



BC JOURNAL

of International Affairs

A large, vintage-style globe is the central focus of the cover. It features a map of the world with landmasses in shades of orange, red, and yellow, and oceans in a light beige. The globe is tilted and has a thin yellow line representing the equator. It is supported by a dark bronze statue of a muscular man, likely representing Atlas, who is shown from the waist up, holding the globe on his shoulders. The statue is positioned at the bottom of the cover, with its arms raised to support the globe. A semi-transparent tan banner is placed over the middle of the globe, containing the title text.

With Great Power Comes
Great Responsibility

THE BOLOGNA CENTER
JOURNAL OF INTERNATIONAL
AFFAIRS

SPRING 2006

PUBLISHED BY
THE JOHNS HOPKINS UNIVERSITY
PAUL H. NITZE SCHOOL
OF ADVANCED INTERNATIONAL STUDIES
BOLOGNA CENTER

The *Bologna Center Journal of International Affairs* was established by the graduate students at the SAIS Bologna Center in 1997 to provide a forum for discussion and dissemination of ideas about current issues in the field of international relations. The BC Journal aims to provide a formal outlet for thought-provoking scholarship from students and faculty at the Bologna Center and other graduate institutions, and welcomes work from experts and practitioners in the field.

The BC Journal is published once a year in May. The views expressed herein are solely those of the authors and do not represent the opinions of the Journal staff. Neither the Bologna Center Journal of International Affairs nor the Bologna Center itself guarantee the accuracy of the data included in this publication and accept no responsibility for the consequences of its use.

INTELLECTUAL PROPERTY RIGHTS

The authors of the articles in this journal maintain the intellectual property rights of the contents. Anyone wishing to reproduce these articles must contact the individual authors to obtain their permission. Please contact the Bologna Center Journal of International Affairs for more information.

SUBMISSIONS

Articles for submission to the Bologna Center Journal of International Affairs are accepted on a continuous basis. Article selection for the print edition takes place in March. Submissions will be judged according to academic merit and relevance to the selected theme. Articles may be submitted to the Managing Editor at the address below or to the e-mail address specified on the website. Authors should include a current biography, contact information and abstract of the article. Our website may be accessed at:

<http://www.jhubc.it/bcjournal>

Copyright 2006, the Bologna Center Journal of International Affairs
ISSN 1592-3436 print/ISSN 1592-3444 online

PUBLISHED BY:

The Johns Hopkins University
Paul H. Nitze School of Advanced
International Studies
The Bologna Center
Via Belmeloro, 11
40126 Bologna
Italy

PRINTED BY:

Baiesi snc
Via Broccaindosso, 2b
40125 Bologna
Italy

Acknowledgements	5
Journal Staff	5
Editor's Note	6
The European Union: America's Rival for Global Governance <i>David P. Calleo</i>	9
The European Integration and its Effects on European Monetary Union Stock Markets <i>Faarnaz Chavoushi</i>	16
Italian Immigration Policies: The Metaphor of Water <i>Rima Al-Azar</i>	35
The American Way of Going to War <i>John L. Harper</i>	52
Between Conscience and Self-Interest: The United States, Sudan and Darfur <i>Alec Barker</i>	78
The Producing Countries: Between Pursuing Development and Meeting Global Demand <i>Giacomo Luciani</i>	92
Sticks Instead of Carrots: Energy Leverage in Contemporary Russian Foreign Policy Toward Ukraine <i>Ryan R. Miller</i>	99
Abuse of Power: The Responsibility of Government to Encourage the Development of Renewable Energy Resources <i>Michael Obeiter</i>	114
Electing Hamas <i>Zecharia Kahn</i>	126
Between Polyarchy and Autocracy: The Delegative Democracy of Carlos Menem <i>Brittany Williams</i>	137
Getting Out of the Intellectual Trap: When Profit-Driven Humanitarian Aid Might Save More Lives <i>Janika Albers</i>	149
The Case for Cash: Delivering Aid Through Direct Payments to the Poor <i>Fatima B. Kassam</i>	164
Human Rights and Sexual Orientation: An Expansion of the Equality Doctrine <i>Elaine Hensle</i>	177

Acknowledgements

The Bologna Center Journal of International Affairs would like to thank:

The Bologna Center Student Body, Faculty and Staff
 The Bologna Center Journal Auction Contributors
 The Bologna Center Student Government
 Professor Erik Jones, Faculty Advisor
 Ryan Haney, Web and Layout Consultant

2006 Journal Staff

Editor-in-Chief	Tarig Hashim Hilal*
Executive Editor	Fatima B. Kassam*
Managing Editor	Mame Soce Sene
Copy Editor	Donna Brutkoski
Web/Layout Editors	Zecharia Kahn & Andrew Duff
Finance Director	Robert Carter Atlamazoglou
Public Relations	John Horton
Event Planning	Leigh Anne DeWine & Nichole Graber

Editors

David Carlisle*	Duncan Burrell
Morgan Bushey	John R. Williams*
Aoife Brophy	Stefan Kahandaliyanage
Cynthia Mar	Leigh Anne DeWine
Gerd Gensbichler*	Sarah Gearen
Colleen Keenan	Sylvie Mahieu
Holger P. Wilms	Tamara Senikidze
Noah Wiss	Jessica Thompson
Nick Haslam	Zecharia Kahn
Donna Brutkoski	Fatima B. Kassam

*Selection Committee

Copy Editors

Hagir Elawad	Morgan Bushey
Colleen Keenan	Aoife Brophy
Kate Ostrander	Meredith Metzler

EDITOR'S NOTE

The world today is dominated by the United States and its allies. It is, however, a rapidly changing world in which new powers are emerging, the gap between the rich and the poor is widening, and conflict and humanitarian crisis continue to afflict the lives of millions. Within this context, the challenges of exercising unparalleled power are proving both complex and unprecedented. It is this power and the responsibilities that come with it that are the focus of this year's edition of the Bologna Center Journal of International Affairs.

The statistics are stunning. The combined GDP of the United States and Europe currently stands at \$22.747 trillion,¹ and the United States alone spends \$455 billion per year on its armed forces – more than the combined military expenditures of the United Kingdom, France, Japan, and China.²

Nevertheless, powerful changes are afoot. The Chinese economy promises a growth rate of 7.2 percent per year³ over the next five years, with many predicting that it will develop an economy of a size that rivals that of the United States by 2020.⁴ China is not alone. A recent study by Goldman Sachs identified three more economies that seem poised to enter the top ranks of economic power within the foreseeable future: Brazil, Russia and India.⁵

Despite this rising prosperity, an estimated 2.9 billion people remain in poverty – almost half of the planet's population.⁶ In its latest report, the International Crisis Group identified no less than 70 existing or potential conflicts around the world.⁷

Nowhere are the challenges and tensions of our changing world more evident than in Africa, where powers old and new jostle for influence against a background of poverty and war. Countries such as Sudan have seen the entire international community descend upon them over the last five years. Development agencies, multinational corporations and the governments of China, India, the United States and several European countries have all painstakingly sought influence, aiming to resolve conflict, encourage development or gain access to resources.

In Africa, the responsibilities of power are thrown into high relief as the difference between success and failure is immediate and visceral. The challenges of this region are, in many ways, a microcosm of what is happening on the broader international scene.

It is in this context that contributors to this year's edition of the Bologna Center Journal of International Affairs have addressed the issues of global governance, interdependence, immigration, military intervention, energy, democratic development, humanitarian aid and human rights. The result is a powerful collection of writings that will provoke and inspire.

Tarig Hashim Hilal

Editor-in-Chief

April 2006

NOTES

¹ “Europe” refers to the EU 15. *OECD In Figures 2005*, Organization for Economic Co-operation and Development <http://www.oecd.org/home/0,2605,en_2649_201185_1_1_1_1_1,00.html>.

² *The 15 Major Spenders in 2004*, Stockholm International Peace Research Institute <http://www.sipri.org/contents/milap/milex/mex_major_spenders.pdf>.

³ Dominic Wilson and Roopa Purushothaman, “Dream With BRICs: The Path to 2050,” Global Economics Paper No. 99 (New York: Goldman Sachs, 2003), <<http://www.gs.com/insight/research/reports/99.pdf>>.

⁴ “The World’s Largest Economies,” *The Economist*, 30 March 2006 <http://www.economist.com/markets/indicators/displaystory.cfm?story_id=6746594>.

⁵ See Wilson and Purushothaman.

⁶ By the \$2 per day poverty measure. William R Cline, “Trading Up: Trade Policy and Global Policy,” Center for Global Development CGD Brief Volume 2, Issue 4 (September 2003), p. 1 <<http://www.gs.com/insight/research/>>.

⁷ “Crisis Watch No. 32,” International Crisis Group, 1 April 2006 <<http://www.crisisgroup.org/home/index.cfm?id=4057&l=1>>.

The European Union: America's Rival for Global Governance

David P. Calleo

*The following is an excerpt from a speech given by Professor
Calleo at the Bologna Center on December 6, 2005.*

In my book of the late 1970s, *The German Problem Reconsidered*, I defined Germany's problem as being too big and dynamic for the traditional European state system to contain. The Cold War, I argued, had provided three solutions simultaneously: Germany's own division, external hegemonies over Europe as a whole and a European confederation. What would happen to these solutions, I asked, when the Cold War finally ended? Answering that question led me to another book, *Beyond American Hegemony*, and ultimately to my most recent book, *Rethinking Europe's Future*, which came out in 2001 and had a new edition with a further conclusion in 2003. The book, which took me a decade to write, was a historical-analytical account of the European Union – of how Europe's old states, sheltered by the Cold War, had resurrected themselves collectively and, in the process, created a new political formula. This new formula was not a federation, or even a federation in the making, but an association of free and still sovereign states, increasing their real sovereign power by cooperating. Together, they were able to achieve national aims that they could never have hoped to achieve alone.

Europe's union is, as I see it, a great advance over traditional state systems. It is a community – a confederacy – that enables rather than diminishes national sovereignty; one that enriches and stabilizes national identity by adding a regional identity. This further identity supplies, for each of its members, a sort of institutionalized superego to discipline national power, on behalf of more general interests. It thus transforms regional interstate relations from a zero-sum game to one of mutual gain within a common community. This union has become, in my view, a vital element in the constitutional structure of each of its members. And, at the same time as the union has been intensifying harmony within its region, it has also provided the collective machinery for more effectively synthesizing and pursuing the external interests of the European states – in relations with neighbors to the east and to the south, and across the Atlantic.

David P. Calleo is Dean Acheson Professor and Director of European Studies at Johns Hopkins University's Nitze School of Advanced International Studies (SAIS) in Washington, D.C. He received his B.A., M.A., and Ph.D. degrees from Yale University; has taught at Brown, Yale and Columbia Universities, and at the Institute d'Etudes Politiques de Paris; was a research fellow at Nuffield College, Oxford; and served as consultant to the US undersecretary of state for political affairs. He has published extensively on political and economic issues concerning Europe and the United States.

This union embodies, in my opinion, a political technology of great promise, not only in Europe but elsewhere across the globe. It is, I imagine, probably the most appropriate model for what I suspect will be the world order of the future: A plural world of several great powers; not only the United States but also China, Russia, Japan; perhaps India and Brazil; eventually some sort of Muslim great power; and so on. The creation of such a European Union in itself implies a plural global system – a world of several great powers. The rise of such a European model, after the terrible experiences of the earlier 20th century, ought to favor belief if not in God, at least in Hegel. There is, however, also a darker side to my analysis of Europe's future, to which I will return. But first, some words about the United States.

The pluralist vision of the future that I try to sketch has never been very popular in the US. It has never caught the American political imagination. Instead, following the collapse of the Soviet Union, that imagination has been bemused by a “unipolar” vision of world order – a global system of interrelated states with one clearly dominant “superpower.” According to that view, America's power is radically superior to that of any other country, or combination of countries, and fated to remain so for the foreseeable future. For many American analysts, and political elites generally, events in the 1990's seemed to confirm this unipolar vision. The vision came in two dimensions, economic and military. Projecting the economic dimension was the work of the Clinton administration. As you know, the mid-1990s saw an unprecedented boom, built on high consumption, low saving, and heavy investment in new technologies, and fueled by huge inflows of foreign capital. For the first time since World War II, America's productivity growth regularly outpaced that of Europe and Japan. Unemployment was at near-record lows and inflation scarcely visible. When President Clinton left office at the start of 2001, the federal budget was pointed toward a large surplus. So much success naturally served to reinforce America's self-image as the avatar of liberal globalization.

True, the United States continued to run a large, and growing, external or current account deficit. America's economy was continuing to absorb – i.e., consume and invest – substantially more than it produced. That external deficit had to be financed by an equivalent inflow of foreign capital. In the 1990s, that seemed not to be a problem. Private foreign investors, attracted above all by the technology boom, regularly flooded the US with more than enough capital to cover the external deficit. Rather than worry about their deficit, Americans could congratulate themselves on the attractiveness of their huge and fast-expanding economy. And although our debts kept growing rapidly, our GDP grew still faster.

Developing the military side of the unipolar vision – the US as the only global superpower – began in Clinton's time but has been the particular specialty of the Bush Administration. Never in modern history, we like to say, has one nation been so militarily predominant as the United States is today. American conventional forces, enabled by satellites, now have something analogous to the “shock and awe” achieved by the German *Blitzkrieg* in the early days of World War II. Transforming the American military into such a force has had great aesthetic and intellectual appeal. And having such military predominance has gradually

encouraged a geopolitical ambition to fit it. Global hegemony has begun to seem an obligation imposed by history. The power to dominate creates the duty to dominate.

As we all know, America's unipolar visions have lately been having painful encounters with reality. To start with our military expectations: Our military transformation to "network-centric warfare" seems not compatible with the requirements of our unipolar global role. The invasion of Iraq, for example, however brilliantly executed at the outset, seems in the end to have replaced a rogue state that was being successfully contained with a failed state that requires an indefinite and bloody occupation. Failed states, societies in chaos obsessed with their grievances, are natural breeding grounds for global terrorism. We feel ourselves highly vulnerable, despite our egregious military predominance. We console ourselves by arguing that terrorism is an illegitimate form of military power. We should not be surprised if the argument carries little weight with the terrorists themselves.

Terrorism is the natural recourse of the weak and dispossessed of this world, the asymmetric populist weapon. Against terrorism our large and spectacularly expensive military establishment is of limited use. It is trained and equipped to attack other military establishments that cooperate by meeting us head-on. But our enemies do not always cooperate. They sometimes refuse to fight the wars for which we are so lavishly prepared but they are not. Instead they turn to terrorism – the form of warfare that suits their weakness and mocks our strength.

The United States is also vulnerable at the other end of the military spectrum – the realm of strategic nuclear weapons. These, like terrorism, produce highly asymmetrical results. As with terrorism, a perverse military logic works against the US. The more America's overweening military power threatens smaller countries with forcible regime change, the more they seek asymmetric weapons to deter us. As they can no longer fight, toe to toe, against our conventional military power, their options are terrorism on the one hand and nuclear weapons, or other weapons of mass destruction, on the other.

Meanwhile, it is no secret that many of our military leaders feel the US is already dangerously overstretched. The greater our military victories, it almost seems, the weaker we have grown. In Iraq and Afghanistan, the US has defeated two undoubtedly evil regimes. But, by itself, the US lacks the military power, as well as the political and moral authority, to bring either victory to a tolerable conclusion. To do so will require the military, political and economic assistance of many countries. To enlist those resources requires legitimacy from a genuine consensus in the international community. Under such circumstances, a militarized American foreign policy, scorning allies and global institutions, is not very promising.

If the evolution of military power belies a Pax Americana, what about the evolution of economic power? Is the world economy "unipolar?" Both recent events and long-term trends suggest otherwise. In 2001, the feverish American stock market bubbled and collapsed. The administration and the Congress reacted with substantial tax cuts and, after the atrocities of 9/11, with security spending

approaching a Cold War scale. Big fiscal deficits returned, while the big external deficits continued growing. Private foreign investment, however, fell sharply, and the dollar with it.

Today, the principal financiers of America's big external deficit are Japan, China, and the OPEC countries. In effect, their central banks finance their own national exports to America. Japan, rich and stagnant, has done this for decades and may well continue. But sooner or later China, still a very poor country with a huge internal market of its own to develop, seems likely to find a more satisfactory use of its savings than subsidizing America's consumption. Without its big Chinese subsidy, the dollar would presumably fall and America's foreign consumption would falter. This might help to revive manufacturing in America, but also would imply a painful drop in living standards. A major breakdown in the world's financial system could very well follow. Certainly, the assumption of unlimited resources, implicit in the vision of the unipolar superpower, would be more difficult to sustain.

Longer-term trends in the world economy also seem in disharmony with the unipolar vision. Respectable projections, for example, see the Chinese GDP overtaking the American [one] in a few decades. Nothing, of course, guarantees China's continuing rapid growth. As has happened so often in modern times, war and revolution may rob this gifted people of the fruits of their labor and talent. But if China does collapse, the rest of the world is unlikely to escape unscathed. China could, however, continue growing rapidly for a long time. There is still a huge reserve of Chinese labor to draw on, plus an enormous gap between Chinese and Western wages. China, moreover, has a phenomenal savings rate – some estimate roughly 40 percent of GDP. And the country is becoming a major technological and scientific power in its own right.

China's continuing success will undoubtedly shake things up in the world. As its GDP eventually overtakes those of the US and the EU, and perhaps goes well beyond either, its competitive prowess also appears to threaten Western living standards. Why should that be? After all, the postwar rise of Japan and Asia's "Little Tigers" also disturbed many Western interests. But the competitive adjustments made Western economies better off in the end. Why should China's rise, or India's for that matter, have different effects? The answer is that both introduce an unprecedented problem of scale. The earlier rising Asian economies – South Korea, Taiwan, Hong Kong, Thailand, Singapore, or Malaysia – had populations of middle-sized or small European states. Even Japan's population is smaller than that of France and Germany combined. As competition drove Japanese and Western wages toward a common average, the pressure on Japanese wages to go up was much greater than on Western wages to go down. Bringing the huge populations of China and India into the global workforce will have a different impact. The new global average for wages will have a long way to fall from current Western standards. Naturally, Western workforces will fight to insulate themselves – to preserve their living standards and their welfare systems. Under such circumstances, wide-scale protectionism seems a likely political imperative.

What do these trends imply for world order, for global governance in the future? Today's world is neither economically nor militarily unipolar, nor is it

moving in that direction. Instead of a closely integrated and US-dominated world economy, we should expect one that is more segmented and politically regulated. Countries relatively compatible with one another will perhaps group into large blocs, perhaps built around a dominant or common currency, or a relatively stable monetary union. Ideally, these blocs will remain reasonably open to each other. But the occasions for severe conflict – both within countries and among them – will be very great. In other words, today's unfolding trends do not promise automatic abundance or peace. Instead, they point to a harsh confrontation between Asian growth and Western prosperity.

If catastrophe is to be avoided in this new century, as the global system grows more plural, and inherently more conflictual, the world's great powers, rising and declining, must learn to conciliate each others' reasonable dreams, and develop machinery to anticipate problems before they grow unmanageable. Under these circumstances, it is difficult not to regard the whole American unipolar perspective on the future as a serious historic misperception, leading to dangerously confrontational policies. In a nuclear world, such policies are – to say the least – dysfunctional.

If the most promising model for global order in a plural world is not likely to be unipolar America's hegemonic fantasy, is there an alternative? If we are lucky, perhaps it will be something like the currently unfashionable European Union. As I wrote in 2003, and was trying to explain a few minutes ago, Europe's "hybrid confederacy" is a "highly creative evolution of the nation-state, a genuinely new political form" — a model for the "new regional systems needed elsewhere" — one that encourages states to moderate their goals and ferret out their common interests. There is, however, a dark side: Europe faces severe challenges and its success seems problematic. On the one hand, the challenge is organizational: Thanks to the Soviet collapse, Europe's union has doubled in size and is still growing. How can the Union embrace so many diverse countries without losing coherence and direction? On the other hand, the challenge is social and economic: In a global economy where European labor will increasingly be competing with the labor forces of China and India, how can Europe preserve its welfare states?

Each challenge is also a dilemma. The promise of enlarging the membership seems the most effective way to stabilize former communist countries, and induce them to rapid and effective transformation toward constitutional governments and market economies. But while enlargement may bring great benefits to the liberated states of Central and Eastern Europe, it also threatens to seriously weaken the EU itself. Some constitutional formula will have to be found to restore the EU's cohesion and direction. Since I wrote all this in 2001 and 2003, a major attempt has been made to redo Europe's constitution – and it has failed. Indeed, the attempt generated a strongly adverse public reaction in two of the European Community's original member states – France and the Netherlands. No one can now think resolving the constitutional challenge will be easy. And yet some solution must be found if the Union is to regain its vigor.

What about Europe's welfare states? They face, I argued, a similar dilemma. Continental Europe's communitarian states are a great advance over the more primitive capitalism being vaunted in Britain and America. But with global markets

growing radically competitive, Europe's economic models will be increasingly challenged to reconcile their humane social values with greater economic efficiency. Improving efficiency was the thrust of the so called Lisbon goals, proclaimed already in March 2000: The idea was to improve Europe's technology and upgrade its labor. The goals were admirable, and necessary, I agreed, but an insufficient response. Given the huge size of the Chinese and Indian labor forces, and their extremely low wages, protectionism – political regulation of trade – seemed to me an inevitable component of any successful European adaptation to the new globalism. Having the EU might, I thought, greatly strengthen Europe's ability to manage such a policy successfully. Protection within an internally liberal but large and diverse bloc would risk far less collateral damage to Europe's competitiveness than protection on a national scale. It would be much harder for special interests to dominate the process – to turn it into a racket.

The real challenge, however, would be whether a protectionist EU could reconcile its own needs with those of a rapidly rising China. China, too, had a right to grow and flourish and would certainly assert that right energetically. China was itself a huge and rapidly expanding market, with particular opportunities for Western industries. Europe could not afford to lose those opportunities – to cut itself off from the most dynamic parts of the new world economy.

In short, balancing the conflicting aims and values of old rich Western states and newly rising giant Asian poor states poses a great intellectual and practical challenge, one requiring a systematic and energetic search for common interests. In the twentieth century, the world had paid a terrible price for the failure of the global system to make room for rising states. Today, China's huge scale – not to mention India's – makes those problems of the 20th century seem comparatively trivial. That is why Europe's rich postwar experience in supranational institution building is so important for the world, as well as for itself. Our new world urgently needs Europe's gift for appeasing rival interests – for transforming conflict into cooperation. In effect, as *Rethinking Europe's Future* argued, Europe's postwar confederal experience is central to managing the problems of our new century.

From this perspective, what happens in Europe now seems critical to what eventually happens in the world. If it is America's proper role to accept the plural world that it has done so much to foster, it is Europe's proper role to precipitate the advent of that world and shape its character. For the United States, this means exorcising the demon of unipolarity. It means finding a new way of looking at the world, one that accepts pluralism and finds a more subtle way to manage it. For Europe, it means consolidating and reinvigorating its union and accepting once more the responsibilities of great power in the world. For both parts of the West, it means finding together the political and institutional imagination needed to embrace China, India, Russia and the rest of the rising world.

For most of modern history, Europe has been a furnace of creativity and energy. The last century shows vividly what can happen when that force fails or is misdirected, and what can be accomplished when Europe is at its best. We are, I suspect, at a critical time for setting the patterns of the future. And Europeans, whether they like it or not, have a determining role to play.

A final thought to close: We should never forget that a strong Europe, with a mind of its own, is very much in the interest – not only of the world in general, but of the US in particular. This is not only because the US – by itself – lacks the resources, legitimacy and skill to manage today's global challenges, but for a more intimate reason: America's hyperactive power abroad threatens its own constitutional balance at home. So much military power and financial wealth, combined with an enthusiastically imperial mindset among America's political elites, threatens to overwhelm our country's old-fashioned system of national checks and balances. Europe and the US have become part of each others' constitutional balance. The more powerful the US becomes, the more a friendly but strong Europe becomes essential not only to limit, refine and reinforce American power in the world, but also to contain it at home. A strong, cohesive and assertive European Union would itself go a long way toward ending American daydreams of a unipolar world. The US needs a strong Europe – to preserve its own sanity. A global superpower, it seems, requires a global dimension to its constitution. Europe, to be sure, also needs the US. There would never have been a European confederacy without strong American support. The US was Europe's silent partner, a guarantor against Europe's Bad Old Times returning.

If Europe and America are to regain their interest and regard for each other, their relationship now urgently needs redefinition and rejuvenation. If we fail, as we failed in the Iraq war, we shall end up defeating each other. Instead of a united and balanced West, there will be an overextended and hysterical America and a fragmented and embittered Europe. In the great Eurasian drama that is to come, we can pray that both Europe and America play their parts well, that each exercises its own best form of leadership, that each lives up to its responsibilities before history.

The European Integration and Its Effects on the European Monetary Union Stock Markets

Faarnaz Chavoushi

The issue of correlations among stock indices is essential for effective diversification strategies of portfolio managers. However, it is often claimed that the economic and financial integration of European markets has increased correlations between stock market indices, making it less attractive for portfolio managers to diversify among European assets. This paper examines the development of the correlation structure between country indices during monetary and economic integration in the European Monetary Union and finds that correlations increased considerably during the sample period of 1979 to 1999, and decreased again after the introduction of the euro. This paper seeks an economically sound explanation for this counterintuitive observation.

INTRODUCTION

The issue of European stock market integration is of considerable importance to both investors and the economy as a whole. Deregulation, elimination of cross-border restrictions on banking and securities transactions, and the abolishment of currency risks are all factors which have contributed to increased cross-border investment activity and accelerated equity capital flows between markets in the European Monetary Union (EMU). With exchange rates no longer a barrier to equity trading in the euro area, the European Commission is now working on harmonizing and eliminating regulatory and structural obstacles such as cross-border trading restrictions, different accounting systems for financial reports, and cross-border trading costs.

The gradual creation of a single market for equities is expected to raise competitiveness levels through the efficient allocation of capital, by mobilizing savings to a larger and more liquid capital market and by disciplining managers. In fact, the creation of an integrated European capital market is seen as one of the strategies established to reach the goals of the Lisbon agenda and to outperform the American economy by 2010. However, although the elimination of currency

Faarnaz Chavoushi is an M.A. candidate at the Bologna Center. Born in Iran, she grew up in the Netherlands. She previously obtained a master's degree in finance and investment, at the Rotterdam School of Management and completed B.A. degrees in economics and international business administration at the Erasmus University.

exchange risk and the high degree of integration probably have substantial positive effects on the competitiveness of the European economy, their implications for the portfolio management industry are somewhat ambiguous.

Until a few years ago, it was common practice among portfolio managers to follow a top-down approach to asset selection. This approach concentrated on first dividing the money over several countries and second on spreading the investment within a country. However, at the end of the 1990s, the argument was increasingly made that the country orientation of the top-down approach should give way, within the euro era at least, to an industry or sector orientation. This proposal is based on the observation that, as stocks become more integrated and move together, the diversification benefits of investing in many countries may be reduced, as correlations among European country indices will increase. Therefore, the focus of diversification should shift from a country level to the industry level across European markets. As a result, superannuation funds and life insurance companies that hold long-term investment portfolios and adopt policies of passive international diversification are increasingly expected to switch portfolio compositions in order to achieve efficient portfolios.

This change in asset allocation strategy is not a minor one. It implies that the teams of analysts, now organized along country lines, are to be reorganized along industry lines. If investment firms were to follow this advice, they would put less time and effort into examining the macroeconomic outlook of a country, and could instead focus on analyzing the prospects of an industry and of specific firms within that industry. Moreover, as the European stock markets become less distinct from one another, asset managers looking to diversify their stock portfolios might increasingly divert part of their assets from the European market to other regions where cross-market correlations are still at a lower level. This would lead to a gradual withdrawal of investment flows away from the European stock markets, which could potentially have substantial implications for the size and liquidity of the European markets, not to mention the ability of continental companies to raise sufficient money for investment projects in the financial markets. As such, the process of financial and monetary integration that the European Union has actively pursued in order to enhance its role as a great economic power in the world can have adverse effects on an important engine of the European economy. Since neither local governments nor the EU as a supranational institution has the possibility of protecting the European financial markets against adverse shocks in investment flows, this could be a serious concern. After all, the only instrument a government has at its disposal to protect its markets from external shocks is the imposition of capital flow restrictions. Capital flow restrictions are not only controversial and deplorable from a liberal-market point of view, but also unsuitable as an instrument in this particular case, since the withdrawal of funds by investment managers would be based on rational economic foundations rather than being driven by short-term investment sentiments.

With the pursuit of great power come many unforeseen consequences, but the issue becomes particularly pressing when one does not have the instruments with which to take responsibility for these consequences. Nevertheless, despite the lack of a responsive instrument for the government at the moment, it is

extremely important to have at least an awareness of the possible future challenges that the pursuit of the process of closer monetary integration can have. The aim of this paper is therefore to assess whether the recent shift of asset allocation strategy would indeed be rationally justified on the basis of financial economics theory. If so, then that would potentially have grave consequences for the financial markets of the European countries. If not, then we can retain the hope that economic rationale will settle the recent discussion on the optimal asset allocation strategies without bringing about a significant change in the direction of investment flows.

This paper evaluates the optimal asset allocation strategy by examining the effects of the process of monetary and economic integration in the EMU on the correlation levels of cross-country indices in Europe, both from a theoretical and empirical perspective. The paper will show that although intuition would lead one to expect that increased monetary and financial integration will lead to higher cross-market correlations, the theoretical and empirical evidence is rather ambiguous. In fact, this study finds that although correlations among the EMU countries substantially increased from 1979 to 1999, cross-country correlations have almost informally dropped since the introduction of the euro. The final section of this paper will discuss a number of theoretical premises that can reconcile our intuitive expectations with the empirical findings.

MARKET INTEGRATION IN THE EUROPEAN UNION AND ITS EFFECT ON STOCK MARKETS

In the portfolio management industry, financial investment and security firms base their optimal asset allocation strategies on the diversification benefits as suggested by the Modern Portfolio Theory. First introduced by Markowitz (1952), Modern Portfolio Theory shows that the risk of a portfolio of assets is not equal to the sum of the weighted average of the variances of individual securities, but also depends on the correlations between the assets. The weaker the correlations between assets are, the greater is the reduction in portfolio risk. Therefore, a well-diversified portfolio (among low correlated stocks) can bring the same return for an investor at a lower risk level for investors. The rationale of international diversification in portfolio management thus highly depends on low correlations between international equity markets. It is therefore not surprising that a lot of research has been conducted on the evolution of correlations among country indices over time. This chapter will discuss the issue of monetary and financial integration within Europe and its possible implications for cross-country correlations and portfolio management.

There are a number of mechanisms through which the strengthening of monetary and economic integration in the EMU can be expected to be accompanied by a substantial increase in cross-country correlations among European markets. The coming section will give a conceptual outline of the link between monetary and financial integration and correlations among stock markets.

THE POSSIBLE EFFECTS OF MONETARY AND ECONOMIC INTEGRATION

The effect of monetary and economic integration on correlations is threefold. First, as a result of the opening of the market and increased levels of trade between nations, the profitabilities of companies across borders will become more interrelated. Another important co-movement mechanism is the convergence of real economies as a result of the increased monetary coordination. Having more similar business cycles and being more interdependent through trade could mean a convergence in expected cash flows and volatilities, resulting in a co-movement of profits and dividends of European firms, and thus a more homogeneous valuation of equities and an increased cross-country correlation in asset prices.

A second outcome of the economic integration is the effect of the convergence of monetary policy on the valuation of companies. As inflation rates and interest rates converge at a European level, dividends and net profitability of companies are discounted at a similar rate, possibly leading to a convergence of stock prices across nations. Furthermore, as exchange rate fluctuations diminish over time, the currency risk factor incorporated in stock prices should be eradicated as well.

Third, exchange rate fluctuations are mainly driven by national economic policies and form an important source of risk priced on capital markets. A more volatile exchange rate of a country raises the national risk premium as investors require a higher return to compensate for the increased uncertainty. In fact, the existence of exchange rate uncertainty can function as an important device for market segmentation. The more volatile and unpredictable exchange rates are, and the more costly hedging against uncertainty is, the stronger the degree of market segmentation and the lower the degree of correlation across markets. Analogously, the reduction or elimination of currency risk, as entailed in the EMU and the introduction of the euro, may raise the degree of financial integration across countries and simultaneously lead to more homogeneous reward to risk ratios across European stock markets.

THE POSSIBLE EFFECTS OF FINANCIAL INTEGRATION

Due to improvements in computer and communication technology, adjustments in delays in international prices are shorter and stock markets have become more synchronized. Furthermore, since controls on capital movements and foreign exchange transactions are relaxed, shocks that have an effect on the valuation of many assets worldwide are more easily transferred across countries in an integrated market. For example, after the accounting scandals concerning Ahold, many Dutch institutional investors temporarily lost market confidence and, as a reaction, sold a large part of their portfolio on a European level. If transaction costs and barriers to trade had been higher (like before the start of the financial integration process), other financial markets in Europe probably would have been less affected by the market specific shock from the Dutch market. Thus, as the barriers in the transmission channel of the financial markets are gradually eliminated (like envisioned in the Financial Service Action Plan), one can expect that shocks to common risk factors affect more countries with the same scope, leading to a higher co-movement between the markets.

PREVIOUS EMPIRICAL RESEARCH

Research has found substantial empirical evidence that correlations are time-varying and that correlations across markets increase during periods of higher economic and financial integration, which is bad news for portfolio managers in the modern era. For example, Goetzmann, Li and Rouwenhorst (2001) collect information from 150 years of global equity market history in order to evaluate the evolution of equity correlation matrices through time. They suggest that the structure of global correlations shifts considerably through time and that periods of free capital flows are associated with high correlations. They find that today, in the beginning of the 21st century, global correlations are at a historical high (though approaching levels of correlation last experienced during the Great Depression). Other researchers have also concluded that the international diversification potential today is very low compared to the rest of capital market history. For example, Rouwenhorst (1999) looks at EMU correlations specifically and finds that both country and sector correlations are increasing over time in the period of 1978 to 1998, although the average increase is not as pronounced for sectors as it is for countries.

Surprisingly, however, more recent empirical research from other authors who examined the correlation structure of European stock markets is not in line with this earlier conclusion. For example, in an early paper, Adjaoute and Danthine (2000) first identified a significant increase in the degree of correlation between national stock indices implying that diversification opportunities have been significantly reduced after the introduction of the euro. However, after revisiting their earlier work in mid-2001 with an extended sample data, they find that more recent data convey a radically different message. They find that the period after the introduction of the euro is characterized by lower correlations than those obtained for the immediately preceding period of the same length. Another example of a recent study that reports a trend in the cross-country return correlation within the EMU, is the paper by Ferreira (2003). Examining the period from 1975 to 2001, it finds that both EMU country and industry correlations reach a peak in the 1995-1998 period and then decrease in the last period.

These findings are surprising since the introduction of the euro is said to have eased the transfer of capital within the European Union, and, based on the analysis of Goetzmann et al, one would expect higher correlations in such periods. In his paper, Ferreira (2003) suggests that the low correlations in the last sub-period of the sample may be related to a reduction in the average stock return of the EMU market. However, Goetzmann et al (2001) have shown convincingly that periods of poor market performance have been associated with high correlations, rather than low correlations.

METHODOLOGY AND DATA

As we have seen, the issue of cross-country correlations is essential for optimal diversification strategies of portfolio managers. An increase in the correlation coefficients could indicate increased market integration over time and a possible reduction of diversification benefits for a mean-variance optimizing

investor. Yet empirical research so far has reached ambiguous and indefinite conclusions on the effects of economic market integration on stock market correlations in the EMU. This paper will therefore take a pragmatic approach and will directly assess the developments of cross-country correlations in the EMU area since the start of the economic integration period.

Toward this purpose, the paper aims to examine the following hypothesis: *Correlations between EMU country indices are rising as the level of economic market integration in the EMU increases over time.*

A comparison of the correlation coefficients among the country indices is made over time to see whether there are significant changes in the four subperiods specified. In order to capture the process of integration in the European stock markets, an analysis was performed over four separate sub-samples, each indicating a different phase of economic integration in the European Monetary Union. Table 1 gives a description of the sub-periods. The sample period is from January 1979 to July 2005. The study employs total market indices composed by DataStream International. Weekly, dividend adjusted, continuously compounded stock returns are used for 12 countries that have signed up for EMU. Unfortunately, data is not available for Finland, Portugal, Spain, Greece, and Luxembourg for the pre-convergence and convergence period. Therefore, data from these countries will only be taken into consideration for the pre-euro and euro era. The sample contains weekly returns on country and industry indices. In order to calculate continuous returns, the price series used in the study are transformed by taking natural logarithms of the raw data. For further ease of comparison, all returns are expressed in the Dutch guilder until January 1, 1999, and afterward in the common euro currency. This method is preferred to expression in local returns, since in practice most investors do not hedge their returns and are subject to exchange risk fluctuations. However, expressing returns in a common currency may introduce a bias toward higher correlations of returns over the years when the EMU exchange rates have become more stabilized and correlated. This makes it more difficult to draw conclusions from the evolution of the correlation matrices, since higher observed correlations can be caused by financial integration as well as a convergence of exchange rate fluctuations. The use of local currency (Dutch guilder) returns are justified; however, since irrespective of the cause of the observed development of correlations over time, the practical implications to local investors are the same.

MAIN EMPIRICAL RESULTS

This section will first give a description of the statistics and the correlation structure of the assets in the period from January 1979 until July 2005. In the discussion chapter, these empirical observations are reviewed and put into perspective by making a comparison to the findings of other studies and by relating the results of this study to the theory of financial and monetary integration.

Table 2 shows summary statistics over the different sub-periods. These include the average annual returns, annual standard deviations and corresponding

risk-reward ratios (from now on referred to as Sharpe ratios). The table illustrates a number of things. First of all, it clearly shows that the mean returns of the capital markets went through different phases during the different sub-periods. For example, the average mean return for the pre-convergence period was 18.3 percent. For the convergence period it was 10.4 percent, for the pre-euro period it was 19.4 percent, and during the euro period it was only 4.4 percent. Although standard deviations also varied over time, the divergence was much smaller. As a result we can see that the standardized returns (Sharpe ratios) vary considerably through time as well. Correspondingly, during the pre-euro period, almost all stock markets experienced a dramatic value increase, only to deflate just as sharply a few years later during the euro era.

When we look within the asset classes in more detail, we see large differences between the individual assets. Ireland has the best absolute performance for countries (average annual return of 13.68 percent) but when corrected for its high average standard deviation, it does not have the best standardized return. The best standard return is achieved by the Dutch index, with a Sharpe ratio of 0.84.

Table 3 shows the average correlations across country indices. The table shows us first of all that the average correlation coefficients for countries with a smaller capital market (Austria, Greece, Ireland and Luxembourg) compared to the rest of the EMU stock markets (0.35) are much lower than average correlations of the larger markets (0.48). Italy is the exception in this trend since it is the only country with a large market capitalization that is experiencing relatively low correlations with the rest of the EMU countries until the last sub-sample period. The results for the correlation coefficients depicted here reconfirm the findings of Yan, Min and Li (2003) who conclude that smaller European markets are less integrated with the rest of the European capital markets, since their small market and concerns of market illiquidity most probably withhold investors from investing in those markets, although they have the same macroeconomic linkages with the other EMU countries.

Table 4 also clearly depicts that correlations have increased considerably through time in order to reach a maximum in the 1994-1998 period, and then decrease slightly in the last sample period. The finding that cross-country correlations have decreased after the introduction of the euro seems rather surprising and in contrast to the intuitive expectations of increased correlations. In order to test that the increase (and later the drop) in average correlation trends is not heavily influenced by the figures of a few countries only, the following graphical analysis is conducted following Adjaoute and Danthine (2001). First, the country or sector index returns are used to compute the unconditional correlation matrices for the different periods. Secondly, the correlation pairs are sorted in ascending order and plotted against the country pairs for each subperiod (i.e. pre-convergence period). Third, the correlation pairs for the subsequent period (i.e. convergence period) are recomputed and plotted along the correlation pairs of the previous period (i.e. pre-convergence) correlation pairs.

The results are visible in figures 1A to 1C. Figure 1A shows the evolution of country pair correlations by comparing the pre-convergence and convergence periods. It is obvious that correlations between all country index pairs were

uniformly larger in the second period than in the first. The average increase of the correlation coefficient was remarkable: 0.324 points. Figure 1B, however, shows a slightly different pattern for the convergence and pre-euro periods. Coefficients for most correlation pairs are larger in the second period than in the first, but the average increase was a mere 0.076 points. Figure 1C on the other hand, shows a radically different picture. In the euro period, it is no longer the case that all correlations are uniformly higher than in the immediately preceding period. In fact, 36 out of 65 correlation coefficients are *lower* in the euro sample than the pre-euro sample. The average decrease in correlations was a slight -0.030 points.

DISCUSSION OF RESULTS:

By illustrating the time-varying pattern of correlations, we have thus seen that cross-country correlations have varied very much in the different subperiods. During the convergence period, there was a drastic increase of correlations among both country and sector indices, illustrating the importance of the free market and the increased trade levels after the Single European Act. However, we have also seen that the upward trend has been broken in the euro era, where cross-country correlations actually declined. This last finding is in contrast to the research by Rouwenhorst (1999) and Goetzmann et. al. (2001) who find that periods of free capital flows are associated with higher correlations. On the other hand, the findings here do correspond to those of other recent studies on this topic, most notably of Adjaoite and Danthine (2000) and Ferreira (2003). This section will attempt to discuss the controversial observation of a plunge of correlation levels in the euro period and try to reconcile it with economic theories described previously.

First of all, despite what is sometimes suggested in the literature, the fact that in the euro period cross-country correlations have decreased does not have to mean that the disappearance of currency risk and the convergence of economic policies have been unimportant for financial markets. Instead, this phenomenon might actually stress the importance of the convergence process *before* the introduction of the euro.

Ever since the establishment of the Single European Act, Europe has been in a process of increasing economic integration. In that transitional phase, the fiscal policies of the member states were more synchronized to a common budget norm for all member states, which put a lot of pressure on the governments of some countries to reform their economic policies. Furthermore, inflation rates were strongly reduced in all member states and interest rates and exchange rate fluctuations were diminished, all of which required deliberate action from the government of the EMU member countries. As a result, this transitional phase has led to large identical macroeconomic shocks in most countries, which have apparently also had a strong impact on the stock markets (judging from the high correlations between national indices in that period).

However, since the establishment of the European Central Bank (ECB) in 1999 and the introduction of the euro, there have been few macroeconomic shocks.

The fiscal policies were already synchronized to the European norms in the 1990s, and after its installation, the ECB has kept the inflation rates at a consistently low level, without much fluctuation among the European interest rates. In the absence of strong macroeconomic shocks, the European stock markets have possibly become less influenced by EU-wide macroeconomic shocks, and relatively more influenced by their domestic idiosyncratic shocks.

