting who played key roles in refining this approach and conducting these analyses, especially Tamara Greenlaw, Donna Mayo, Ritu Sharma, Alan Graham, Tamara Kett, Beth Farrelly, Whitney Pickels, Erik Waldron, and Alma Rico.

↑ Back to the top

---

[1] Reports of the DARPA-sponsored regional stability study (historical Mozambique situation used as example - conducted in 2008) and a predecessor DARPA study examining insurgency dynamics in Iraq

(conducted in 2006) are available to US Government audiences upon request. Contact Dr. Brian Pierce, Deputy Director Strategic Technology Office, DARPA, for further information.

[2] A description of the System Dynamics methodology is covered in the classic book by the founder of the field: Jay Forrester, Industrial Dynamics, Waltham, MA: Pegasus Communications, 1961.

[3] From "Understanding the dynamics of low intensity operations in developing nations – identifying key strategic challenges", PA Consulting Final Technical Report for DARPA Strategic Technology Office, April 2008, ARPA Order No. X225/00, Program Code: 7P30

[4] A description of the Scenario Planning approach is covered in the classic work: Peter Schwartz, The art of the long view: Planning for the future in an uncertain world. New York, Doubleday, 1996.

Appendix A. Acronyms

A2C2	Adaptive Architectures for Command and Control (program at Office of Naval Research)
ACH	Analysis of Competing Hypotheses
ACM	Association for Computing Machinery
AI	Artificial Intelligence
AKO	Army Knowledge Online
ASW	Anti-Submarine Warfare
C2	Command and Control
CCT	Collaboration Consulting Team
CDC	Centers for Disease Control and Prevention
CEO	Chief Executive Officer
CIA	Central Intelligence Agency
COCOM	Combatant Command
COI	Community Of Interest
COTS	Commercial-Off-The Shelf
CPU	Central Processing Unit
CRM	Crew Resource Management
DDNI	Deputy DNI
DDR&E	Director, Defense Research and Engineering (in OSD)
DHCP	Dynamic Host Configuration Protocol
DHS	Department of Homeland Security
DIME	Diplomacy, Information, Military, Economic - The four instruments of national power
DIMEFIL	DIME + Finance, Intelligence and Law Enforcement
DISA	Defense Information Systems Agency
DKO	Defense Knowledge Online
DNI	Director of National Intelligence
DoD	Department of Defense
DOS	Department of State
DPI	Deep Packet Inspection
DTI	Directed Technologies, Inc.
EUCOM	US European Command
FBI	Federal Bureau of Investigation
FC	Fiber Channel
FEMA	Federal Emergency Management Agency
GMU	George Mason University
GWU	George Washington University
HHS	Health and Human Services
HIPAA	Health Information Portability and Accountability Act
HPLT	High Performing Leader Teams
HRO	High Reliability Organizations
IC	Intelligence Community
ICES	Intelligence Community Enterprise Services

ICS	Incident Command Structure
ICT	Information, Computers, Telecommunications
IED	Improvised Explosive Device
IM	Information Management
INR	Bureau of Intelligence and Research (DOS)
IP	Internet Protocols
IT	Information Technology
JIATF	Joint Inter-Agency Task Force
JIIM	Joint Interagency, International, Multinational
JTF	Joint Task Force
KM	Knowledge Management
LOE	Limited Objective Experiment
LSU	Louisiana State University
LTX	Leader Team Exercises
MBA	Master of Business Administration
MSEHP	Model State Emergency Health Powers (Act)
NAD	natural or anthropogenic disaster
NAS	National Air Space
NASIC	National Air and Space Intelligence Center (USAF)
NCES	network-centric enterprise services
NDU	National Defense University
NFCE	Nexus Federated Collaboration Environment
NGA	National Geospatial-Intelligence Agency
NJOIC	National Joint Operations Intelligence Center
NRF	National Response Framework
NSI	National Security Innovations, Inc.
NTI	National-to-Tactical Integration
NVOAD	National Voluntary Organizations Active in Disaster
ODNI	Office of the Director of National Intelligence
OLPC	One Laptop Per Child
OODA	Observe, Orient, Decide, and Act
OSC	Open Source Center
OSD	Office of Secretary of Defense
OSS	Open Source Software
P4CMI	Patterns for Computer-Mediated Interaction
PNNL	Pacific Northwest National Laboratory (Department of Energy)
RRTO	Rapid Reaction Technology Office (OSD/DDR&E)
RSS	Really Simple Syndication - a family of Web feed formats used to publish frequently updated works
SARS	Severe Acute Respiratory Syndrome
SATCOM	Satellite Communications
SBU	Sensitive But Unclassified
SIS	Strategic Information System
SKA	Skills, Knowledge, and Attitudes
SMA	Strategic Multilayer Assessment
SOP	Standard Operating Procedures

TCS	Task, Condition, and Standard
ToL	Teams of Leaders
TRADOC	Training and Doctrine Command (US Army)
TTPs	tactics, techniques and procedures
UNO	University of Nebraska Omaha
USAFA	US Air Force Academy
WHO	World Health Organization
WMD	Weapons of Mass Destruction
WoG	Whole of Government
WREN	Worldwide Rare Event Network

↑ Back to the top