Although financial market integration in the EMU has accelerated in the last subperiod, higher financial integration in itself does not necessarily result in higher correlations. Financial integration merely opens the channels through which international (EU-wide) shocks are transmitted, but in itself it does not amplify correlation levels. After all, in a completely integrated market, returns on different assets may be divided into a common component and an idiosyncratic one. In the absence of strong European shocks, the existence of the idiosyncratic component may be sufficient as to render *ex post* correlations rather low.

Alternatively, cross-country variation in investor sentiment drives a wedge between the return of firms that are in the same sector but located in different countries. This indicates that differences in institutions across countries affect the transmission of global shocks to asset values. Differences in interest sensitivity of aggregate demand (due to, for example, heterogeneous industrial structures) and the importance of the 'credit channel' (differences in the structure of the banking system, business firms' reliance on bank loans, and financial stability) are potential determinants of the country-specific economic effect of global shocks (Mihov 2001). As such, even if a common monetary policy is carried out, as long the transmission of monetary policy to economic activity remains different among EMU countries, business cycles may not synchronize and correlations among EMU country indices may remain low.

Thirdly, the literature that predicts the waning of country factors is based on the assumption that the monetary and financial integration brings a convergence of risk-free rates and of risk premia which implies that stocks will have a tendency to be priced closer together. However, although the pricing components of equity returns have indeed converged, the objects being priced might have changed over time, introducing increasing divergence in country returns. In particular, if a country's industrial structure becomes more specialized, the fundamentals of country indices become increasingly dissimilar leading to less synchronized returns (Adjaoute and Danthine 2004).

Fourth, as is argued by Beckers et al (1992), low correlations between different stock markets may be perfectly consistent with complete market integration. Different industries may play a crucial role in each market: pharmaceuticals in Switzerland, oil in the United Kingdom, or forestry and paper in Finland. Stocks within a given industry may be similarly impacted by global events: oil companies around the world in general should react in a similar way to changes in oil prices or the invasion in Iraq; financial institutions may broadly be affected in the same way by the overall level of interest rates regardless of the EMU country they are situated in; textile companies may react similarly to GATT decisions on production and import quotas. Thus, the evolution of the correlation

between two indices could be caused either by the industry or the country factors of each index return.

Random shocks that selectively affect different industries would therefore naturally lead to low correlations among nations when particular industries are more present in one country than another. It is, for example, well known that particularly the information technology (IT) sector was severely hit by the market crash in the period after 2001. It might be that companies from the IT sector are more heavily represented in the market index of some countries (for example Nokia in Finland has a weight of 64 percent in the market index). As a result, the returns of the national indices diverge more from other countries when these particular sectors show strong fluctuations. Of course, that would also have implied that correlations should have been lower in the previous subperiod when the IT sector was experiencing extraordinary positive returns. But, assuming that industry effects have become more important than country effects in the recent years (as is suggested by much of the contemporary literature on this topic¹), the combination of the increased importance of the sector effects and the recent collapse of the IT bubble can give a plausible explanation for the decrease in cross-country correlations.

CONCLUSIONS

Investments from the asset management industry constitute a very significant portion of the total investments that go into national stock markets. As such, they are an important source of financing for large companies that attempt to finance their investment projects through their national stock markets. A substantial decrease in the investment flows of the asset management industry in the European continent can thus have grave consequences for not only the national stock markets but for the vitality and functioning of the European economies as a whole. Asset management firms base their decisions for their investment flows on the basis of modern portfolio theory, which stresses the importance of low correlations in the portfolio selection process. The lower the correlation among stock markets, the more a portfolio manager can reduce the total risk of the portfolio and thus the higher the incentive to diversify the portfolio by investing in these stock markets.

The problem that has been identified is the fact that the economic and financial integration of European markets has the potential of having increased cross-country correlations through a number of mechanisms, making it relatively less attractive for portfolio managers to diversify their portfolios by spreading investments over the various European stock markets. However, as we have seen, higher levels of market integration are neither a necessary nor sufficient condition for higher cross-country correlations. This paper therefore takes a practical approach and directly examines the development of the correlation structure

¹ See, for example, Cavaglia, Brightman and Aked (2000), Isakov and Sonney (2002), Carrieri, Erruza & Sarkissian (2000), Gerard, Hillion & de Roon (2002), or Adjoute & Danthine (2001).

between country indices during the process of monetary and economic integration in the EMU. The main empirical finding is that correlations among country indices have increased considerably during the period of closer economic cooperation and integration starting from 1979 until 1999. However, the study also finds that after 1999, intra-country correlations have *decreased* substantially and that consequently, a geographic diversification strategy is optimal for most investors even after the introduction of the euro.

Unfortunately, it is difficult to say whether the upswing in dispersion and the recent phenomenon of lower correlations among EMU countries is a temporary, permanent, or cyclical event. Unless the recent decrease in correlations is a permanent and continuing phenomenon, the contemporary plunge in the euro era in itself does not bring the correlations back to the levels experienced in the era before the process of integration in Europe started. This means that diversification benefits for mean-variance optimizing investors are not as substantial as they used to be, and a case can indeed be made for the restructuring of the asset allocation grids of portfolio management firms on an industry rather than a country basis. Also, asset management firms would increasingly be tempted to move away part of their portfolio from some of the European indices and look for better diversification opportunities at stock market indices in other regions of the world where cross-country correlations have not yet reached similar high levels. Interestingly enough, however, this study also finds that the smaller European markets, thanks to their low liquidity and small market size, are still subject to low cross-correlations compared to the rest of the EMU countries. As a result, instead of being unfavorably affected by the process of monetary integration, the smaller EMU stock markets may in fact have become increasingly *attractive* as international diversification options. The increased appeal of these small stock markets for the asset management industry can, in the long-term, result in an increase in the size and liquidity of these small markets as portfolio managers seek their refuge in these markets for portfolio diversification purposes. Hence, the process of monetary and financial integration in the EMU can substantially affect the size and liquidity of worldwide stock markets even beyond the increased cross-country correlation levels observed in this study. As such, dealing with the distortion of financial market structures is one of the many responsibilities the European Union will have to undertake as a result of its pursuit of great power ambition and the pioneering policies associated with it.

REFERENCES

- Adam, K., T. Jappelli, A. Menichini, M. Padula and M. Pagano, "Analyze, Compare and Apply Alternative Indicators and Monitoring Methodologies to Measure the Evolution of Capital Market Integration in the European Union." Centre for Studies in Economic and Finance, Department of Economics and Statistics, University of Salerno, January 28, 2002.
- Adjaouté, K. and J.P. Danthine, "EMU and Portfolio Diversification Opportunities." Centre for Economic Policy Research Discussion Paper, no. 2962

(2000).

- Adjaouté, K. and J.P. Danthine, "Portfolio Diversification: Alive and Well in Euroland!" Centre for Economic Policy Research Discussion Paper, no. 3086 (2001).

- Adjaouté, K. and J.P. Danthine, "Equity Returns and Integration: Is Europe Changing?" FAME Research Paper, no. 117 (2004).

- Aked, M., C. Brightman and S. Cavaglia, "Increasing Importance of Industry Factors." *Financial Analyst Journal* 56, no. 5 (September/October 2000): pp. 41-45.

- Beckers, S., R. Grinold, A. Rudd and D. Stefek, "The Relative Importance of Common Factors Across the European Equity Markets." *Journal of Banking and Finance* 16 (1992): pp. 95-96.

- Brooks, R. and M. Del Negro, "The Rise in Co-movement Across National Stock Markets: Market Integration or IT Bubble?" Federal Reserve Bank of Atlanta Working Paper, no. 2002-17a (2002).

- Eiling, S., B. Gerard and F.A. DeRoos, "Asset Allocation in the Euro-Zone: Industry or Country Based." Center Discussion Paper, Tilburg University, 2004.

- Ferreira, M.A., "The Importance of Industry and Country Effects in the EMU Equity Markets." SSRN Working Paper (March 2004).

- Fratzscher, M., "Financial Market Integration in Europe: On the Effects of EMU on Stock Markets." *International Journal of Finance and Economics* 7 no. 3 (2002): pp. 165-193.

- Gerard, B., P. Hillion and F.A. de Roos, "International Portfolio Diversification: Industrial Structure, Country and Currency Effects Revisited." SSRN Working Paper (March 2002).

- Goetzmann, W., L. Li and G. Rouwenhorst, "Long-Term Global Market Correlations." NBER Working Paper no. 8612 (2001).

- Hardouvelis, G.A., D. Malliaropulos and R. Priestley, "EMU and European Stock Market Integration." Centre for Economic Policy Research Discussion Paper no. 2124 (1999).

- Heckman, L., S. Narayanan and S. Patel, "Country and Industry Importance in European Returns." Salomon Smith Barney Working Paper (February 11, 1998).

- Heston, S.L. and K.G. Rouwenhorst, "Industry and Country Effects in International Stock Returns" *The Journal of Portfolio Management* (Spring 1995): pp. 53-58

- Isakov, D. and F. Sonney, "Are Practitioners Right? On the Relative Importance of Industrial Factors in International Stock Returns." SSRN Working Paper, HEC-University of Geneva no. 2003.04 (2002) and FAME Research Paper no. 72.

- Longin, F. and B. Solnik, "Is the Correlation in International Equity Returns Constant: 1960-1990?" *Journal of International Money and Finance* 14, no. 1 (1995): pp. 3-26.

- Longin, F. and B. Solnik. "Extreme Correlation of International Equity Markets." *Journal of Finance* 56, no. 2 (2001): pp. 649-675.
- Markowitz, H., "Portfolio Selection." *Journal of Finance* 7, no. 1 (March 1952): pp. 77-91.
- Markowitz, H., "Portfolio Selection: Efficient Diversification of Investment." Yale University Press, New Haven, CT, 1959.
- Rouwenhorst, K.G., "European Equity Markets and the EMU." *Financial Analyst Journal* 55 no. 3 (1999): pp. 57-64.
- Sharpe, W.F., "The Sharpe Ratio." *Journal of Portfolio Management* (Fall 1994): pp. 49-58.
- Yang, J., Min, I., Li, Q., "European Stock Market Integration: Does EMU Matter?" *Journal of Business Finance and Accounting* 30 (2003): pp. 1253-1276.

APPENDIX I

Table 1 Description of subperiods

The table shows the data range and observations for the whole sample as well as the sub-periods. The data provided by Datastream is ranging from January 4, 1979 until the end of July 2005 (1385 observations).

Sample Period	Significant Developments	Data Range	Observations	Notes
Pre-convergence period	Developments following the establishment of the European Monetary System in 1979 aimed at reducing the fluctuation of intra-European exchange rates and at safeguarding price stability	Jan 4, 1979 - Dec 26 1985	364 observations	No data available for Finland, Portugal, Spain, Greece, or Luxembourg
Convergence period	In February 1986, the Single European Act imposed a deadline, the end of 1992, for the complete abolition of all obstacles to free cross-border movement of labor, capital, goods and services. At the same time, the regulatory structure of the financial markets was harmonised with the First and Second Banking Directive and a set of other legislative and institutional measures.	Jan 2, 1986 - Dec 30, 1993	417 observations	No data available for Finland, Portugal, Spain, Greece, or Luxembourg
Pre-euro period	In February 1992, the Maastricht Treaty was signed, which became effective in November 1993, and opened the possibility for a monetary union among those EU countries that would satisfy the convergence criteria on inflation, interest rates, the public finances and exchange rates.	Jan 6, 1994 - Dec 12, 1998	260 observations	
Euro period	In 1999, a common monetary policy conducted by the European Central Bank became effective across EMU and the common currency; the euro was introduced in giral form.	Jan 7, 1999 - July 21, 2005	341 observations	

Table 2A. Descriptive statistics

Shows the mean annualized returns, standard deviation and the respective risk return ratio of the country and industry portfolios. The weekly returns are calculated in the local (Dutch) currency and are in percentages. Risk is measured as annualized standard deviations. The risk return ratio is simply calculated by dividing the industry or country index by the respective standard deviation, thus assuming a risk-free interest rate of 0%.

In order to calculate continuous returns, the price series used in the study are transformed by taking natural logarithms of the raw data. For further ease of comparison, all returns were expressed in the Dutch guilder until 1st January 1999, and afterwards in the common euro currency.

	Average Whole Sample Jan 1979 - April 2004		Pre-Convergence Jan 1979 - Dec 1985		Convergence Jan 1986 - Dec 1993		Pre-Euro Jan 1994 - Dec 1998		Euro Jan 1999 - July 2005v						
	Return	Risk	Return	Risk	Return	Risk	Return	Risk	Return	Risk					
THE NETHERLANDS	13.461%	16.003%	0.841	21.174%	14.258%	1.485	12.337%	13.578%	0.909	21.157%	16.740%	1.264	0.729%	19.463%	0.037
FRANCE	12.963%	19.147%	0.677	16.301%	19.723%	0.827	13.991%	18.857%	0.742	15.076%	17.752%	0.849	6.532%	19.939%	0.328
GERMANY	8.996%	17.197%	0.523	15.470%	11.692%	1.323	6.493%	17.903%	0.363	14.460%	15.961%	0.906	0.976%	21.593%	0.045
BELGIUM	12.330%	15.930%	0.774	14.647%	15.127%	0.968	12.473%	16.446%	0.758	20.820%	13.771%	1.512	3.204%	17.559%	0.182
AUSTRIA	12.671%	16.924%	0.749	20.305%	14.034%	1.447	10.500%	22.968%	0.457	1.856%	14.626%	0.127	15.430%	11.902%	1.296
ITALY	12.451%	24.288%	0.513	27.990%	27.354%	1.023	2.122%	23.888%	0.089	18.437%	25.809%	0.714	3.921%	19.547%	0.201
IRELAND	13.681%	19.609%	0.698	12.553%	16.343%	0.768	15.110%	23.303%	0.648	23.398%	17.720%	1.320	5.723%	19.247%	0.297
FINLAND	10.888%	30.661%	0.355							32.184%	26.170%	1.230	6.343%	37.543%	0.169
PORTUGAL	5.407%	16.856%	0.321							20.789%	19.192%	1.083	-0.523%	15.421%	-0.034
SPAIN	10.306%	21.141%	0.488							22.973%	20.188%	1.138	4.585%	18.770%	0.244
LUXEMBURG	12.237%	17.916%	0.683							14.535%	11.543%	1.259	2.921%	22.351%	0.131
GREECE	13.288%	29.171%	0.456							27.270%	25.772%	1.058	3.447%	26.400%	0.131
Average	11.557%	20.404%	0.590	18.349%	16.933%	1.120	10.432%	19.563%	0.567	19.413%	18.770%	1.038	4.441%	20.811%	0.252

Table 2B. Extremes of EMU weekly returns

Countries	Pre-Convergence		Convergence		Pre-Euro		Euro	
	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest
	Netherlands	France	Netherlands	Italy	Netherlands	Austria	Austria	Portugal
	1.485	0.827	0.909	0.089	1.264	0.127	1.296	-0.034

Table 3 Average Cross Country Correlations

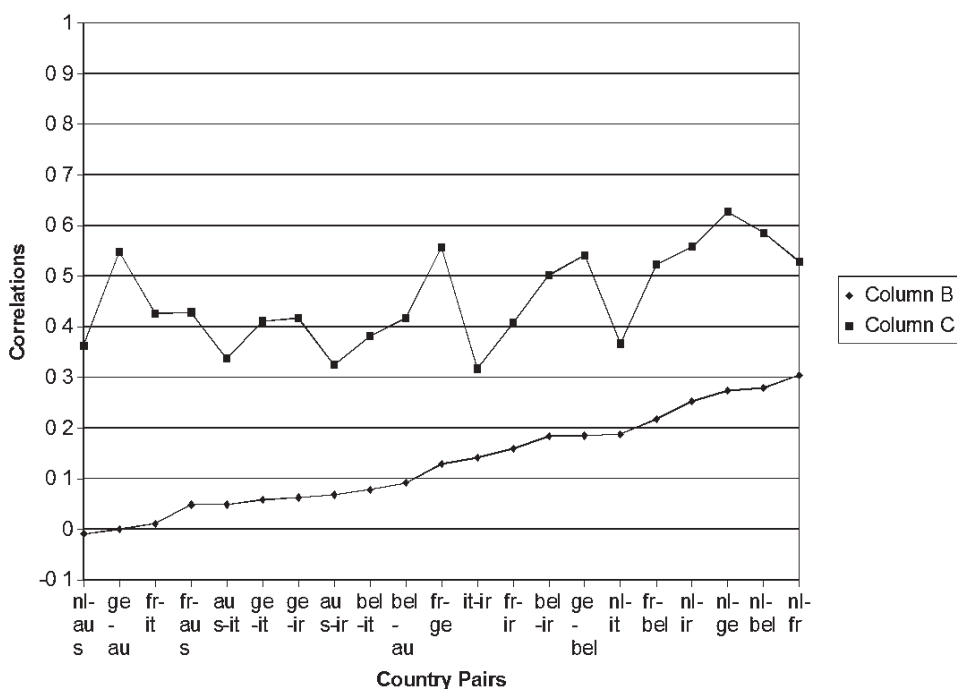
Shows the average correlations of a country index with the other country indices (eg the average correlation of the Dutch stock index with the other 11 EMU countries). Please note that the correlation of each index with itself is excluded from the average numbers in the columns.

	Pre-Convergence Jan 1979 - Dec 1985	Convergence Jan 1986 - Dec 1993	Pre-euro Jan 1994 - Dec 1998	Euro Jan 1999 - April 2004	Whole Sample 1979-2004
NETHERLANDS	0.184	0.433	0.580	0.582	0.537
FRANCE	0.124	0.410	0.548	0.605	0.490
GERMANY	0.102	0.443	0.575	0.604	0.532
BELGIUM	0.148	0.422	0.543	0.497	0.482
AUSTRIA	0.036	0.346	0.484	0.292	0.333
ITALY	0.075	0.320	0.459	0.563	0.389
IRELAND	0.124	0.361	0.496	0.430	0.414
FINLAND			0.496	0.390	0.443
PORTUGAL			0.445	0.392	0.419
SPAIN			0.522	0.549	0.535
LUXEMBURG			0.367	0.310	0.339
GREECE			0.333	0.305	0.319
Average Countries	0.132	0.456	0.532	0.501	0.436

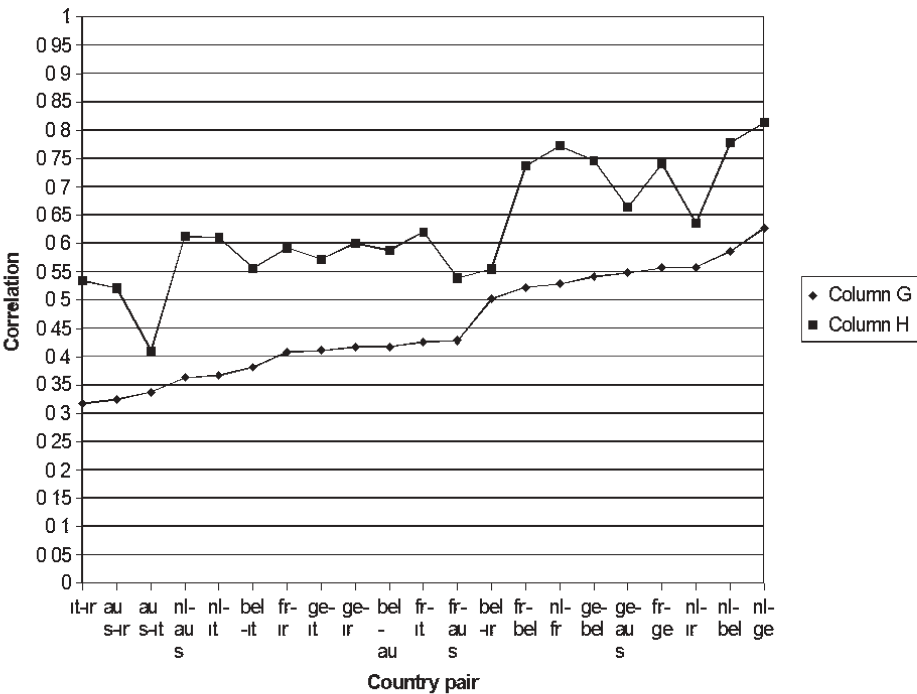
FIGURE 1

These figures show the evolution of country pair correlations in different sub-periods. Period 1 is the pre-convergence period (Jan 1979 - Dec 1985), Period 2 is the convergence period (Jan 1986 - Dec 1993). Period 3 is the pre-Euro period (Jan 1994 - Dec 1999), while Period 4 corresponds to the Euro era (Jan 1999 - April 2004). The methodology is as follows: First, the country or sector index returns are used to compute the unconditional correlation matrices for the different periods. Second, the correlation pairs are sorted in ascending order and plotted against the country pairs for each sub-period. (e.g. pre-convergence period). The correlation pairs for the subsequent period (e.g. pre-convergence period) are re-computed and plotted along the correlation pairs of the previous period (pre-convergence) correlation pairs.

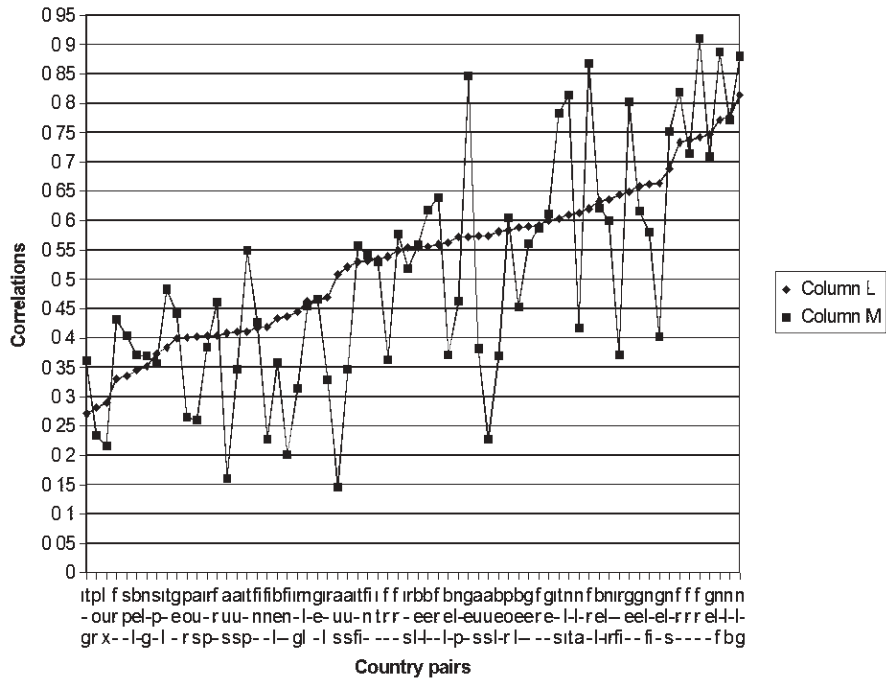
**A. Evolution of country pair correlations
period 1 and 2**



B. Evolution of country pair correlations
period 2 and 3



C. Overview of country pair correlations
period 3 and 4



THE DEMOCRATIC JUDGE

Bertolt Brecht about Los Angeles

In Los Angeles, before the judge who examines people
 Trying to become citizens of the United States
 Came an Italian restaurant keeper. After grave preparations
 Hindered, though, by his ignorance of the new language
 In the test he replied to the question:
 What is the 8th Amendment? falteringly:
 1492. Since the law demands that applicants know the language
 He was refused. Returning
 After three months spent on further studies
 Yet hindered still by ignorance of the new language
 He was confronted this time with the question: Who was
 The victorious general in the Civil War? His answer was:
 1492. (Given amiably, in a loud voice). Sent away again
 And returning a third time, he answered
 A third question: For how long a term are our Presidents elected?
 Once more with: 1492. Now
 The judge, who liked the man, realized that he could not
 Learn the language, asked him
 How he earned his living and was told: by hard work. And so
 At his fourth appearance the judge gave him the question:
 When was America discovered? And on the strength of his correctly answering
 1492, he was granted his citizenship.

Bertolt Brecht

From *Writing Los Angeles*, a Literary Anthology, The Library of America, 2002.

Italian Immigration Policies:¹ The Metaphor of Water

Rima Al-Azar

How should Italy address its rising tide of immigration? This paper compares the phenomenon to the flow of water; it can be blocked or it can be channeled to bring about positive outcomes. Italian regulations since the 1980s have restricted and only mildly directed the flow of immigrants. Economically, Italy as a whole has benefitted greatly from the entrepreneurial spirit brought by new arrivals. Culturally, however, the nation has yet to find the means to fully integrate its foreign-born population. To address the issue, the cost of immigrating illegally must be raised, while the cost of doing so legally must be lowered. The European Union, national and local governments should strengthen social services, increase access to the banking system, sponsor skills training in countries that send immigrants and link trade and migration together.

INTRODUCTION

The tide of dark, dangerous, poor people washing up on the Italian shores. The flux of foreigners. The waves of clandestine *extra-comunitari*. This is the common way of describing immigration in the media – a destructive, uncontrollable, body of water that invades a country, changes its society, corrodes its culture, and drowns the jobs available to local people.

However, like any body of water, when channeled properly, it can equally be a source of life and growth – as opposed to destruction. Immigration can increase creative diversity, rejuvenate an aging population, fill in labor gaps in the industrial, agricultural, health, and service sectors, and be a bridge of development across the Mediterranean.

Should a wall be built to protect Italy from this assault coming from the sea, or should the flow of immigrants be properly managed to channel their capacity where it is most needed and in the most beneficial way for the community, the

Rima Al-Azar, a candidate for the M.I.P.P. degree, has worked extensively on international development issues with various multilateral organizations and NGOs. She specializes in assisting vulnerable groups at the community level. Her experience spans the Middle East, Africa, the Balkans, Central Asia, India, and the United States.

immigrant, and the sending country? In order to answer the question whether immigrants are a burden or a blessing and to formulate a realistic national immigration policy, we must examine the profile of immigrants, the national laws that regulate immigration, and local practices. In addition, in order to ensure the formulation of a practical immigration policy and to reassert the right to monitor and control who enters Italy, rather than naïve openness or rigid restrictionism, or, for that matter, knee-jerk responses to the latest crisis, the impact of immigration on political, social, economic, and demographic factors needs to be analyzed.

IMMIGRANTS IN ITALY – A PROFILE

A current snapshot of the profile of immigrants in Italy renders a multinational, multireligious, younger-than-average group of people present predominantly in the center-north of the country.

RECENT INFLOWS

According to the Forum Internazionale ed Europeo di Ricerche sull'Immigrazione (FIERI), the 1981 census exposed an unexpectedly high number of foreign residents (320,778). The first big influx, however, happened later. Between 1984 and 1989, approximately 700,000 to 800,000 people entered Italy. Of these, it is estimated that 300,000 to 350,000 entered or remained in Italy without a valid residence permit. The most recent Caritas report on immigration indicates that in the last four years, the number of legal foreign residents has doubled to reach 2.6 million. Currently, 1 in 22 persons living in Italy is foreign-born.

Even taking this into account, the sudden increases in the number of registered foreigners in Italy brought about by the four amnesties in the last 12 years, there are two significant features of immigration in Italy: rapid flows with substantial volumes and a high proportion of undocumented immigrants.

DIVERSE NATIONALITIES AND RELIGIONS

There are no overwhelmingly predominant nationalities among the immigrant population in Italy. By December 2003, the largest national group was that of Romanians, who, however, made up only 10.9 percent of the total foreign population in Italy. The next most numerous nationalities were Albanians (10.6 percent), Moroccans (10.4 percent), Ukrainians (5.1 percent), Chinese (4.6 percent), Filipinos (3.4 percent), Polish (3.0 percent), Tunisians (2.8 percent), Americans (2.2 percent), and Senegalese (2.2 percent) (Source: Caritas, 2004.).

Neither can we speak of a prevalent religion. In December 2002 – figures from Caritas and the Ministry of the Interior (re-elaborated by Caritas) – estimate that 24.1 percent of foreigners were Catholic, 21.6 percent were other Christian sects, and 36.6 percent were Muslims.²

If we consider 2003, the latest year for which data is available, most foreigners living in Italy were European (almost 44 percent) and have come to Italy for work-related reasons (40.3 percent) or family reunification (32.2 percent). Table 3 shows the type of visas issued in 2003 per continent of origin.

PREDOMINANT NORTHERN URBAN PRESENCE

Foreigners reside predominantly in the larger urban areas. According to the 2001 census, 32.9 percent of foreigners lived in the communes with more than 100,000 inhabitants. A comparison between the 2001 and 1991 censuses shows a wider spread of immigrants into the smaller communities. Table 4 shows the concentrations of non-Italians in the different communes according to population size.

Finally, immigrants tend to gravitate to the north of the country. Caritas has described roughly the geographic distribution of foreigners in Italy as 6-3-1. In the north, where 60 percent of foreigners reside, they are to be found predominantly in Lombardy. In the center, where a third of immigrants live, Lazio is the province that attracts the greatest number of them. Finally, Campania is the region in the south that hosts the most immigrants.

ITALIAN IMMIGRATION LAWS – A REVIEW³

Like many of its southern European neighbors, Italy has struggled to find the right tone and approach toward immigration. In the 1980s, almost overnight, Italy found itself transformed from a country of emigrants — providing a larger number of immigrants to the Americas in the late 19th and early 20th centuries than any other European country — to a net receiver of immigrants. This inflow increased during the 1990s. In fact, it was only in 1997 that the balance of remittances became negative.⁴

Prior to the 1980s, Italy did not have any laws to address the legal existence of foreigners. The unique norm was that found in the public security law of 1931, which required foreigners to declare their presence to the authorities. When the country established its first amnesty program for illegal aliens in the late 1980s, immigration policy became a matter of national concern. Since then, several laws have been passed to regulate immigration in Italy.

The first of these laws (no. 943), passed in 1986, regulated immigrants' access to the labor market. Following the union-led protest in 1989 which forced immigration on the government's agenda, Law no. 39 (known as the Martelli Law) was passed in 1991 and recognized both the rights and obligations of immigrants. Italy's first comprehensive immigration legislation was set in motion. The new law aimed to attract multilateral attention to Italy's growing immigration concerns and to increase "burden-sharing" to help Italy manage its increasingly porous borders. While implementation and enforcement activities were criticized, the law was the government's response to ease public discomfort and deep European skepticism about Italy's ability to manage its long seacoast. This was especially important since other European countries viewed Italy as the unsecured door through which immigrants were entering other European countries.

The main achievements of the Martelli Law were a broad-based amnesty that included employers and the establishment of an annual quota system with

input from unions. It is also to be noted that it was during that time that the first (and last) national conference on immigration was held to discuss the phenomenon of immigration in Italy. However, Law no. 39 failed to define a real procedure for legal entry. This slowly led to an increase in illegal immigration because both the Italian economy required a greater number of workers and Italian families a larger number of domestic helpers.

During the 1995-1996 period, the center-right government, under pressure from the anti-immigrant Lega Nord to increase police powers to deport illegal immigrants, passed the Dini Decree. Unions led a demonstration with 150,000 protesters against the threat of increased police powers, which the decree envisaged. The final decree was revised to include an amnesty and its most restrictive aspects dropped.

The bulk of the legislation that currently regulates immigration and integration matters in Italy is the result of two laws. The Single Act no. 286 of July 25, 1998, which was essentially based on Law no. 40 of March 6, 1998, called the Turco-Napolitano Law⁵ and Law no. 189 of July 30, 2002, called the Bossi-Fini Law.⁶

In 1998, Italy came under further pressure to restrict illegal immigration in order to become a full member of the Schengen Agreement before the April deadline. The Italian Law no. 40, of March 6, 1998, on immigration and foreigners, which entered into force on March 27, 1998, is a complex and detailed document. The law's objectives were to improve efficiency in managing the flow of immigrant labor; increase prevention and containment of illegal immigration; and expand measures for effective integration of legal foreigners. The new law also provided annual planning for the immigration flow on the basis of an appropriate quota established by the government. The Turco-Napolitano Law was extremely strict in the matter of prevention and containment of illegal immigration. It dealt not only with new arrivals at the border, but also with those already illegally in the country. All these new restrictive measures had the effect of bringing Italian policy in line with Schengen. The illegal immigration enforcement part of this act was actively pursued and led to an increase in deportations, primarily of other Europeans and Africans.

The Turco-Napolitano Law was the first time that a law was proposed which delineated a procedure to become a legal resident in Italy. Secondly, it also allowed foreigners to come to Italy even before they found employment. Though, this could have reduced dramatically the numbers of illegal immigrants, this aspect of the law was rarely applied due to a fear that other European countries would accuse Italy of being too lax. Another innovation was the law's recognition of immigrants' social rights, including access to health services and family unification. The law separated for the first time humanitarian and refugee issues from immigration policy matters. It also included an open category for third world immigrants and a reduced waiting period for permanent residency to five years.

Italy's immigration picture changed further with the victory in 2001 of Silvio Berlusconi. Berlusconi's cabinet, which included members from the far-right Lega Nord (which has made its opposition to immigration central to its electoral agenda)

and the formerly neofascist National Alliance, has been seeking ways to curtail immigration into Italy and to deploy a range of enforcement and control mechanisms. In August 2002, the government passed legislation to regulate immigration and, in September of that same year, adopted a decree to provide for the regularization of undocumented immigrants already in the country.

The new Law no. 189, also known as the Bossi-Fini Law, amended the 1998 Immigration Act and introduced new clauses. Some of the most significant changes included: immigrant quotas, mandatory employer-immigrant contracts, stricter illegal immigration deportation practices, amnesty for illegal immigrants who have worked and lived in the country for over three months, and new provincial immigration offices to help manage immigrant worker and family reunification cases. The law also provided for the legalization of two types of irregular immigrants: those employed either as domestic workers and home-helpers or as dependent workers. These individuals could qualify for regularization, provided that they had not received a deportation order.

The Bossi-Fini Law repealed the sponsorship that had been introduced by the Single Act of 1998. The Turco-Napolitano Law had envisaged a job-seeker visa, providing for the allocation of an annual quota of residence permits to foreigners seeking employment in Italy. These potential workers could enter the country sponsored by private individuals, regions, municipalities and associations listed in a register. Sponsors were required to deposit an economic guarantee, offer appropriate accommodation, and pay the contributions for public health insurance.

In contrast, the new law tightened the link between the work contract and residence permit by bringing them together under one single *contratto di soggiorno-lavoro* (residence-employment contract). The residence permit for work was made dependent on a combined residence and employment contract. The residence permit was valid only for the same duration as the employment contract and could be for no more than nine months for seasonal workers; no more than one year for temporary workers; and no more than two years for non-temporary workers. Finally, it modified the 1998 law by requiring immigrants to have job contracts before entering Italy.

Both trade unions and employers' organizations have criticized aspects of the new legislation, arguing that they could ultimately harm the national economy. Trade unions objected at the new mandatory employment contracts, fearing that they would be simply another barrier to entry and would divert potentially legal flows toward illegal and irregular channels. Employers' organizations were especially opposed to the provision that denied immigrant workers regularization if they had received a deportation order. They noted that many firms that have employed these workers would be left without replacements, especially in regions of high employment. Also, equally affected would be the many small and medium firms in northern Italy and farms in the south that rely on foreign labor. In addition, the large and growing number of families who depend on Philippine or Sri Lankan home help would feel the impact of this new law.⁷

The Bossi-Fini Law also introduced a linkage between the quota allocated

to certain countries and their cooperation in stemming the flow of people at the source.

By pursuing this twin-track approach, the Italian government hoped to demonstrate to illegal immigrants that it was not worth taking the risk; and it wanted to make deals with governments of countries from which they set out. However, immigrants were not deterred by a law doubling the length of time that illegal entrants can be held after detention and imprisoning those caught re-entering. The reality is that illegal immigrants usually enter without papers and refuse to give their nationalities, to avoid deportation. Most “expulsions” from Italy, except to Albania, are notional; the expelled simply go underground, or to another European country, after their release from detention.⁸

Other sectors within Italy, however, viewed the new legislation positively. Employers must now sign formal contracts that guarantee immigrant workers housing and return travel expenses, while also fixing wages and length of employment. Furthermore, the stricter visa issuance policy provided for a more selective immigration process, especially for immigrant workers. Those in favor asserted that Law no. 189 provided a major innovation with regard to immigrants’ living standards and ultimately benefited Italy’s business sector by filling their ever-changing needs with a pool of better-qualified immigrants. However, most immigration experts have viewed the Bossi-Fini Law as a restrictive law that denied immigrants some fundamental social rights.

All the above-mentioned progressively more stringent laws and amnesties have been stopgap measures, introduced in an emergency atmosphere. Until now, Italy has failed to design and implement a comprehensive immigration policy that is based on political, economic, social, and demographic realities that take into account the long-term needs of Italy and the benefits that a regulated immigration policy would bring to Italian society. In addition, providing immigrants’ political and social rights, and assisting them to better integrate into Italian culture would ensure a more harmonious and secure society. Fortunately, the Italian government can be influenced by both its own regions as well as by the European Union in terms of policy formulation.

THE LOCAL INITIATIVES

The previous section reviewed the national Italian legislation governing immigration. However, the devolution of certain administrative and legislative powers to the regions has allowed the regions to pass laws on the international recruitment of foreign labor and cooperation with sending countries (Chaloff, in OECD 2004). In several of the Italian regions, these laws address the need for foreign workers in the context of international development projects. Most Italian regions emphasize in their regional laws the importance of training for the return migration of workers who have been temporarily in Italy.

The different regions’ strategies have five common objectives for linking immigration and economic development.⁹ These can be summarized as follows: (1) socioeconomic cooperation for the re-entry of migrants to their home countries;

(2) information and multicultural training to raise awareness regarding the phenomenon of migration through cultural exchanges with institutions in the sending countries; (3) economic and financial integration through immigrants' resources and capacities; (4) programs to manage the flow of immigrants; and (5) international cooperation for local development linking economic development with migrants' flows.

The Veneto region is at the forefront of this kind of initiative, for two reasons (Barin in OECD 2004). First, Veneto is the home of numerous small and medium-sized enterprises, and is experiencing the most acute labor shortage in Italy. Second, local authorities are interested in regulating labor migration as part of their policy priorities. On an even more local level, there are provincial activities in Veneto for facilitating immigration (e.g., a Chilean-Argentinean information desk in the province of Padua and a Brazilian recruitment center in the province of Vicenza).

Another example is that of the Friuli-Venezia-Giulia region, which has been very active in Romania and has opened its third international desk in Bucharest to recruit nurses. Some Italian regions have also worked with the Romanian government to provide language and professional training for Romanian workers wanting to migrate temporarily to Italy (Barbin in OECD 2004). In some cases, regionally facilitated recruitment exists, in the absence of international agreements. For example, to ensure support of its regional tourist industry, the Trentino-Alto Adige region has taken the unusual step of placing its own staff in Italian consulates in central Europe to monitor and streamline administrative procedures for issuing seasonal work permits.

Local initiatives, in addition to regional development agreements with sending countries, include innovative programs providing services to immigrants. For example, the commune of Rome supports immigrants in establishing small- and medium-scale enterprises by providing them with soft loans at favorable interest rates. In Treviglio, the Cassa Rurale has established a new service targeting immigrants and offering them competitive rates for sending remittances to their home countries. Finally, numerous NGOs, such as Caritas, have developed different projects for assisting immigrants at the community level.

Several of these subnational- level migration and cooperation policies are quite promising. This is true despite a number of serious structural limits, including a lack of coordination among different levels of government and the lack of resources necessary to effectively link migration and development. With a greater degree of institutional harmonization and cooperation, as well as with a review of lessons learned from all these local initiatives, there is an opportunity not only to formulate a better-designed Italian immigration policy but also to share Italy's experience with the European Union and contribute to a pan-European immigration procedure.

THE SUPRANATIONAL RESPONSE

Italy, being a member of the European Union, has also to abide by the EU's guidelines and regulations. The Council of Europe's Strasbourg Convention, which

was held in February 1992 and ratified by Italy in 1994, provided guidelines and incentives for member countries to “guarantee to foreign residents, on the same terms as to its own nationals, the classical rights of freedom of expression, assembly and association.”¹⁰ In addition, the Strasbourg Convention called on all signatories to “make efforts to involve foreign residents in processes of consultation on local matters.”

Increasingly, the European Commission is addressing the issue of immigration in order to face “a set of apparent contradictions in migration-related issues. The need to control borders in a post-September 11 security context is often at odds with the need to welcome visitors, promote trade, gain from orderly and selected migration, and strengthen the Union’s absolute commitment to free movement in a newly enlarged Union of twenty-five Member States”.¹¹ In 2004, the EC’s DG Employment and Social Affairs published a green paper on economic migration, which is expected to be followed by a public hearing in 2005. However, it should be noted that even if a fully harmonized set of EU rules is eventually achieved, it would unlikely ensure that all EU countries share a proportionately equal burden of immigrants and asylum-seekers. That is because the laxness or otherwise of national regimes is only one factor determining where people migrate and where asylum seekers make their claims: family and cultural ties, work opportunities and language are equally important.

Compared to traditional countries of immigration – such as the USA, Canada, and Australia – Europe remains a closed and hostile continent. In an article, the *Economist* recounted the poem inscribed on America’s Statue of Liberty (“Give me your tired, your poor/Your huddled masses, yearning to breathe free”) and wondered how the European Union would welcome immigrants. “A similar monument if it were to be built in Brussels, what message would it carry? Perhaps something like: ‘We have vacancies for a limited number of computer programmers and will reluctantly accept torture victims with convincing scars. Migrants looking for a better life can clear off’.”¹²

The EU’s leaders have made it clear that controlling illegal immigration is a priority. This urgency stems partly from the recent rise of populist anti-immigration parties across the Union, including France’s National Front, the Freedom Party in Austria, the Pim Fortuyn List in the Netherlands, the Northern League in Italy, and the Danish’s People’s Party.¹³ European politicians are worried that if they are not seen to act against illegal immigration, the far right will gain even more ground.

In particular, the EU is concerned with the ease in which illegal immigrants can be smuggled into Europe across the Mediterranean. Many European governments who, in the past, have doubted Italy’s political will and capacity to manage its borders welcomed the restrictions brought about by the Fini-Bossi Law. However, some critics continue to question certain elements of the new law. They believe that creating bilateral agreements that focus on smart border management and economic development in sending countries is the only answer to Italy’s illegal immigration problem.

In addition to the combined driving forces of responding to right-wing local politicians and addressing security concerns, the EU is increasingly interested in

looking at immigration as a way to maintain its competitive edge vis-à-vis the United States which is capable of attracting a greater number of highly-skilled immigrants. Also, US immigration policy is one of the reasons why American society, on average, is younger and more dynamic than Europe's. Currently, the United States' median age is 35.5, while Europe's is 37.7. But according to Bill Frey, a demographer at the University of Michigan, by 2050 the median age for Europe will have risen to 52.7, compared to a US median age of 36.2. This difference is accounted for almost entirely by the dramatic aging of the European population, and these shifting demographic patterns have enormous fiscal consequences. By 2050, it is estimated that the number of people over 65 will be equivalent to 60 percent of the working-age population in Europe, thus increasing the "dependency costs".¹⁴

The same immigration issues that Italy is grappling with – namely, social and political integration, access to social services, impact on demography and labor market – are also of concern to the EU as it attempts to formulate a common strategy for all its member countries – a strategy that is at once pan-European, profoundly national and even local in nature.

IMMIGRATION AND POLICY-MAKING

As mentioned earlier, immigration policy is closely related to four main aspects of policy: political, economic, social, and demographic. Placing these issues on a continuum, political issues fall in the realm of short-term decisions, whereas those that are economic in nature are more medium-term, followed by social and, finally, demographic policies that address long-term concerns. While immigration is a hotly disputed political issue, this paper will address the economic and social aspects in more depth.

ECONOMIC ASPECTS OF IMMIGRATION

As mentioned previously, politics and economics push the government in opposite directions when it comes to formulating a coherent immigration policy. Politics argue for being seen to clamp down on immigrants. But economics argue for encouraging immigration. Not only does the receiving country benefit economically from this self-selected, entrepreneurial group of people, but also the sending nation reaps benefits from the remittances sent home.

In the receiving country, immigration expands the economy by increasing the workforce. Businesses also benefit from immigration because immigrants can fill specific labor shortages. Though the fear is that immigrants "steal" the jobs of native workers, it is, however, the other immigrants who arrived earlier that are the most affected by the more recent inflow of foreigners. Most of the research on the impact of immigration on wages has been carried out by the US National Research Council which reported that overall immigration reduced the wages of groups competing with immigrants – predominantly low-paid people – by 1 to 2 percent.¹⁵

According to economists, the potential economic benefits to the world of liberalizing migration dwarf those from removing trade barriers. The benefits are

even greater in countries, such as Italy, where populations are aging and economies are sluggish. Immigrants, unlike natives, move readily to areas where labor is in short supply, thus easing bottlenecks. They also bring a just-in-time supply of skills. According to Caritas, in 2003 foreigners in Italy were employed in the following sectors: 7.4 percent in agriculture; 21.7 percent in industry; 27.2 in services; and 43.7 percent in domestic work. The 12 most important fields of work by order of importance were: domestic work, construction, hotels and restaurants, agriculture, real estate activity, metallurgic industry, transport, commerce (retail and bulk), food sector, textiles, and public services.¹⁶

Another less noted and publicized factor of immigration is that of the risks immigrants take. Not only are they willing to do the work that Italians are not interested in, but they also pay a heavier price in terms of work-related accidents – that could be deadly at times. Whereas, work-related misfortunes suffered by Italians are decreasing, those happening to foreigners are on the rise. According to the Istituto Italiano di Medicina Sociale, the number of work-related mishaps happening to foreigners increased from 73,778 in 2001, to 92,014 in 2002, to 106,930 in 2003 (of which 129 were fatal).¹⁷

Finally, another misguided fear commonly found in receiving countries is that immigrants are a burden on the public social services. In fact, it has been found that in many cases, immigrants pay more in taxes than they cost in public spending.¹⁸ In Italy, the “*badante*” are themselves an informal social service system since they provide assistance to families by taking care either of the very young or the elderly when the state itself cannot provide these social services.

From the sending country perspective, the importance of remittances sent home by workers residing abroad has, until recently, been underestimated. According to the World Bank, worldwide remittances to low-income countries are almost double total Overseas Development Assistance provided by the developed countries.¹⁹ In addition, remittances’ impact on development is even greater than development assistance since they are not associated with an obligation to be repaid with interest; they are sent directly to the people for whom they are intended; and they cannot be intercepted and squandered by corrupt governments. There are no specific figures as to how much immigrants in Italy send back to their countries but it is presumed to be substantial. The lack of statistics is partly due to the use of informal channels for sending money back home, which cost immigrants a great deal. Seizing an opportunity some socially-minded banks (e.g., Banca Etica, Cassa Rurale di Treviglio) have started providing cheaper services for immigrants to send their remittances. Ironically, Italy itself is one of the top 15 countries with total remittances received. In 2001, Italians working overseas sent home €2.266 million in remittances.²⁰

In addition to remittances, immigrants can have a positive impact on their home countries by setting up small-scale businesses upon their return. This could lead to a transfer of skills they acquired while living and working abroad. This link between immigration and development has only recently been explored. And, as mentioned earlier, several Italian regions are at the forefront of setting up projects and initiatives to ensure the transfer of skills and to combat the brain drain caused

by immigration.

In conclusion, the issue of migration is more a political challenge than an economic one. Migration undoubtedly presents a marvelous opportunity for advancing human welfare, but this clash of economics and politics and weighing its costs and benefits is very difficult. From a liberal economic perspective, the benefits are clear, however, when viewed with immediate political gains, the costs increase.²¹

Politicians and Italian businesses along with NGOs favoring immigration should unite to lobby for the easing of immigration rules and to raise public awareness regarding the economic benefits of immigration for the current generation and even more for the following one.

Kenneth Prewitt, the former head of the US Census Bureau, argues that “in the struggle to find workers to support growing economies, nations that are more hospitable to immigrants will have an advantage.”²² “Immigrants will go where they have friends and family to welcome them and help them find employment. Where will they find a more hospitable welcome – Europe or America?” If left to the potential immigrants, a determining factor in their choice of country to immigrate to is the ease with which they can socially integrate, based on religious, linguistic and family ties, along with cultural affinity.

ACCESS TO SOCIAL SERVICE AND CULTURAL INTEGRATION

Italian immigration law differentiates between legal and illegal immigrants and their right to access social services. The law also includes provisions for addressing cultural diversity.

Undocumented immigrants are given basic rights – essential public healthcare and state education. The 2002 center-right reform changed undocumented immigrants’ rights marginally. Undocumented immigrants are still entitled to all “essential services and treatment even if long-term,” not just to treatment in the event of emergency, pregnancy and for children, as is the rule in other countries. Undocumented immigrants are given a special anonymous public health card. They are asked to pay the normal contribution or to declare they are unable to pay it. For undocumented minors, public education is not only free, but also compulsory.

The 1998 law limited temporary accommodation to legal immigrants only, but the Bossi-Fini Law specified that *all* “social integration measures” are limited to legal immigrants and therefore denied private bodies (for example, Catholic associations such as Caritas) the possibility of assisting and giving shelter to undocumented immigrants. Not all organizations have actually observed this regulation however. The Bossi-Fini Law also no longer allowed local councils to provide any kind of shelter in the event of exceptional illegal influxes, as was allowed under the Turco-Napolitano Law.²³

FIERI – a think tank specializing in immigration issues – argues that pluralism would be better achieved by passing the Religious Liberty Bill.²⁴ This bill would not only provide more freedom and benefits to religious minorities in

Italy, but it would also allow the government to sign conventions with Islamic associations, which are currently under excessive foreign control. The signing of such agreements would however send a strong signal and acknowledge public recognition of Muslim minorities in Italy.

Italians' attitudes towards Islam have changed after September 11, 2001. Prior to the terrorist attacks of September 11, Italians' expressed attitudes towards Islam were relatively tolerant. According to a comparative survey conducted by the European Union in 2000,²⁵ only 10 percent of Italians were against the entry of workers from Islamic countries (against a European average of 18 percent), 30 percent declared themselves ready to welcome them without restrictions (against a European average of 17 percent). This is comparable to Italians' attitude towards mainly Christian foreigners from Eastern Europe: 31 percent were ready to welcome them, and only 9 percent did not want them. The more hostile feelings towards Muslim foreigners – Moroccans, Albanians, Tunisians – seemed determined more by the high crime rates in their immigrant communities than by their religion. Since the Sept. 11, 2001, terrorist attacks, a third of Italians have declared feeling more anxious towards people of Arab origin (ISPO, October 2001).²⁶ However, according to a more recent poll, Italians are more willing to accept other religions: 70 percent would be against a law similar to the French one that forbids wearing the veil in public schools.

Examining Italian attitudes towards specific nationalities, an ISPO survey conducted in 2000 revealed that the two most positively regarded groups are the Filipinos and Senegalese. However, the Roma are the least accepted, with only 24.3 percent of Italian responding that they "liked" them; slightly better off were the Albanians who were accepted by 37.8 percent of Italians.

Lastly, the issue of second-generation immigrants is one that is critical – especially for the future social stability of Italian society. The fact that the law does not grant citizenship to those born in Italy – not even after having resided in the country for an extended number of years – means that they will suffer from an identity problem. A tell-tale indicator of the problem of integration is the number of children born to immigrants attending school. According to Nazzarena Zorzella, 30 percent of students in Bologna's primary schools are foreigners.²⁷ But their enrollment decreases at the high school level and is almost nonexistent at the university level. This is a symptom of both a lack of educational and linguistic support as well as a reflection of the precariousness of immigrants. Why should immigrants invest in the education of their children when their residence status can be at any moment be lost when they are no longer employed?

Italy has yet to find the means to fully integrate its foreign population and to benefit from the cultural variety and richness that they bring to society. Though it is difficult to measure the value that diversity brings to a society, there is no doubt that immigrants are a self-selecting group of people who are unusually entrepreneurial. Will Italy opt for the multicultural *laissez-faire* approach or put in place proactive integration policies, will it choose the "mosaic" method or the "melting pot" process to deal with immigration? Or will it come up with a uniquely creative Italian way with dealing with the issue? One conclusion is certain, for

long-term social stability and harmony, immigrants need to be better integrated. Integration is a two-way responsibility: Newcomers should make an effort to “when in Rome do (as much as you can) as the Romans do,” but the public authorities can also design programs to decrease prejudice against the “vu’ comprà” and the “marrochini.”

THE LONG-TERM PERSPECTIVE: DEMOGRAPHY AND POLICY-MAKING²⁸

As Auguste Comte, a 19th century French philosopher, once said, “Demography is destiny.” Translated in a different way, this means that when formulating an immigration policy that is – even if only partially – based on demography, the long-term perspective needs to be considered. Though demographic trends can change over the decades, demographic effects last longer, and have a wider impact than most social and/or economic forces.²⁹

Italy has the lowest birth rate in Europe. It currently stands at around 1.8 births per woman when it is necessary to have a rate of 2.1 to keep population size constant. It is expected that by 2050 the number of Italians may have fallen from 57.5 million in 2000 to around 45 million. In addition to a shrinking population, life expectancy is on the rise. It is expected that by 2050, based on present demographic trends, the ratio of pensioners to workers will be one-to-one in Italy. Since pensions in Italy are prepaid out of current tax revenue, the implication is that taxes will have to be significantly increased to fund the generous pensions. However, average earners in Italy are already paying close to 33 percent of their wages into the state pension scheme.

The impact of immigration on demography is a double one: not only does the number of people increase but immigrants usually have a higher fertility rate than the native population thus also contributing to an increase in the number of inhabitants. In addition, immigrants are a relatively young group. According to ISTAT, immigrants in Italy have a median age slightly over 30 years (30.4 for men and 31.4 for women) in contrast to a median age for the Italian population of 41.7 (40.1 for men and 43.1 for women).

However, immigration alone will not be sufficient to resolve the demographic challenges that Italy faces. The OECD calculates that immigration might have to be between five and 10 times its current level just to neutralize the economic effects of Europe’s ageing populations. Persuading Italians to have more children is the obvious alternative answer. Part of the problem may be what Italians call the “partial emancipation” of women, who are free to pursue a career but are still expected to fulfill the responsibilities of the traditional “casalinga,” or housewife, by bringing up children, looking after the grandparents, and doing the housework. Making family life easier and/or less expensive might help keep up the population. For example, France, which has some of the most extensive state-funded child care in Europe, also has one of its highest birth rates.

The impact of immigration on demography is clear. Foreigners who are more youthful than the native population and have a higher fertility rates not

only produce a larger population, but a society that is younger, more mixed ethnically, and on balance more dynamic.

CONCLUSION

Any immigration policy – or lack thereof – will be a reflection of how the Italian government and Italian society at large view immigration and immigrants. Will Italians think of foreigners as an uncontrollable, destructive tsunami wave? A water tap that can be opened and closed depending on internal societal pressures? An unpredictable rain that depends on external factors such as wars, or economic depressions? Or a water channel that can be designed and directed to where it is most needed to ensure a flourishing society and continued economic and social growth?

Regardless of how immigration is viewed, it is clear that a large number of Italian public and private institutions—ranging from businesses and civil society to the government—have become deeply implicated in migration. Economic and labor demands, social and security interests, colonial history, and political priorities create the conditions that allow and even encourage migration. Reducing this complex relationship to its simplest form, each and every consumer in Italy is implicated in migration by the mere fact that he or she benefits enormously from the work of immigrants.

In an informal policy note prepared by the World Bank and the Hamburg Institute of International Economics advising the EU on immigration policies, the authors provide the following recommendations for a win-win European immigration policy between sending and receiving countries.³⁰ Italian policy-makers could find that these recommendations are equally applicable to the Italian context.

1. Bilateral exit and entry controls and joint recruitment strategies: The objective of such measures is to increase the entry costs of irregular immigrants while simultaneously decreasing them for legally admitted ones. This entails a review of the visa types, visa allocations, and labor permits by linking them to skill levels, knowledge of language, and other required features.
2. Social protection and access to social services for migrant workers: Though reduced access to welfare provisions is sometimes conceived as a policy tool to lessen the attractiveness of receiving countries, immigrants should have the right to access healthcare, education, unemployment insurance, and pension systems. Sending and receiving countries should develop mechanisms guaranteeing transferability of social welfare benefits. Such mechanisms could even encourage return migration.
3. Secure and enhance remittance flows and access to the banking system: Governments should attempt to reduce the fees associated with remittance flows. The European experience of reducing transfer fees within the euro zone could be a good practice that could be replicated. In addition, immigrants should have easier access to banking services, in particular access to credit. This would help them to establish new businesses and expand existing ones. It would also improve

the economic integration and performance of immigrants, because, on average, they are more entrepreneurial than the native-born population.

4. Strengthening transnational communities: Enabling migrants to keep ties with their countries of origin would increase economic, social, and cultural ties between Europe and a larger number of countries, thus enhancing economic and other opportunities.

5. Skill formation and co-development: The EU could sponsor training specifically needed skills in schools and universities in sending countries. Support of educational facilities in neighboring regions would counterbalance potentially negative effects of brain drain for sending countries.

6. Trade-migration link: Migration is sometimes described as globalization from below. Negotiations on future bilateral and multilateral trade and cooperation agreements between the EU and developing countries could include both import-export and migration-related issues.

In addition to this supra-national perspective, the challenges that one particular Italian region – Emilia-Romagna – is facing could also shed light on the general Italian context and provide further guidance to the design of any future immigration policy.

In conclusion, a new and different approach to immigration policy formulation would recognize the complex nature of the issue and the multitude of stakeholders. It would search for a win-win approach that would benefit the receiving country, without harming the sending one, while guaranteeing the political and social rights of the immigrant. In order to find such a solution, the management of immigration might be turned into a national project that would involve civil society, employers, government and other public institutions alike.

As indicated previously, there are already several elements that could become the first building blocks of such a new immigration policy. Different regional initiatives linking economic development and migration, discussed earlier, as well as local good practices of improving remittance flows (the example of the Cassa Rurale in Treviglio) and access to credit (the Comune of Roma) could be replicated across Italy. The fact that Italians, compared to other Europeans, are more open to immigrants is also a factor that would help to create an enabling environment for a more comprehensive policy. Finally, the government could tap into Italians' experience of being immigrants themselves to increase voters' acceptance of a more politically and socially oriented immigration policy. As the American judge in the Bertolt Brecht poem took into consideration the socio-economic context of the non-English speaking Italian immigrant and granted him his political rights, similarly, an Italian immigration approach could be more flexible and understanding. By remembering the Italians who were sent in the 1940s and 1950s to work in the Belgian mines³² and were abused and mistreated, any current immigration policy should protect the social rights of foreign workers and protect them from being exploited.

The first step in this direction could be to hold a national conference with representatives from all the major Italian parties as well as civil society, religious

institutions, the private sector, and immigrant associations to discuss a future Italian immigration policy that would guarantee the flourishing of Italian society in the decades to come, ensuring economic growth, social stability, and moderate population growth. As Caritas points out, this is the only way to guarantee a “società aperta, società dinamica e sicura”.

NOTES

¹ This paper does not analyze refugee and asylum types of migration.

² “Main Features of Italian Immigration Stocks and Flows,” (September 2004) Forum Internazionale ed Europeo de Recherche sull’Immigrazione (FIERI), Torino, 2004 <<http://www.fieri.it/pagInterna.cfm?liv=98&pag=attivita>>

³ Nazzarena Zorzella, “Dall’Europa delle migrazioni all’Europa dell’integrazione” come cambiano le strutture di accesso all’informazione ed ai servizi” (lecture, Infopoint Europa, Bologna, Italy, 10 December 2004).

<<http://www.migrationpolicyinstitute.org/>>

<<http://www.fieri.it/>>

⁴ Jonathan Chaloff, “From Labor Emigration to Labor Recruitment: The Case of Italy” *Migration for Employment: Bilateral Agreements at a Crossroad* (2004) OECD, Paris, 2004

⁵ After the names of the then center-left Government Ministers of Social Affairs and of the Interior.

⁶ After the names of the Minister for Devolution and the Deputy Prime Minister of the current center-right government, the leaders of the Lega Nord and Alleanza Nazionale respectively.

⁷ “Stemming the Flow,” *The Economist* 13 June 2002.

⁸ “Men in boats,” *The Economist* 23 October 2003.

⁹ *Regioni e province autonome tra cooperazione e immigrazione* (2002) Centro Studi di Politica Internazionale (CeSPI), Rome 2002

¹⁰ <<http://conventions.coe.int/treaty/EN/cadreprincipal.htm>> (May 20, 2003)

¹¹ *Migration Policy Institute: Conference on Future European Cooperation in the Field of Asylum, Migration and Frontiers* (September 2004)

Ibid: “Conference Summary Report” in Netherlands’ European Union Presidency Conference on the Future EU Cooperation in the Field of Asylum, Migration, and Frontiers (Amsterdam, 31 August - 3 September 2004)

¹² “Huddled masses, please stay away,” *The Economist* 13 June 2002.

¹³ Ibid.

¹⁴ “Half a billion Americans?,” *The Economist* 22 August 2002.

¹⁵ “Who Gains from Immigration?,” *The Economist* 27 June 2002.

¹⁶ *XIV Rapporto sull’immigrazione, Caritas/Migrantes, Immigrazione – Dossier Statistico*

(2004) Caritas, Rome 2004

¹⁷ Ibid.

¹⁸ "Opening the door," *The Economist* 31 October 2002.

¹⁹ "Monetary Lifeline," *The Economist* 29 July 2004.

²⁰ *Remittance Data*, Migration Information Source, 1 June 2003, Migration Policy Institute <http://www.migrationinformation.org/feature/print.cfm?ID=137>

²¹ "Migration and Development," *Economist.com* 6 May 2004 <http://economist.com/displaystory.cfm?story_id=E1_NGDRDTJ>

²² "Half a billion Americans?"

²³ FIERI, 2004.

²⁴ Ibid.

²⁵ Ibid.

²⁶ The IPSO survey was conducted in October 2001, only a month after the terrorist attacks. One wonders if a survey were to be carried out now what would the results be? Also, Italians' negative feelings are more toward Arabs and not Muslims – Senegalese enjoy a rather positive impression.

²⁷ Zorzella

²⁸ "Europe's Population Implosion," *The Economist* 17 July 2003.

²⁹ "A tale of two bellies," *The Economist* 22 August 2002.

³⁰ "Challenges and Opportunities of International Migration for the EU, Its Member States, Neighboring Countries and Regions A Policy Note," *Social Protection Discussion Paper Series*, Robert Holzmann and Rainer Münz, 0411 (2004). Washington, DC, The World Bank, 2004.

³¹ "Comparing US and European Approaches to Issues of Immigration" in *Contemporary Issue Series*, 16 (symposium, University of Bologna, Dickinson College, Clarke Center, Bologna, 20 – 21 June 2003).

³² In 1946, the Italian and Belgian governments signed an agreement which stipulated that for each day an Italian worker toiled in the mines of the Walloon region, 200kgs of coal would be sent to Italian state enterprises. This gave an incentive to the Italian government to encourage migration and institutionalize the mechanisms of recruitment and transport of these workers. A turning point was the 1965 tragedy in Marcinelle, Belgium when 262 miners died in an explosion; 136 of them were Italian, including many adolescents. (Chaloff, 2004).

The American Way of Going to War

John L. Harper

America's "unnecessary wars" adhere to a basic pattern. They have been fought in the name of the broader mission that many Americans believe Providence has chosen their nation to carry out, but have been characterized by a prewar "fog" of incomplete or flawed information. They are the handiwork of a small but determined "war party," and the US political system often acts as a stimulus to the use of force, rather than a check on it, as opposition politicians join the call for fear of being branded unpatriotic. Finally, from the War of 1812 to the invasion of Iraq in 2003, unnecessary wars have more often than not failed to advance the interests of those who pursued them. These lessons of the past are cause for serious reflection; the least that can be said after reviewing these wars is that the benefit of the doubt should never be given to those who urge military action.

Historians and other scholars have spilled much ink over the subject of 'the American way of war.' Developed during the Civil War, perfected during the Second World War, partially revived in the 1991 Gulf War after the frustrating limited conflicts of the Cold War, the 'American way' exploits the wealth and technological superiority of the United States and relies on overwhelming force to annihilate the enemy.¹ Less attention of a systematic kind has been paid to the American way of *going to war*. This is no doubt partly because Americans prefer to think that the dynamics of going to war in their history have been relatively simple. For the most part, America has unsheathed its 'terrible swift sword' when others have forced it to do so, when its honor and security have been at stake, and as a last resort.

A distinguished historian has recently contended that, on the contrary, pre-emptive war and annexation of territory are an old American strategy, one that developed in reaction to the British attack on Washington during the War of 1812.² This argument exaggerates both the similarities between 24 August 1814 (a reprisal

John L. Harper is Professor of American Foreign Policy and European Studies at the Bologna Center. He is the author, most recently, of American Machiavelli: Alexander Hamilton and the Origins of US Foreign Policy.

for an American attack on Toronto) and 11 September 2001, and the impact on the alleged father of pre-emption, Secretary of State John Quincy Adams, of the British raid. Gaddis makes much of Adams's bid to wrest the Floridas from Spain in 1818. But some Americans had been trying to take the Floridas long before the burning of the White House, just as some had been hoping to intervene in Iraq long before the felling of the Twin Towers. Expansion and interventionism on the periphery were driven at times by defensive considerations, but mainly by the belief in a continental destiny, the desire to open transportation routes, sectional ambition, and hunger for more land.³ Until the Bush administration, the United States had never embraced preventive or pre-emptive war, and on several occasions presidents had rejected such a course. In the late 1790s, President John Adams frustrated Alexander Hamilton's design to seize New Orleans. In 1950, President Truman repudiated his navy secretary's call for preventive war against the USSR. The Eisenhower administration rejected a US Air Force plan advancing a related scheme.⁴

Nevertheless, the manner in which, and the reasons for which, the United States went to war against Iraq in March 2003 were *not* a radical departure. History shows that the United States has had a strong propensity to become involved in conflicts which, though it would be misleading to call them 'wars of choice,' were unnecessary wars.⁵ This essay will try to explain why. Some of what will be said undoubtedly applies to other nations; indeed, those in a position to judge may well conclude that there is little that is exceptional in the American way of going to war. But this is not a comparative study – my focus is necessarily on the United States.

'War' is understood as a major armed conflict in which the United States has committed ground forces against a sovereign enemy for the purpose of defeating it.⁶ An 'unnecessary war' can be defined against its opposite, a 'war of necessity.' A war of necessity is a war in response to an unprovoked attack on one's territory or citizens or those of a friendly country, or a war to preserve the nation's political independence and way of life. This definition encompasses pre-emptive war in the face of what the political theorist Michael Walzer calls a 'sufficient threat,' that is, the existence of 'a manifest intent to injure, a degree of active preparation that makes that intent a positive danger, and a general situation in which waiting, or doing anything other than fighting, greatly magnifies the risk.'⁷ It also encompasses collective action under the auspices of the Security Council to deal with a serious breach of international peace.

Preventive war against another state based on the presumption of a future threat, a staple of the classic balance-of-power system, cannot be deemed unnecessary as a category. One need only imagine such a war against Nazi Germany, circa 1936. But there have been far more instances of unnecessary or counterproductive preventive wars than of hypothetical preventive wars which hindsight shows should and could have been fought.⁸ The gravity of the supposed threat is more often than not exaggerated and there is little guarantee that such a war will lead to the desired outcome. Third parties (in the recent case, North Korea and Iran) do not necessarily receive the intended message, while the moral

authority of the power starting the war is likely to suffer as a result.

‘Necessary’ does not exactly correspond to ‘just.’⁹ All necessary wars are just, but not all just wars are strictly necessary. Just but unnecessary wars, arguably, might include wars to assist a people in overthrowing a foreign or even a local tyrant, to stop genocide, or to spread democracy. The problem is that such wars tend to dissipate precious resources in dubious enterprises, to be used by outside powers and local actors for the purpose of self-aggrandizement, and to prove unhelpful to those they are supposed to assist. As Walzer observes, ‘foreign intervention, if it is a brief affair, cannot shift the domestic balance of power in any decisive way toward the forces of freedom, while if it is prolonged ... will itself pose the greatest possible threat to the success of those forces.’¹⁰ While one should not make a fetish of the so-called legalist paradigm, as a general rule there is superior wisdom in a conservative position: going to war to protect oneself and one’s allies while resisting the temptation to do so in order to foster the kind of world that one might like to create.¹¹

That said, the following were wars of necessity: the War of Independence, triggered by pre-emptive British action against the nascent American confederation in April 1775, the Civil War, triggered by pre-emptive Southern action in April 1861, entry into the Second World War, which even without Pearl Harbor (a not altogether unprovoked attack), would have been necessary to defeat Nazi Germany, the initial phase of the Korean War, triggered by North Korea, and the Afghanistan War, following the 11 September 2001 attacks. By contrast, the War of 1812, the Mexican War, the Spanish–American War, the second phase of the Korean War, the Vietnam War, and the 2003 Iraq War were instances where the United States could have avoided war without seriously endangering itself or its allies, or used methods other than military force. There are a pair of borderline cases, the First World War and the First Gulf War. On balance, the former (as will be shown) was an unnecessary and the latter a necessary war.¹²

AMERICA’S MISSION

No two wars begin and end in the same way or for the same reasons. Each is unique. Nonetheless America’s unnecessary wars adhere to a basic pattern. First – although the point applies to nearly all of America’s wars – they have been fought in the name of the broader mission that Providence has allegedly chosen the United States to carry out, or of universal principles that the United States, as so many Americans believe, has the responsibility to defend. The notion of mission has always served to consecrate policies whose practical effect, though not necessarily their conscious purpose, is the preservation or expansion of American wealth and power.

The characteristic American outlook on America’s purpose in the world, though neither Hegelian¹³ nor Marxian, is strikingly deterministic: it holds that events unfold according to an underlying design or logic. This has much to do with the influence of Protestant eschatology on the American view of history; specifically, with the twin notions that events are moving toward an end-state, in

time, when Good will triumph over Evil, and that America is the agent through which this design is being carried out.¹⁴ Thomas Jefferson embraced a secularized version of this millennialist thesis. On the effects of the American and French Revolutions, he wrote, 'This ball of liberty, I believe most piously, is now so well in motion that it will roll around the globe.' According to the New York journalist John L. O'Sullivan,

All this will be our future history, to establish on earth the moral dignity and salvation of man — the immutable truth and beneficence of God. For this blessed mission to the nations has America been chosen and her high example shall smite unto death the might of tyranny of kings, hierarchs, and oligarchs.

For Woodrow Wilson, 'America had the infinite privilege of fulfilling her destiny and saving the world.' For George W. Bush, history 'has a visible direction, set by liberty and the Author of Liberty ... America, in this young century, proclaims liberty throughout the world, and to all the inhabitants thereof.'¹⁵ A powerful complementary thesis, derived from a reading of nineteenth-century British history and bolstered by 'hegemonic stability theory,' is that without a leader, the international political and economic systems will descend into anarchy. America, by virtue of its power and ideals, is the world's redeemer nation by default. It is 'bound to lead'.¹⁶

The conviction that, in going to war, the United States is pursuing a historical mission or defending universal principles has been particularly important in mobilizing elites — editorialists, clergymen, liberal (and today neo-conservative) intellectuals, politicians — who wish to justify action to themselves and others on the basis of idealistic motives, or who see America's ideals as a form of power. The notion of mission has often been reduced to a formula or slogan that radically simplifies a complex question and endows it with 'moral clarity'. These slogans have galvanized opinion while also tending to obscure the self-seeking motives bound up with American action. The question is not whether such moralizing is sincere or instrumental; it is clearly both.

In 1845, John O'Sullivan launched the slogan 'manifest destiny,' expressing the idea that Providence had ordained America's dominion from sea to shining sea. Journalists and politicians from all sections and parties argued that like-minded neighboring peoples such as the Texas Republic should be incorporated into the federation, that America should regenerate the backward peoples of the hinterland, and that in disputes with foreign powers (over Oregon and California, for example), true title to the land belonged to those (namely, Americans) able to occupy it and develop its potential as God had intended. When the United States occupied large parts of Mexico, the James Polk administration claimed that it was extending 'the area of freedom,' a slogan revived in recent years.¹⁷

A form of 'manifest destiny' re-emerged in the 1880s and 1890s, reinforced by the Social Darwinism and 'navalism' of Alfred Thayer Mahan, Theodore Roosevelt and Henry Cabot Lodge. Mahan believed that there was no choice but to compete with the other great powers for markets and colonies. Together with

Britain, America was destined to lead the coming decisive struggle between Eastern and Western civilizations. To that end, it should control the Caribbean, build a canal across the Isthmus of Panama, and project its power into the Pacific by annexing Hawaii. The Reverend Josiah Strong, an influential Congregationalist minister, declared that 'the Anglo-Saxon is divinely commissioned to be, in a peculiar sense, his brother's keeper' and to spread his beneficent influence around the globe.¹⁸

In 1917, America went to war 'to make the world safe for democracy,' Wilson pledged that

we shall fight for the things that we have always carried nearest to our hearts — for democracy, for the right of those who submit to authority to have a voice in their own governments, for the rights and liberties of small nations, for a universal dominion of right.¹⁹

The Fourteen Points (January 1918) reflected Wilson's belief that the time had arrived for America to put itself at the service of humanity and fulfil its mission of bringing peace and freedom to the world.

Proclamations of America's mission during the Cold War drew on Wilson's vision of a world where the freedom of small states would be protected, on the lesson of the 1930s that aggression must be promptly confronted, as well as the presumed lesson of the pre-1914 world that peace and prosperity depended on the constant efforts of a powerful liberal state. For Truman, the United States must assist free peoples to work out their own destinies in their own way.' Aggression in Korea must be dealt with because it was 'part of the attempt of the Russian Communist dictatorship to take over the world, step by step.' The authors of the secret study NSC 68 (April 1950) were franker in stating that

America's mission must be to 'foster a world environment in which the American system can survive and flourish', and that even if the USSR did not exist the United States would face the fact that 'in a shrinking world, the absence of order among nations is becoming less and less tolerable.' By defining the Cold War as a contest of 'freedom' against 'slavery,' however, they created a caricature – for themselves in the first instance – of that world. According to the logic of NSC 68, the war of Vietnamese nationalist- communists against a US-sponsored dictatorship in Saigon became a test of America's capacity to learn the lessons of Munich and defend liberty around the globe.²⁰

In the preface to the Fourteen Points, Wilson had declared the dawn of a new age: 'the day of conquest and aggrandizement is gone by.' The spread of democracy in East Asia and Latin America in the 1980s, China's economic liberalization, and the sudden collapse of the Soviet Empire appeared to some to be a belated vindication of Wilson's view. According to the notion of 'the end of history', the Western model had demonstrated its inherent superiority over all competing models. The year 1989 marked the fading of grand ideological conflict

and the advent of Jefferson's secular millennium: the adoption of democracy and free markets around the world.²¹ According to a widely-held complementary argument, the United States should seize the opportunity to consolidate its overwhelming military superiority and its political hegemony over Europe, the Far East and the Middle East.²² The two arguments lie at the heart of the George W. Bush administration's view of America's mission. For Bush, there is but a 'single sustainable model of national success: freedom, democracy and free enterprise', and the United States must extend its benefits everywhere. What is more, the United States 'has and intends to keep military strengths beyond challenge, thereby making the destabilizing arms races of other eras pointless.'²³ The remarkable confidence with which Bush has viewed the task of spreading democracy in the Middle East springs in part from his quasi-millennialist assumptions. With the world heading toward 'the end of history', the United States need do little more than kick away the rotten scaffolding of the old order in a country like Iraq in order to release the forces of democracy and free markets to play their appointed part. In a real sense, Iraq was a faith-based war.

THE FOG BEFORE THE WAR

Much has been made of Clausewitz's notion of the 'fog of war.' A second consistent feature of America's unnecessary wars has been the 'fog before the war.' Policymakers make honest mistakes on the basis of flawed or incomplete information. Sometimes they deliberately deceive the public. Moreover, as an acute observer has written, 'it may be true that the most effective way to deceive others is first to deceive oneself.'²⁴ At the heart of the American decision to go to war there has often been a self-deception or a delusion of one kind or another: a fixed, mistaken conception of reality not readily susceptible to correction through examination or reasoning. A typical pre-war delusion is the belief, in spite of evidence to the contrary, that the United States is the innocent party and that the other side has little or no reason to feel threatened or aggrieved. Often this is connected with a dramatic incident or incidents whose dynamics are distorted by a kind of 'political solipsism.' Political solipsism is the inability to see, and the natural tendency to deny, the selfish and power-seeking component in one's own behavior, and the ways in which one's actions may condition the actions of others. As another acute observer put it, 'Nations, particularly great nations, are usually too proud to understand that their power might be a peril to other nations.'²⁵ To this could be added that the greater the power, the greater the sense of peril and the inclination to resist, on the part of those who feel subject to that power.

The delusion according to which the United States is in the right while others are wilfully aggressive or malevolent has been essential to creating a climate of righteous indignation and rallying the general public behind America's unnecessary wars. The nationalist, populist and 'Jacksonian' elements of the body politic have typically been less concerned than the liberal elites about justifying American action in the name of universal principles, but are sensitive to questions of national sovereignty and honor and quick to take offense.²⁶ American history bears out

Alexander Hamilton's observation, 'Wars more often proceed from angry and perverse passions than from cool calculations of Interest.'²⁷

An incident that inflamed American opinion was the Battle of Tippecanoe (November 1811) between US forces and Indians in the Indiana Territory. American opinion had longed blamed British agents for arming and inciting the tribes of the Northwest. To put an end to this *ante litteram* example of state-sponsored terrorism, and to punish British maritime behavior, Americans called for regime change north of the Great Lakes. In fact, British officials were attempting to discourage Indian attacks on Americans for the simple reason that they put Canada under threat from the United States. The Indian attack had been provoked by Governor William Henry Harrison's encroachment on territory President Madison had told him to remain out of. On the impact of this incident, one historian justly comments: 'History is often influenced as much by erroneous conviction as by truth.'²⁸

The American conquest of California and New Mexico in the 1840s confirms the point. President Polk and his cabinet acted partly on the assumption that the British were trying to take California. Though London did not favor US acquisition of California, there was not much evidence that it had designs of its own and the opening of the British archives in the twentieth century showed that this theory was a canard.²⁹ In 1845, the US consul in Texas convinced the Polk administration that the British, having failed to prevent the annexation of Texas by the United States, intended to goad Mexico into a war with its northern neighbor. This theory was also at variance with the facts.³⁰ The incident preceding the US invasion of Mexico in 1846 occurred when US forces occupied the area of southern Texas between the Nueces and Rio Grande Rivers to which the United States had a shaky legal claim. When the Mexicans sent troops north of the Rio Grande and a clash occurred, Polk assured the nation — and seems himself to have believed — that Mexico had 'invaded our territory and shed American blood upon American soil.'³¹

The incident sparking the Spanish–American War was the explosion of the US battleship *Maine* in Havana harbor on 15 February 1898, killing 266 crewmen. Well before a US inquiry reported that a Spanish mine or torpedo had caused the disaster, many had reached Theodore Roosevelt's conclusion: the *Maine* had been sunk 'by an act of dirty treachery on the part of the Spaniards'. With the release of the report, war hysteria swept through the United States: 'neighborhoods, suburbs, small towns, and rural counties simply caught fire. No section, no type of community, no occupational group was immune.'³² Though the Spanish had afforded the *Maine* every courtesy during its three-week stay in Havana, it was not implausible that local elements of the Spanish military considered its presence a provocation and had planted a mine. But it was equally plausible that Cuban revolutionaries had acted to trigger US intervention. The least plausible hypothesis was probably that the Spanish government wished to provoke a war. Under US pressure, and facing a winless conflict with Cuban rebels, the Liberal Party cabinet in Madrid had recalled the notorious General Valeriano ('the butcher') Weyler and announced that it would grant the island autonomy along Canadian lines.

During the crisis following the incident, Madrid accepted US demands that it end 'reconcentration' of the rural population and grant an armistice to the rebels.

Understandably, Madrid rejected the US demand that it give up sovereignty over Cuba, and war was the result. American experts today believe that the *Maine* was destroyed by spontaneous combustion inside the ship.

America's righteous indignation in 1917 arose from incidents occurring in the German-declared war zone in the vicinity of the British Isles, in particular, the sinking of the British liner *Lusitania*, killing 128 Americans, in 1915, and the attack on the French liner *Sussex* in 1916. The US declaration of war followed the German resumption of unrestricted submarine warfare (suspended after the *Sussex* incident) in the zone, and interception of the 'Zimmermann telegram' revealing Berlin's plans to entice Mexico into the coming war with the United States. The German military had concluded that the United States as a belligerent could not provide much more help to the Allies than it was already providing as a 'neutral,' and that before US power could be mobilized, German submarines could knock Britain out of the war. This proved to be a fatal miscalculation.³³ At the same time, in the words of a well-known historian,

the American people, in pointing an accusing finger at the submarine, were prone to overlook their own responsibility for what befell them. Their large-scale assistance to the Allies in munitions and other contraband, coupled with Washington's acquiescence in the British blockade, drove the Germans to desperate expedients that ultimately involved the United States.³⁴

From early on, the British gave an arbitrary definition to contraband (items that could be confiscated from neutrals), compelled American vessels bound for Europe to undergo extended searches in British ports, and created a heavily mined exclusion zone in the North Sea to enforce a blockade of Germany. Washington protested such unprecedented interference with neutral shipping, but thanks to the efforts of pro-Ally officials (Wilson's adviser Colonel Edward House, US Ambassador to London Walter Hines Page, Secretary of State Robert Lansing) the protests were never pushed too far. In the meantime, the United States became a major supplier of the Allies, the value of its trade with them nearly quadrupling in the first two years of the war.³⁵ The Germans' unhappy choice was to allow munitions, often carried by passenger ships, to reach Britain, or to attack such shipping without warning to enforce their own blockade.³⁶ The *Lusitania* carried some 4,200 cases of ammunition and the Germans had publicly warned Americans against sailing on Allied ships. After the incident, some in Congress thought it would be prudent for the US government to issue a similar warning. Wilson, though he declined to press US rights against British violations, insisted on the right of US citizens to travel as they wished. It is hard to avoid the conclusion that a self-serving and far-from-even-handed policy, together with the failure to grasp how American behavior affected German behavior, led the United States into the war.³⁷

The second phase of the Korean War, following General Douglas MacArthur's successful counter-offensive in September 1950, is a classic case of how a delusion, subscribed to chiefly by MacArthur but shared to a degree by most top policymakers, led to a disastrous turn of events. Reflecting the Truman

administration's previously adopted aim of destroying North Korea's armed forces, the Joint Chiefs of Staff prepared orders authorising UN forces to cross the 38th parallel, though only if there had been 'no entry into North Korea by major Soviet or Chinese Communist forces, no announcement of intended entry, nor a threat to counter our operations militarily in North Korea.' Approved with minor changes by President Truman, Defense Secretary Marshall and Secretary of State Acheson, these became MacArthur's orders on 27 September 1950. Marshall advised him in a separate communication that he should 'feel unhampered tactically and strategically to proceed north of the 38th parallel.'³⁸

The delusion consisted in the belief that the Chinese would not see themselves sufficiently threatened by an all-out offensive in the direction of their industrial heartland, that they lacked air and naval support, and that they would not be willing to make the necessary economic sacrifices to commit themselves to a major war with the United States. The delusion, by definition, persisted in the face of abundant contrary evidence: major, observable troop movements into Manchuria and across the Yalu River; warnings from Indian, Dutch, British and other sources; Zhou Enlai's 2 October 1950 statement through the Indian ambassador to Beijing, the day after South Korean troops crossed the parallel, that if US troops followed, China would intervene; and subsequent, explicit warnings in the Chinese press. Those who claimed to understand the 'oriental mentality' (MacArthur and his entourage) argued that the Chinese were bluffing.³⁹ Anyone who remembered that the Japanese had used northern Korea as a base from which to attack Manchuria in the 1930s, and who realized that the Chinese were unlikely to believe US pledges to stop at the Chinese border (the US had initially stated that it would stop at the 38th parallel), took a different view.

A series of dramatic incidents involving US forces paved the way for the escalation of the Vietnam War: the clash in the Gulf of Tonkin in August 1964, the attack on Ben Hoa air base in October 1964, the attack at Pleiku in February 1965. The Johnson administration's reaction to the Gulf of Tonkin incident indicates that it did not so much misconstrue reality as deliberately deceive Congress to advance its objectives. Even as he assured the nation that he 'sought no wider war', Johnson solicited a blank check from Congress in the form of the Gulf of Tonkin Resolution to be able to pursue just such a war after the 1964 elections, if he chose. The administration knew that the United States had been conducting covert (Oplan 34-A) attacks against North Vietnam for a number of months and that US naval patrols monitored the North Vietnamese reaction. To win support for the resolution, however, the administration presented North Vietnamese attacks on US destroyers in the Gulf of Tonkin as 'unprovoked' and part of a pattern of 'naked aggression.'⁴⁰

The delusion (as opposed to deception) consisted in the belief of Johnson, close advisers and the US military command in Saigon that a viable South Vietnamese state and army could be built in the shadow of an invasive US presence, and that support at home would hold in the face of significant casualties. This flew in the face of historical evidence. Until April 1965, Secretary of Defense Robert McNamara and Ambassador to Saigon Maxwell Taylor strongly resisted a US combat role, the latter predicting that, eventually, 'like the French, we would be

occupying an essentially hostile country.' Undersecretary of State George Ball argued in 1964–65 that a negotiated settlement would be welcomed by most US allies, that sending US combat forces would be like giving cobalt treatment to a terminal cancer patient, and that history's great commanders had resisted the temptation to throw good money after bad. Johnson's shrewd political adviser, Clark Clifford, predicted that Americanization would lead to 'catastrophe'. The president's trusted Senate friend, Richard Russell, conveyed a similar message. Yet even after the 1968 Tet offensive had exploded the assumptions underlying US strategy, Johnson was prepared to support the military's request for an additional 206,000 troops and expansion of the war beyond South Vietnam. Johnson's was the delusion of the gambler who, rather than cut his losses, goes deeper and deeper into debt in the belief that a miraculous payoff will save his fortune and reputation in the end.⁴¹

While few would deny that the US-led war in Afghanistan in 2001 was necessary, the Bush administration subsequently opted to exploit the anger and sense of vulnerability spawned by 11 September to pursue its prior objective of toppling Saddam Hussein. To that end, the administration deliberately misled the public by linking al-Qaeda and Iraq. To an even greater degree, however, the administration misled itself.

The American delusion concerning Iraq was multi-layered. It did not consist of the belief that Iraq had weapons of mass destruction. On the basis of various pieces of evidence and of Iraq's past behavior, it was not unreasonable to think this was the case. The delusion was, first, to believe that Iraq, having suffered a crushing defeat in 1991, with armed forces less than half the size they had been before the Gulf War, and with much of its airspace under US control, represented a growing threat. It was to believe, moreover, that because Iraq had used WMD against unarmed civilians it would use them without provocation against Israel or the United States. While there was a remote theoretical possibility that Islamic terrorists might obtain WMD from the secular Iraqi regime, an invasion seemed as likely to trigger such a transaction as to prevent it, and it was far more reasonable to assume they would seek such weapons from Russia, North Korea or Pakistan. At the heart of the Bush administration's delusion was the belief that the Middle East Gordian knot was somehow ripe for cutting by US military power: Iraq's reconstruction and transformation to democracy would be rapid and largely self-financing; Iran, fearing Iraq's fate, would abandon its nuclear ambitions; the Palestinians, seeing the light, would accept a deal like the one they had rejected at Camp David; terrorists everywhere would lose heart. To this could be added a seemingly wilful denial of one of the basic lessons of Vietnam: not only is a large foreign military presence no substitute for legitimate and self-reliant local forces; it is one of the obstacles to the emergence of such forces in Iraq.

ALL THE PRESIDENT'S MEN

The third characteristic of America's unnecessary wars is that they are not the handiwork of the people nor of a single dominant leader but typically of a small, determined 'war party'. Its members subscribe to the notion of America's

mission and see the United States and its friends as in the right but are also driven by the desire to gain (or retain) personal power and recognition. Often it is a relatively closed, partisan circle where there is a reluctance to challenge the assumptions underlying policy and where personal loyalty tends to trump intellectual honesty. The party has always included the president, though not necessarily as its motor force. It pushes, but at a certain point is itself borne along by events. Elements of the press often act as auxiliaries to the war party. They do so out of genuine conviction, the lure of profits, and the wish to curry favor with those at the top. In the case of America's unnecessary wars, only rarely has the war lobby included members of the industrial and financial elite.

In 1812, the war party was a group of young, ambitious congressmen (most famously Henry Clay and John C. Calhoun) known as the 'War Hawks'. Mainly from the South and West, they believed in the British-Indian conspiracy theory, and that just as gentlemen must be prepared to duel in private life, upholding American honor on the high seas required war. Less enthusiastic congressmen and senators supported the War Hawks' call for the raising of an army in the hope that it would lead to British concessions. Once such preparations were underway, however, it became difficult to back down. An anti-war congressman remarked that the war party and its followers had 'advanced to the brink of a precipice, and not left themselves room to turn.'⁴² At a certain point American — and their personal — credibility was at stake.

President Madison himself was not a charter War Hawk. In 1810, Congress (through legislation called 'Macon's Bill #2') removed US restrictions in place on trade with France and Britain, promising that if one of those powers ended its violations of American rights the United States would then impose total- non-intercourse on the other. In an attempt to embroil the United States and Britain, Napoleon's government stated its intention to revoke edicts that harmed American commerce. Madison chose to accept these flimsy assurances, gambling that it was in the French interest to follow through. In early 1811, Madison supported a law prohibiting British exports to the United States until the Orders in Council were revoked. If Britain refused, the only way to uphold America's 'absolute rights' would be war.⁴³

Even after it became obvious that the French had not changed the substance of their policy, Madison clung to the fiction that they had, and insisted that Britain do the same. In June 1812, after receiving news from London suggesting that British policy would not change, and facing a presidential election later in the year, he invited Congress to declare war. The final irony is that in late June, the British government, unaware of the American declaration of war and pressed by British exporters, revoked the notorious Orders in Council. But when this news reached Madison in August, he dismissed it as a trick. Once more he would not admit his mistake.

The Mexican War is a case where the president and his inner cabinet were the driving force. Polk was no intellectual of the Madison type but the over-achieving protégé of fellow Tennessean Andrew Jackson. After pushing through the annexation of Texas in 1845, Polk was determined to fix the US-Mexican border

on the Rio Grande and to acquire California and New Mexico. With one hand he offered cash (plus the assumption of damage claims allegedly owed by Mexico to US citizens). With the other he pointed a pair of pistols: an American army poised on the bank of the Rio Grande opposite the Mexican port of Matamoros and a naval squadron off Vera Cruz. After learning that Mexico had refused to receive his minister, John Slidell, he decided to ask Congress to declare war, even before receiving news of the clash with Mexican troops north of the Rio Grande in April 1846.

From the outset of the First World War, the Republican imperialists of 1898 (Mahan, Roosevelt, Lodge, Root) argued that the United States must intervene on the side of Britain and France to punish German aggression against Belgium and to help preserve the balance of power. Key administration officials (House, Page, Lansing, Secretary of War Newton Baker, Assistant Secretary of the Navy Franklin D. Roosevelt) adopted a similar point of view. So did industrialists (the Du Ponts) who supplied the Allies, and bankers (the House of Morgan) who had advanced around \$2.3 billion in cash and credits to them by 1917. Though he sympathized with the Allies, Wilson did not favor intervention and was re-elected in 1916 as the man 'who kept us out of war'. Immediately after the election he undertook to mediate an end to the carnage and was deeply disillusioned with the Allies when they declined to go along. With some reason, he saw little to choose in rapaciousness between the two sides. In early 1917, he warned against 'a peace forced upon the loser' that would 'leave a sting, a resentment, a bitter memory', and called instead for a 'peace without victory'. By that point, however, much like Madison in 1812, he had trapped himself in a policy that tolerated the abuses of one belligerent and thus made war probable with the other. Specifically, after the Sussex incident in 1916, Wilson had warned the Germans that if they did not alter their method of submarine warfare, the United States would cut off diplomatic relations. The Germans had agreed, but on condition that there be a relaxation of the British blockade. When no such relaxation occurred and the Germans resumed their attacks, Wilson had little choice but to take sides in the war.

At the center of the war party in 1950 was the charismatic MacArthur, his syncophantic staff in Tokyo (above all, General Charles Willoughby) and supporters in the Pentagon like General Charles L. Bolté. MacArthur was also the champion of right-wing Republican congressmen and senators (William Knowland, Joseph McCarthy, Walter Judd and others) who had excoriated the Truman administration for 'losing China' and shared his goal of rolling back Communism in the Far East. In the State Department, MacArthur's course was backed by Asia specialists John Allison and Dean Rusk. The president's role was ambiguous, but crucial. Truman was eager to share credit for the liberation of North Korea and feared the political price he would pay at home if he were to rein in the 'viceroy of the East.' Secretary of Defense Marshall, a former soldier, declined to limit the freedom of a field commander. Secretary of State Acheson deferred to the judgement of the Olympian Marshall, his former boss. MacArthur himself had grown seriously worried about the presence of Chinese forces in North Korea by November 1950, but argued that halting his drive north would be seen by the world as a new Munich. Despite reservations, Washington gave the green light for his 'home-by-Christmas' offensive

on 24 November.⁴⁴

Shortly before his death, President Kennedy approved plans to withdraw 1,000 US advisers from Vietnam by the end of 1963, and to remove all 16,000 by the end of 1965. Almost immediately, his successor approved increased pressure on North Vietnam (Oplan 34-A) and declared to his advisers: 'I am not going to lose Vietnam.'⁴⁵ Johnson was an ambitious domestic reformer determined to shield himself from the kind of right-wing attacks that had weakened Truman. But he was also a literal-minded believer in America's mission to defend the Free World, a conviction reinforced by several advisers inherited from Kennedy, in particular Walt Rostow and Dean Rusk. A crucial component of the war party were the Joint Chiefs of Staff under General Earl Wheeler, and the field commander, William Westmoreland, who mirrored the thinking of conservative Republicans in pressing Johnson for a more aggressive war against the North. Even more crucial were those members of Johnson's inner circle who had recommended withdrawal to Kennedy in late 1963 – McNamara and Taylor. A combination of fear of the collapse of Vietnam, loyalty to the new president, and a desire to retain credibility with the Joint Chiefs led them to swallow their doubts about an American ground war in Vietnam.

The war party in 2002–03 included a group of fundamentalist believers in American power (Richard Cheney, Donald Rumsfeld, Paul Wolfowitz) whose drive and determination recalls the War Hawks of 1812. As in 1898, a well-organized lobby of exiles played an important part. George W. Bush's role also has an obvious precedent: James Polk in 1846. An historian's description of the Polk type applies well to Bush: 'No paralysing scrupulosity or forecast of future danger holds them back; and woe to the land if they be misguided, for they do things.'⁴⁶ Condoleezza Rice paraphrased Bush's view of his presidency: 'The country could sit on its unparalleled power and dispense it in small doses, or it could make big strategic power plays that would fundamentally alter the balance of power. Bush planted himself in the visionary camp.'⁴⁷ When the history of the war is written, Bush's domestic advisers may well emerge as members of the war party, encouraging Bush to believe that a quick, successful war would make him unbeatable in 2004. Colin Powell was the Robert McNamara of the story. After warning Bush in August 2002 of the probable dire consequences of a US invasion, he swallowed his doubts and went along.⁴⁸ Even if Powell had decided to oppose military action, by early 2003, with large US forces poised to attack Iraq, Washington's credibility was seen to be on the line and it was too late to turn back.

DEMOCRACY GOES TO WAR

Much has been written on the question of democracy's relationship to military action. What emerges from the study of America's unnecessary wars is that congressional opposition is weak and the two-party democratic competition more often than not acts as a stimulus to the use of force. Opposition politicians have tended either to mute their criticism out of fear of being branded as unpatriotic, or to jump on the war bandwagon. Their reluctance to challenge the 'war party' is sometimes based on what are seen as the lessons of the past. By the same token,

the party in power typically calculates that successful military action will pay political dividends, while showing weakness will lead to defeat at the polls. Few friendly countries can be counted on to have the nerve to express their doubts about American action. Those who do are usually dismissed as pusillanimous, or having crass commercial motives, or wishing to undermine the United States.

During the pre-war debate in 1812, the leader of the Federalist opposition in the House of Representatives, Josiah Quincy, recommended that his party keep a low profile and try to avoid the damaging charge of Anglophilia. The Federalists assumed that with the country unprepared and largely opposed to war, the War Hawks and their followers were bluffing, but they had no clear alternative of their own. During the Mexican War, articulate spokesmen from both parties (the Democrats Calhoun and Albert Gallatin; the Whigs John Quincy Adams and Abraham Lincoln) denounced Polk's policy as a reckless war of conquest, but the opposition found itself at a basic disadvantage. The Democratic majority in the House attached the formal declaration of war to an army supply bill, forcing the Whig minority either to vote yes or be accused of giving aid and comfort to the enemy. (The House vote was 174 in favor, 14 opposed, 35 abstaining; in the Senate, 40 in favor, 2 opposed.) The Whigs had not forgotten that their predecessors, the Federalists, had paid a high price for their refusal to support the War of 1812. Though loathing Polk and his policies, the Whigs voted for troops and supplies throughout the war.⁴⁹

Senior leaders from both parties (former Democratic President Grover Cleveland, the Republicans Carl Schurz and Speaker of the House Thomas B. Reed) considered war with Spain a national disgrace, but few politicians were willing to resist the fever of 1898. At the end of March 1898, Democratic Party leader William Jennings Bryan announced that he favored intervention in Cuba, a signal that the Democrats would exploit the crisis if the Republicans did not. The House vote on the ultimatum to Spain was 324 in favor, 19 opposed. The war gave rise to the Anti-Imperialist League which denounced plans to annex the Philippines and Cuba as a violation of American principles. Its members included such luminaries as Andrew Carnegie, William James and Mark Twain. McKinley ignored the League and went on a speaking tour to test opinion in the 'swing states' of the Middle West. As he revealed in an interview with Methodist ministers, before deciding to support annexation of the Philippines, he had gone down on his knees and 'prayed Almighty God for light and guidance more than one night'. The message he received was that the Filipinos could not be left to Spain, to some other power, or to themselves. 'There was nothing left for us to do but to take them all, and to educate the Filipinos, and uplift and Christianize them, and by God's grace do the very best we could by them, as our fellow-men for whom Christ also died.'⁵⁰

There was little or no opposition to MacArthur's drive north in 1950, a classic case in which the two-party competition acted as a stimulus to reckless military action. Skeptical Democrats were silenced by the fear that halting MacArthur would increase their vulnerability to attacks in an atmosphere already poisoned by charges of Communist infiltration of the Roosevelt and Truman administrations. MacArthur's offensive began a month before mid-term elections in which the

Democrats hoped to limit their losses. The Gallup poll (13 October 1950) indicated that 64 percent of Americans favored crossing the parallel, while 27 percent were opposed.⁵¹ In such circumstances, eloquent dissenting voices in the State Department (George Kennan, Charles Bohlen, Edward Barrett and Livingston Merchant, among others) had no effect. British Foreign Secretary Ernest Bevin, the one foreign statesman with the weight to exert some influence on the US administration, took the position that it would be dangerous to press the Americans to delay their offensive to give an opportunity to diplomacy. 'If Truman deferred to such an appeal and military problems arose later on, the British would take the blame.'⁵²

In Vietnam, the US political system once again stacked the deck in favor of escalation. When Johnson asked Congress in August 1964 for the authority 'to take all necessary measures to repel any armed attack against the forces of the United States and to prevent further aggression', the Senate concurred by a margin of 88–2, the House by 416–0.⁵³ Democrats uneasy about granting a blank check to Johnson had not forgotten the early 1950s. They voted yes in part to undermine Republican charges in the electoral campaign that their party was weak on defense. (There is no way to know if Kennedy would have reduced the US presence to a minimum after the 1964 elections,⁵⁴ but there is no question that it would have been easier for him politically to do so. He had proved his mettle as a Cold Warrior and would not have been running for re-election in 1968.) In 1964–65, as in 1950, the United States was ill-served by a British Labor government that hesitated to press its doubts about escalation. Washington pointedly rejected the wise advice (to seek the neutralization of South Vietnam) offered by the French.⁵⁵

In 2002–03, elder statesmen from both parties (former President Jimmy Carter, former national security adviser Brent Scowcroft) were appalled by the rush to war, but congressional opposition to the administration's policy was nearly as feeble as it had been in 1846, 1898, 1950 and 1964. Some Democrats shared the administration's analysis in good faith. Many (including those running for president) appear to have been cowed by the fear that they would be charged with weakness in the 'war on terror' if they tied the president's hands on Iraq. In 1990, in spite of the clear case of Iraqi aggression, 47 senators had opposed a resolution authorising the use of force; in 2002, in spite of the lack of an Iraqi provocation, only 24 did so. One reason for the difference is that the 1990 vote took place in December, after the mid-term elections, while the 2002 vote was held in October, in the middle of the campaign. Liberal newspapers like *The Washington Post* and *The New York Times* were hesitant and ineffective in questioning the case for war. This time, the leaders of the British Labor Party not only abandoned their initial doubts, but became strong supporters of the war. The well-founded criticisms of the French and Germans were dismissed as malevolent (Vice-President Cheney at one point asked the French ambassador whether France was an ally or an enemy) or beneath contempt.⁵⁶

WARS ARE EASY TO START...

The final feature of America's unnecessary wars, including those that have ended in victory, is that they exhibit a kind of law of unintended consequences. More often than not, they have failed to advance the interests of the individuals and political parties who have pursued them. America's unnecessary wars bear out Machiavelli's warning, 'Because anyone can start a war when he wants to, but not finish it, before taking on such an enterprise, a prince must measure his strengths, and govern his conduct on that basis.'⁵⁷

The War Hawks' assumption that an American army would conquer Canada in a few weeks in 1812 proved to be a fantasy. Advocates of war failed to foresee the reversal of Napoleon's fortunes in 1812-14, allowing the British to commit additional naval and military forces against the United States. The conflict emptied the national treasury, brought foreign trade to a standstill, and nearly broke up the union as much of New England refused to support it. The peace settlement did not bring territorial gains or address the original maritime controversies. Among the war's few significant results was that it launched the career of America's nationalist-populist icon, Andrew Jackson, who defeated the British at the (unnecessary) Battle of New Orleans.

After hostilities with Mexico began in April 1846, Polk was confident they could 'be speedily terminated.'⁵⁸ When, though beaten on the battlefield, the Mexicans refused to make peace and the American popular press began to clamor for the annexation of 'all Mexico,' the administration faced the prospect of a long and costly occupation of the entire country.⁵⁹ Polk seized on the exit strategy offered him by his peace envoy, Nicholas Trist, whose authority he had revoked. Luckily for Polk, the man he called 'an impudent and unqualified scoundrel' had disobeyed orders, negotiated a treaty with the Mexican government, and sent it to Washington in February 1848.⁶⁰ The war added enormously to American wealth and power at the expense of a weaker neighbor. It also reopened the question of slavery (which had lain dormant since the Missouri compromise of 1820), when northerners of both parties insisted (the Wilmot Proviso) that slavery be banned from the new territories. Polk attempted to paper over the question, but it was soon to explode. The Whigs recruited a war hero, General Zachary Taylor, as a presidential candidate and ousted Polk's party from the White House. Exhausted by the strain of the war, Polk went to an early grave.

The United States went to war in 1898 with an army and supply system sadly unprepared for hostilities. Theodore Roosevelt, an eyewitness, wrote that 'the mismanagement has been beyond belief ... The lack of transportation, food and artillery has brought us to the very verge of disaster.'⁶¹ Fortunately for the United States, the Spanish army had been brought to the verge of defeat by disease and the local rebels. Yet the American invaders were greatly disillusioned by the Cubans they had come to liberate, finding them slovenly and politically immature. Instead of gratitude, the Americans encountered widespread suspicion of their intentions (some felt they had been fooled by the Cuban exiles). Though the ultimatum voted Congress had promised that the United States would not acquire

'sovereignty, jurisdiction, or control' over Cuba, Washington excluded the Cubans from the peace negotiations and imposed a protectorate status on the island in the form of the Platt Amendment. With the collapse of Spanish authority over the Philippines, the United States had little choice but to annex the islands. In so doing it inherited not only a bloody war against the local independence movement but a territory of limited commercial value and indefensible against Japanese attack.⁶²

The case can of course be made that American entry into the First World War was necessary—not to protect neutral rights, to crush German militarism or to foster a new world order, but to preserve the European balance of power. According to this line of argument, the United States should have intervened on the side of the Allies in 1914, but without insisting on Germany's unconditional surrender. But this is a purely academic argument. American democracy would never have permitted such a step. As it was, US entry in 1917 guaranteed Germany's defeat and imposition of a 'peace forced upon the loser,' that would rest 'as upon quicksand', that Wilson had warned against.⁶³ In the course of rallying the American people to war, Wilson himself embraced the view that not only the elites but the German people must be punished. With US approval, the Paris settlement was riddled with contradictions of the principle of self-determination and the Versailles Treaty included Article 231 saddling the main loser with sole responsibility for causing the war and the consequent destruction of lives and property.⁶⁴ Given the evils that flowed from the collapse of Germany and Austria-Hungary and the peace of 1919, it is difficult to believe that the United States, and the world in general, would not have been better off had Washington stuck to strict neutrality and had the conflict ended in compromise or a narrow German victory.

Continuing the Korean War beyond the 38th parallel brought the worst battlefield defeat in US history and cost thousands of American and hundreds of thousands of Korean and Chinese lives. 'MacArthur's war' led to a serious crisis of confidence in American leadership around the world. It was in the context of a shooting war with China that the United States tied itself to a bankrupt regime in Taiwan and concluded that 'Indochina is of great strategic importance in the general international interest rather than the purely French interest, and is essential to the security of the Free World.'⁶⁵ Had the United States and China avoided war and relations been less bitter and confrontational, Washington would surely have seen its stakes in Vietnam in a different light. Truman's failed gamble cost the Democratic Party at home, and the Republicans took control of the White House in 1952.

Escalating the Vietnam War, rather than cutting US losses, in 1964–65 led to an unprecedented political–military defeat. Nearly 60,000 Americans died, as well as an estimated 3,000,000 Vietnamese. The war put a severe strain on the US economy and threatened to tear apart the social fabric of the United States. Among the domestic casualties was Johnson's Democratic Party, whose internal unity was shattered by the war.

Apart from the initial campaign, little has gone as intended in Iraq. The problems anticipated by the war's architects (a refugee crisis, destruction of the oil fields) did not materialize, while those that did (systematic looting, the collapse of basic services, a guerrilla insurgency) took them by surprise. The war distracted

attention and resources from the campaign against al-Qaeda and its allies and won new recruits and sympathizers for jihad against the West.⁶⁶ As in Vietnam, the local regime became dependent on US military power to maintain its position, yet the invasive US presence, and in particular the abuse of detainees and the killing of civilians, has undermined the regime's legitimacy and fuelled radicalism and resistance. The January 2005 elections (insisted on not by Washington but Shia religious leaders) do not guarantee a transition to a stable, united and democratic Iraq, let alone a democratic Middle East.⁶⁷ As for the war's other objectives, it may well be that US action prompted North Korea and Iran to redouble their efforts to obtain a nuclear deterrent against the United States. Bush won re-election in spite of, not because of, the war. Iraq remains a sinkhole for American blood and treasure during his second term.

A CAUSE FOR REFLECTION

A recent, cogent argument suggests that the durability of America's post-Second World War leadership has been based not only on its enormous power and the existence of a common enemy, but on 'strategic restraint', the capacity of the United States to take the interests of others into account and to fashion a kind of hegemony based on consent.⁶⁸ The purpose here has been to understand how and why the United States, over the course of its history, has frequently acted without restraint. The explanation begins with the habit of America's elites – though it is not unique to America's – to justify self-seeking and aggressive action in the name of a predestined mission or defense of universal principles. This habit, in the case of America's necessary wars has undoubtedly resulted in much good, and not only for the United States. But it is a two-edged sword. All too often the politically and morally reassuring notion that one is pursuing a mission has absolved consciences, numbed rational analysis and hastened the unnecessary use of force.

Hamilton, a critic of facile historical determinism, once observed that 'evil is seldom as great, in the reality, as in the prospect'.⁶⁹ America's unnecessary wars demonstrate that reality in pre-war periods is rarely what it is believed to be. Policymakers are prone to discount the self-seeking component in their own behavior and to delude themselves about the intentions and capabilities of the other side. It should be a cause for serious reflection when contemplating military action in the future that the premises on which the United States decided to go to war in 1812, 1846, 1898, 1917, 1950, 1964–65 and 2002–03, were mainly false.

Hamilton's contemporary Thomas Paine wrote that 'in the early ages of the world there were no kings, the consequence of which was there were no wars.'⁷⁰ Paine's point was inaccurate but suggestive. No war is the work of a single person, but small groups of determined and well-placed individuals have played an indispensable role in starting most of America's unnecessary wars. With several exceptions (1846, 1964 and 2003) these bellicose groups have not initially included the president, but no war has been started without presidential approval. More often than not, presidents embrace war because they see advantages to their careers and reputations in doing so, and dangers if they do not.

Speaking of American behavior in the First World War, George Kennan wrote: 'History does not forgive us our national mistakes because they are explicable in terms of our domestic politics ... A nation which excuses its own failures by the sacred untouchableness of its own habits can excuse itself into complete disaster.'⁷¹ The two-party democratic competition, with the strong incentive it entails to support military action and brand enemies as unpatriotic, has contributed on a number of occasions, if not to disaster, to inexcusable foolishness. Proponents of the 'democratic peace' theory might answer that this fact does not invalidate their argument. None of America's wars, they might contend, has been fought against a fellow liberal democracy, and the more democracies there are in the world the fewer wars there will be. Even if this is true (which is not self-evident⁷²), in the meantime (likely to be a long time), American democracy will retain its propensity to produce Vietnams and Iraqs. And what about the country most likely to enter America's cross hairs in the near future? Regardless of whether it is authoritarian or democratic, China will be a burgeoning power, determined to assert superiority in its own neighborhood, and subject to moods of nationalist paranoia and exuberance. In some ways it is much like the United States in 1898.

America's unnecessary wars, finally, have led to the useless expenditure of lives and resources and to unexpected and undesired consequences. To this it can, of course, be objected that not all of them have been unsuccessful. The Mexican War, the Spanish-American War and the First World War added considerably to American wealth and power. The United States did not suffer permanent damage and recovered its position relatively quickly after Korea and Vietnam. Leaving the aside the moral question of lives thrown away for no reason, the answer to the first point is that American demographic and economic growth would sooner or later have produced comparable levels of wealth and power (assuming that is what truly matters) even without war. The fact that the United States recovered relatively quickly from Korea and even Vietnam does not mean that those wars were not dreadful mistakes that no one would care to repeat.

Bismarck is supposed to have said that God in his mercy looks after drunks and the United States. Perhaps he had heard the story of how Nicholas Trist's unlikely treaty provided a fortuitous exit from Mexico for James Polk. Indeed, the American experience – thus far – has been relatively fortunate compared to that of most of the other, far older, great nations of the world. But it would be unwise to count too much on God or luck to ensure America's future safety from self-inflicted injury. The least that can be said is that the benefit of the doubt should never be given those who urge military action. In the twenty-first century, going to war unnecessarily could prove to be an even more expensive habit than in the past.

NOTES

¹A classic treatment is Russell F. Weigley, *The American Way of War: A History of United States Military Strategy and Policy* (New York: Macmillan, 1973). Weigley observes (p. xxii): 'In the history of American strategy, the direction taken by the American conception of war made most American strategists, through most of the time span of American history, strategists of annihilation. At the beginning, when American military resources were still slight, America made a promising beginning in the nurture of strategists of attrition; but the wealth of the country and its adoption of unlimited aims in war cut that development short, until the strategy of annihilation became characteristically the American way in war'.

² John Lewis Gaddis, *Surprise, Security and the American Experience* (Cambridge, MA: Harvard University Press, 2003), chap. 2

³ Ibid. As Gaddis himself points out, Adams warned against the temptation of foreign interventionism and opposed the annexation of Texas and the war on Mexico. Gaddis's argument about pre-emption as a motive in the case of the Floridas seems more valid when applied to the period *before* the War of 1812 than the years following the Peace of Ghent. Writing of the 'lure of the Spanish colonies' during the years 1783–1812, Julius Pratt observed that 'land and the control of transportation routes had been the principal aims of the expansionists, but with these had been mingled other motives – the dread of Spanish influence among the Indians, the fear that Florida might be used by Spain or some other power as a base of operations against the United States, and the desire to secure full control of the Gulf of Mexico'. See *The Expansionists of 1812* (Gloucester, MA: Peter Smith, 1957), p. 60.

⁴ Hamilton did not actually propose an attack on New Orleans before a formal declaration of war, but he hoped to bring about a break with France and Spain that would allow the United States immediately to move against the strategic city. It could be added that there have been several failed attempts at the preventive annexation of territory. In 1843, the Senate rejected the Tyler administration's argument that Texas must be annexed to foil alleged British designs to control it. In 1893, President Grover Cleveland rejected the Republican argument that Hawaii must be annexed to keep it out of the hands of some other major power.

⁵ It would be misleading to call them 'wars of choice' because those entering into them believed that they had little choice but to do so.

⁶ This article does not deal with the subject of so-called 'small wars', including US campaigns against the Indians and interventions in Latin America, the Caribbean, Lebanon, Somalia, Bosnia, Kosovo and elsewhere. On that subject, see Max Boot, *The Savage Wars of Peace* (New York: Basic Books, 2002).

⁷ See Michael Walzer, *Just and Unjust Wars* (New York: Basic Books, 1977), p. 81.

⁸ The former category would include the South's war against the North in 1861, Germany's war against Russia and France in 1914, and Japan's against the United States in 1941. The latter category would include the war that might have been fought against Nazi Germany.

⁹ 'Historically the *jus ad bellum* has developed around a set of seven principles on how to justify resort to war: the requirement that, to be justified, a war must have a *just cause*, be waged by *proper authority* and with a *right intention*, be undertaken

only if there is *reasonable hope of success* and if the total good outweighs the total evil expected (*overall proportionality*), be a *last resort*, and be waged for *the end of peace*. See James Turner Johnson, *Morality and Contemporary Warfare* (New Haven, CT: Yale University Press, 1999), p. 27. The key of criterion of just cause was understood, according to traditional Catholic doctrine, to include protection of the innocent, retaking people or property wrongly taken, and punishment of evil. Contemporary positive international law as embodied in the UN Charter, or what Walzer refers to as the 'legalist paradigm', interprets just cause more narrowly, namely, as limited to self-defence against territorial aggression. The trend since the end of the Cold War has been to broaden the meaning of just cause to include 'humanitarian intervention' in the internal affairs of sovereign states, thus harkening back to the broader traditional definition of just cause.

¹⁰ Walzer makes this point in the context of his discussion of the views of John Stuart Mill. See *Just and Unjust Wars*, p. 88. Mill's classic argument on foreign intervention makes a distinction between cases in which an oppressed people is struggling for freedom against a foreign despot (or a local despot backed by foreign power) and one in which it is struggling against a purely local despot. In the former case, he argues, intervention is justified to offset the power of the foreigner and to level the playing field. In the latter case, however, intervention is not justified because if the freedom fighters do not possess sufficient strength, dedication and popular support to overcome the despot on their own, it is unlikely that the overthrow of the despot by a foreign power will lead to a durable free government. Mill observes, 'When a people has had the misfortune to be ruled by a government under which the feelings and virtues needful for maintaining freedom could not develop themselves, it is during an arduous struggle to become free by their own efforts that these feelings and virtues have the best chance of springing up. Men become attached to that which they have long fought for and made sacrifices for.' J.S. Mill, 'A Few Words on Non-Intervention' (1859), in *Dissertations and Discussions* (London: Longmans, Green, Reader and Dyer, 1867), p. 175.

¹¹ On the legalist paradigm, see note 9, above.

¹² A case can be made that the Gulf War was unnecessary: the United States had abetted Saddam Hussein's rise to regional power in the 1980s by tilting in his favor during the Iran-Iraq War (specifically through the provision of satellite intelligence, the granting of large credits for the purchase of American food, and the protection of Kuwaiti ships against Iranian attacks). The United States declared a neutral position in the dispute over the British-drawn Iraq-Kuwait border, thus conceding that Iraq's claims were not without merit. Once Iraq had invaded Kuwait, the United States (as Colin Powell argued at the time) had options other than war, namely, sanctions and containment from the air. Washington's shift to an offensive military posture in late 1990 not only imposed a deadline for resolution of the crisis but probably made Saddam Hussein less likely to back down. This argument is ultimately unconvincing since sanctions have often proven ineffective and the United States and its allies could not run the risk of allowing Saddam Hussein to augment his power by absorbing Kuwait.

¹³ Francis Fukuyama's study advancing a universal, directional history of mankind

based on Alexandre Kojève's interpretation of Hegel is surely the exception rather than the rule. See Francis Fukuyama, *The End of History and the Last Man* (New York: Free Press, 1992).

¹⁴ See Ernest Lee Tuveson, *Redeemer Nation: The Idea of America's Millennial Role* (Chicago, IL: University of Chicago Press, 1968). Tuveson draws a distinction between the Roman Catholic (Augustinian) view according to which salvation in the temporal world is a contradiction in terms, the fundamentalist Protestant or 'millennarian' view (entertained by some Evangelical Christians today) according to which Christ will literally return to reign on earth, and the more mainstream Protestant, or 'millennialist', interpretation of the Book of Revelation, according to which the Reformation signified the downfall of Babylon (embodied by the Roman Church), to be followed, in due course, by the reign of the just on earth. He observes: 'We make history ourselves though we are under the inspiration of grace. The modern world, to a greater extent than we may appreciate, has been motivated by this kind of conviction, whether its form is Christian apocalypticism or 'Marxist dialectic'; both set forth predictions and general guides, both call for the most strenuous efforts to destroy the opposition and to progress toward the happy time sure to come, by a path marked out in advance' (pp. 47–8).

¹⁵ Jefferson to T. Coxe, 1 June 1795, in *The Papers of Thomas Jefferson*, vol. 28, pp. 373–4; John O'Sullivan, 'The Great Nation of Futurity', *The United States Democratic Review*, vol. 6, no. 23, November 1839, pp. 426–30. Lincoln's Second Annual Address to Congress, December 1862; Wilson speech, 13 July 1916, and Cheyenne speech, September 1919; George W. Bush, Second Inaugural Address, 20 January 2005. One could add many other examples including Herman Melville: 'We Americans are the peculiar, chosen people – the Israel of our time; we bear the ark of the liberties of the world'. Hermann Melville, *White-Jacket; or, the World in a Man-of-War* (London: Oxford, 1929), pp. 143–4.

¹⁶ See Joseph S. Nye, *Bound to Lead: The Changing Nature of American Power* (New York: Basic Books, 1990).

¹⁷ For the first use of the term, see John O'Sullivan, 'Annexation', in *United States Democratic Review*, vol. 17, July–August 1845, p. 5. See also Frederick Merk, *Manifest Destiny and Mission in American History* (New York: Knopf, 1963), chap. 2; John Schroeder, *Mr. Polk's War* (Madison, WI: The University of Wisconsin Press, 1973), p. 161.

¹⁸ Josiah Strong, quoted in Ernest May, *Imperial Democracy: The Emergence of America as a Great Power* (New York: Harcourt, Brace & World, 1961), p. 8. See also A. T. Mahan, 'The U.S. Looking Outward' (1890), in his collection of essays, *The Interest of America in Sea Power* (New York: Harper Bros., 1897). Lodge warned that since the great powers were 'rapidly absorbing for their future expansion and their present defence all the waste places of the earth, the United States must not fall out of the line of march'. Henry Cabot Lodge, quoted in Julius Pratt, *Expansionists of 1898* (Baltimore, MD: Johns Hopkins Press, 1936), p. 2.

¹⁹ Wilson war message, quoted in John Milton Cooper, *Breaking the Heart of the World* (Cambridge: Cambridge University Press, 2003), p. 23.

²⁰ Truman Doctrine speech, 12 March 1947; NSC 68, April 1950; Truman's State of

the Union Address, 8 January 1951.

²¹ See Fukuyama, *The End of History*. To be fair to Fukuyama, he did not predict how quickly or how easily the historical process he identified would be brought to fruition, and noted (pp. 234, 238) the possibility not only of new international rivalries along cultural lines but of 'new, illiberal ideologies'.

²² See the discussion of the Pentagon's 1992 Defense Planning Guidance (published in part over the signature of Secretary of Defense Dick Cheney) in James Mann, *The Rise of the Vulcans: The History of Bush's War Cabinet* (New York: Viking, 2004), chap. 13. As Mann points out, the Clinton administration, while putting economic globalization at the center of its agenda, did not challenge the basic arguments of the document.

²³ US National Security Strategy, September 2002; Bush speech at West Point, 1 June 2002.

²⁴ Robert W. Tucker, *The Just War: A Study in Contemporary American Doctrine* (Baltimore, MD: The Johns Hopkins University Press, 1960), p. 4.

²⁵ Reinhold Niebuhr, *The Children of Light and the Children of Darkness* (New York: Scribner's, 1944), pp. 179–80. Psychologists have observed the same phenomenon in individuals: 'We are prone to alter our perceptions of causality so as to protect or enhance our self-esteem.' Albert Hastorf et al., cited in Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976), pp. 343–4.

²⁶ On the 'Jacksonian' tendency, see Walter Russell Mead, *Special Providence* (New York: Knopf, 2001), chap. 7.

²⁷ Alexander Hamilton letter to President George Washington, 14 April 1794, Papers of Alexander Hamilton, vol 16, pp. 266–79.

²⁸ Perkins, *Prologue to War*, p. 286.

²⁹ Merk, *Manifest Destiny*, observes that local British officials feared US intrigue and favored British acquisition of the province, but the idea received no encouragement from London (p. 77).

³⁰ Consul A. J. Donelson's fears arose in part from the activities of his British counterpart, Charles Elliot. In 1845, Elliott attempted to arrange a deal whereby Texas would remain an independent state (rejecting annexation) and in return be recognized as such by Mexico (which still considered Texas a rebellious province). When this demarche became known and Texas rejected the deal, Donelson concluded that the British move had been a ploy to anger Mexico and to goad it into a war with the United States. According to a recent study of the question, 'Donelson's fears that the British were working actively behind the scenes to subvert annexation at any cost had no basis in fact. The U.S. chargé had greatly exaggerated the influence of his British counterpart, whose mission to Mexico had never been authorized by Her Majesty's government. Mindful of the wave of anti-British sentiment in Texas and the United States which Elliot's activities had caused, [Foreign Secretary] Lord Aberdeen censured his freewheeling diplomat and informed the U.S. minister in London that Elliot had acted on his own initiative. Donelson's suspicions that the belligerent posture assumed by the

Mexican government had been instigated by the British were equally unfounded. Far from encouraging Mexico to embark on a war against its northern neighbor, Her Majesty's government had notified Mexican leaders that they could expect no aid from Great Britain should an invasion of Texas be attempted, a warning which they would repeat on numerous occasions in the months ahead.' Sam W. Haynes, "'But What Will England Say?'" Great Britain, the United States, and the War with Mexico', in Richard V. Francaviglia and Douglas W. Richmond, *Dueling Eagles: Reinterpreting the U.S.-Mexican*

War, 1846-1848 (Arlington, TX: University of Texas Press, 2000), p. 23.

³¹ Quoted in *ibid.*, p. 88.

³² Theodore Roosevelt, quoted in Louis A. Pérez, *The War of 1898* (Chapel Hill, NC: The University of North Carolina Press, 1998), p. 58. On the war hysteria, see May, *Imperial Democracy*, p. 147.

³³ For one thing, it failed to take into account the fact that US naval power already in being would help to counter the submarine threat to Britain.

³⁴ Thomas A. Bailey, *A Diplomatic History of the American People* (New York: Appleton-Century-Crofts, 1964), p. 593.

³⁵ *Ibid.*, p. 570.

³⁶ Surfacing to issue a warning put the lightly-armed submarines at risk from armed merchantmen and passenger ships. The *Lusitania* was under secret orders to ram attacking submarines.

³⁷ Bailey, *A Diplomatic History of the American People*, p. 593.

³⁸ See William Stueck, *The Korean War: An International History* (Princeton, NJ: Princeton University Press, 1995), pp. 89, 94.

³⁹ On MacArthur's reaction, see *ibid.*, p. 95.

⁴⁰ For a lucid account, see Daniel Ellsberg, *Secrets: A Memoir of Vietnam and the Pentagon Papers* (New York: Penguin, 2002), chap. 1. The second of the two reported North Vietnamese attacks, on 4 August, probably never happened. There were serious doubts at the time.

⁴¹ Maxwell Taylor, quoted in M. Young, *The Vietnam Wars* (New York: HarperCollins, 1991), p. 134. See also George Kahin, *Intervention: How America Became Involved in Vietnam* (Garden City, NJ: Anchor, 1987), chap. 15.

⁴² John Randolph of Virginia, quoted in Perkins, *Prologue to War*, p. 433.

⁴³ The administration newspaper wrote in May 1811 that 'our rights are absolute, not contingent'. Quoted in *ibid.*, p. 257.

⁴⁴ See Stueck, *The Korean War*, pp. 111-19.

⁴⁵ Quoted in Howard Jones, *Death of a Generation: How the Assassinations of Diem and JFK Prolonged the Vietnam War* (New York: Oxford University Press, 2003), p. 444. On JFK's withdrawal plan see *ibid.*, chaps 15-16.

⁴⁶ George Garrison, quoted in Merk, *Manifest Destiny*, p. 146.

⁴⁷ Quoted in Bob Woodward, *Bush at War* (New York: Simon & Schuster, 2002),

pp. 281–2.

⁴⁸ On Powell's meeting with Bush and Rice at the White House on 5 August 2002, see *ibid.*, pp. 332–4.

⁴⁹ Merk, *Manifest Destiny*, chap. 4.

⁵⁰ William McKinley, quoted in May, *Imperial Democracy*, pp. 252–3.

⁵¹ Cited in Bailey, *A Diplomatic History of the American People*, p. 822.

⁵² Stueck, *The Korean War*, p. 95.

⁵³ H. J. Res. 1145, 7 August 1964 (the Gulf of Tonkin Resolution).

⁵⁴ There can be little doubt that this was his intention at the time of his death. See Jones, *Death of a Generation*, pp. 452–3.

⁵⁵ On this point see Frederik Logevall, *Choosing War: The Lost Chance for Peace and the Escalation of War in Vietnam* (Berkeley, CA: University of California Press, 1999).

⁵⁶ On Cheney, see Mann, *The Rise of the Vulcans*, p. 355.

⁵⁷ Niccolò Machiavelli, *Discorsi sopra la prima deca di Tito Livio* (Torino: Bollati Boringhieri, 1993), vol. 2, chap. 10, p. 266.

⁵⁸ See Allan Nevins (ed.), *Polk: The Diary of a President, 1845–1849* (New York: Longmans, 1952), p. 87.

⁵⁹ Polk also feared that Congress (controlled by the Whigs as of early 1847) would sooner or later cut off funding for the war, forcing him to withdraw US forces from California and New Mexico as well as central Mexico.

⁶⁰ For Polk on Trist, see Nevins, *Polk: The Diary of a President*, p. 309.

⁶¹ Theodore Roosevelt, quoted in Pérez, *The War of 1898*, p. 92.

⁶² *Ibid.*, chap. 4. Language of the joint resolution of April 1898 quoted in May, *Imperial Democracy*, p. 159. On the latter point see Mark Stoler, 'War and Diplomacy: Or, Clausewitz for Diplomatic Historians', *Diplomatic History*, vol. 29, no. 1, January 2005, p. 19.

⁶³ On this point, see George F. Kennan, *American Diplomacy* (Chicago, IL: University of Chicago Press, 1984), pp. 66–8.

⁶⁴ The article read: 'The Allied and Associated Governments affirm and Germany accepts the responsibility of Germany and her Allies for causing all the loss and damage to which the Allied and Associated Governments and their nationals have been subjected as a consequence of the war imposed on them by the aggression of Germany and her allies.' According to a perceptive account of Wilson's role in Paris, 'it was precisely on those questions that the Germans objected to most vigorously – questions that had a punitive character and involved national honor and prestige, such as exclusion from the League of Nations, the trial of the emperor, the delivery of war criminals, and of course, the war-guilt clause of Article 231 – that Wilson was not to be moved'. See Manfred F. Boemeke, 'Woodrow Wilson's Image of Germany, the War-Guilt Question, and the Treaty of Versailles', in Manfred F. Boemeke, Gerald D. Feldman and Elizabeth Glaser (eds), *The Treaty of Versailles: A Reassessment after 75 Years* (Cambridge: Cambridge University Press, 1998), p. 613.

⁶⁵ NSC 124-2, 25 June 1952.

⁶⁶ The head of the CIA, Porter Goss, admitted as much in testimony to Congress in February 2005. See D. Jehl, 'Intelligence Officials cites Wide Terror Threats', *The New York Times*, 17 February 2005.

⁶⁷ Though the Bush administration has claimed that 'freedom and democracy are on the march in the Middle East', it is difficult to see a connection between the transition in Iraq and movement toward the resolution of the Israel-Palestine issue. What limited progress one has seen is linked to Sharon's decision to withdraw from Gaza (in order to preserve a Jewish majority within Israel and to tighten Israel's hold on parts of the West Bank), and to the death of Yasir Arafat. Nor is there an obvious connection between Iraq and events in Lebanon (triggered by French and US pressure on Syria to leave the country after Damascus's manipulation of the Lebanese constitution, and by the assassination of ex-prime minister Hariri). The one clear connection is Mubarak's announcement of multi-party presidential elections, but the fact that a Mubarak-controlled parliament will vet opposition candidates induces a certain scepticism about the possibility of change.

⁶⁸ See G. John Ikenberry, *After Victory* (Princeton, NJ: Princeton University Press, 2001).

⁶⁹ Alexander Hamilton to George Washington, 15 September 1790, *The Papers of Alexander Hamilton*, vol. 7, pp. 52-4.

⁷⁰ Thomas Paine, *Common Sense and Other Political Writings* (New York: The Liberal Arts Press, 1953), p. 10.

⁷¹ Kennan, *American Diplomacy*, p. 73.

⁷² Anyone thinking otherwise should re-read the final chapter of Fukuyama, *The End of History*.

Between Conscience and Self-Interest: The United States, Sudan and Darfur

Alec Barker

What is the United States' "vital national interest" in stopping the killing in Darfur? How can the United States best contribute to ending the bloodshed in Sudan? This essay examines the many American concerns in Sudan, from oil to the prosecution of crimes against humanity. But it settles on one decisive interest that tips the balance in favor of supporting intervention: the maintenance of international support for American policies as a primary source of US power. The United States should therefore offer serious support to an African-led intervention – an approach that will resolve the Darfur conflict without exacerbating already diminished perceptions of American legitimacy.

The ongoing mass murder in Darfur is a heartbreaking crime perpetrated by a Sudanese government desperately gripping power. But what does it mean to an American superpower fighting a most unconventional world war? How should the United States respond? Firstly, this paper asks what the United States' strategic interest in stopping the atrocity in Sudan is. Although this appears to be a cold-blooded question, it gives an encouraging answer: the superpower is compelled to act by its need to return to an internationally respectable foreign policy inclusive of moral reasoning and careful analysis. Secondly, the paper asks how the United States can best achieve this interest, seeking in particular a practical approach to the problem. Ironically, the answer lies in dispelling the myth of American omnipotence and acknowledging that strong support to an African-led international intervention is, although imperfect, preferable to both inaction and direct intervention.

In an era when the international community debates the idea of unipolarity, the viability of international norms and institutions, and the limits of sovereignty, the United States debates the effects of its foray into self-interested nation-building in Iraq. The Bush Administration has backpedaled from faith in pre-emption and unilateral action but has simultaneously increased rhetoric about the still-elusive democratic peace. However the words and respective deeds do not correspond. While American intelligence collaborates with the Government of Sudan (GoS) in the war on terror, the US executive has both applauded the southern peace and condemned the Darfur atrocity. The US Congress has demanded an end to the

Alec Barker is an M.A. candidate at the Bologna Center, concentrating in strategic studies. He is a former United States Army officer.

violence but has also reneged on a \$50 million pledge to support the African Union (AU) intervention. The government's confused policies toward Sudan have, in effect, accommodated this ongoing crime, further calling into question the United States' faltering position as an international leader. But with some Darfur policy adjustments, America can seize the opportunity to reassert itself as a just *and* powerful world leader.

In the broader context of right and wrong, Darfur represents a chance to build upon lessons learned at high costs in Srebrenica and Rwanda.¹ In the decade following the end of the Cold War, principles of collective security were expanded upon to protect against violations of human rights. But a comprehensive, prescient, and executable doctrine did not exist to implement the new mandate. Although peace-seeking international intervention remains an imperfect practice, we have learned that it takes will, money and above all conscience to end crimes against innocent non-combatants.

I. WHAT IS THE UNITED STATES' INTEREST?

It seems heartless to ask, understanding the despicable nature of the Darfur quasi-genocide,² why the United States should consider intervening to stop the violence. But as with all questions dealing with exercising national power, leaders must define the "vital national interest," or lack thereof. Previously, the coldly calculated economic costs of intervention were far greater than the costs of indifference. But now the cost of indifference is too high and the potential payoff to a smart intervention is too great.³

There are five general categories of strategic interests pertinent to the United States in Sudan. The first four are subordinate to the fifth and will be discussed in the following section. The four subordinate categories of interest are regional stability in Africa, economic concerns, issues related to the "long war" (Pentagon-speak for the war on terror), and problems of international law and conflict management.

A. SUBORDINATE INTERESTS

Continued instability in Darfur in particular and Sudan in general may lead not only to the continuation of the quasi-genocide but also to the expansion of conflict and instability to Chad, the Horn, Great Lakes, and Nile regions. Specifically, the conflict has the potential to topple the Chadian president, Idriss Deby; catalyze renewed war between Eritrea and Ethiopia; fuel murder, rape and slavery in Uganda; and revive tensions between Sudan and Egypt.⁴ The crisis runs the risk of upending the GoS itself or returning it to radicalism, both of which could result in a failed state that jeopardizes international security. Finally, the US would benefit from a successful AU intervention because it would strengthen a new outlet for conflict management in the troubled region.

While Chinese and Indian corporations ignore the situation in Darfur in their pursuit of energy, the humanitarian disaster in Sudan restricts American companies from contracting the drilling of oil that Chevron discovered some 28

years ago. At a time when the American economy is searching for new sources of energy, yet cannot afford to compromise principles long ago forgone by China and India, the unresolved crisis in Darfur is an obstacle to American development of Sudan's oil resources.⁵

There are few lines of economic communication as strategically important as the Suez Canal. Maintenance of this arterial passageway is almost a clichéd American interest. Sudan owns more than 300 miles of Red Sea coastline and is therefore vitally important to securing the southern outlet of the Suez Canal. A civil war between the GoS and the Beja-Rashaida Eastern Front could jeopardize control of this shipping artery. Given the recent increase in maritime piracy in the Indian Ocean and Gulf of Aden, as well as radicalism in Yemen, stability near Sudan's shores is an essential US interest.

Since the September 11, 2001, terrorist attacks on the United States, cooperation with foreign intelligence services has become vital to the long war. Indeed, Sudan became extremely important after 9/11 because it had information about Bin Laden and Al Qaeda; it could collect intelligence where the CIA could not; and it was trying to shed its image as a pariah state. Close collaboration between the CIA and the Sudanese Mukhabarat enables the American offensive against terrorists, but makes them unlikely allies.⁶

Americans have a military interest in Sudan which is undeniable. With the US military fully deployed in the long war, and its Army and Marine Corps particularly stretched thin, it is clearly in the United States' interest *not* to be in another large protracted engagement. Most American soldiers and marines have served one to two combat deployments since 9/11, and some have spent more than 36 months at war. Many units not historically intended for combat have been sent to Iraq, and Reserve and National Guard units have been relied upon heavily to distribute the burden.⁷ The US military simply cannot afford another large scale commitment. Any response to Darfur must account for this reality.

A cornerstone of the national security strategy is control of the proliferation of weapons of mass destruction (WMDs). This principle applies not only to rogue states, but also to stateless terrorist organizations such as al Qaeda. But a recent European intelligence assessment claims state-owned enterprises in Sudan have served as a conduit for nuclear program equipment and materials en route to third party nuclear weapons programs.⁸ The network of disgraced Pakistani nuclear scientist A.Q. Khan is the suspected culprit, and Iran is believed to have been the customer.⁹ Although the purchases seem to have ended in 2001, the US cannot permit Sudan to be a black market broker for WMDs.

Americans, though some disagree or would prefer not to admit it, have an inherent interest in seeing violations of human rights and international law stopped and punished. Some, like Robert Kagan, argue that the internationalist tendency towards laws and rules is a position taken up by weak states, lacking the physical power to influence world politics.¹⁰ But as we have seen in Taliban Afghanistan, the Iraq insurgency, and present-day Somalia, lawlessness is a precondition for terrorism. The United States must support the norms that comprise international law, though it may be required to swallow some bitter pills in the process.

American interest in Darfur is complicated by what Alan J. Kuperman calls the “moral hazard of intervention.”¹¹ To some extent, this has already occurred in Sudan, in that the diplomatic pressure placed on the GoS to finalize the Comprehensive Peace Agreement (CPA) has encouraged rebels elsewhere in Sudan to take up arms. American effort to resolve the Darfur crisis must be aware of this principle and act to reduce its effect by pressuring the GoS to include all marginalized people in any Darfur peace settlement.

B. PARAMOUNT STRATEGIC INTEREST

The above are admittedly tired arguments. Advocates of humanitarian intervention have invoked similar reasons before, trying to prod western powers to stop crimes against humanity. As cold as they may seem, they are not cynical enough to spur intervention, especially in Darfur. They have failed in the past, and were they the only concerns here, they would probably fail again.

1. A CATEGORICAL IMPERATIVE?

The question lying at the heart of international debate over humanitarian intervention is the theoretical categorical imperative that a superpower must stop a foreign state from massacring its own people, irrespective of its territorial sovereignty. This norm is simply not influential in Washington. Although contested by isolationists, this principle has often been presented as a great power’s responsibility or burden. Human rights and international law advocacy groups, such as the International Council on Human Rights Policy, have proposed a government’s “responsibility to protect.” This asserts that “outside states have a duty to act when a government is unwilling to prevent violations of rights in their territories or is itself the cause of those violations... Their duty is to act, within their capacity, to assist and protect those whose rights are being violated.”¹²

The UN World Summit in September 2005 reaffirmed the emerging norm of a “responsibility to protect” from genocide, war crimes, ethnic cleansing, and crimes against humanity.¹³ But the UN produces many well-intended resolutions that cannot be enforced. It is clear that national governments make decisions in consideration of the vital interest at stake rather than the survival of a theoretical norm. The notion of a superpower’s inherent or supranational responsibility is an obstacle to US involvement in humanitarian intervention. This is seen as a foreign attempt to control the superpower. But the US should adjust its perspective and recognize the *strategic* imperative to adopt a moral foreign policy in order to help maintain the current condition of American primacy.

2. A STRATEGIC IMPERATIVE

If the United States is to sustain the “unipolar moment,” it must treat the termination of ongoing massacres like the one in Darfur as an element of national strategy.¹⁴ Permitting the continued slaughter and displacement of innocent civilians suggests to some foreigners that American democracy is hypocritical and therefore unfit to lead. This undermines international support for American policies. International consensus-building is far more important to the hegemon in a unipolar world than it was in bipolar competition with the Soviet Union. The

lesson of the decision to invade Iraq without majority international agreement is that raw force has proven a flimsy way to support American objectives in today's world, and that the ideological battleground is more important than was assumed.

Prior to the Iraq insurgency, the US had assembled a military machine far superior to any other whose strength no one doubted. But it has since revealed weaknesses and simultaneously alienated some of its key supporters. The priority of the second Bush Administration should have been the restoration of legitimacy as a superpower, but despite some recent positive developments in conciliatory language and diplomacy, it has wholly failed to generate international support. This has devastated US foreign policy.

There is talk in the United States of an "Iraq syndrome," or a renewed, Vietnam-like public reluctance to allow its commander-in-chief to pursue vague policy goals in deploying the armed forces abroad. The real Iraq syndrome in the future will not be characterized by American popular reluctance to support foreign adventures, but will instead describe the reluctance of international partners to cooperate with US policy initiatives or to participate in American-led military operations.¹⁵ It will include their desire to constrain American power and their pursuit of a counter-balance to unipolarity. Some suggest existing alliances are unnecessary vestiges of a bygone era. Although there need not always be an unqualified alliance in, say, American-European relations, there certainly cannot be an enduring rift.

Therefore, the United States has a large strategic interest in demonstrating it is a moral hegemon, a legitimate and wise authority to be trusted and admired. It must also convince its partners that it is capable of advancing carefully reasoned approaches to thoroughly analyzed problems.

The term "moral" is not used here in the sense in which it is often popularly used to connote adherence to a particular system of norms assumed to be superior by the person employing the term. In this parlance, the many contradictory systems of supposedly "moral" behavior can be used to rally various constituencies to widely divergent causes. Rather, it is used here to recall a rational philosophy that acts on full understanding of a problem, making value judgments where necessary.¹⁶ This gives a way to derive policy approaches that are neither purely pragmatic (and indifferent to human suffering) nor purely idealist (and recklessly ignorant of associated costs and unintended consequences).

"Moral," in this context, is more akin in definition to "well-reasoned." A basis for a moral foreign policy would be found in the rational Augustinian approach to the just war, later developed by Hugo Grotius. The Augustinian tradition balances the idealism of the categorical imperative and the pragmatism of the realist. It asks not only if combat is philosophically just because of its cause, but also what the potential costs and likelihood of achieving that justice are. If these costs are too high, the risks too great, and victory unlikely, a war could be considered unjust no matter how revered the cause.

The current American leadership appears to have failed to deliberate in this way, seemingly substituting determination to fight for sound information and

analysis. If the US adopted such an approach now, it could not only accomplish ethically just objectives, but also employ more efficient methods. In Darfur a moral foreign policy demands for the killing to be stopped, although traditional forms of intervention are not appropriate.¹⁷

II. ACHIEVING THIS INTEREST

We have seen in Sudan that a coldly reckoned interest is also a moral one. In Darfur, the means that America chooses are as important, if not more so, than the end. The use of soft power, cooperation with allies and a thorough understanding of the crisis are essential. The US cannot intervene alone or as a leader, but it cannot afford to do nothing to help stop the atrocities in Darfur. It must wholeheartedly support an African-led intervention with tough diplomacy, financial assistance, equipment, and logistics focused on the specific shortcomings of the mission in Sudan.

A. A THOUGHTFUL APPROACH

The US should recognize the need to play a supporting role in Sudan. It already suffers from a conflicted mélange of policies and should not further complicate things by either a full intervention or leadership role in any international effort. In addition, an all-out US intervention would further exacerbate the perception of American incursions on “Muslim territory,” a problem which has inflamed anti-American aggression worldwide.

A comprehensive American-led military intervention is inappropriate for three reasons.¹⁸ First, a weighty American presence would become an obstacle to peace because it would attract Jihadists, invoke accusations of petro-imperialism, and rouse anti-American, Islamist hard-liners in the GoS, destabilizing al Bashir. Second, Secretary Rumsfeld’s transformation is incompatible with effective peacekeeping operations. The lighter, faster and cheaper hi-tech philosophy may pay lip service to peacekeeping, but neglects the reality that successful peacekeeping operations are technologically basic and require overwhelming traditional capabilities.

But the reason trumping all others is that the US is in fact precluded from militarily intervening in Sudan by its involvement in Iraq and Afghanistan. The Pentagon is simply overextended and cannot now project overwhelming conventional force anywhere else.¹⁹

B. REGIONAL ORGANIZATION

This brings us to the importance of regional organizations, particularly the newly reborn AU and its African Mission in Sudan (AMIS). They are better informed, more culturally aware, and more familiar with the operational environment. They will not be confronted as arrogant outsiders seeking imperial influence.

One critique of the regional peacekeeping operation is that it provides a way for self-interested parties to intercede in a neighbor’s crisis. This does not

appear to be a factor in AMIS. The risk that they may not be completely impartial and unbiased in carrying out its mission is minimal and an acceptable one, provided the AU maintains the current force structure. The fact that none of the troop-contributing nations shares a border or long-standing disputes with Sudan is an encouraging factor.²⁰

It is well understood that regional organizations have a mixed record in peacekeeping operations.²¹ Despite this, the US must invest in AMIS. As one of the AU's first two peacekeeping operations, the international community is watching its progress closely. If AMIS is able to prove that the African Union is capable of peacekeeping, then it may begin to become the strong stabilizing force African leaders envision. Perhaps the most important reason the United States should support AMIS is the increasing likelihood that as humanitarian crises continue in Africa, the US will be increasingly militarily unable and politically unwilling to project intervening force on the Continent. The long war will likely be a generational war and will exact a high price from the all-volunteer service. Investing now in the AU's capability will pay long-run dividends.

This is an extension of earlier (still valid) arguments for the African Crisis Response Initiative (ACRI) and its successor, the Africa Contingency Operations Training and Assistance (ACOTA) program. President Clinton proposed ACRI "to help African nations to respond to humanitarian crises and peacekeeping missions in their region."²² President Bush's ACOTA has trained peacekeepers from 15 African States and has scheduled the training of an estimated 40,000 more over the next 5 years.²³ It has facilitated force interoperability, command and control, and contributed to more capable African peacekeeping.

Moreover, American support for a decisive peacekeeping effort can prevent an intervention from becoming part of the problem. An intervention that does not distinguish between victim and aggressor sustains ambiguity and insecurity, permitting the bloodletting to continue unabated.²⁴ While American rhetoric must not undermine an AU mission by sending conflicting signals about the international community's partiality or impartiality, it should be applied behind closed doors to convince the AU to take a clear stance against state-sponsored murder. Without a clear challenge to the perpetrator, AMIS may become more a party to the problem than a solution, like UNAMIR in Rwanda or UNPROFOR in Srebrenica.

C. EMPLOYING REAL STRENGTH: MONEY, MATERIEL AND TRANSPORT

The popular misconception that the American military is omnipotent suggests that its supposedly awesome capabilities require the US to intervene in Darfur based on its "responsibility to protect." This is a dangerously influential notion that has, at times, caused blame to shift from the very criminals responsible to the non-intervening superpower. The idea that the US is compelled to act by virtue of its "power" is short-sighted; it historically has very little relative power in humanitarian crises. Somalia, Bosnia-Herzegovina and Rwanda prove its military has no extraordinarily pre-eminent expertise in stability operations. Therefore, the United States is not compelled to lead a traditional intervention by

the “power brings responsibility” argument. However, it is “powerful” with respect to its position to fund, equip and provide logistical support to an intervention, and should therefore treat this as a moral responsibility.²⁵

Several studies of the AU’s effort in Darfur have determined that AMIS is ill-equipped, poorly trained and unprepared to fulfill its purpose. In November 2005, the Brookings Institution’s University of Bern Project on Internal Displacement and the humanitarian NGO, Refugees International, published separate analyses of AMIS and reached a similar conclusion: that significant reinforcement is required for AMIS to become an effective deterrent to violence in Darfur.

The United States should pursue the following on behalf of AMIS:

1. An increase in personnel commensurate with the expansive territory it must patrol, to between 12,500 and 20,000 peacekeepers. This increase should be manned by AU member states, preferably those least partial to the GoS or the Darfur rebels, and not the NATO troops Bush has called for. But AMIS’s recent payroll troubles show it would be hardpressed to fund such an expansion.

2. Delivery on the \$50 million pledge made in April 2005. This funding is critical. AU Peace and Security Commissioner Said Djinnit has warned that AMIS only has enough funding to sustain itself through April 2006, despite a December 2005 donation by the European Union of 70 million euros, or \$84 million.²⁶

3. A strengthened mandate. The current mandate places the primary responsibility for protecting civilians on the GoS. This forces AMIS to abstain from intervention in any situation where government forces assert sovereignty, even with clear indications that a criminal act is taking place. The mandate also requires AU troops to fight only in self-defense and reduces their effectiveness to a weak deterrent. A revised mandate should empower AMIS with the clear responsibility to actively and forcefully protect civilians and humanitarian agencies.

4. Increased capabilities in transport, logistics, intelligence, command and control, and civilian police.²⁷ It should match the Canadian donation of armored personnel carriers with a large fleet of 2.5 and 5 ton trucks outfitted for cargo, personnel transport, and convoy escort (including anti-vehicular weaponry), as well as the parts needed to fix them. AMIS needs intelligence training, support, and equipment, in addition to the clear authority to collect intelligence as necessary to monitor and prevent attacks against civilians. The force also needs long-range tactical radios with encryption features to enable its transport and security operations. These are simple examples of the many ways in which the mission lacks basic capabilities.

5. More firepower. This issue is a very controversial one, as enhancing a peace operation’s offensive capability seems counterproductive. But, no peace will endure until the *janjaweed* are deterred by a robust and determined foreign presence. AMIS needs heavy machine guns, basic anti-armor weapons, helicopter gunships, and fixed wing attack aircraft to provide credible deterrence and to defend itself.

6. Increased pressure on GoS. This should begin with an AU-declared and

internationally supported no-fly zone over Darfur. The GoS must stop paying lip service to ending the atrocities and demonstrate that it is truly ending its foray into indiscriminate counter-insurgency by disarming the *janjaweed*.

7. Coordinated information operations. There is a basic tendency for those within a conflict to misunderstand the goals and mandate of intervening forces, leading to unrealistic demands, low morale and a deficit of trust. In the worst cases, these misunderstandings flare into violence. A clear message delivered in the colloquial language would help mitigate this problem.

The aforementioned are just a sampling of some of the basic changes necessary to improve AMIS.²⁸ The list is by no means comprehensive, but the bottom line is that the US and the international community must enable the AU Mission in Sudan, unlike previous peace keeping operations, to truly protect and not simply to monitor or serve as a witness to crime.

D. LONG-RUN CONSIDERATIONS

No effort in Sudan will succeed unless it addresses the root causes of the conflict, of which there are many that initially appear intractable.

Any solution must overcome the fundamental problems in Sudanese center-periphery relations. Feasible political solutions include a consociational democratic constitution in either the letter or the spirit of the Comprehensive Peace Agreement (CPA), partition, and/or independence. The unity option must provide simultaneous redress for all of Sudan's marginalized people. Accordingly, US Deputy Secretary of State Robert Zoellick has insisted that any solution to Darfur should be addressed within a CPA framework and should include the Eastern Front.

A power-sharing political arrangement could begin to facilitate an end to unequal development in the outlying regions. GoS representatives have prevaricated in dealing with development for decades, making stale promises of potable water and paved roads to pacify the demands of the neglected. Only a truly representative power-sharing arrangement or self-government will address these needs.

Competition for and extraction of Sudanese resources by foreign investors will continue to provide incentive for the GoS to hold on to total power and thereby benefit from revenue. Removing this incentive will be difficult, because it will mean engaging the governments of China and India in a conversation about their human-rights-ignorant energy policies. The international community has heretofore been unsuccessful in persuading China to leverage the link between Sudanese oil and the Darfur crisis.²⁹ In the long run, international organizations must ameliorate this condition.

To a lesser extent, the GoS must address ethnic and religious divides. Theoretically, a government of national unity should make public reconciliation a goal. However, until there is a concerted effort to recognize and place value on the differences among Sudan's many fractured groups, the government cannot achieve reconciliation. It will take a combination of dynamic leadership,

transparency and effort in good faith. The current composition of the GoS makes the prospects for such reconciliation unlikely in the short term. But fortunately, ethnicity is not the major cause of violence and it is probable that if the required political and economic changes follow a demobilization of the *janjaweed*, the Sudanese people will return to an ethnically intermingled lifestyle.

The United States should get used to enabling peacekeeping in Africa. To some extent, the Africa Contingency Operations Training and Assistance program demonstrates it already has. But this is only an investment in human capital, a highly perishable commodity. When the time comes, the US must be prepared to fund and equip African (hopefully AU-led) peacekeeping operations. This capitalizes on US relative strengths, reduces the need to use its own troops, and reinforces its role as a well-deserved world leader.

Finally, a long term solution must deal with the dilemma of conflict displacement. This suggestion implies that an intervention may succeed in stabilizing Sudan, but may also succeed in destabilizing the surrounding region in the process. While there is no panacea to such a complex problem, strong support of AMIS now might validate the AU's overall mandate and thereby reinforce the young organization, providing a potential future outlet for conflict prevention and mediation on the continent.

III. CONCLUSION

With respect to Darfur specifically and humanitarian interventions in general, American leaders should be neither true believers nor cynics, but rather seek the most effective solution to this horrific problem and simultaneously seize an opportunity to reaffirm their nation's role as a powerful and just world leader.

What does Darfur really mean to the US? Although it represents many tangible and short term interests, the conflict's overarching significance is a chance for the US to recapture legitimacy as a benevolent hegemon.³⁰ Thankfully, this opportunity can be fulfilled with a relatively small investment in an African-led peacekeeping operation.³¹ Successful execution of this approach requires a middle-of-the-road policy. It must resist the temptation towards idealistic justifications and strategies for intervention, but must also guard against purely self-interested and heartless indifference. Furthermore, a balanced assessment of costs reflects both the material expense of total intervention and the political damage caused by diffidence to confront atrocity. Careful evaluation of internal political sensitivities to the situation in Darfur demonstrates that the American contribution to an intervention must avoid the historic tendency to dominate. Only this type of reasoned approach to foreign policy will justify and maintain international support for the US in a unipolar framework.

NOTES

¹ The public has yet to fully realize the enormity of the recent atrocities in the Democratic Republic of the Congo (DRC), from which there are also many other important lessons to be drawn.

² A detailed analysis of the controversy over the definition of genocide in Darfur is contained in "Understanding the Darfur Crisis," an unpublished paper available from the author. "Quasi-genocide" is Gerard Prunier's apt description of the atrocity in Darfur. There is a great deal of controversy over defining the genocidal character of the Darfur conflict, but we lack the ability to properly address the topic here. Briefly, we can say this controversy variously informs *and* confuses the real issues causing violence in Darfur: deficiencies in center-periphery relations, uneven distribution of political and economic power, and the GoS's sophomoric decision to use unofficial militia and ethnicity as tools of counter-insurgency.

³ An essential analysis of national interest and the politics of genocide is captured in the Pulitzer Prize-winning contribution of Samantha Power, *A Problem from Hell: America and the Age of Genocide*. (London: Flamingo, 2002).

⁴ There are also less likely scenarios in which the very complex situation in the eastern DRC could be inflamed.

⁵ Despite this fact, Sudan's oil is not even a partial answer to the US's energy problems. Technological development away from the current dependency on petroleum is the long run solution.

⁶ Although intelligence cooperation is critical to the long war and offers Sudan a potential way to leverage the American economic sanctions begun during the Clinton years, neither Bush nor al Bashir is eager to publicize the relationship. Al Bashir is opposed by hardliners within his government who denounce any cooperation with the US. And for Bush Darfur makes rewarding Sudanese cooperation with the end of sanctions and its removal from the list of states sponsoring terrorism unpalatable. A senior US government official familiar with terrorist threats in the region said Khartoum was not at present a state sponsor of terrorism. "These are not all nice guys, but they have gone way past a passing grade on counter-terrorism cooperation and don't technically belong on the list," he said. "The reason they are still there is Darfur, which is not related to state-sponsored terrorism but makes lifting sanctions now politically impossible." See Silverstein, Ken, "Official Pariah Sudan Valuable to America's War on Terrorism." *Los Angeles Times* 29 April 2005.

⁷ It might be hard to conceive of a uniformed military unit "not intended for combat deployment," but there are several in the US Defense Department such as the 11th Armored Cavalry Regiment, 1/509th Parachute Infantry Regiment, 1/3rd Infantry Regiment, and Combat Training Center Operations Groups that have exclusively training or ceremonial missions and have deployed since 9/11.

⁸ Traynor, Ian and Ian Coban, "Clandestine Nuclear Deals Linked to Sudan." *The Guardian* 5 January 2006.

⁹ *Ibid.*

¹⁰ Kagan, Robert, *Paradise and Power: America and Europe in the New World Order* (London: Atlantic Books, 2003).

¹¹ This principle suggests foreign powers which intervene on behalf of a minority in a sovereign state's internal conflict offer incentive to other groups to instigate a

crisis in the hope they too will be supported by external saviors, thereby making the problem worse. See Kuperman, Alan J. "Humanitarian Hazard: Revisiting Doctrines of Intervention," *Harvard International Review*, vol. 26, no. 1 (Spring 2004). Also, "Transnational Causes of Genocide, or How the West Exacerbates Ethnic Conflict" in *Yugoslavia Unraveled: Sovereignty, Self-Determination, Intervention*, Raju G. C. Thomas, ed. (Oxford: Lexington Books, 2003).

¹² International Council on Human Rights Policy, "Duties Sans Frontières: Human Rights and Global Social Justice" (2003): p 51.

¹³ *UN World Summit Outcome*, A/RES/60/1, 15 September 2005: paragraphs 138-39.

¹⁴ Is the unipolar condition desirable, for the US or the rest of the world? That is the topic of an entirely different debate which cannot be examined here. We assume American unipolarity for the indefinite future is desirable, acknowledging this is a highly debatable premise to be discussed elsewhere.

¹⁵ To some unilateralists this may be a good result, but they fail to see unipolarity is unsustainable without international cooperation.

¹⁶ By "full understanding," I refer to an understanding based on a decision-making process that considers divergent points of view, that seeks information from unfiltered primary sources, that challenges its assumptions, and comprehensively evaluates costs.

¹⁷ To be sure, the United States' foremost foreign responsibility is to leave Iraq in the hands of a capable, democratically elected government equipped with the basic tools necessary to develop its economy and defend itself. But ending the mass murder in Sudan is linked to America's recovery from its failures in the Iraq war.

¹⁸ And in the possible framework of a US-led United Nations intervention, there is a fourth reason. Any UN proposal to forcibly intervene in Darfur faces the likelihood of a Chinese veto in the Security Council. China may avoid alienating the GoS to protect its energy interests.

¹⁹ James Kurth writes, "Instead of pursuing humanitarian interventions, the United States has engaged in two wars, one in Afghanistan and one in Iraq, that the Bush administration justified in human rights terms. This is especially true in the case of Iraq, but the real impact of that war has been to make humanitarian intervention by the United States elsewhere impossible. This radically reduces the prospects for successful humanitarian interventions in the future, while improving the prospects for undeterred and uninhibited ethnic massacres or genocides, such as has been occurring in western Sudan." See Kurth, James, "Humanitarian Intervention After Iraq: Legal Ideals vs. Military Realities." *Orbis*, vol. 50, no. 1, (Winter 2006): p. 88.

²⁰ *Ibid.*, p 121. Rwandan initiative is critical to the success of AMIS. Paul Kagame's government was not only one of the first to respond to the AU's call for help in Darfur, but its troops are highly motivated by personal memories of their own genocide and the inability of UNAMIR to stop it. Rwandan soldiers make AMIS far different from previous attempts at regional peacekeeping and must be supported.

²¹ Only two operations are considered anything but failures. The 1965 Organization of American States intervention in the Dominican Republic restored order and made a timely departure, but is criticized as a veneer for US imperial interventionism. The 1990 Economic Community of West African States (ECOWAS) effort in Liberia may have mitigated an all-out catastrophe, but it failed to enforce a cease fire and in the long term, more unrest and war led to a UN intervention. See Diehl, Paul F., *International Peacekeeping* (Baltimore: Johns Hopkins University Press, 1999): pp. 121-122.

²² According to ACRI's US State Department website, accessed 8 January 2006. <<http://usinfo.state.gov/regional/af/acri>>

²³ White House press release: "Fact Sheet: United States and G8 Renew Strong Commitment to Africa," 8 July 2005.

²⁴ Matthew Krain discusses the weaknesses of peacekeeping models that are either impartial or offer the presence of a international witness as the lone deterrent to mass killing, arguing instead for interventions against the perpetrating party. See also the report of the Lakhdar Brahimi panel: "No failure did more to damage the standing and credibility of United Nations peacekeeping in the 1990s than its reluctance to distinguish victim from aggressor." *Report of the Panel on United Nations Peace Operations*, A/55/305 S/2000/809, 21 August 2000, p. ix.

²⁵ The stories of UNAMIR's failures in Rwanda demonstrate it lacked several empowering characteristics which could have helped prevent or stop the genocide. Canadian Lieutenant General Romeo Dallaire's account almost gives a formula for poor peace operations: insufficient troop numbers, a weak mandate, unfulfilled pledges of financial support, inadequate transport, bad logistics, nonexistent intelligence, and the presence of spoilers. We see all of these characteristics in AMIS and Darfur today, yet no remedy has been offered. This is perhaps due to a pervasive general pessimism in the west about sponsoring African peace operations. It seems the calls for a UN takeover of AMIS began as soon as AMIS was authorized. See Dallaire, Romeo A. and Brent Beardsley, *Shake Hands with the Devil: The Failure of Humanity in Rwanda* (New York: Carroll & Graf Publishers, 2003).

²⁶ "Sudan: AU Mission in Darfur Running out of Cash," Reuters Alert Net, 16 December 2005 (accessed at www.alertnet.org on 17 December 2005). This covers all but \$51 million of a \$135 million projected AMIS shortfall. The United States should read between the lines: the EU's decision to leave approximately \$50 million unfunded is a thinly veiled challenge to Capitol Hill's decision not to honor its McDoom, Opheera, "Arabs Pledge to Fund African Darfur Troops." Reuters Alert Net, 28 March 2006 (accessed at www.alertnet.org on 8 April 2006).

²⁷ The US did provide air transport of AMIS troops into Sudan, using two aircraft to move Rwandan and Nigerian units. Office of the Press Secretary (White House), "Statement on the Expanded African Union (AU) Mission in Sudan," dated 18 October 2004 (accessed at www.whitehouse.gov on 14 January 2006).

²⁸ For more on these suggestions, see the Brookings-Bern and Refugees International reports.

²⁹ China has promised a badly needed combat medical company to UNMIS,

but it has not yet deployed to Sudan.

³⁰ The United States was most recently recognizable as a benevolent hegemon at the end of the Persian Gulf War.

³¹ Relative to the costs of the current efforts in Iraq.

The Producing Countries: Between Pursuing Development and Meeting Global Demand

Giacomo Luciani

*Remarks given by Giacomo Luciani of the European University
Institute Florence at the Third Joint OPEC-IEA Workshop in
Kuwait City on May 15, 2005.*

GLOBAL DEMAND: A DEPENDENT VARIABLE

As we address the issue of whether OPEC investment is sufficient to create the capacity and produce enough crude oil to satisfy world demand, I believe we should consider the latter a dependent, rather than an independent variable.

Producers have frequently complained about demand uncertainty. They feel they are called to take the risk of investing in new capacity, only to discover that it is not required, because demand has not grown as expected. Producers therefore seek risk mitigation, an assurance that their capacity will be utilized, if they invest in creating it.

But should we assume that demand growth is exogenous and that producers must simply conform to it? Are we taking demand for sacrosanct, even if not entirely predictable? Or shall we discuss desirable demand trends?

Inevitably, we shall need to question demand patterns. A first major reason for this is the impact on the environment and emissions of greenhouse gases: the danger of climate change is serious enough and needs to be addressed urgently. A further reason is that even if we were able to capture and master greenhouse gas emissions, the fact remains that our model of consumption of fossil fuels is unsustainable in the long run.

The per capita primary energy consumption of the United States (at 277 million BTU per head) is 60 percent higher than that of the European Union and Japan (at 174 million BTU per head). This gap is not mirrored by an equivalent

Giacomo Luciani is professor of political economy and co-editor of the Mediterranean Program at the University Institute. He frequently lectures at the Bologna Center. Most recently he co-authored and co-edited Regime Change in Iraq: The Transatlantic and Regional Dimensions.

gap in the standard of living or per capita income¹: it simply reflects the absence of a serious effort at saving energy, and especially transportation fuel. In fact, even the pattern of consumption, which is characteristic of the EU or Japan, may not be sustainable, in the sense that it could not become universal.

The major component of demand growth today originates from China (which consumes 30 million BTU per head, 1/10 of the US consumption level) and India (12 million BTU per head, or less than 1/20 of the US level), two very large and rapidly developing countries, which aspire to bridging the gap that separates them from the industrial countries. This will require huge and essentially unstoppable fossil fuels demand growth: should this be met only through increased capacity and production, or also through a measure of containment, if not decline, in the demand growth of the industrial countries?

Resources are scarce: creating capacity to meet a rapidly growing demand requires huge investment, large and scarce human resources, and time. What we see at present is competition among net importers to acquire their needed supplies, leading to upward pressure on prices. It may be convenient to blame the OPEC producers that are not increasing capacity fast enough, but perhaps we should move more decisively to contain the industrial countries' demand.

I believe OPEC should not simply accommodate any demand growth, but factor in and anticipate a policy of demand restraint on the part of the industrial importing countries - simply because this is rational and necessary in order to achieve a demand pattern that is sustainable in the long run. Accommodating demand only temporarily hides the nature of the problem, but does not solve it.

SHOULD PRODUCERS ANTICIPATE GLOBAL DEMAND?

From the producers' point of view, is it advisable to anticipate or accommodate demand? It would seem that the best course is rather to trail demand growth somewhat - i.e. not to the point of creating serious scarcity or prices rising beyond control. A condition of relatively tight supplies and gently rising prices seems to be what is required to send the right message to the world: demand should be restrained and non-conventional supplies expanded.

Is there a danger that prices may become "too high"? OPEC of course is influenced by the historical memory of the early 1980s, when the group's market share was eroded and the system of posted prices had to be abandoned. However, conditions today are profoundly different: there is no major province coming on stream with such prolific supplies so as to marginalize OPEC by compensating for the declining production of an ever longer list of mature basins.

There will be fuels originating from non-conventional sources, such as the bitumen and heavy oils of Venezuela, or the tar sands of Canada; from natural gas (gas-to-liquids) or coal; and even from agriculture (corn, rapeseed, sugar cane and many more). But all of these require very large investment and long lead times before they generate a stream of products: they are, so to speak, long seen in coming.

We are talking of a different time scale from projects in the major oil producing countries, which can add hundreds of thousands or even more than a million barrels per day in five years or less. Hence, the major producers will have plenty of advance notice and will be able to influence prices in order to make sure that their capacity is not left underutilized.

Generally, major oil producers have preached against the development of non-conventional fuels and have tended to see them as a danger: here I am proposing exactly the opposite. It is in the interest of the major producers to “encourage” some development of non-conventional fuels and to manage prices so that this development does not become too rapid (a very unlikely occurrence). The “long-term equilibrium” price of crude oil is one that allows for development of alternative energy sources at a pace compatible with the producers’ preference for the pace of their reserves’ draw down. Hence in a sense it is up to the producers to establish their long-term production policy based on their preferences for reserves’ life; and to let the price encourage alternative fuels sufficient to fill the likely gap.

ARE PRICES TOO HIGH?

Can we say that prices today are high, or even too high? The evidence is less clear-cut than it might appear at first sight.

True, the prices of WTI and Brent have increased substantially, recently passing the \$60 per barrel threshold – about five times the level they were five years ago. However, neither WTI nor Brent are representative of the average crude oil, and they do not play the role of marginal supply equalizing global demand and supply.

The quality differentials between light and sweet crude oils, as both WTI and Brent are, and the heavier and sourer streams that constitute the bulk of global supplies – certainly of *marginal* global supplies – has considerably widened. Hence on average the price of crude oil has not increased as much as the price of WTI and Brent.

Secondly, we should distinguish between nominal and real prices – a price of \$60 per barrel in today’s dollars remains below the peak historical price of \$32 per barrel in 1980 dollars. In addition, for quite some time the dollar has been depreciating relative to other currencies, notably the euro: it is only in recent weeks – at the time of writing – that this tendency has reversed. The European consumer is therefore feeling the added burden of the depreciation of the euro (meaning that effectively in recent weeks the price of crude oil to the average global consumer has increased more than the nominal price in dollars would indicate).

Finally, in Europe and Japan taxation creates a huge wedge between the price of crude oil and the price of products to the consumer: for every euro spent on gasoline, 75 cents go to the taxman and only 25 pay for the product.

The huge incidence of taxation has two unwelcome consequences. Firstly, it “buffers” the consumer from proportionate product price increases, as the tax is

at least in part computed as a fixed-sum excise rather than as ad-valorem proportional to the industrial price of the product. Secondly, and possibly more importantly, it indicates that the final consumer is willing to pay much more than the industrial price of the product – meaning that the “market value” of the product is well above the cost of buying crude oil at the current high prices, and refining it. The high level of taxation, in other words, is a constant reminder to OPEC that consumers are willing to pay a lot more for the convenience of using fossil fuels.

Nevertheless, the private car remains the cheapest mode of transport - even in Europe, where it is most heavily taxed. It is normally cheaper than regularly scheduled air transport, and even in many cases cheaper than travel by train, notwithstanding the fact that whoever travels by car pays a lot of taxes and supports the government in the process, while generally public transport passengers are net beneficiaries of government subsidies.

A lot has been said to decry the American habit of driving SUVs, but the latter have become increasingly popular in Europe as well, notwithstanding the much higher gasoline and diesel prices. We have a technology that can provide the same comfort and driving experience with much reduced consumption: the hybrid engine. But do we see a dash to buy hybrid cars, or more fuel efficient ones? One can hardly detect a trend. In short, the consumer may complain about the increasing price of fuel, but he/she is hardly changing his/her habits or preferences.

It is possible that we cannot yet see a demand response because it will react with some delay. If the demand response is retarded, it is possible that prices may overshoot the “equilibrium” level. Yet changes in demand patterns take time and are visible at the margin well before they can make a huge difference on the average. Hence, producers will be able to monitor the symptoms of a change in demand behaviour well before it bites: they just need to be flexible and prudent – as they failed to be in the early 1980s.

In this sense, producers have an interest in price “stability”, i.e. in avoiding sudden and large increases in prices, which might trigger an unpredictable delayed response from the system – demand and supply. In this sense, prices have increased too rapidly for comfort in the past five years, and producers should strive, as they are doing, to contain or even somewhat reverse the increase. A period of, say, five years of price stability is required before we can say with certainty what the response to higher prices might be. It is however not clear that the market will allow for that.

SHOULD PRODUCERS MAXIMIZE PRODUCTION?

A further point, which needs to be discussed, is whether it is in the best interest of the producing countries to produce more, i.e. draw down their available resources more rapidly in order to accommodate rising global demand.

It has become fashionable to repeat that the Stone Age did not end because of a dearth of stones, or the coal age because of the exhaustion of coal deposits. What these statements imply is that the oil age might end because oil is made

irrelevant by some fundamental discovery, which will provide an alternative source of energy, cheaper and more convenient.

However, no such leap is anywhere in sight, and scientists called to discuss the potential for it have concluded that oil will remain relevant for the foreseeable future. Any fundamental discovery would require long lead times before it becomes commercial, and even longer lead times before its use becomes widespread. By the way, the age of coal has in fact not ended at all: coal remains very much in use to this day.

The opposite preoccupation – that oil resources might be exploited too rapidly, before the country possessing them has had a serious chance to develop and diversify its economy – is much more serious. Development is not bought in brightly coloured packages from the supermarket shelves: it is increasingly associated with the accumulation of human capital, and the latter requires generations to be achieved.

The official Saudi strategy presently envisages reaching a capacity of 15 million barrels per day and maintaining it for 50 years: should we conclude that this is too conservative? I do not think so. In the end, it is a question of whether the producing country would have good use to make of the additional money in case it produced more – assuming that it would increase its revenue by increasing production, which is far from certain.

The issue of what to do with the money should not be belittled, and is one that the importing countries must recognize if they wish to encourage the major exporters to produce more. An increase in oil revenue is a potent temptation to also increase expenditure: however, the danger that by so doing the government will set off the infamous “Dutch disease”² and damage diversification is very real indeed.

In the 1970s and earlier, governments massively invested in infrastructure and in industry. Recent research confirms that such investment, if well directed, may serve to cure or prevent the Dutch disease. Due to the weakness of the private sector, governments felt they were required to step in directly.

However, development requires a strong private sector. In fact, private business enterprises have grown much more capable in the past 30 years in many major producing countries (though certainly not in all) and today the government must refrain from decisions which might “crowd out” the private sector – not from the sources of financing, but from profitable investment opportunities. Private enterprises are constantly claiming for new investment opportunities and complaining that the government is subtracting opportunities from them. In infrastructure, industry, and public services, governments are advised to leave investment to private enterprise and to privatize companies that were established by the state. Hence, governments in turn find fewer opportunities in which it is justified for them to step in and invest directly.

Accumulating financial assets in order to invest internationally – that is, creating a “reserve” or “stabilization” fund along the lines adopted by a growing number of producing countries and advocated by the World Bank and IMF – is a

possibility, but raises the question of assets management. Unfortunately, money can also be lost in the management of these funds: a consideration that points in the direction of very conservative investment, i.e. primarily government paper from the major industrial countries – the US first and foremost. But to what extent does it make sense to extract more oil and increase revenue only to lend it to the US government to finance the federal budget deficit? One is left to wonder.

The stereotype, according to which the governments of the producing countries need larger and larger oil revenues in order to satisfy the expectations of their rapidly growing populations, implicitly assumes that these countries are unable to diversify their economies or spend wisely. In many cases, it is a very poor approximation of reality.

ADD/OPTIMIZE VALUE

Rather than necessarily maximizing crude oil production, the major exporting countries should focus their attention on adding value to crude oil and optimizing value creation downstream. Refining returns have become much better in recent times. Indeed, the shortage of conversion capacity globally is considered to be one of the causes of higher Brent and WTI prices and is also one of the reasons why increasing the supply of heavier and sour crude oils does not necessarily help easing them. Producers are, therefore, announcing ambitious plans to invest in refining and petrochemicals, and they are very wise to do so.

The development of refineries and petrochemical plants in huge industrial estates in conjunction with power generation, water desalination, aluminum smelting and other energy-intensive industries (including gas liquefaction or transformation into liquid fuels, and the production of hydrogen for export, whenever a demand for it will emerge) will create conditions ideal for the capture and sequestration of CO₂ in the same oil and gas fields. The technology for doing this exists – it just needs to be combined and applied on a scale that is unheard of for the time being. It is a direction in which the producing countries should massively invest rather than just sinking more wells in the ground.

We should, in other words, possibly shift our focus from the availability of crude oil to the availability of the required final products – fuel and non-fuel. The required fuel products may be new, as in the case of hydrogen, or old – gasoline, kerosene, diesel etc. They may be produced out of crude oil distillation and conversion, or out of coal, gas or agriculture. We shall need to revise our vocabulary: I deliberately avoid using the term “petroleum products,” because they may not come out of petroleum at all.

In a world in which the required fuels are produced from a variety of sources through a variety of technologies – all well established and mostly commercial at current prices – we may progressively see an increase in the relative importance of final products, and a decline of the weight of crude oil in international trade. This, I would suggest, would be a welcome development.

At present, the price of crude oil is the major factor determining the price of products. However the “fundamental” demand is for the latter: crude oil prices

are determined by a market structure whose defects have been underlined and extensively documented, and which is affected more by financials than by fundamentals. A world in which the production of fuels is more diversified, and trade in fuels is more important, may be one in which final demand plays more of a role, and the price of crude oil is more responsive to market conditions for fuels, rather than vice versa.

NOTES

¹ Each dollar of US GDP requires close to 11 thousand BTU, against 7 thousand BTU for each dollar of Western European GDP and 5 thousand BTU for each dollar of Japanese GDP; all figures are taken from the EIA's International Energy Outlook 2004 and refer to 2001.

² This term, coined in 1977 by the Economist to describe the decline of the manufacturing sector in the Netherlands after the discovery of natural gas in the 1960s, is a concept that tries to explain the apparent relationship between the exploitation of natural resources and a decline in the competitiveness of the manufacturing industry.

Sticks Instead of Carrots: Energy Leverage in Contemporary Russian Foreign Policy Toward Ukraine

Ryan R. Miller

This paper explores the long-standing energy tensions between Russia and Ukraine, examining developments in the politico-diplomatic role of Russian oil and gas pipelines in order to evaluate their effectiveness as instruments of intimidation. As the analysis will highlight, when Russia has used its energy exports as a “stick” with the goal of blackmailing leaders in Kiev into agreeing to Moscow’s foreign policy objectives, its leverage has been limited. Conversely, in the cases where Russia has used its energy exports as a “carrot,” Moscow has actually seen returns for its efforts. Thus, the Russo-Ukrainian energy relationship is interesting because it calls into question traditional perceptions of relative power.

INTRODUCTION

In today’s global political economy, power comes in many forms, and is not simply a function of a nation’s armed forces. No country truly wields power and influence without economic instruments of statecraft. Realists have traditionally held that any diplomacy not linked to the potential use of military force is effectively sterile. In the same vein, one could argue that diplomacy detached from the potential extension (or withholding) of commerce will be plagued by its own impotence. One area where the importance of such commercial power becomes particularly evident is the field of international energy. National economies need fuel in order to grow, and populations without access to electricity can be condemned to cold and darkness. Consequently, energy-exporting countries are well-positioned, at least in theory, to demand that their views and preferences be taken into account by their customers.

Ryan R. Miller is a research assistant and M.A. candidate at SAIS in Washington. Ryan is pursuing a concentration in European studies and his research interests include transition economies, Russian foreign policy, and international energy security. He holds a B.A. in international relations and French from Juniata College in Huntingdon, Pennsylvania, and has studied at Sciences Po in Strasbourg, France, as well as the St. Petersburg State Technical University in St. Petersburg, Russia.

At present, the world is witnessing the emergence of a global energy tsar, the Russian Federation. While Russia still claims to be a key player in traditional military terms, its great power status in the coming decades will be measured as much (if not more) by its hydrocarbon deposits and pipeline networks as by its nuclear warheads.¹ Will Moscow use this power responsibly in the coming decades? Only time will tell, but the Russian track record is anything but clean.

As the year 2006 began, a gas dispute between Ukraine and Russia grabbed international headlines, throwing Russo-Ukrainian energy tensions back in the spotlight. The two countries' energy relationship is especially interesting because it calls into question traditional perceptions of relative power. In point of fact, Ukraine, the weaker of the two, is actually the stronger. Indeed, the case of Russia's energy leverage over Ukraine in the post-Cold War era provides an example of the limitations of energy as a foreign policy instrument.

During the Soviet period, Ukraine depended enormously on heavily subsidized oil and natural gas from Russia. After the USSR's collapse, Ukraine's dependence on Russian energy nonetheless persisted as domestic production has never been close to meeting demand. In 1991, Ukraine's energy dependency was at its height, when Moscow provided 88 percent of the country's oil and close to 100 percent of its natural gas.² Over the past fifteen years, this dependence has only slightly diminished.³ In principle, therefore, it might seem that Moscow has been in a very favorable position to use energy as a "stick" (either by unilaterally raising the price of oil and gas or by withholding supplies) to force political concessions from the Ukrainian government.

So how effective is energy as a Russian tool of diplomatic coercion over Ukraine? To answer this question, one must look at the historical record. This paper thus examines both past and current developments in the politico-diplomatic role of Russian oil and gas pipelines in order to evaluate their effectiveness as instruments of intimidation. As the present analysis will highlight, when Russia has used its energy exports as a "stick," with the goal of blackmailing leaders in Kiev into agreeing to Moscow's foreign policy objectives, it possessed limited leverage. The reason is that such bullying often caused a political backlash in Ukraine, and that the latter took advantage of its position as the gas bridge to Europe to undermine the utility of Russia's stick by raising transit fees, or siphoning off gas destined for Russia's European customers. Conversely, in the cases where Russia has used its energy exports as a "carrot," Moscow has actually seen returns for its efforts.

This paper has four sections. The first surveys Russia's leading foreign policy objectives with regard to Ukraine since 1993. The second provides an historical account of the occasions when Moscow attempted to use the interruption of energy supplies to bully Ukrainian leaders, highlighting the limitations of this energy lever. The third describes the "carrot" approach of the Putin administration, and how such a strategy paid off. The concluding section offers some perspectives in light of Ukraine's recent "orange revolution" and the most recent gas dispute that occurred last winter.

RUSSIAN ENDS: PRESERVING INFLUENCE IN THE FORMER SOVIET EMPIRE

By 1993, Russian foreign policy assumed a more aggressive and nationalistic character.⁴ Although cooperation with the West continued on multiple fronts, Moscow elites began to perceive that Russia's best hope for maintaining itself as an important player in the international system would be to solidify its sphere of economic, political, and military influence in the post-Soviet space.⁵ The main vehicle for realizing this scheme was the Commonwealth of Independent States (CIS) – the one international organization Russia dominated, and the one available framework for manipulating events in the country's "near abroad." In fact, many elites had early on imagined, and even preferred, that the CIS serve to re-establish structures similar to the make-up of the Soviet Union.⁶ By enjoying exclusive control within the former Soviet space, Russia potentially could once again be considered a great power with influence beyond its borders, and so distinguish itself from other powers such as China or India.⁷

More so than any other former Soviet republic, Ukraine has been central for the realization of this neo-imperialist dream. Ukraine enjoys a geostrategic position on the Black Sea, the second largest population of the former Soviet states, and large concentrations of heavy industry – particularly in the military-industrial sphere – in the Eastern part of the country. Another, more psychological factor, is that paternalistic "slavophiles" in Russia typically view Ukraine as a younger brother with a leading role in the early formation of the Russian state. From this perspective, as some commentators have noted, "history requires that Russia's relations with Ukraine be qualitatively different from its relations with other foreign countries."⁸

If Russia could succeed in "reintegrating" Ukraine within the CIS, Kremlin elites thought, this would go a long way toward reconstituting the Soviet empire in the form of a modern-day sphere of influence. This did not mean that Ukraine would have been subjugated directly, in the traditional sense. Instead, as Taras Kuzio explains, the Russian neo-imperialism of this period has aimed "to penetrate [CIS states] economically and thereby gain political influence with the help of forward military bases."⁹ The language of an early Russian draft of a Russo-Ukrainian "friendship treaty" confirms this line of thinking. In the document, the Kremlin envisioned a unified defense policy, Russian military installations in Crimea (where the Russian Black Sea Fleet remained stationed) and throughout Ukraine, as well as economic integration.¹⁰

But Ukraine has been opposed to CIS integration if this means subservience to Russia along a Belarusian model. In contrast to Moscow's pretentious vision of the organization as a vehicle for institutionalizing Russian hegemony, Kiev has largely viewed the CIS as a necessary short-term instrument to aid Ukraine's economic transition.¹¹ Specifically, Ukraine has needed both raw materials, especially energy inputs, and foreign markets for its export sectors. Russia and the CIS were (and still are) the only realistic providers of both.

What Ukraine has been unwilling to do is surrender its autonomy or neutrality to Moscow under the guise of regional integration. Consequently, Ukraine only agreed to become an “associate member” of the CIS Economic Union, making it the only state with such half-hearted status.¹² Furthermore, Ukraine voiced its fierce opposition to Russian proposals for united CIS security and defense structures, or even anything that might lead down such a road.¹³

This divergence of visions concerning Ukraine’s future – one of reintegration and subordination, the other of autonomy and equality – helped to lay the groundwork for a series of political clashes during the 1990s. On several occasions, Russia’s energy supplies entered the diplomatic fray. Janusz Bugajski notes how, at critical junctures, state-owned energy companies like Gazprom act as arms of the Russian foreign policy apparatus by manipulating energy supply, price, and credit for political blackmail.¹⁴ Some Russian politicians have openly encouraged the Kremlin to use Gazprom in this way. In 1992, Vice Premier Mikhail Polotoranin declared that “Russian foreign policy vis-à-vis the former Soviet republics must be conducted through the use of such levers as natural gas and oil.”¹⁵ His wish was soon granted. As it turned out, Moscow would largely decide to use energy as a stick, rather than a carrot, in order to pressure its Slavic brother.

RUSSIAN MEANS: ENERGY AS A STICK

As Moscow became increasingly frustrated with Ukraine’s resistance to reintegration, the energy lever was brought into play on a number of occasions to try to force Ukraine to pay for its oil and gas with political concessions. In January 1993, some of the main bones of contention in Russo-Ukrainian relations included Ukraine’s nuclear status, the future of the Black Sea fleet, and Ukrainian sovereignty in Crimea. Agreements to resolve these issues faltered, tensions grew, and what started as political disputes manifested themselves in fisticuffs of economic warfare. Still, on balance Ukraine emerged from these ordeals bruised but in a better position than Russia in many respects.

1993: THE OPENING SHOTS

In February, Moscow disregarded earlier promises to keep the cheap energy flowing and declared that henceforth it would charge Ukraine the average world price for natural gas. Leaders in Kiev complained that Moscow was pressuring them with the reduction in oil supplies in order to punish them for Ukraine’s position on political issues. In making their case, they pointed to the fact that Belarus, behaving like an obedient vassal state, was charged only half the price that Ukraine was paying for Russian gas and petroleum.¹⁶

But Russia’s energy leverage was not what many in the Kremlin would have hoped, as Ukraine did not take this abuse lying down. The fact was that Ukraine served as the main transit route (via a vast pipeline system) for Russian energy to the lucrative markets in Western and Central Europe. Indeed, up until the late 1990s, about 90 percent of Russian oil and most of its gas exported to Europe traveled through pipelines crossing Ukrainian territory.¹⁷ Taking advantage of this situation, Kiev retaliated by increasing the transit fees for gas. Ukrainian

transit for 1000 cubic meters of gas per 100km rose to \$3.00, more than double the average world transit price.¹⁸ Moscow objected, but there was little that it could do – it relied on Ukraine for the continued flow of its energy exports to its hard currency customers. The President of Gazprom threatened to stop pumping gas to Ukraine, but nothing ever came of this warning.¹⁹ All the while, Ukraine's oil and gas supplies continued to be bought on credit, and Ukraine's debts to Russia only grew. These ballooning debts should be seen not only as a sign of Ukraine's weakness, but also as an indication of its strength – Ukraine continued to consume Russia's energy without paying, while constraining Russia's ability to retaliate.

SUMMER 1993: ENERGY AND THE BLACK SEA FLEET

In the meantime, Moscow continued to assert its position on long-standing points of disagreements with Kiev. In June, Russian Foreign Minister Andrei Kozyrev, when visiting the Crimean Peninsula, proclaimed his support for Russo-Ukrainian reunification.²⁰ Moreover, elites in the Moscow foreign policy establishment argued that Russia should acquire “unlimited access to Ukraine's market of goods, services and capital, and the creation of an effective political-military alliance.”²¹ Ukrainian President Leonid Kravchuk however did not accede to Moscow's demands, and insisted upon a Russo-Ukrainian rapport based on equality and mutual recognition of each country's sovereignty.

One month later, Moscow started to restrict further the supply of its gas to Ukraine, and at the same time began to bring its oil prices nearer to world levels.²² The consequence of this move was that Ukraine incurred massive debts that it could not afford to service. Prime Minister Leonid Kuchma thus flew to Moscow to sign a protocol establishing “uninterrupted” energy deliveries from Russia to Ukraine until at least January 1, 1994.²³ But this promise was less than solid; Russia in August decided to further curtail deliveries because of the debt issue.²⁴ For a period of one month, Russia even shut off the gas and oil completely in response to non-payment of these debts.²⁵

To be sure, Ukraine was pushed into a corner, but Kiev nonetheless found room for maneuver. The Ukrainians responded by seizing for itself huge amounts of Russian gas passing through their territory to Europe. According to Gazprom, the company had in January failed to deliver 11 million cubic meters of gas to its European customers, and at least 30 percent of this missing gas was explained by stealing on the part of the Ukrainian authorities.²⁶ Moreover, Moscow's supply cuts caused sudden shortages in Central European markets. This experience taught the Russian authorities that they cannot shut off Ukraine's oil and gas supplies without disrupting the flow of energy to their major clients.²⁷ Overall, Russia's stick was ineffective.

In the autumn, Yeltsin and Kravchuk met at Massandra to discuss the status of the Black Sea Fleet. Prior to this meeting, both sides had informally agreed to divide the fleet equally. However, the Russian position had changed in the interim, and representatives from Moscow began to insist that Kiev hand over part of its fleet in exchange for partial settlement of Ukraine's energy debts. Yeltsin warned that if Kravchuk refused to accept these terms, then Ukraine's much-needed fuel would be cut off.²⁸ The Ukrainian President was unable to offer firm resistance at

the time, and left Massandra apparently having acquiesced to Moscow's demands.

But at this point Moscow began to encounter yet another complication with the use of its energy lever: Ukrainian nationalism. Anatol Lieven observes that many citizens in Southern and Eastern Ukraine typically distrust Ukrainian nationalists, "but if the Russian state tries to push its own interests at the expense of Ukraine, then even these 'pro-Russians' find themselves firmly on the Ukrainian side."²⁹ Indeed, a number of ethnic Russian and typically pro-Russian figures took up the banner of patriotism and voiced their opposition to the deal at Massandra.³⁰ Needless to say, the rest of the Ukrainian public was also furious at Kravchuk for selling out. In the face of such popular hostility, President Kravchuk renounced his agreement with Yeltsin, and the issue thus remained unresolved. Russia's energy threats were once again of only limited utility, and may have been counter-productive in the long-run when judged against the country's foreign policy objectives.

1993-1994: THE WINTER OF SELF-RESTRAINT

During the winter of 1993-1994, Moscow used its energy lever selectively – with the goal of trying to influence Ukrainian electoral politics. At this time, the cold weather sent Ukraine's energy consumption through the roof, causing its debts to worsen, and Kiev's delinquencies on payments became even more pronounced. But the Ukrainian presidential elections were just a few months away, and Moscow did not want to weaken the hand of the "pro-Russia" candidate, Kuchma.³¹ Although Kuchma won due to a variety of factors, it is reasonable to suppose that an energy supply shock in the middle of winter could have seriously hurt his popularity, given that he was campaigning in part on a platform of cooperation with Russia.

So as not to undercut Kuchma's hand, the Russian authorities abstained from withholding any gas or oil to a large degree. Ukrainian representatives reached an agreement with Russian Prime Minister Victor Chernomyrdin on future gas and oil deliveries, with an agreed amount and lower price for deliveries to be made in 1994.³² Moscow even decided not to intervene as Kiev siphoned off gas from the Druzhba pipelines at a rate of about 40 million cubic meters per day.³³ As it happened, Kuchma did win the elections that summer, but he did not prove to be as "pro-Russian" as some in the Kremlin hoped.

1994-1999: SUBTLE SKIRMISHES

From late 1994 to 1999 (Kuchma's first term), there was a general lull in Russo-Ukrainian tensions, as Moscow became distracted by other concerns.³⁴ Relations were still far from cordial, but Russia did not immediately link energy with outstanding Crimean issues, and energy debts were rescheduled in 1995. Another factor helping Ukraine was that Kuchma began to cultivate ties with the West, gaining in the process a stronger hand in dealing with Moscow.³⁵ An important consequence of Kuchma's improved position was the fact that Russia began to acquiesce to Ukrainian independence from CIS structures. Kuchma declined participation in the customs union, as well as joint guarding of borders and full membership in the economic union.³⁶ In the end, Moscow had to settle

for a watered-down, “multi-speed” CIS, with only Russia, Belarus, and Kazakhstan agreeing in 1995 to form a customs union.³⁷

To be sure, Russia continued to pressure Ukraine with energy in more subtle ways. To begin, Gazprom’s strategy during this period focused on acquiring outright ownership of Ukraine’s energy facilities.³⁸ Russia also used an excise duty on its energy products, forcing Ukraine to pay higher than world prices. The goal of this action was to pressure Kiev into joining the CIS customs union. The Yeltsin administration pledged to leave the tax in force until Ukraine joined the common economic structures.³⁹ But Ukraine found a way to hit back. Up until 1996, there was a double-standard concerning pipelines. Russia delivered oil to Ukraine at world prices, while the fees paid to Ukraine for the transit of oil were based instead on bilateral agreements between Moscow and Kiev.⁴⁰ This arrangement started to change after Ukraine, demanding higher transit fees, closed its portion of the Druzhba pipelines for a few hours in January 1996. This action stopped the flow of oil to Central and Western Europe temporarily, but long enough to make Russia grudgingly accept *de facto* higher transit fees (closer to world prices) by the end of the year.⁴¹ Again, leaders in Kiev could act unilaterally and disrupt Russia’s energy stick because of Ukraine’s geoeconomic position.⁴²

One of the first instances where energy was effectively used as a carrot was the 1997 Friendship and Cooperation Treaty. According to the treaty’s provisions, Kiev received for the first time Russian recognition of Ukraine as a separate political entity, but the text also reflected the influence of a Russian energy carrot.⁴³ Indeed, Moscow was able to negotiate a twenty-year lease on Sevastopol as a naval outpost in exchange for a partial write-off of Ukraine’s now massive energy debt.⁴⁴ Yaroslav Bilinsky views this outcome in a sinister light, noting that this linkage of energy to a military presence mirrored in some respects Russia’s tactics for securing basing rights in Belarus.⁴⁵ Whatever the merits of such a parallel, the treaty nonetheless represented an example where Russia maintained its long-standing goal of continued Sevastopol presence, due in large part to the incentive of debt relief – a carrot – rather than the threat of withholding vital supplies.

1999-2000: ANOTHER OPEN CONFRONTATION

The return of the stick came in late 1999, for both political and economic reasons. Russia’s stated problem was that Ukraine’s transit fees remained relatively high at \$1.09 per 1,000 cubic meters per 100km (Belarus’s fees were \$0.55 for the same service).⁴⁶ Furthermore, the 1997 treaty had stipulated that Ukraine could not impede the flow through its territory of Russian goods (i.e. oil or gas) or raise transit tariffs unilaterally.⁴⁷ But the reality was that Ukraine continued to make a mockery of Russia’s energy monopoly and the restrictions imposed by the treaty by siphoning off Russian energy from the vast pipeline network.⁴⁸ This theft persisted while Ukraine continued to make purchases of Russian oil and gas on credit. Ukraine’s “new” energy debts to Russia accumulated after 1997 (in addition to those restructured previously) amounted to \$1.8 billion by 1999.⁴⁹ Nevertheless, political considerations also played a major role in the 1999-2000 dispute. In particular, Gazprom wanted to use these debts as leverage for maneuvering into a dominant position of the Ukrainian energy sector.⁵⁰ Moreover, Kasyanov hinted

that Ukraine would find it easier to avoid such gas difficulties in the future if it joined the CIS customs union.⁵¹

Therefore, acting Prime Minister Kasyanov declared, as winter was coming, that there would be no further oil or gas deliveries until Ukraine entirely halted diversion of Russian exports. This came at a time when Turkmenistan had earlier in the year stopped gas deliveries to Ukraine due to lack of payment.⁵² According to some calculations, Ukrainian theft of Russian natural gas equaled 185 million cubic meters per day.⁵³ Ukraine's authorities cracked down somewhat on their illegal theft, but Russia continued to withhold energy supplies and restated its earlier warnings.

But in the end, Moscow could not mimic Turkmenistan's example of cutting off the gas completely because of fears that Russia would lose important credits in its balance of payments, and possibly cause further instability and anti-Russian sentiment in Ukraine.⁵⁴ Eventually, Russia lifted its oil blockade of Ukraine in February 2000.⁵⁵ To compensate for the lost supplies, Ukraine also was able to siphon off illegally approximately 13 billion cubic meters of Russian gas in the first five months of 2000.⁵⁶ In a nutshell, Moscow suffered yet another humiliation. The difference seemingly was that Russia's new leader, Vladimir Putin, understood the limitations of his country's traditional bullying tactics.

THE WITHDRAWAL OF THE STICK AND THE EMERGENCE OF A CARROT

It seems that in order to move away from a strategy of bullying Ukraine, Putin needed some sort of diplomatic victory in order to save face. As such, energy deliveries resumed only after a summit in April between Putin and Kuchma, and only after a second summit in August was a mechanism put in place to oversee the resolving of Ukraine's debt problem.⁵⁷ Ukraine made a minor concession to Russia by joining one of the CIS's new security structures, the anti-terrorism center.⁵⁸

With this face-saving settlement in place, Putin gradually changed Russia's strategy for influencing Ukraine's geopolitical course. James Sherr describes the shift:

Putin has been more skeptical than Yeltsin toward the traditional Soviet concept of "integration"... He has intensified covert means of influence, placed greater emphasis on internal Ukrainian allies, and has endeavored to turn weakness, division, and even scandal inside Ukraine to Russia's geopolitical advantage.⁵⁹

The key move was from a "coercive" to "covert" means of manipulating the country's external relations. For instance, such subterfuge previously involved a policy of using spy craft and media propaganda to marginalize Ukrainian nationalists. Ukrainian Prime Minister Victor Yushchenko, appointed in late 1999, was kicked out two years later with the help of what some suspect was Russian influence.⁶⁰ Some also speculate that Ukrainian authorities launched their judicial proceedings against then Deputy Prime Minister Yulia Tymoshenko in consultation with Moscow.⁶¹ Both political figures were wary of Russian influence. While targeting would-be adversaries, Putin simultaneously sought to cultivate good

relations with potential allies. Most notably, he succeeded in bringing about a rapprochement with President Kuchma. In the 2002 parliamentary elections, Putin's Ambassador to Kiev openly backed the pro-Kuchma parties.⁶² To further warm up with Kuchma, Putin apparently arranged for his Ukrainian ally's election as chairman of the Council of CIS Heads of State in January 2003.⁶³ Later, Putin established close relations with Prime Minister Viktor Yanukovich.⁶⁴

It is worth noting that this turn toward Russia has not come about because of petro-bullying. For all of its huff and puff, Moscow did not succeed in pressuring Ukraine into being a Russian satellite by withholding energy supplies. If anything, such economic pressure only strengthened the hand of the Ukrainian nationalists such as Yushchenko, and provoked Ukrainian disaffection with Russia (even among ethnic Russians living in Ukraine and those in favor of close ties with Moscow), making a mockery of Russian aspirations to lead any pan-Slavic confederation. Instead, Russian tactics in the battle for influence over Ukraine's position in the post-Cold War shifted away from curtailing oil and gas supplies to battles within Ukrainian politics itself. Putin's move away from a stick approach is consistent with his general prioritization of economics in foreign policy, and it is conceivable that Putin, from 2000-2004, realized the negative consequences for Russia of cutting off energy to Ukraine.

With the stick withdrawn and a policy in place that emphasized building relationships with select Ukrainian politicians, Putin began to put his country's energy exports to use as a carrot. To breathe new life into CIS integration, Putin organized a draft treaty for a planned Single Economic Space (SES) in August 2003. The plan envisioned a free trade area, a joint economic policy, a customs union, coordinated accession to the WTO, and a central supranational commission with the power to impose decisions on member states.⁶⁵ In exchange for Ukraine's tentative acceptance, Putin promised his carrot, cheaper Russian energy.⁶⁶ Such actions fit within the context of an intended softer approach to the near abroad. In July 2004, he stressed at a meeting of Russian ambassadors that Russia should avoid calling for exclusive Russian leadership over fellow CIS members, noting that Russia had not yet cultivated "sufficiently well the historical credit of trust and friendship, the close ties that link the peoples of [the CIS] countries."⁶⁷

The way in which Russian energy was to be made cheaper was through the repeal of the VAT on Russian oil and gas exports to Ukraine and other CIS states. In September 2004, Putin gave Kuchma a gift: a donation of \$800 million in the form of such an exemption.⁶⁸ This VAT repeal was part of a basket of carrots that Moscow offered, one that included assistance to the electoral campaign of Kuchma's preferred successor, Yanukovich. Indeed, as the Ukrainian presidential elections neared, Putin seized his opportunity to shape policy in Kiev. Gazprom supplied a third of Yanukovich's election funds, Russian television ("managed" by the Kremlin) campaigned heavily in Yanukovich's favor, and the Russian authorities opened 400 polling stations for Ukrainian expatriates living in Russia (most of whom are ethnic Russian and were thus more likely to vote for Putin's favorite).⁶⁹

What did Moscow get in return? The Russians gained geopolitically in a number of areas. First, Russia gained Kuchma's signature to the SES agreement.⁷⁰ Second, according to some observers, Yanukovich promised Putin that he would

end Ukraine's flirtation with NATO, promote an open border and dual citizenship regimes for Russians and Ukrainians, and consider elevating the legal status of the Russian language in Ukraine.⁷¹

The third area where Moscow saw returns pertained to the direction of the flow of oil within Odessa-Brody pipeline. In 2003, the Russians had proposed that the recently constructed pipeline be used in the reverse direction – from Brody to Odessa – in order to provide another means for transporting Russian energy to the Black Sea, and from there to the world market through the Turkish straits.⁷² Early on, the Ukrainian government rejected the idea, realizing the long-term benefits of diversifying its energy supply sources and of becoming a transit point for Caspian Sea oil exported to Europe.⁷³ But as the pipeline went unused, due in part to diminished outside interest in the project, Kuchma began to look East for other partners to provide the needed oil.

In July 2004, Prime Minister Viktor Yanukovich approved the reverse use of the pipeline in favor of Russian companies. Putin's strategy of maintaining allies in Ukraine on an individual level – with Kuchma and Yanukovich – appeared to have paid off. Kuchma announced the long-term use of the pipeline for Russian oil in August during a summit with Putin in Sochi. The next month, as highlighted above, the Russians removed the VAT to sweeten the deal.⁷⁴ Thus, the energy card had worked when it had been played properly.⁷⁵

BACK TO THE FUTURE? THE ORANGE REVOLUTION AND THE REVENGE OF THE STICK

As has been shown, oil and gas have primarily only been an effective means toward enhancing Russia's influence in the former Soviet space when these strategic products were not used in a coercive manner. The reasons for this was that any coercive approach strengthened anti-Moscow sentiments in Ukraine, and encouraged leaders in Kiev to take advantage of their country's position to cause Moscow elites significant headaches. Ironically, Moscow hoped that its pipelines would serve to keep Ukraine integrated within Russia's geopolitical orbit, but when Russia tried to use these links to bully its younger brother, they became a means for Kiev to express its independence from Russian influence. Only when Moscow spoke softly and acted in a positive manner – using its “soft power,” as Fiona Hill describes – has energy been effective from a politico-diplomatic perspective.⁷⁶

Apparently, Kremlin leaders today still fail to understand this point. Judging from the events earlier this year, some in Moscow are smarting from the “orange revolution,” and want to see how far they can push the envelope in their relations with Yushchenko's Ukraine. Gazprom briefly shut off the gas to Ukraine because of a pricing dispute, but once again Ukraine emerged victorious. Ukraine had been receiving Russian gas at a price of \$55 per 1,000 cubic meters at a time when most of Gazprom's customers had been paying over \$200 or had made arrangements to gradually bring their gas prices closer to market levels. Under such conditions, it was inevitable that Ukraine would have to pay more. The

question was how much, and under what circumstances. Gazprom declared its intention to quadruple Ukraine's price, bringing the price of gas to \$220. Kiev wanted any price hike to be phased in over a matter of years.

On New Year's Day, Gazprom technicians turned down the flow of natural gas to Ukraine. Though the conflict was essentially a commercial one, there were arguably political considerations in the background. Putin, some experts say, sought to punish Ukraine for its aspirations for EU and NATO membership and demonstrate to Yushchenko that life outside of Moscow's orbit can be cold and dark.⁷⁷ This is perhaps the worst tactic Russia could have used. As the *Financial Times* highlighted last January, "like a nuclear deterrent, [energy supplies] should never be used to punish."⁷⁸ Some observers also surmised that Putin was trying to use its energy leverage to influence the Ukrainian parliamentary elections in March, but the move seemingly only strengthened the hand of Yushchenko.⁷⁹

But Ukraine had warned that it would retaliate by quadrupling transit fees, and also threatened to confiscate 15 percent of the gas destined for Europe in lieu of transit fees.⁸⁰ When Russia finally did shut off the gas, shortages were quickly reported in European countries.⁸¹ At the time of this writing, Ukraine still maintains that it did not siphon off supplies. Instead, Kiev insisted, Russia caused the shortages in Europe by reducing pressure in the pipelines.⁸² Regardless of who caused Europe's gas crisis, the fact remains that Ukraine's position as Gazprom's chief transit state continues to saddle Russia with an artificial inferiority. Only a day after turning off the gas, Russia had to back down and return pressure to Ukraine after scaring the daylight out of its European gas customers.

Moreover, the gas settlement reached on January 4 should be considered on balance a victory for Ukraine. Through a complicated gas-pooling arrangement with a Swiss-registered Gazprom affiliate, Ukraine will buy a combination of Russian and Central Asian gas at \$95 per 1,000 cubic meters. Though higher than what Ukraine previously was paying, this price is allegedly close to the target for which the Ukrainian authorities had been holding out.⁸³ Indeed, if one takes a look at the prices paid for Russian gas around Central and Eastern Europe, \$95 is a steal. Furthermore, the agreement stipulated that Ukraine may to increase its transit fees from \$1.09 to \$1.60 for each 1,000 cubic meters of gas per 100 kilometers. When this is taken into account, Ukraine is only paying a price of gas \$20 higher than before!⁸⁴ In short, Russia shot itself in the foot last January, and only confirmed that energy threats are still of limited utility for Russia in dealing with Ukraine. All in all, David has again defeated Goliath.

Looking toward the future, it is unclear whether this pattern will continue to be the norm, for is all a matter of relative leverage. On the one hand, Ukraine plans to increasingly diversify its sources of energy. Besides the Odessa-Brody pipeline, one major energy diversification project includes plans for the creation of a gas corridor linking Ukraine with Turkmenistan's and Kazakhstan's gas supplies. Yet another plan would link the Ukrainian market to the gas fields of Iran via a massive pipeline.⁸⁵ The less Ukraine depends on Russia for its energy, the less effective Russia's stick becomes, and the less likely Moscow will threaten Kiev with it. But Russia meanwhile will seek to diversify its transit routes. Specifically, Moscow appears to be moving away from the monopolized Druzhba

to a pipeline running through the Baltic.⁸⁶ There is also still the possibility of constructing a Yamal-II pipeline running from existing lines in Poland and link up with the other end of Duzhba in Slovakia, effectively bypassing Ukraine. In such a situation, Ukraine may be unable to make demands on the terms of Russian energy supply if Moscow has other pipeline options.

In summary, both the Russian and Ukrainian monopolies on energy and energy transit, respectively, are showing signs of diminishment. If this trend continues, one will speak less and less about energy “power” on either side. Assuming both processes of diversification proceed in sync and at the same speed, this scenario has the potential to be a positive development, reducing the likelihood of energy disputes and cooling tensions in the region. Alternatively, if one process pulls ahead of the other, then history will repeat itself and Ukraine’s spats with Gazprom will continue.

NOTES

¹ Russia’s proven reserves of natural gas, estimated at 1,680 trillion cubic feet, are the largest in the world. See: Keith C. Smith, *Russian Energy Politics in the Baltics, Poland, and Ukraine: A New Stealth Imperialism?* (Washington: CSIS, 2004) p. 9.

² Gregory Krasnov & Josef C. Brada, “Implicit Subsidies in Russian-Ukrainian Energy Trade,” *Europe-Asia Studies* (July 1997) p. 826.

³ Russia now supplies about 70 percent of both Ukraine’s oil and gas needs, and Kiev is heavily indebted to Russia’s energy monopolies.

⁴ See: Nichole Jackson, *Russian Foreign Policy and the CIS: Theories, debates, and actions* (London: Routledge, 2003) pp. 69-70.

⁵ *Ibid*, pp. 97-99.

⁶ See remarks by Head of Foreign Intelligence Yevgeniy Primakov in 1994. Roman Wolczuk, *Ukraine’s Foreign and Security Policy, 1991-2000* (London: Routledge, 2003) pp. 58-59.

⁷ Sherman W. Garnett, *Keystone in The Arch: Ukraine in the Emerging Security Environment of Central and Eastern Europe* (Washington: Carnegie, 1997) p. 61.

⁸ Quoted in: Paul Kubicek, “Russian Foreign Policy and the West,” *Political Science Quarterly* (Winter 1999-2000) p. 557.

⁹ Taras Kuzio, “National Identity and Foreign Policy: The East Slavic Conundrum,” in Kuzio (ed.) *Contemporary Ukraine : Dynamics of Post-Soviet Transformation* (Armonk: Sharpe, 1998) p. 229.

¹⁰ Adrian Karatnychy, “The Nearest Abroad: Russia’s Relations with Ukraine and Belarus,” in Uri Rara’an and Kate Martin (eds) *Russia: A Return to Imperialism?* (New York: St. Martin’s Press, 1995) p. 76.

¹¹ Kuzio, “National Identity,” p. 229.

¹² Rawi Abdelal, *National Purpose in the World Economy: Post-Soviet States in Comparative Perspective* (Ithaca: Cornell UP, 2001) p. 65.

¹³ Buszinski, *Russian Foreign Policy After the Cold War*, p. 105.

¹⁴ Janusz Bugajski, *Cold Peace: Russia's New Imperialism* (London: Praeger, 2004) pp. 16, 86.

¹⁵ Quoted in: Taras Kuzio, *Ukrainian Security Policy* (Washington, D.C.: CSIS, 1995) p. 71.

¹⁶ Oles M. Smolansky, "Ukraine's Quest for Independence: The Fuel Factor," *Europe-Asia Studies* (Spring, 1995) p. 75.

¹⁷ Roman Wolczuk, *Ukraine's Foreign and Security Policy, 1991-2000*, p. 42.

¹⁸ Oles M. Smolansky, "Ukraine's Quest for Independence: The Fuel Factor," *Europe-Asia Studies* (Spring, 1995) p. 75.

¹⁹ Ibid.

²⁰ Karatnychy, "The Nearest Abroad," p. 77.

²¹ Kuzio, "National Identity," p. 229.

²² Karatnychy, "The Nearest Abroad," p. 77.

²³ Smolansky, "Ukraine's Quest for Independence," p. 81.

²⁴ The monthly payments demanded by Moscow came to about \$600 million. See: Ibid, p. 82.

²⁵ Karatnychy, "The Nearest Abroad," p. 77.

²⁶ Smolansky, "Ukraine's Quest for Independence," p. 82.

²⁷ Michael S. Lelyveld, "Russia Seeking Ukraine Gas Pipelines," *Journal of Commerce* (March 12, 1998) p. 10A.

²⁸ Karatnychy, "The Nearest Abroad," p. 77.

²⁹ Anatol Lieven, *Russia and Ukraine: A Fraternal Rivalry* (Washington, D.C.: US Institute for Peace, 1999) p. 104.

³⁰ Karatnychy, "The Nearest Abroad," p. 78.

³¹ Ibid, p. 79.

³² Smolansky, "Ukraine's Quest for Independence," pp. 82-3.

³³ Ibid, p. 83.

³⁴ There were three main distractions for Moscow during the 1994-1995 period: (1) the war in Chechnya; (2) an upcoming CIS summit where leaders were to examine the Russian proposal for a customs union; and (3) the rise to power of the Republican Party in the US Congress, which began to champion NATO expansion.

³⁵ Kubicek, "Russian Foreign Policy and the West," p. 560; "Between East and West," *The Economist* (September 28, 1996) p. 58.

³⁶ One exception was a 1995 agreement for Ukraine to participate in common CIS air defense consultations. See: Olexiy Haran, "Between Russia and the West: Prospects for Ukrainian Foreign Policy Choices," *The Final Report Prepared in the Framework of 'Democratic Institution Fellowships' in Kiev* (1998) p. 35.

³⁷ Abdelal, *National Purpose in the World Economy*, p. 62.

³⁸ Bugajski, p. 86.

³⁹ Margarita Mercedes Balmaceda, "Gas, Oil, and Linkages Between Domestic and Foreign Policies," *Europe-Asia Studies* (March, 1998) p. 260.

⁴⁰ *Ibid*, pp. 268-9.

⁴¹ *Ibid*.

⁴² Matthew Kaminski, "Ukraine Exploits its Energy Pipelines Monopoly," *Financial Times* (15 March 1996) p. 2.

⁴³ Marta Dyczok, *Ukraine: Movement Without Change, Change Without Movement* (Amsterdam: Harwood, 2000) p. 121.

⁴⁴ *Ibid*, p. 127.

⁴⁵ Yaroslav Bilinsky, *Endgame in NATO's Enlargement: The Baltic States and Ukraine* (London: Praeger, 1999) p. 59.

⁴⁶ Arkady Moshes, "Russian-Ukrainian Relations After Ukraine's Elections," *PONARS Policy Memo No. 82, Institute of Europe* (October 1999) p. 1.

⁴⁷ Balmaceda, "Gas, Oil, and Linkages Between Domestic and Foreign Policies," p. 265.

⁴⁸ Gazprom did cut deliveries by 20 million cubic meters in July 1997 in a bid to force payment of overdue energy bills, but the government avoided direct meddling, so this episode should not be considered a government stick.

⁴⁹ Moshes, p. 1.

⁵⁰ Tor Bukkvoll, "Off the Cuff Politics: Explaining Russia's Lack of a Ukraine Strategy," *Europe-Asia Studies* (December 2001) p. 1144.

⁵¹ James Sherr, "Russia and Ukraine: A Geopolitical Turn?" in Lewis (ed.), *The EU and Ukraine: Neighbors, Friends, Partners?* (London: TEPSA, 2001) p. 163.

⁵² Moshes, pp. 2-3.

⁵³ There is the possibility that Gazprom siphoned off its own gas in the first place. See: Sherr, p. 162.

⁵⁴ Moshes, p. 3.

⁵⁵ "Ukraine Energy Information," found at: <www.ukrainepower.com>

⁵⁶ *Ibid*.

⁵⁷ Sherr, p. 162.

⁵⁸ *Ibid*, p. 163.

⁵⁹ *Ibid*, p. 161.

⁶⁰ Bugajski, p. 93.

⁶¹ Taras Kuzio, "Russian President Gives Ukrainian Counterpart a Helping Hand Against Opposition," *Radio Free Europe / Radio Liberty, Newslines* 6, no. 153 (15 August 2002).

⁶² Bugajski, pp. 81-82.

⁶³ *Ibid*, p. 82.

⁶⁴ Sherr, p. 167.

⁶⁵ Bugajski, p. 89.

⁶⁶ Ibid.

⁶⁷ Fiona Hill, "Energy Empire: Oil, Gas, and Russia's Revival," *The Foreign Policy Center, London* (September 2004) p. 20.

⁶⁸ "Ukraine, Russia Sign Series of Oil, Gas Accords," *Current Digest of the Post-Soviet Press* (15 September 2004).

⁶⁹ Jackson Diehl, "Putin's Unchallenged Imperialism Moves to Ukraine," *Christian Science Monitor* (October 28, 2004) p. 9.

⁷⁰ Arina Sharipova, "36 Billion for Friendship," *Gazeta, no. 4* (4 August 2004) p. 7.

⁷¹ Diehl, p. 9.

⁷² Natalya Slobodyan, "UkrTransNafta Set to Pump Crude Oil Through Odessa-Brody Pipeline," *Ukrainian Weekly* (21 November 2004) p. 12.

⁷³ Ibid.

⁷⁴ "A Blow for Caspian Exports," *Petroleum Economist* (October 2004) p. 1.

⁷⁵ It should of course be noted that Ukraine's "orange revolution" last winter dealt a major blow to Putin's successes. Since Yushchenko assumed the presidency, Ukraine has announced in March 2005 that the pipeline's flow will revert to its original direction. To ensure that the oil flows west, the EU has agreed to fund the pipeline's extension to Poland. With regards to the SES, the new government in Kiev has maintained that it does not want CIS integration to prejudice Ukraine's aspirations for EU and NATO membership. Because of Ukraine's lukewarm support for the SES, Putin appears to have resigned himself to promoting economic integration among Russia, Belarus and three of the Central Asian republics.

⁷⁶ Fiona Hill, "Russia's Newly Found Soft Power," *The Globalist* (26 August 2004).

⁷⁷ Neil Buckley, "Moscow has more than just the price of gas on its mind," *Financial Times* (3 January 2006). p.2. There is also the view that the gas conflict with Ukraine stemmed from a battle in the Kremlin for private profits.

⁷⁸ "Putin's Gas Logic Is In Short Supply," *Financial Times* (5 January 2006) p.15.

⁷⁹ Ibid.

⁸⁰ "Russia-Ukraine Gas Row Heats Up," *BBC News Online* (27 December 2005).

⁸¹ Thomas Catan et al, "Russia Gas Dispute Cuts Supply to Europe," *Financial Times* (3 January 2006) p. 1.

⁸² "Under pressure, Russia Restores Gas," *Newsday* (3 January 2006) p. A20.

⁸³ Tom Warner and Neil Buckley, "Settlement of Dispute Allows Both Putin and Yushchenko to Claim Victory," *Financial Times* (5 January 2006) p. 6.

⁸⁴ Ibid.

⁸⁵ "Ukraine Proposes a Project for Construction of Iran-Europe Gas Pipeline," *The Russian Oil and Gas Report* (July 27, 2005); "Iran-Europe gas pipeline to bypass Russia," *RIA Novosti* (26 July 2005).

⁸⁶ "Moscow and Kiev: Different Views on Ukrainian Transit," *RIA Novosti* (3 March 2005); Tom Parfitt, "Russo-German Pipeline Deal Raises tensions," *The Guardian* (8 September 2005).

Abuse of Power: The Responsibility of Government to Encourage the Development of Renewable Energy Resources

Michael Obeiter

As economies develop, they increasingly rely on energy to fuel their growth. This energy has historically come from traditional power sources – notably oil and coal – as they are relatively cheap and easily exploitable. Developed economies have derived much of their strength from the use of these energy sources, at the expense of environmental quality. These economies now have the responsibility to lead the world in advancing the utilization of renewable energy; the deregulation of the electric industry in many of these countries offers the perfect opportunity for governments to encourage the promotion of renewable energy through a variety of policies.

The last years of the 20th century brought with them a considerable move towards deregulation in the United States and liberalization in Europe. This new tendency to favor free market enterprise over government management manifested itself in many industries, including perhaps the most well-known example – the electric industry. The liberalization of the electric industry in Europe – in which the markets in Belgium, Finland, the Netherlands, Norway, Sweden, the United Kingdom, Germany, Italy, Spain and France have participated – came close on the heels of a wave of privatization of many national industries, a trend that started in England under Margaret Thatcher and eventually included the telecommunications, automobile, banking, and energy sectors of many European nations.¹

Michael Obeiter is an M.A. candidate at the Bologna Center, where he is concentrating in international environmental and energy policy. Before coming to SAIS, he was a utility analyst at Harvard University, where he was involved in decision-making processes related to the university's energy needs. He holds a B.A. in mathematics and history from Williams College.

Many US states turned to deregulation to, ostensibly, lower the cost of electricity generation for consumers through increased competition and efficiency gains (electricity transmission, a clear natural monopoly situation, remains regulated). In both the US and European examples, opening up the electricity market to competition has allowed for the entry of new firms, and in nearly every instance has indeed lowered the cost of electricity for the end-user.

In examining these newly competitive markets, one can easily overlook the effect that deregulation and liberalization have had on the renewable energy (RE) industry. An oft-marginalized segment of the energy sector under the regulatory apparatus, deregulation has threatened its further marginalization as the price of conventional electricity generation decreases.² Under state-controlled regulation or nationalization schemes, governments at the state and national level could dictate a requirement for the inclusion of RE in the mix of electricity supplied to the grid. With the advent of deregulation, however, the state loses its power to enforce such a requirement, and the utilization of renewable resources is left to the discretion of the market. Deregulation is a process in which “the institutions associated with vertically integrated, highly regulated electric utilities...are being replaced by new institutions, such as independent system operators and power pools and privately owned, distribution only companies;” there are opportunities for the government to get involved in this process and actively work to influence the supply and demand of renewables.³ As electricity markets are opened to competition, there is necessarily an influx of new producers supplying energy from easily exploitable and readily available conventional sources, notably the traditionally employed fossil fuels. This should not come as a surprise, as the infrastructure at every stage of development (including oil rigs, coal mines, and gas wells in the exploitation stage; and the tankers, trucks and trains, and pipelines in the transport and delivery stage) is already in place. All that remains for the new supplier is to site a power plant, interconnect it with the grid, and ensure delivery of the fuel of choice – admittedly difficult tasks, but unlike nearly every RE project, securing the necessary financing is not a Sisyphean struggle. The electric generation industry thus transforms from a regulated monopoly to a market structure that more closely resembles perfect or oligopolistic competition. In this manner, the price differential between RE and conventional energy – already quite large in the case of solar power and other less efficient systems – increases, as competition drives down the price of conventional energy while funding and siting difficulties keep the supply of RE systems low and prices relatively high.

In such a situation, there is a very clear market failure, as the positive externalities of RE (especially the mitigation of air pollution and the diminution of the effects of global warming) and the negative externalities of conventional sources (air and water pollution, greenhouse gas emissions, radioactive waste) are not internalized. This presents the opportunity for government intervention to reduce the sizable entry barriers that prevent the effective penetration of renewables into the electric supply market. Through a variety of measures, most of them fiscal in nature, state and national governments on both sides of the Atlantic have the ability to at least partially rectify the market imperfections inherent to a deregulated or liberalized electricity market. In presenting many of the means

through which governments can help the RE industry to overcome entry barriers, this paper intends to argue that these governments have the responsibility to adopt any or all of these measures to ensure that renewable market penetration is non-negligible, and that the market share of RE sources grows to the extent that the world's energy future is cleaner and more secure.

Environmental concerns, including RE and conservation, were largely non-factors in the debate over deregulation, and were not part of the rationale used to justify the course of action taken in many states and countries; what is more, no studies were performed to gauge the potential environmental effect of liberalization in Europe.⁴ Because RE was not considered during the debates over how deregulation should proceed, which were largely political in nature, governments must consider them after the commencement of deregulation in order to neutralize or reverse many of the negative effects deregulation has on the industry.

There are a number of ways that RE is hurt by deregulation. First, as the market shifts away from its traditional monopoly structure, research and development (R&D) is expected to decrease in the short run as firms try to cut costs and gain market share; R&D under deregulation becomes tied to business considerations, and no longer takes policy considerations into account.⁵ The benefits of energy R&D are well established, and include many environmental attributes such as reduced emissions and improved efficiency of renewable technologies. Many renewable technologies, including solar and fuel cells, are still in the early stages of development, and are considered high-risk, high-reward investments. As utilities begin to compete with one another, there will no longer be jointly conducted R&D projects, and the funding that is still devoted to R&D will likely shift to shorter-term projects with less risk. Spending on R&D is already declining in many countries: the electric industry in the US devoted 0.3 percent of sales revenue to R&D in 1994, and this figure is expected to decrease; from 1984-1998, energy R&D decreased in every US state; Dutch utilities are decreasing their investments from 0.7 percent to 0.6 percent of sales; and Spanish utilities, currently obligated to spend 0.85 percent of sales on R&D, will likely reduce this figure by 0.5 percent.⁶ These cuts in energy-related R&D are doubly damaging, as public sector investment often encourages private sector investment, and so as government funds devoted to R&D are reduced, so too will flows of private money dry up.⁷ Although empirical evidence suggests that oligopolistic markets have the greatest incentive to innovate, in a deregulated marketplace most of this innovation will likely be in slight improvements to current technologies, not advances in new technologies (including RE), as firms move away from higher-risk investments. Unless governments act to increase public investment in long-term energy R&D (and thus implicitly encourage more private investments), the RE sector will stagnate and will continue to be burdened with high costs. This means that "though utility customers may be short-term winners from the deregulation of utilities by paying less for energy, a national failure to invest in energy R&D is very likely to have long-term negative impacts on national energy sectors, the economies, and environmental well-being."⁸

Long-term prospects for RE could be hurt further by dismantling or cutting back policies that had been previously provided by utilities under the regulatory

framework. These policies that were beneficial to RE are considered inappropriate in a deregulated market, as they give preferential treatment to some suppliers. Renewables could also be hurt by new rules and requirements brought about by deregulation, such as increased transmission costs and the need to provide ancillary services. If decisions regarding the post-deregulation framework are made without considering the effects on renewable technologies, the RE industry may be disadvantaged.⁹

There are arguments, however, that restructuring the electric industry could potentially be a boon to the renewables industry. First, with the advent of competition in the market, consumers are now free to choose their supplier, and the more environmentally conscious among them may be willing to pay a small premium for electricity generated in a way that minimizes environmental impact. In addition, deregulation often brings with it increased access to the transmission grid, which in turn provides more opportunities for RE systems located far from load centers. Given the nature of many renewable resources – especially wind and hydroelectricity – the best sites are often far from population centers; deregulation could thus ease the grid interconnectivity for these systems.¹⁰ Lastly, and most appropriately for the purposes of this paper, deregulation offers state and federal governments the prospect of researching and implementing new policies to promote RE. If governments take full advantage of this opportunity, RE could emerge as a significant contributor to the electricity pool for the states and countries that encourage its development with far-reaching and forward-looking policy solutions.

Before delving into the various forms these solutions can take, the question of whether government intervention is necessary in a deregulated market should be addressed. As is the case with most economic issues, there is no definitive answer, although the presence of market failures is largely regarded as beyond question. Those who argue in favor of the efficacy of markets in providing the socially desired outcome generally shun the necessity of interventionist policies. Government intervention is often far from perfect, and does not come cheaply. Even when some fiscal policies could be considered helpful – such as pollution taxes, government R&D, and removal of subsidies to conventional energy suppliers – it can be argued that technology-specific policies are not necessary.¹¹

However, it is the position of this author that private investment alone cannot adequately fund the development of a sizable market share for RE, due to the aforementioned market failures, which include the following:

- Uninternalized negative externalities stemming from the environmental costs of electricity generation through fossil fuel combustion or nuclear power, as mentioned above.
- Uninternalized positive externalities stemming from the mitigation of these negative environmental effects through the increased utilization of RE, which are difficult to measure as the benefits accrue in a large part to future generations.

- The presence of public goods, most notably energy R&D, that are historically under-funded by free markets.
- Imperfect information, which prevents the market from running as efficiently as it otherwise could. While perfect information is an unattainable goal, government intervention could alleviate many of the problems associated with this limitation of the market.
- Unequal tax treatment and subsidies given to the more well-established, traditional sources of electricity generation, which in turn lead to price distortions that favor conventional sources of energy to the detriment of RE.

Many existing policies have helped to overcome some of these market failures, and should continue.¹² The processes of deregulation and liberalization of electricity markets present opportunities for the re-evaluation of existing policies, in addition to the formulation of new strategies to increase market penetration of RE. It has been argued that a small price differential of less than 20 percent between RE and conventional sources of energy, in conjunction with a campaign to inform customers as to the true costs of energy, could drive demand for RE.¹³ Too high of an initial demand, however, will lead to high RE prices, which will serve to discourage customers in the long-run from switching to green energy. On the other hand, if initial demand is too low, there is less incentive to develop new capacity, leading to higher unit costs of RE generation, less competitiveness on the part of suppliers, and the need for government support to keep the industry afloat. The goal of government policies should therefore be to ensure that demand for RE grows steadily, and that supply continues to grow to keep pace. The following policies, to be used in various combinations with one another depending on circumstances intrinsic to the market in which they are being considered, are suggested as means through which governments can play an active role in maintaining both the steady growth rate of demand for RE and the balance between that demand and concurrent supply.

RENEWABLE PORTFOLIO STANDARD

The concept of a renewable portfolio standard (RPS) has been widely adopted and praised as an effective policy for increasing RE generation and consumption. Under an RPS, regulators or lawmakers dictate a certain percentage of the state or country's power that must come from RE. In a regulatory environment, such a requirement would force the regulated utility to invest in RE projects in order to increase the share of green electricity in their power mix. In a deregulated market, the same requirement serves to increase demand for RE on the part of suppliers, all of whom must now take their power mix into account. Flexibility is provided through the introduction of renewable energy credits (RECs), which are generated by wind farms, photovoltaic panels, landfill gas plants and other RE projects at the rate of one REC for each kilowatt-hour of electricity generated. The sale to, and trade between, suppliers of these RECs is made possible

through the unbundling of the electricity generated by renewable systems from the environmental attributes (such as the mitigation of greenhouse gas emissions and other positive externalities mentioned above) these projects produce. Except in the case of distributed generation (described below), the electricity generated goes into the grid, while the attributes are sold and traded on the auxiliary REC market. This market functions in much the same way as the market for SO₂ permits, which has worked to curb air pollution and limit acid rain since the early 1990s. Whether by investing in their own RE resources or buying them from other projects (both existing and those in the process of being built), retail suppliers must meet the RPS requirement, thus creating market opportunities for new projects.

Many US states already have, or are considering, RPS requirements. Massachusetts, for example, enacted its RPS in 2002 and it took effect beginning in 2003 with the initial requirement of 1 percent a year, to be increased by 0.5 percent each year until 2009, and by 1 percent each year thereafter (until the annual increase is suspended by the Massachusetts Division of Energy Resources).¹⁴ If a retail supplier is unable to purchase enough RECs to meet its requirement, it must purchase Alternative Compliance Credits at the prevailing REC market price. The revenues from the sale of these alternative credits are used by the Massachusetts Technology Collaborative, an organization created expressly for this purpose, to encourage RE project construction.

There is a similar policy in place in the United Kingdom to create opportunities for the construction of new RE projects. The Non-Fossil Fuel Order (NFFO) of 1995, later changed to the Renewables Obligation in 2002, was put into place with the stated objective of creating 1,500 MW of new renewable generation, representing approximately 2.5 percent of Britain's electricity needs, by the year 2000.¹⁵ In each successive round of NFFO advancement, the intention of a convergence in the price of RE and conventional energy sources was reiterated, so that "in the not-too-distant future the most promising renewables can compete without financial support."¹⁶ The goal was increased in 2004 to 15 percent of Britain's electricity needs by 2015. Unlike an RPS, which relies on the creation of a new market for RECs to encourage the growth of the RE industry through normal market mechanisms, the NFFO in the UK subsidized investment in RE through a tax on consumers' electric consumption.

DISTRIBUTION SURCHARGE PROGRAMS AND EDUCATIONAL INITIATIVES

There is a host of programs aimed at increasing RE market penetration that can be funded through a small surcharge added on to the distribution portion of a consumer's electric bill. Charges are volumetric and non-bypassable, and are similar to the tax on electricity consumption used in the UK to fund the NFFO. The money raised in this way can be used by governments to directly fund the construction of new RE projects, as under the NFFO. Or it can be used to help correct various market imperfections, thereby allowing markets to function more efficiently, to the advantage of the RE industry. The surcharge could go towards funding long-term R&D, for example, or towards fiscal programs that serve to level the playing

field between the RE and conventional energy camps. It could fund studies to develop new permitting and siting restrictions that would be specific to RE projects, or provide incentives to manufacturers of RE equipment. It could provide low-interest loans to project developers, or give production incentives to renewable production.¹⁷ Alternatively, surcharge revenues could encourage green power marketing, to be discussed in more detail below.

In addition, through the support of educational and informational programs aimed at end-users, a distribution surcharge could help eliminate the market imperfection created by incomplete information, in this case the lack of understanding of the true costs of electricity generation on the part of consumers. There exists empirical evidence in the economics literature that consumers' willingness to pay for electricity increases in conjunction with their awareness of the environmental problems caused by its generation.¹⁸ In the transition period from regulation to free competition, end-users will likely continue to buy electricity from their usual suppliers as the lack of information about alternatives will lead to uncertainty and reluctance to change, especially among risk-averse firms.¹⁹ Early adopters are likely to be environmentally-minded, and any spillover from this group to other groups of consumers is unlikely. Therefore, RE should be targeted to different groups in different ways, so that initial demand rises.²⁰ In this way, the positive network externality normally associated with RE (i.e., since most RE projects are not built until they have enough customers subscribed, low numbers of subscribers keeps supply low and RE prices high) can be proactively addressed.

Consumer preferences help determine the "technological trajectories" of the electricity market; these trajectories become more entrenched as time passes, so it is important to "give renewable energy consumption a boost at the beginning of the liberalized market."²¹ This demonstrates the need for an active government-supported educational campaign, perhaps including the instruction of the benefits of RE in schools, so that consumers are better prepared to shoulder the onus of demanding technological change and progress that is placed on them.

Electricity generation is currently on a trajectory that has long favored non-renewable resources, and market-based approaches to diffusion of RE are unlikely to suffice for the long-run.²² What are needed, therefore, are policies that encourage the optimization of current technologies in the short-run while simultaneously promoting change and development in the electric market in favor of RE for the long-run. Renewables are not yet ready to meet the demands of the system, and RECs are likely to foster hybrid solutions that easily fit into the existing generation infrastructure (such as combined heat and power and biomass generation).²³ More revolutionary technologies that buck the trend – such as wind, solar, and fuel cells – require public support, and the best way to drum up that support is through the dissemination of information that educates the public as to the environmental benefits of RE and the environmental costs of conventional fuel sources.

GREEN POWER MARKETING

As has already been mentioned, generating an appropriate level of demand at the time of deregulation will have a lasting effect on the success of various

policies over the longer term. With the transition to a deregulated marketplace, prices of electricity should fall across the board, but with greater declines in the prices of conventional energy sources. Renewable generation will necessarily be inelastic in the short-run, even with an RPS or similar policy in place, due to the (often considerable) amount of time needed to secure investment, site, plan, develop, and build renewable projects. This inelasticity will likely dampen the demand for RE, and increase the tendency to free ride that could emerge even among well-informed consumers that do not want to pay more for RE despite their knowledge of the true costs of their electricity use (with the hope that others are willing to pay the premium). The aim of policies that deal with the marketing of green power, therefore, should be the creation of initial demand low enough to keep prices low and reduce the temptation to free ride, yet high enough to stimulate investment in the RE industry, and with the potential for future demand growth.²⁴

The limited possibility of rapid growth in RE supply in the short-run, and the implications this causes for the inability of RE to capture increasing returns to scale, could continue to pose barriers to market entry for years after deregulation is initiated. Economies of scale are generated as RE capacity increases, due to the fact that this increasing capacity brings about a decline in the difference between average and peak generation requirements. This extra capacity from RE removes the need for the most expensive generators ("peakers") to be active during peak transmission times – as well as increased supply reliability.

The price premium on RE should be enough of a deterrent on initial consumption to keep demand low; the government, without pushing demand too far above initial supply, should endeavor to grow demand to meet consistently growing supply. There are a number of ways that government intervention in the electricity market can assist RE companies to market their products. In addition to the public education programs mentioned above, which alert consumers as to the availability and benefits of RE, there are also monetary incentives that can be given either to customers that voluntarily purchase renewables, or to producers themselves.²⁵ These fiscal incentives can take the form of price discounts or various types of subsidies to encourage the consumption and generation, respectively, of RE. Furthermore, some states have given funds collected through a distribution surcharge to green power suppliers in the form of rebates, which decrease as demand increases; these rebates are then used to offset the companies' marketing costs, both raising the profile of RE and lowering costs to consumers.²⁶

Other policies governments can pursue in order to support green power marketing include fuel source disclosure or standardized green power certification, so that consumers are aware of the fuel mix used in the generation of the electricity they are using; and government purchases of RE, to buoy demand. The success of green power marketing as a means through which to encourage RE production and consumption depends on the emergence of retail suppliers of green power; some argue that the materialization of these suppliers eliminates the need for continued government intervention, while others say a successful government RE policy requires active market intervention to sustain this niche market.²⁷

Market forces alone would fight an uphill battle to convince any but the most environmentally conscious consumers to pay a significant premium on their

electricity use and switch to green energy. Government policies to support green power marketers can help in lowering the price of RE, but only the most dedicated policies will reduce the difference between RE and electricity from conventional energy sources to a negligible amount. In many cases, however, there are also non-governmental actors working in the public interest to promote increased utilization of RE. One of these, the World Resources Institute's Green Power Marketing Development Group, works with some of the world's largest corporations to increase their RE purchases. The proposition is viewed as win-win, as large purchases serve to increase global demand for RE through better media coverage of the RE market and increased customer awareness, while for the company, they often serve as "an investment to develop, improve, or reinforce their corporate image."²⁸ Some states, including Rhode Island and Massachusetts, are encouraging their governments to buy RE or giving incentives for large customers to purchase RE and promote those purchases, thereby further raising the profile of RE and creating a stronger overall market for renewables.²⁹

ENCOURAGEMENT OF DISTRIBUTED GENERATION

Distributed generation (DG) is the term applied to electricity generated close to the load it serves. In recent years, DG has increasingly taken the form of renewable microprojects, be they rooftop solar photovoltaic panels, isolated wind turbines, geothermal facilities, or the like. With the deregulation of electricity markets, demand for DG has increased rapidly, due to its increased utilization as a means through which consumers reduce their peak load requirements; the overload of the transmission grid; widespread resistance to the construction of new power plants; growing demand for more environmentally sound generation; and what one author calls "horribly botched deregulation and restructuring of the electricity industry."³⁰ Whether or not the savings attained through peak-shaving make DG economical in some instances, the fact remains that by its very nature, DG cannot take advantage of scale economies. Therefore, government policies to support renewable DG projects is imperative for their continued development.

Distributed generation projects face significant market entry barriers that discourage their construction, stemming from both technical and non-technical issues. On the technical side, not only are there scale economies attained by larger projects, but often the efficiency of the technology itself is not even viable on a smaller scale. Wind turbines, for example, require a minimum wind speed to make them efficient enough to warrant their construction, and often times this minimum speed isn't reached below heights that would preclude their construction in many areas. On the non-technical side, entry barriers often emerge from the application of rules meant for larger projects to much smaller DG projects. In Massachusetts, for example, a one megawatt project was burdened with "tariffs, transmission costs, operating charges, and penalties for losses levied by ISO-NE [Independent Service Operator – New England], the region's grid operator." This problem was eventually solved by treating this particular DG as negative load instead of generation.³¹ Policies that work to incentivize DG, through tax credits or subsidies for example, can help raise the image of RE and promote its diffusion. Governments

at the state and national level can even work together to foster larger markets for DG technologies.³²

CONCLUSIONS

Each of the aforementioned policies to increase the market penetration of RE has its advantages and disadvantages. An RPS, for example, is a market-based, administratively simple mechanism that ensures a minimum amount of renewable generation while integrating RE into the competitive power market. At the same time, however, it brings with it uncertain costs and to some extent benefits existing renewable systems to the detriment of new projects (especially if the mandated quantity of RE in the power mix increases too slowly or not at all). A distribution surcharge, on the other hand, includes explicit cost containment and maximizes the exploitation of renewable resources within set cost limits, but could be perceived as a tax (making it somewhat politically unstable) and could marginalize RE in a competitive market.³³ For these and other reasons, a combination of policies is often more viable and more suitable than any one policy, and the effectiveness of the policies developed during the restructuring process depends on their duration and stability. There exists concern that public support for efficiency and R&D programs, for example, will decline with the restructuring of the electric industry in many areas; RE and environmental advocates therefore need to work together and articulate the need for continued, effective support for RE.³⁴

What is needed more than anything else, perhaps, is a paradigm shift in the electric industry that serves to shed new light on new and existing policies and uses the newly-available free market mechanisms to address the inertia that has stagnated innovation in the industry over the last century. Several years have passed since electric markets were liberalized in many European countries and US states, and other countries and states continue to debate costs and benefits in their own decisions to open their markets to competition. When considering what policy actions, if any, to take to support RE, these countries and states should surely consider the experiences of the UK, California and other pioneers and adapt the most successful policies to their individual situations.

They must also realize, however, the tendency to regress to the norm in situations involving change and uncertainty. For consumers, this means a reluctance to change suppliers; for producers, it means an aversion to risk as they try to establish their footing in the market. This is understandable, of course, but it remains the responsibility of the government to help the market overcome this industrial indolence, through informing customers of their options and providing them with incentives to switch to RE, and to shoulder much of the long-term risk inherent with energy R&D.

Electricity generation has evolved slowly throughout its history, generally seeming to follow a punctuated equilibrium model of evolution, to borrow a concept from biology. Long periods of relative inactivity in which existing technologies are refined and improved are interrupted by "punctuations" of more rapid change and drastic transitions; nuclear power and the increased reliance on RE after the oil crises constitute some of these punctuations, with the former more

of an exclamation point and the latter more of an ellipsis. We have arrived at a point in history that seems appropriate for the continuation of the evolution of electricity generation that seems to have lost momentum over the past several decades. Present are both the need – global warming threatens our very survival, and RE is one of the most promising solutions – and the opportunity. All that is needed is the catalyst, in the form of proactive government policies that can overcome market failures, to open the door to a new era in electricity generation.

NOTES

¹ J.J. Dooley, “Unintended Consequences: Energy R&D in a Deregulated Energy Market,” *Energy Market* 26.7 (1998): p. 550.

² For the purposes of this paper, renewable energy will include electricity generated from wind, solar, landfill gas, small-scale hydropower, and other sustainable resources. Conventional sources include oil, coal, natural gas, and nuclear power (nuclear power produces waste and is not generally considered renewable).

³ Bruce Edward Tonn and Martin Schweitzer, “Institutional and Programmatic Suggestions for Satisfying Public Policy Responsibilities in a Retail Competitive Electricity Market,” *Energy Policy* 25.1 (1997): p. 31.

⁴ Niels I. Meyer, “Distributed Generation and the Problematic Deregulation of Energy Markets in Europe,” *International Journal of Sustainable Energy* 23.4 (2003): pp. 217-219.

⁵ Dooley, p. 551.

⁶ Ibid.

⁷ Ibid, p. 549.

⁸ Ibid, p. 547.

⁹ Ryan Wiser et al, “Renewable Energy Policy and Electricity Restructuring: A California Case Study,” *Energy Policy* 26.6 (1998): p. 466.

¹⁰ Ibid.

¹¹ Ibid, p. 467.

¹² Ibid.

¹³ Roger Fouquet, “The United Kingdom Demand for Renewable Energy in a Liberalized Market,” *Energy Policy* 26.4 (1998): p. 281.

¹⁴ Massachusetts Renewable Energy Portfolio Standard Regulations <<http://www.mass.gov/doer/rps/225cmr.pdf>>.

¹⁵ Fouquet, p. 287.

¹⁶ Ibid.

¹⁷ Wiser, p. 468.

¹⁸ Fouquet, p. 284.

¹⁹ Doris Fuchs and Maarten Arentsen, “Green Electricity in the Market Place: The Policy Challenge,” *Energy Policy* 30 (2002): p. 534.

²⁰ Ibid, p. 533.

²² Ibid, p. 528.

²³ Ibid, p. 529.

²⁴ Fouquet, p. 281.

²⁵ Wisser, p. 468.

²⁶ Mark Bolinger et al. "States Emerge as Clean Energy Investors: A Review of State Support for Renewable Energy." *The Electricity Journal* November 2001: 85.

²⁷ Ibid.

²⁸ Fouquet, 286.

²⁹ Bolinger, 93.

³⁰ Thomas F Armistead, "Distributed Generation Seeks a Place in the Sun," *Engineering News-Record* 246.14 (2001): p. 45.

³¹ Ibid, p. 46.

³² Bolinger, p. 94.

³³ Wisser, p. 471-472.

³⁴ Ibid, p. 465-466.

Electing Hamas

Zecharia Kahn

With the election of Hamas, the Bush Administration's democracy promotion policy in the Middle East appears to be a failure. However, an in-depth review of the theory, motivations and actions leading up to the election of Hamas in the Palestinian Authority parliamentary elections shows a more complex picture. The history of the Hamas movement proves that it is a pragmatic and forward thinking organization that has been able to adapt to the modern electoral system with great skill; furthermore, in its election may lie the seeds for a lasting peace through the democratic process.

Interviewer: Can you clarify your concept of the Palestinian state?

Sheikh Ahmad Yasin (founder of Hamas): The Palestinian state must be founded on every inch of Palestine that we liberate, but without conceding the rest of our rights.

Interviewer: Do you recognize Israel?

Yasin: If I recognized Israel, the problem would be finished, and we would have no rights left in Palestine.

Interviewer: But if Israel withdraws from the West Bank and the Gaza Strip, would you recognize it?

Yasin: When it withdraws, I will say.

Interviewer: But at that time should it be recognized?

Yasin: I leave this matter to the representatives of the Palestinian people.

Interviewer: Who are they?

Yasin: Those whom the Palestinian people will elect.¹

INTRODUCTION

On January 25, 2006, a democratic earthquake struck the Middle East when the Islamic fundamentalist group Hamas won an absolute majority in the Palestinian parliamentary elections. The implications of the election for peace and stability in the region are unclear but by looking back at the recent history leading up to this historic election we can gain some much needed perspective

Zecharia Kahn is an M.A. candidate at the Bologna Center. He previously attended Colby College in Waterville, Maine, and has worked for the OSCE-PA in Copenhagen, Denmark, and in the United States Senate.

and perhaps come to a better understanding of Hamas and how they are likely to act in the future.

While the Bush Administration's rhetoric proclaims Iraq as the vanguard democracy in the heart of the Middle East, it is in the Palestinian Territories that this democratic revolution is taking place. Hamas' electoral victory seemed to call into question the basic foundation of American foreign policy in the region. This was evident through US President George W. Bush's declaration of the goal of US policy to promote democracy in the Arab world in an effort to transform the region into a more peaceful, stable and terrorist-free place in the long term. The election of an organization officially labeled as terrorists by both the US and EU, whose stated goal is the elimination of the State of Israel and who have killed hundreds of people in terrorist acts over the last 20 years would seem to call into question the wisdom of such a policy. Based on the assumption that democratic elections inherently help spread American values and support US security, these results can be seen as a horrendous backfiring of a misguided policy.

However, there appears to be no reason to make this assumption nor was the election necessarily a failure. While there is no reason to believe that democratic elections should bring about a result that automatically aligns with American interests, there is some evidence to support the argument that free and open elections leading to true representation are a necessary step towards moderation and accountability for radical groups. Though a Hamas victory appears to be a step backward and a move away from a possible peace, it could in fact be a necessary step forward. This may be an essential element in leading to a real and lasting Arab-Israeli peace since any deal made by the Palestinian Authority (PA) would require the backing of the broadest possible support from the Palestinian people. This means the Islamist groups would need to be included.

Hamas is a complex organization, one that is extremely flexible and pragmatic despite its fundamentalist ideology. Paradoxically, at the same time as it is forward thinking and modern, it is conservative and radical. Even though a PA dominated by Hamas runs the risk of extinction and at the very least particularly hard times to come, there are reasons to look at this as an opportunity as opposed to an unmitigated disaster. The Israelis had already been unilaterally imposing an attempt at peace upon a divided and corrupt PA still ruled by the remnants of Yasser Arafat's Fatah party. Prime Minister Ariel Sharon's government was not negotiating with the PA and under his successor, Ehud Olmert; this will not change any time soon. Instead, it gives the Palestinians the opportunity to get their house in order. Of course, it also gives them enough rope to hang themselves by, but this is the gamble that has been taken and there is no backing out now.

THE INTEGRATION GAMBLE

For years many Middle East experts, pundits and analysts have been stressing the need to allow and encourage Islamists and Islamic political parties to pursue their goals through the confines of the legal political system. By repressing Islamist movements, the argument goes; governments legitimize and empower

the violent and radical tendencies that exist within such movements. What is needed instead, is a more nuanced and suitable response. First, there should be a differentiation between the radical and moderate elements of an Islamist movement. Second, a dual approach should be used, marginalizing the radicals and providing the moderates with incentives and rewards for playing a constructive role politically and socially.²

According to this argument, blind repression by a government only empowers the radical elements of the movement and discredits the moderates. What should be done instead is work to support the moderate Islamists as they develop a substantial stake in the existing government and institutions. Once this bond with the system is fully established, the Islamist movement will have been effectively co-opted and the radical and violent tendencies discredited. Once in power, the theory goes; the Islamist movement will be restrained by the reality of governing. Gamal al-Banna, now age 85 and the youngest brother of the founder of the Egyptian Muslim Brotherhood, Hamas' parent organization, clearly describes the situation as: "The real test of the Brotherhood is to let it enter politics. They will be in a different situation when they confront the necessities of ruling, and there are only two possible outcomes. They will have to compromise or fail."³ It was this logic that led the PA to allow Hamas's electoral participation and eventually to the dramatic election results.

Every situation is complex and has its own unique characteristics. This is certainly true for the Palestinian-Israeli conflict, with its incredibly high international profile and systemic fragility, where a single violent act can upset the entire peace process. Within this framework an infantile Palestinian state is emerging and under the current constraints and influences is attempting to move towards a democratic form of government. In order to do this, the Palestinian President Mahmoud Abbas decided to gamble on integrating the Palestinian Islamist movements. As Abbas stated shortly after his election in January 2005:

Of course they should be converted into a political party... [I]t is good for us. We are talking about national unity. ... [N]ow Hamas and Jihad [Palestinian Islamic Jihad] are running for the elections and what does it mean? It means that in time they will be converted into political parties.⁴

Abbas's statement proves that he had bought into the integration argument. This, and other statements, proves he believed that the more their role becomes political, the more moderate Hamas and Islamic Jihad would have to become. By supporting the moderate voices within the jihadist movements, internal divisions would grow and faced with the practical choices of attempting to govern, the violent and radical tendencies within the movements could be restrained. As Hamas, the largest and most popular Palestinian Islamist group, takes on the responsibilities of power, they would specifically be more prepared to acknowledge the need to support the two-state solution and become more accountable to 'ordinary Palestinians.' This theory follows the belief that 'ordinary Palestinians' are more interested in peace and economic security, and polls have indeed shown that most Palestinians prefer a secular to an Islamic state.

Hamas seems to have other ideas. They see as their model the Lebanese Hizbullah movement. Hizbullah never gave up their armed resistance and have still successfully participated in Lebanese politics while seemingly forcing an Israeli withdrawal from Southern Lebanon. As Khaled Masha'al, one of the most senior members of Hamas's external leadership put it, "negotiating without resistance leads to surrender but negotiation with resistance leads to real peace."⁵ Or, as Hashem el-Masri, the elected Hamas acting-mayor of the West Bank town of Qalqilya stated, "[Hamas] is trying to create a balance between political and military issues," and will not trade in one for the other. "Hamas is a resistance movement that will enter these elections to promote reform, but we remain a resistance movement."⁶ For his part, Abbas had no plans to confront Hamas with arms, declaring that he had no intention of causing a Palestinian civil war. Indeed, even if they had wanted to, Abbas and the Fatah-run PA were not strong enough to disarm Hamas. Instead, they hoped that once a part of the government, Hamas' violence would have been curtailed to a greater degree by the new legislature and public opinion.

Critics, on the other hand, argued that this was a foolish assumption. Hamas's participation in elections certainly did not, and has not, made it inherently peaceful. Further, they were concerned that once Hamas had solidified its position within the PA's institutions, it would resume its war of terror on Israel and seek to control the Palestinian Territories through force. They pointed out that the stated goal of Hamas is still the destruction of Israel and a state based on Islamic law. Also, allowing Hamas to participate without requiring them to disarm continues to allow them the option of using force to disrupt political negotiations they disagree with. Finally, the suspension of violence allows Hamas to rebuild its military force and plan future terrorist attacks.

Haim Malka, a fellow at the Middle East Program at the Center for Strategic and International Studies, wrote in *The Washington Quarterly* that "despite the inherent risks, proponents of expanding Hamas's role in Palestinian national politics...point to the fact that Hamas's leaders have long called for transparent and accountable governing institutions and have demonstrated political pragmatism, suggesting that the group could accept less than its absolutist demands. Continuing to marginalize Hamas, which represents a significant portion of Palestinian society, will ensure that the movement will continue terrorist activities, thwarting any future political agreement with Israel and dashing any prospect for the creation of a vibrant Palestinian democracy."⁷

Allowing Hamas, which represents a major portion of the Palestinian people, into the electoral process was an absolutely necessary step if there was going to be a long-term independent and democratic Palestinian state and Abbas realized this. Yet by taking this gamble, Abbas risked Hamas gaining a significant position in the Palestinian Parliament. As *The Economist* magazine noted in November 2005, "everyone is wondering what to do then: the United States and European Union, who classify Hamas as a terrorist organization; Fatah which stands to lose its cozy, corrupt monopoly over the PA; and not least Israel, for whom Hamas is the standard-bearer of an Islamist Palestinian nationalism that seeks to destroy the Jewish state."⁸

ORIGINS AND GOALS

Like many Islamist movements, Hamas has become as much a nationalist movement as it is a religious one. That being said, Hamas is first and foremost an Islamic religious fundamentalist organization. Hamas is an abbreviation of its full name, *Harakat al-Muquawama al-Islamiyya*, which means 'Islamic resistance movement,' and can be translated as 'zeal.' It first emerged as a force in Palestinian politics shortly after the start of the first Intifada, or uprising, in 1987. Hamas was founded by the leadership of the Muslim Brotherhood in the Gaza Strip, which is still its traditional stronghold of support, as a direct challenge to the Palestinian Liberation Organization's (PLO) monopoly on political power within Palestinian society. As the PLO evolved into the PA, Hamas remained its main challenger for power and, in many ways; it is opposition to the PLO, and later the PA, that helped define Hamas.

The PLO (and then the Fatah dominated PA) stood for a secular nationalist Palestinian identity. The ultimate goal was a political program, achieved through military and terrorist means, which would lead to the establishment of a territorial Palestinian state. In order to differentiate itself, Hamas appropriated parts of the PLO's nationalist narrative and put it into an Islamic context. By combining an Islamic-national vision with community activism, Hamas is able to mix religious doctrine with daily concerns. This community activism was already a part of Hamas activities in its previous incarnation as the Muslim Brotherhood of Gaza. In the Muslim Brotherhood, or Society of the Muslim Brothers (*Jama'iyyat al Ikhwan al Muslimin*), we can find much of the organizational foundations of Hamas that help explain its popularity, evolution and current actions. As a *New York Times* article about the Brotherhood's attempts to run in recent Egyptian elections put it, "the Muslim Brotherhood is often seen in two ways — as the fusty great uncle of Islamic politics, content to bide his time, or as the womb of all subsequent Islamic terror movements. Both views are true."⁹

The members of the Muslim Brotherhood in Gaza who created Hamas in 1987, led by their political and spiritual leader, Sheikh Ahmad Yasin, founded it upon the basis of an Islamic-nationalism, ostensibly without the need for compromise. The Hamas charter envisions a Palestinian state brought about not through negotiation but through holy war, a state that would exist in all of Palestine, not in part of it, and one that would be based on shari'a, or Islamic law, not a secular state. Hamas defined itself as an alternative power to the PLO with a similar goal, a Palestinian nation, but a superior method of achieving its objectives: the Islamization of the struggle.

PRAGMATISM, OSLO AND TERRORISM

With its combination of Islamic charity organizations and violent resistance to Israeli occupation, Hamas soon became extremely popular in the Gaza Strip. The high profile of the Palestinian-Israeli conflict and the success of some military operations soon raised Hamas to the status of an internationally known terrorist organization. The public face of Hamas was that of a fundamentalist organization

committed only to violence and the destruction of the State of Israel. However, within the complex world of the Israeli-Palestinian struggle, Hamas has had to walk a fine line between pragmatism and ideology, between fulfillment of its perceived Islamic duty of holy war against Israel and its awareness of the boundaries and constraints of the political and social environment in which it operates.¹⁰ Several challenges have required a fine-tuning of the Hamas ideology: the Oslo Accords, the establishment of the PA and the first general election, and finally, the continual question of whether to accept a temporary, peaceful settlement or to continue the jihad against Israel.

One of Hamas's most important moves was to distinguish between its short-term goal of a Palestinian state in the West Bank and Gaza based on the 1967 borders, and its long-term objective of the ultimate destruction of the State of Israel and the establishment of a Palestinian Islamic state in all of Palestine. By differentiating between short-term and long-term goals, Hamas was able to achieve political flexibility while not seeming to compromise its ideological credibility. In this way Hamas has been able to maintain its unique position of being a Palestinian national movement while holding fast to its stated Islamic-nationalist principles. By declaring that any settlement with Israel was justified in the short term as a step towards an ultimate long-term goal, Hamas seemingly overcame the necessity for negotiation and compromise.

This flexible ideology can be seen in many Hamas actions. One technique is the way in which the movement uses a 'hudna,' or a military truce, as well as another type of ceasefire called a 'tahdiya,' which usually refers to a 'cooling off' or 'period of calm.'¹¹ Neither the hudna nor the tahdiya is seen as supporting peace or accepting the legitimacy of Israel, but rather as endorsing a pause in conflict. In most cases the ceasefires have been offered in attempts to avoid overwhelming Israeli retaliation against the group's leadership, usually in response to a Hamas terrorist act or following a Hamas-PA confrontation. Whenever they are used, they are always presented or accepted at a time when Hamas needed a break to organizationally regroup, consolidate their gains and prepare for new attacks.¹²

In addition to ceasefires, Hamas uses other methods such as controlled violence, calls for Palestinian solidarity, cunningly ambiguous statements, political independence and a calculated participation with the PA to maintain its position and influence. Hamas is also reliant on outside forces to provide it with popularity. Specifically, Hamas' popularity thrives on three things. First, as the vanguard of the Palestinian resistance; they always benefit from Israeli aggression, whether or not it is in response to Hamas actions. Like many traditional Islamic movements oppressed by their governments, any military aggression serves to support Hamas as a symbol of the resistance. However, in Hamas' unique case, the oppressive governments included Israel, the United States and the PA. Second, the rampant corruption in the Fatah movement has played right into the hands of Hamas, with their reputation of Islamic piety and honesty. This may be the single most important factor in Hamas' current rise in popularity. Finally, Hamas thrives on poverty and desperation. With extensive charitable organizations and a spiritual outlook,

Hamas offers assistance where no one else does. This is particularly true with life in the poverty stricken refugee camps that have always been one of its greatest centers of support. Traditionally, most of Hamas' militants are recruited from among the young urban poor.

Hamas has managed to nimbly negotiate its way through the minefield of the conflict, adjusting to challenges, picking up opportunities and allies along the way, only to lose them, suffer setbacks and somehow survive. Many of the movement's leaders, including two of the group's most senior leaders, Yasin and Abd al-Aziz Rantisi, have been killed in targeted attacks by Israel. Despite such setbacks, their inherent strengths and flexible pragmatic politics, along with the corruption and ineptitude of the PA and Fatah, have placed Hamas in as strong a position as they have ever been.

BEYOND THE SECOND INTIFADA

As noted above, the current surge in electoral popularity for Hamas is as much, if not more, a reflection of Palestinian disappointment with Fatah, than it is about Hamas. Khaled Duzdar, a Palestinian analyst at the Israel/Palestine Center for Research and Information put it more bluntly: "The success of Hamas is only due to Fatah's failures."¹³ In life and in death, Yasser Arafat, the founder of the PLO and its main Fatah faction, seems to dominate Palestinian politics. Today, he is perhaps more conspicuous in his absence than he ever was, holed up in his Ramallah compound during his last few years. It was Arafat's decision to endorse a second Intifada in the fall of 2000 brought an official end to the Oslo peace process. In addition, it was he who created the problems, particularly the corruption, which unraveled his Fatah party's dominance within the Palestinian territories.

For its first four months, the Second Intifada was a Fatah-led insurgency. Headed by Marwan Bhargouti (now serving five consecutive life sentences in an Israeli prison), the *tanzim*, a new military wing of Fatah, carried out most of the attacks. It was not until the election of Ariel Sharon in February 2001 that Hamas stepped up its own level of violence. Soon it was Hamas suicide bombings within Israel that came to symbolize the uprising. Hamas did not declare a new official organizational goal for the increase in its operations; it was simply a continuation of its resistance to peace with Israel. At the same time however, it did appear to be following a new strategy based on the Hizbullah model of military resistance. In October 2002, Rantisi, then the Hamas political leader in Gaza, was quoted as saying "The intifada is about forcing Israel's withdrawal from the 1967 territories... [b]ut that doesn't mean the Arab-Israeli conflict will be over."¹⁴

While the Second Intifada did not force Israel to withdraw to its 1967 borders, it did bring great gains in popularity for Hamas. According to reliable surveys, Hamas increased its popularity by as much as 60 percent in the first three years of the Intifada, emerging as a power equal to Fatah in parts of the West Bank and far out-polling it in Gaza.¹⁵ This rise in popularity was due to a combination of traditional factors from which Hamas gained support: armed resistance to Israel, the collapse of PA police forces and divisions in Fatah sown by Israel's West Bank

and Gaza invasions, and the visceral appeal to Palestinians of its suicide attacks inside Israel. The extensive charitable and welfare services continued to be as important and stood in stark contrast to the inefficiency and collapse of the PA ministries. The result was that by late 2002, Hamas was less a party in opposition to the PA and Fatah than an independent national force bent on establishing, in the words of former PA Culture Minister, Ziad Abu Amr: “a political, social and military alternative to the existing Palestinian order.”¹⁶ Now once again, Hamas had an important choice to make. What to do with its newfound power?

The Hamas leadership made a strategic decision in 2004 to focus domestically, specifically on gaining control of the PA through the electoral process and facing the risks that that entailed. There were a number of restraints and influences, both internal and external that helped convince the leadership to make this strategic choice. First of all, there was the new regional order that was ushered in by the US response to the events of September 11, 2001. The US war on terror and the resulting focus on the Middle East meant that Hamas had to be very careful if it did not want to end up being labeled as a global jihadist group and lumped in with Al Qaeda. Already classified as a terrorist organization by the US, Hamas was starting to feel politically and economically isolated. The US, EU and PA had frozen bank accounts and disrupted Hamas's flow of funds and, “according to one Egyptian intelligence official, by 2005 Hamas' funding from Arab and Islamic states, with the exception of Iran, had all but dried up. In September 2003, the European Union put the whole of Hamas (rather than just its military wing) on its ‘terrorism’ blacklist, a huge political setback for a movement that ha[d] striven to be recognized internationally as an authentic Palestinian party, and a further crimp on its fundraising abilities.”¹⁷ In addition, by August 2003 Israeli military actions had devastated parts of the Hamas infrastructure and it needed a break from fighting. Without financial support, under assault from Israel and with world political attention focused on stopping terrorism, Hamas proved its pragmatism once again by adjusting its goals. There were also internal reasons for the shift. Palestinian public opinion had turned with the Second Intifada. A majority of Palestinians now wanted peace and more importantly economic security. The Intifada had left the infrastructure of the territories destroyed, the PA devastated and the economy in ruins. Finally, with Sharon imposing a unilateral withdrawal from Gaza and the northern West Bank, Hamas saw an opportunity it could not miss to cash in on the political capital its armed resistance and welfare programs had earned.

In order to do this, Hamas turned once again to its strategy of accepting short-term gains without fully relinquishing its long-term objectives. Just weeks before he was killed in March 2004, Sheikh Yasin presented a new platform consisting of three carefully crafted parts. First, Hamas would temper its attacks on Israel for the duration of the withdrawal. This was further extended into a hudna signed in Cairo in March 2005, and extended to the end of the year in August. Second, Hamas would escalate its resistance efforts within Gaza in an attempt to try to show politically that Israel was withdrawing under fire. Finally, and most importantly, the new platform called for working alongside the PA towards the goal of a Palestinian state within the 1967 borders, with Jerusalem as its capital.¹⁸

It was this fundamental shift towards working alongside the PA leadership, for essentially the same goals, that led Hamas to participate in the PA electoral process.

POWER AND RESPONSIBILITY

In December 2005, about a month before the parliamentary elections, Hamas won big electoral victories in the West Bank towns of Nablus and Jenin, once Fatah strongholds. Despite Hamas's previous success in March and July elections and high poll ratings all year, the victories still came as a shock to Fatah leaders, foreshadowing what was soon to come. "Issam Abu Baker, Fatah's leader in Nablus, told The Associated Press: 'We didn't think for a moment that Hamas would win so many votes. The earth shook under our feet and this will have an effect on the Parliament.'"¹⁹ As problematic as this was for Fatah, Hamas's electoral success was also a serious problem for Hamas itself who did not appear to be seeking leadership of the PA but instead to become the dominant force in Palestinian politics.

Again, foreshadowing future events, the electoral victory for Hamas meant they were now responsible for the problems of the people they governed. Evidence of this can be seen in the experience of the West Bank city of Qalqilya, where in May 2005 Hamas won every seat on the city council, as well as the mayoral race. In its short tenure, the Hamas administration has had some success in reforming city management and budgeting, but has lost many supporters by refusing to allow the use of city property for a traditional Palestinian cultural festival and by barring a group of intellectuals from meeting in a city building unless they first apologized for criticizing Hamas. Worst of all, Hamas, which ran on a platform of reform and lower prices, especially for electricity, has been hurt by an increase in electricity prices. While Israel provides Qalqilya its electricity, and prices have gone up everywhere because of higher oil costs, the blame falls to Hamas as the party in power:

"Hamas promised to lower electricity prices, but increased them," said one resident, Youssef Sharwar. "They promised reform, but I think they were lying."

Qalqilya "is already under siege from the Israeli wall," said Mr. Zaid, the poultry merchant, referring to the separation barrier - here mostly a 10-yard-high concrete wall - that cuts off the city from Kfar Sava and curves around the Israeli settlements that overlook Qalqilya. "But this vote has created another siege by the world." Omar Rai voted for Hamas, but now regrets it. "I worked in Hamas's campaign," he said, citing Fatah's reputation for lassitude and corruption. "But they want to bring us back to Afghanistan!"²⁰

Despite its insistence that it can reconcile short-term compromise with its long-term objectives, the step that Hamas took by participating in the elections cannot be underestimated. Since its inception, Hamas has held fast to three important principles that are directly challenged by its participation in the PA

electoral process and threatens to fundamentally change the nature of the organization. First is the goal of the destruction of Israel, second, the right to political violence and third, the establishment of a state based on Islamic law. By taking part in the governing institutions of the PA, Hamas is participating in a government that was founded on the acceptance of the two-state solution; one that has accepted Israel's right to exist. If Hamas hopes to dominate this organization, to help run the Palestinian economy and have a hand in negotiating the future of the Palestinian people within the framework of the PA, it will have to acknowledge Israel. It will need to allow the PA to have a monopoly on violence and give up the demands for an Islamic state, a generally unpopular notion with the Palestinian people. This would fundamentally change the character of Hamas; in essence accomplishing Abbas's goal and transforming the terrorist jihadist Islamic movement into a legitimate political party.

Hamas leaders had intended to put off facing these decisions for as long as possible. In fact, they probably didn't expect to have the electoral success that they've had and assumed instead they would be able to maintain their position as the opposition, with a minority in the parliament and the threat of military action as leverage. Now they are faced with the challenge of having an absolute majority in the Parliament and having to form the government. As Menachem Klein, a political scientist at Bar-Ilan University told Israeli radio before the elections: "If Hamas wins a majority, it will have to conduct a policy that it does not agree with and conduct foreign relations on an ideological basis that it rejects, formally. It will have to change. This will not be simple at all."²¹

CONCLUSION

Hamas's history is far from simple and clear. When examined in detail a fundamental pragmatism and flexibility in the face of adversity appears. In order to survive Hamas has had to adapt its ideology as well as its methods. This flexibility should be seen as a warning for those who assume they can predict Hamas's behavior or effectively bully and threaten it. It is also a positive sign for those who hope that Hamas will become more moderate as it adapts itself to the responsibilities of governing and the reality of being held accountable in elections for its actions. However, there is a long road ahead, as Hamas remains armed and defiant. In addition, whether or not this is the best case, Hamas is now the test case in the Middle East for electing Islamists.

If Hamas fails, it would be a spectacular failure with devastating results that would be a huge setback for both Islamists and democracy in the Arab and Muslim world. Such a failure would mean the Palestinian Territories would become the prime example of the danger of allowing Islamists to participate in the legal political system. The US and EU would then be inclined to provide whatever support is necessary to 'friendly' autocratic Middle Eastern governments to crush similar movements. This would be a great loss since there will never be real progress in the Middle East without greater pluralism and some kind of accommodation between religion and politics.

NOTES

¹ R. Scott Appleby ed., *Spokesmen for the Despised: Fundamentalist Leaders of the Middle East* (Chicago: University of Chicago Press, 1997), p. 245. Citation from Schiff and Ya'ari, *Intifada*, p. 225.

² Guilan Denouex, "The Forgotten Swamp: Navigating Political Islam," *Middle East Policy* (June 2002): pp. 56-82.

³ Neil MacFarquhar, "Will Politics and Success at the Polls Tame Egypt's Muslim Brotherhood?" *New York Times* 8 December 2005.

⁴ Haim Malka. "Forcing Choices: Testing the Transformation of Hamas." *The Washington Quarterly* (Autumn 2005): pp. 37-54.

⁵ Ibid.

⁶ Steven Erlanger, "Voted In, Hamas Sets a West Bank City Astir." *New York Times* 4 November 2005.

⁷ See Malka.

⁸ "The Hamas Conundrum." *Economist* (12 November 2005): p. 50.

⁹ See MacFarquhar.

¹⁰ See Mishal and Sela, p. 2.

¹¹ See Malka.

¹² Jeff Cary, "Hamas Ceasefire Proposal: Peace or Pause?" *The Washington Institute for Near East Policy, Peace Watch #447*, 16 January 2004, <http://www.ciaonet.org/pbei/winep/peace_2004/2004_447/index.html>,

and Seth Wikas, "The Hamas Ceasefire: Historical Background, Future Foretold?" *The Washington Institute for Near East Policy, Peace Watch #37*, 3 January 2002. <http://www.ciaonet.org/pbei/winep/peace_2002/2002_357.html>.

¹³ Steven Erlanger. "In Era After Arafat, Islamic Militants are Edging into Power," *New York Times*, 18 December 2005.

¹⁴ Graham Usher, "The New Hamas: Between Resistance and Participation," *Middle East Report Online*, 21 August 2005. <<http://www.merip.org/mero/mero082105.html>>

¹⁵ Ibid.

¹⁶ Ibid.

¹⁷ Ibid.

¹⁸ Ibid.

¹⁹ Steven Erlanger, "Hamas Surges in West Bank; Blow to Fatah." *New York Times* 17 December 2005.

²⁰ Erlanger, "Voted In, Hamas Sets a West Bank City Astir."

²¹ Erlanger, "Hamas Surges in West Bank, Blow to Fatah."

Between Polyarchy and Autocracy: The Delegative Democracy of Carlos Menem

Brittany Williams

Seemingly cursed by the legacy of the *caudillo*, government under a strong executive has come to characterize the structure of the modern Argentine state. Since the administration of Carlos Menem in the 1990s, questions have arisen as to the progress Argentina has made with regard to the consolidation of its democracy. However, while Menem set a precedent for directly challenging democratic institutions during his presidency, has history justified his unilateral decision-making as the only means of overcoming the barriers that obstructed Argentina's political and economic development?

Debate continues to rage between theorists seeking to characterize the nature of the democracy established in Argentina following the collapse of the bureaucratic authoritarian state in 1983. According to the principles established by Robert Dahl, Argentina fulfills the basic conditions necessary to be classified as a *polyarchy*. Officials are chosen by the populace in free and open elections. The public possesses access to independent sources of information, and the people retain the freedom of expression as well as the right to associational autonomy.¹

Controversy arises, however, concerning the extent to which the democracy in Argentina is consolidated and representative, as multiple grievances attest to the illiberal nature of the state in spite of the presence of democratic institutions. Guillermo O'Donnell has attempted to clarify the inconsistencies in the process of Argentine democratization through his model of *delegative democracy*. In such a system, a stable democracy exists, but it is neither consolidated nor institutionalized, often forgoing a final effort to create a truly representative government.

O'Donnell argues that delegative democracies are typically led by legitimately elected presidents that perceive themselves as "the custodian of the national interest," using their popular election as a mandate "to govern as [they

Brittany Williams is an M.A. candidate at the Bologna Center, pursuing a degree in Western Hemisphere/Latin American studies and international economics. She graduated in 2005 from Case Western Reserve University with a B.A. in political science, international studies and Spanish. Brittany has conducted independent research on the 2003 Argentine presidential election.

see] fit." In the absence of institutionalized horizontal accountability between the branches of government, the executive is able to exert considerable influence over the affairs of the state. The prevalence of unilateral action in delegative democracies alienates the legislature and the judiciary, ultimately undermining the legitimacy and the representative nature of the state. The executive seemingly assumes full responsibility for policy decisions in the eyes of the citizenry, typically leading to open public disillusionment with the government and the political process in general as it is perceived that the state is only representative of the will of the executive, not the interests of the populace.²

In Argentina, the model of delegative democracy can clearly be applied to the regime of Carlos Saúl Menem, who served as president of Argentina from 1989 to 1999. Throughout the 10 years of his presidency, Menem exploited the weaknesses of the Argentine democracy to establish a personalistic system that bent to his will. Through the use of decree powers, he marginalized Congress (*Congreso de la Nación Argentina*), suppressing its constitutional powers to exert accountability over the actions of the executive. Menem compromised the integrity of the judiciary, successfully "packing" the Supreme Court (*Corte Suprema de Justicia de la Nación*) in order to exercise control over legal affairs. And he manipulated the party system, exploiting the feeble institutions of the *Partido Justicialista* (PJ) to eliminate dissent to his political agenda and to ensure the continuing support of the party and masses. But while the divisive nature of Menem's actions has become infamous, was the establishment of a delegative democracy in Argentina a gross abuse of power or a necessary evil?

EXECUTIVE DECISION

Perhaps Menem's most flagrant application of delegative democracy was his frequent use of executive decrees in the place of legislative action for the realization of policy. According to Argentine law, the president was granted the power to enact executive orders, or *decretos*, in "states of emergency marked by events out of the ordinary, exceptional situations, unpredictable circumstances, or predictable but unavoidable circumstances."³ These decrees were binding until repealed by Congress or declared unconstitutional by the Supreme Court. Typically utilized with extreme caution, only 20 of these orders were enacted by constitutional governments in Argentina from 1853 to 1983.⁴

However, during Menem's first administration alone, over 336 executive decrees were proclaimed,⁵ addressing issues ranging from the prohibition of mass strikes by labor unions to the donation of cement to Bolivia.⁶ He argued that the need to enact multiple decrees of "urgency and necessity" was due to the "glacial pace of legislative activity" which resulted from high levels of absenteeism experienced within Congress. The failure to reach a quorum on multiple occasions had created an accumulation of pending legislation which complicated Menem's efforts to promptly address the troubles facing the nation. While lawmakers in the Chamber of Deputies focused upon drafting proposals to fine truant representatives, Menem bypassed the legislature completely, enacting his economic and political agenda by utilizing executive decrees.⁷

Menem also displayed a proclivity for attempting to manipulate electoral rules, though many of his efforts were soundly defeated. Seeking to secure a party majority in the Chamber of Deputies for the *Partido Justicialista*, he proposed altering the system of representation in Congress in 1992, shifting from a multi-member proportional system to a single-member plurality system. When this was rejected by the legislature, Menem declared that census results reported that 23 additional seats were required in the Chamber of Deputies, of which eleven would be allocated to the province of Buenos Aires, a PJ stronghold. This effort was also defeated. Finally, in 1993, Menem announced that the government was investigating the logistics of lowering the voting age to 16, allowing voters from all parties to vote in primary elections, and replacing the closed-list ballot system with one in which voters would simply cross out the names of candidates for whom they did not want to vote. Considered by many to be yet another attempt to empower the PJ, the proposal was largely ignored.⁸

Menem did succeed, however, in obtaining a significant constitutional reform during his administration that allowed him to pursue a second consecutive term as president. Capitalizing on a weakened Radical party, Menem began pressuring Congress for the reform in 1993, unabashedly insisting that his Radical presidential predecessor, Raúl Alfonsín, and other party leaders collude with PJ members in order to amend the Constitution. By forging the controversial Olivos Pact (*Pacto de Olivos*) between the PJ and the Radical Party, Menem secured his reform in 1994 and announced his candidacy for the 1995 presidential election.

JUDICIAL CONSENT

Menem quickly resigned the judiciary to a position of inferiority during his first term, realizing early in his presidency that the influence of the judiciary must be curbed in order for the power of the executive to remain unconstrained. Under Alfonsín, the Supreme Court had established itself as a formidable and independent entity, delivering decisions which obstructed several of Alfonsín's social and economic programs. Intent on limiting the Court's influence over presidential affairs, Menem began to employ strategies to strip the judiciary of its status even before he was officially in office. Seeking first to simply reorganize the Supreme Court through appointments, Menem implicitly encouraged the resignations of the existing justices by offering them ambassadorships and other prominent positions in his administration. Thwarted by the resolve of the unwavering justices, Menemists in the Chamber of Deputies made open threats of impeachment against them. Ultimately, however, Menem's efforts were in vain as only one judge, Chief Justice José Severo Caballero, resigned from his post in 1989.⁹

Publicly admitting his determination not to be "the only president in 50 years who hasn't had his own court," Menem proposed to increase the size of the Supreme Court from five to nine members in 1990. In spite of protest from the existing members of the Court, the controversial measure passed through the Argentine legislature with surprising ease due to successful lobbying by the horde of Menemist representatives in Congress. In addition to Caballero's resignation, Justice Jorge Bacqué resigned in protest shortly after Menem's proposal was ratified,

ultimately permitting the president to appoint a total of six justices to the newly enlarged Court, securing his desired majority. The subsequent neutrality of the justices chosen for the Supreme Court by Menem was dubious at best. Seeking a judiciary that would fully support his presidential agenda, Menem rewarded his close political associates with the positions. Justice Julio Nazareno was a partner in Menem's law office. Justice Eduardo Moliné O'Connor, allegedly, was his tennis partner.¹⁰

By creating a Supreme Court whose prerogative was, "to go along with the politics of the president," according to the newly appointed Justice Rodolfo Barra,¹¹ Menem's delegative rule was given "constitutional legitimacy" through often distorted legal logic. The Supreme Court firmly upheld Menem's constitutional right to act as the "supreme chief of the nation," fully granting him the power to oversee "the general administration of the country" ("*jefe supremo de la Nación... responsable político de la administración general del país*").¹² According to the Court, this allowed Menem to exert ultimate authority "over any obstacles originating from partial interests," given that, as president, he was defending "the interests of the nation as a whole."¹³ Ultimately, by widening his scope of authority, the Court sanctioned Menem's sweeping presidential decrees.

A "NEW" PERONISM

Aside from abusing the democratic institutions of the state in order to suppress dissention, Menem impeded upon Argentina's party system, specifically by manipulating the feeble structure of his *Partido Justicialista* as a means of ensuring unwavering support for his political endeavors. The PJ, also referred to as the Peronist Party, began a considerable transformation shortly after Menem's first election as president in 1989. Since its conception in 1945, the PJ typically maintained strong connections with the working class of Argentina, allowing its leaders to rely upon a stable base of support in spite of sweeping volatility in times of political crisis. Traditionally a bastion of labor-based politics, the platform of the party significantly conflicted with the neoliberal economic reforms enacted by Menem, a former party president. But as Menem rapidly disassembled the corporatist political and economic state created by Juan Perón in the 1940s, the PJ remained quiet, failing to openly criticize the president as he violated his campaign promises to protect organized labor and workers' interests. And by the time of Menem's reelection in 1995, the National Council Executive Board of the PJ had openly pledged their support of his agenda.¹⁴

According to Steven Levitsky, this acceptance of a clearly contradictory political program was made possible by the low degree of routinization within the institutional structure of the PJ.¹⁵ The party is plagued by the legacy of Perón who, like other populist leaders, utilized it for little more than a personalistic means of amassing popular support. Centralized around the designs of Perón, the PJ was routinely reorganized in accordance to his political needs, preventing the establishment of strong internal mechanisms such as a central bureaucracy and a means of horizontal accountability among party leadership. Even in the years following Perón, this weakness was never fully addressed, making the PJ extremely vulnerable to domination by the interests of a single figure such as Menem.

Without an organized bureaucracy, transfers of power within the PJ are unpredictable and easily manipulated. There are no established career paths or tenure tracks and, in spite of a system for internal elections, PJ leadership has never been selected through institutional means. This lack of a hierarchy has created a general sense of disdain among members of the PJ seeking to advance their political careers as opportunities for advancement provided by the party are severely limited and unreliable.

Many Peronists thus turn to the state to seek positions of authority and status. And Argentine political history has shown a distinct overlap between PJ executive leadership and state leadership, a trend that continues to exist in the modern era. Aware of this correlation, many Peronists seeking positions of higher authority within the state will frequently “[defect] to internal factions [within the PJ] that hold (or are about to hold) state power.”¹⁶ This creates a bandwagon effect within the party, as members in search of political advancement will adapt or completely discard their own political beliefs for the chance of securing a position through the patronage-based system.

Menem quickly seized upon this trend as a means of co-opting members of the PJ in order to secure both party and popular support for his neoliberal economic reforms in the 1990s. Throughout his first term, Menem’s deviation from the populist presidential platform of his election brought sharp criticism from many within the Peronist party. However, while surveys demonstrated that less than a quarter of the party agreed with his reform efforts, the PJ never openly condemned Menem or his actions.¹⁷ This was due to a wave of “conversion” to *Menemism* shortly after Menem secured the party’s presidential nomination in 1988. Many leaders of the PJ, such as Chamber of Deputies Majority Leader and future Interior Minister José Luis Manzano and future PJ president Roberto García, joined ranks with Menem, ultimately securing their positions within the government and within the echelon of PJ elite.

The formation of a coalition of loyal Menemists throughout the PJ not only meant that Menem could rely on allies scattered throughout the government to support his reforms, but it also granted Menem the ability to quell the voices of minority dissent in the PJ. Outnumbered and marginalized, Antonio Francisco Cafiero remained a staunch anti-Menemist throughout his term as president of the PJ, but he rapidly recognized that any effort to censure Menem would be fruitless, ultimately resigning from the post in 1990. In an interview with Levitsky, Cafiero stated, “We produced reports and declarations... [that] contradicted what the government was doing. So we were ignored... There was no way to make [Menem] understand that there was another authority at his side.”¹⁸ García echoed this sentiment after becoming PJ president in 1993. “In the first phase of my presidency,” García commented, “I drew up the party communiqués and got them approved by the government ... In the second phase, the government sent me the communiqués ... In the third phase, I read about the communiqués in the newspapers.”¹⁹

WAKING A SLEEPING GIANT

While Menem clearly undermined the democratic institutions established within Argentina during his presidency, it is arguable that his actions were indeed justified. When first elected as president of Argentina in 1989, Menem faced a nation in chaos. His predecessor Alfonsín was the first democratically elected president since the disintegration of the bureaucratic authoritarian state. However, during his administration, he had faced a wave of obstacles that prevented him from establishing order and stability in Argentina, from the continued threat of military intervention to an economy that was in shambles. Resigning before the end of his term, Alfonsín handed the reins to Menem, who now faced a more desperate situation. But by crafting a delegative democracy, he was unhindered by the institutional impediments that had thwarted Alfonsín and others. Before the end of his first term, Menem had successfully enacted a comprehensive and controversial program of economic restructuring and political transformation, that halted Argentina's descent into absolute disorder.

What is most striking about Argentina's economic transformation under Menem is the manner in which reform was enacted in comparison to that in other Latin American nations. Radical economic transition was accomplished without extreme repression, such as in Bolivia, or without the establishment of a truly authoritarian regime, as was the situation in Chile under General Augusto Pinochet and in Peru under Alberto Fujimori. And, unlike the sluggish economic transitions that occurred in Costa Rica, Uruguay, Venezuela, and Brazil, where reform was inextricably tied to the maintenance of democracy, Argentine economic restructuring was sweeping and complete, enacted in less than one presidential term.²⁰

According to Menem's Minister of Economics (*Ministro de Economía de la Nación*) Domingo Cavallo, only 20 percent of Argentina's economic reforms would have been realized without Menem's combined use of legislative approval and executive decrees.²¹ Menem marked his first administration with a comprehensive series of neoliberal economic policies, intent on stabilizing the financial crisis while also integrating Argentina into the world economy. The staunch Peronist seemingly reversed his political platform, calling for the immediate privatization of nearly all nationalized firms, such as the railroad and subway systems as well as the oil, coal, and electricity industries. By 1994, nearly \$24 billion worth of government controlled firms had been sold or converted into private ventures.²²

Menem's fiscal policy included overhauling the tax system in Argentina, eliminating Alfonsín's export and energy taxes and replacing them with a series of value-added taxes and income taxes. The new system increased yearly government revenue by more than \$10.7 billion and, from 1991 to 1993, the state ran a small budget surplus. With regard to trade and economic integration, Menem reduced nominal tariffs on imported goods and, in 1991, he enacted a "mega decree" that abolished import quotas and regulatory boards for assorted goods, deregulated price restrictions, simplified customs regulations, and removed limitations on business hours. Finally, in 1992, he signed the MERCOSUR free trade agreement with Brazil, Uruguay, and Paraguay, directly binding Argentina's

economy to others in Latin America.²³

Possibly the most notorious of Menem's economic reforms was the introduction of the "convertibility law" in 1991 which instituted a one-to-one exchange rate between the Argentine peso and the American dollar. In combination with the additionally imposed reforms, the currency board program caused inflation to plummet from 1,832 percent in 1990 to 4 percent by 1994. The nation experienced an increase in gross national product of approximately 8 percent annually. Foreign investment in Argentina was renewed, with direct investment from the United States rising to \$2 billion in 1994. And the resulting economic boom from 1989 to 1993 resulted in an overall reduction in poverty for the Argentine population.²⁴

Politically, Menem's practice of delegative democracy and unilateral decision making allowed him to swiftly and efficiently address threats to Argentine democratization. In response to the instability plaguing the nation upon his election, Menem acknowledged the valid threat of military intervention in government affairs. Realizing that direct attacks upon the military elite would provoke further volatility, he sought to weaken the organization's political power while still allowing its institutional existence, a different approach than that of Alfonsín, whose platform included the open destabilization of the military as punishment for its crimes. Discreetly, Menem limited the operational budget of the military, cutting defense expenditures in his initial plan to reduce the deficit in 1989. Unable to financially maintain a force of 200,000 soldiers, the draft was eliminated and enrollment in the armed services dropped to 100,000 by 1990, with the army standing at only 30,000 men. Subsequent internal reorganization of personnel was thus necessary, and many officers were demoted to fill vacant low-level positions.²⁵

Most noted, however, was Menem's controversial decision in 1990 to pardon military officers accused of crimes committed under the dictatorships, including human rights abuses associated with the "Dirty War" as well as crimes stemming from the ill-fated Malvinas War. In spite of popular opinion polls stating that 60 percent of the nation was against such a measure, Menem had pardoned more than 280 officers, including former military president Jorge Videla, and fully assumed control over judicial proceedings against the military. While this caused national and international outrage and divided the government, Menem understood that it was a step toward creating an environment in which the government would possess the advantage over the military. The effects of this measure were almost immediate. In December 1990, an armed rebellion of noncommissioned officers seized army headquarters in Buenos Aires, demanding an increase in the military's budget and the resignation of the army chief of staff. Menem grasped the opportunity to exert executive power, publicly ordering the military elite to immediately restrain and discipline the defiant subordinates. Within hours, the army responded with tanks, suppressing their own rebellion.²⁶ And Colonel Mohammed Ali Seineldín, the chief conspirator behind the insurgence, was sentenced to life in prison.²⁷

From that moment forward, Menem implemented a final attack on military authority in Argentina. He further reduced the military's budget, from 18.2 percent of government expenditure in 1989 to less than 10.6 percent by 1993, and privatized

military-owned enterprises, requiring that all armaments manufacturing firms be liquidated immediately. It was decreed that the Ministry of Economy and Production (*Ministerio de Economía y Producción*) be granted sole authority over matters of military spending and that the power to deploy the military was now the responsibility of the Ministry of Foreign Affairs (*Ministerio de Relaciones Exteriores, Comercio Internacional, y Culto*). Military officers were expunged from the cabinet and autonomous proclamations from the armed forces were forbidden. Menem had accomplished what Alfonsín could not. He had defeated the military in Argentina.²⁸

With regard to his impact upon the political institutions of Argentina, Menem arguably consolidated some aspects of Argentine democracy through his practice of delegative politics. Though criticized for pressuring Congress to adopt the constitutional amendment to ensure his chance for re-election, the additional reforms passed in conjunction with the presidential term extension advanced the state of Argentine democracy. Under the 1994 Constitution, senators would now be elected through a direct popular vote instead of through appointment by provincial legislatures, increasing accountability between the upper house of Congress and the general public. The reforms also instituted a dual ballot system for presidential elections in which plurality winners from the preliminary round would advance to a runoff election that would ultimately decide the presidency. This would provide the victor with greater legitimacy as the mandate would now reflect a solid majority of popular vote instead of a marginalized portion obtained in a race against multiple candidates.²⁹

While he transformed the composition of the judiciary, Menem surprisingly did not seek to curb its constitutional authority as it was in his best interests to respect the actions of the Supreme Court. Menem allowed the justices to maintain and even expand upon the Court's judicial sphere, especially in cases which upheld the power of the executive. One such example of judicial expansion arose in 1990 concerning a case involving the privatization of Aerolíneas Argentinas. With the approval of the Menem administration, the nationalized airline was to be sold to a group of investors associated with Iberia Airlines, but an injunction was filed to prevent the sale of the firm to a foreign company. Knowing that the delay could jeopardize the deal, the Supreme Court voted to grant itself *per saltum* in the case, a power not constitutionally granted to the Argentine Court. Under *per saltum*, jurisdiction was immediately seized from the lower courts on the grounds that the issue required urgent attention and, within a day, the matter was resolved and the transaction finalized.³⁰

By seeking to preserve his "mandate" from the public, Menem also contributed to the stability of democratic governance in Argentina, inspiring a desire for continuity and a sense of commitment among the electorate. Through the *Partido Justicialista*, he was able to provide direct economic support to the poor, establishing territorial outposts of the party in Argentina's provinces. By creating soup kitchens, distributing consumer goods, and organizing social and cultural events, the urban and rural PJ networks immediately helped to suppress social volatility, specifically the mass looting and rioting that marked the latter days of the Alfonsín administration. In the long-term, this aid renewed a sense of loyalty

to Menem and the PJ, drawing lower-class voters away from the radical and far-left movements that were calling for an end to his reform agenda. By co-opting a larger portion of the electorate, Menem was able to ensure continued PJ dominance in Congress and, consequently, the legislative approval of his programs. The legions of loyal Peronists had become indistinguishable from the Menemists, voting in 1991 and 1993 to protect the president's majority in the legislature and, inherently, the democratic legitimacy of his administration.³¹

LASTING LEGACY

By the time Menem began his second term in 1995, Argentina was already suffering from the fallout of his unilateral management, specifically with regard to his overly ambitious neoliberal economic agenda. Unemployment rates soared due to the initial push towards privatization and budget reduction. More than 85,000 of the approximately 246,000 workers within public industries were dismissed. And they joined the 217,000 unemployed civil servants that were fired by Menem between 1989 and 1992 in an attempt to curtail the expanding budget.³² The trend only worsened, as, by 1995, the unemployment rate had reached 18.6 percent and the underemployment rate was at 11.3 percent. Argentina possessed the second-highest unemployment rate in all of Latin America, barely trailing behind Nicaragua.³³ The public debt doubled and economic inequality consequently worsened, as a population of "new poor" began to grow in Argentina. This group was mainly made up of members of the middle class that had lost their economic status as the overall standard of living rapidly declined. By 1996, 20 percent of households in Greater Buenos Aires fell below the poverty line³⁴ and nearly half of all Argentina's children were living in poverty by 1998.³⁵

Although Menem had combated hyperinflation through a currency board with the American dollar, his plan marked a severe overvaluation of the Argentine peso, leading to a trade deficit of more than \$6 billion in 1994.³⁶ Dollarization prevented the Menem administration from adjusting its monetary policy to compensate for the downturn. Unable to devalue the peso, the government was unable to further stimulate growth. In combination with a deepening fiscal crisis due to a mounting debt, which comprised 41 percent of gross domestic product by 1998,³⁷ and decreasing tax revenues, Argentina slid into a recession in 1999 which ultimately served as a precursor for its economic collapse in 2001.

Institutionally, Menem's penchant for issuing decrees and bending rules set a precedent for his presidential successors, essentially demonstrating that it was possible to shape the Argentine government to reflect the will of the executive. The mere premise under which the constitutional restructuring of 1994 occurred clearly demonstrated the power Menem held over Congress and his opposition. Whatever positive reforms were made, the fundamental reasoning behind amending the constitution was to ensure Menem's personal political gain by granting him the ability to be re-elected for a second term. And though the reforms of 1994 intended to curb executive authority by, for example, systematically defining what constitutes a "public emergency" when issuing decrees, the damage had been done as Menem had already established a precedent of executive contempt for constitutional boundaries and the separation of powers.

Menem's politics also destroyed public confidence in the Argentine democratic system. The lack of accountability created a system based upon unparalleled levels of corruption as the Menem administration begged, bribed, and threatened government officials, the media, and the public in order to garner support for its agenda. Frustrated with the overt manipulation and the strong degree of presidential power, public opinion surveys reflected a sense of disillusionment with democracy, with a Latinobarómetro poll indicating that, by 2001, only 58 percent of Argentines surveyed agreed that, "democracy is preferable to any other kind of government," plummeting from 71 percent in 1996.³⁸ This disenchantment paved the way for the collapse of the following administration of Fernando de la Rúa (1999 – 2001) during which public anguish materialized in rioting and rebellion against the government under the slogan "*¡que se vayan todos!*" ("They all must go!")

While Menem's practice of delegative democracy brought about positive change in Argentina by temporarily stabilizing the economy, eradicating the threat of a military coup d'état, and institutionalizing fundamental principles of democracy, the long-term negative consequences of his presidentialism considerably outweigh the immediate constructive outcomes of his administration. Ultimately, Menem's presidency joins the regimes of Fernando Collor de Mello in Brazil and Fujimori in Peru as a historical illustration of the eventual consequences that delegative democracy will inflict upon a state. The paradox of delegative democracy is that, though it allows the executive to implement often vital changes to a nation's policy, these decisions are often made in great haste, forgoing the natural revision and analysis process that would occur were the decisions approved through the legislature and the judiciary or subject to public opinion. Unfortunately, it is only in retrospect that the flaws of delegative democracies are identified. By then, the future is already paying the price of the transgressions of the past.

However, the lessons of Menem's administration are seemingly forgotten in Latin America as democratically elected leaders continue to accelerate the consolidation of power around their executive authority. Hugo Chávez's "Bolivarian Revolution" continues to undermine the balance of power within Venezuela's democracy. In 1999, Chávez sought to revise the constitution as part of his promise to transform the Venezuelan state. But the reforms served mainly to empower the executive branch over the legislature, granting the president the power to dissolve Congress, eliminating congressional oversight in military promotions and limiting its role in the appointment of judicial officials, and creating the office of a vice-president to be selected by the president.³⁹ And in Argentina, Néstor Kirchner's "K Style" politics, described as a "volatile, unpredictable approach that blends the moral fervor of a crusader with the irascible tough-mindedness of an old-fashioned political boss," characterize government action.⁴⁰ Recent shuffling of cabinet personnel has seen the curious replacement of competent ministers with outspoken supporters of Kirchner, most conspicuously noted in the ousting of Minister of Economics Roberto Lavagna, coordinator of Argentina's astounding post-debt crisis economic recovery, and the subsequent appointment of Felisa Miceli, a lesser-known economist but self-described "Kirchnerite soldier."⁴¹

As the struggle for liberal democracy in Latin America continues, it appears that the region simply cannot escape the legacy of government under strong executive leadership. And while delegative democracy is ostensibly a positive alternative to the authoritarian dictatorships that mark Latin America's past, it is still, according to O'Donnell, "far from the boring beauty of consolidated democrac[y]." Ultimately, only time will show the true nature of these regimes and the intentions of the presidents at their helms. But if current trends are accurate indicators of future direction, these leaders would do well to note that history has failed to vindicate their predecessors.

NOTES

¹ Robert Dahl, *Democracy and its Critics* (New Haven: Yale University Press, 1991).

² Guillermo O'Donnell, "Delegative Democracy," in *Counterpoints: Selected Essays on Authoritarianism and Democratization* (South Bend: University of Notre Dame Press, 1999).

³ James McGuire, *Peronism Without Perón: Unions, Parties, and Democracy* (Stanford: Stanford University Press, 1997).

⁴ Peter Smith, *Democracy in Latin America: Political Change in Comparative Perspective* (New York: Oxford University Press, 2005)

⁵ *Ibid.*, p. 164.

⁶ McGuire, p. 255.

⁷ *Ibid.*, p. 255.

⁸ *Ibid.*, p. 256.

⁹ Christopher Larkins, "The Judiciary and Delegative Democracy in Argentina," *Comparative Politics* (July 1998): pp. 428 – 432.

¹⁰ *Ibid.*, pp. 428 – 432.

¹¹ *Ibid.*, pp. 428 – 432.

¹² *Constitución Política de la Republica Argentina de 1853*, Capitulo III, Artículo 99.

¹³ *Rolón Zappa v. Estado Nacional*.

¹⁴ Steven Levitsky, "Organization and Labor-Based Party Adaptation: The Transformation of Argentine Peronism in Comparative Perspective," *World Politics* 54, pp. 27 – 56.

¹⁵ *Ibid.*, pp. 27 – 56.

¹⁶ *Ibid.*, pp. 27 – 56.

¹⁷ *Ibid.*, pp. 27 – 56.

¹⁸ *Ibid.*, pp. 27 – 56.

¹⁹ Steven Levitsky, Interview with Roberto García, Buenos Aires, June 23, 1997.

²⁰ Steven Levitsky, "Argentina: From Crisis to Consolidation (and Back)," in Jorge I. Domínguez and Michael Shifter, *Constructing Democratic Governance in Latin*

America, 2nd Ed. (Baltimore: Johns Hopkins University Press, 2003).

²¹ Smith, p. 164.

²² McGuire, p. 218.

²³ Ibid., p. 218.

²⁴ Ibid., pp. 218 – 220.

²⁵ Gary Wynia, *Argentina: Illusions and Realities*, 2nd Ed. (New York: Homes & Meyer, 1992).

²⁶ Ibid., p. 216.

²⁷ McGuire, p. 253.

²⁸ Levitsky, p. 250.

²⁹ McGuire, p. 253.

³⁰ Larkins, p. 432.

³¹ Levitsky, p. 252.

³² McGuire, p. 218.

³³ Ibid., p. 222.

³⁴ Smith, p. 101.

³⁵ Peter Winn, *Americas: The Changing Face of Latin America and the Caribbean* (Berkeley: University of California Press, 1999).

³⁶ Thomas Skidmore, *Modern Latin America* (New York: Oxford University Press, 2005).

³⁷ “Survey of Argentina: Becoming a Serious Country,” *The Economist* (June 5 2004).

³⁸ “Democracy’s ten-year rut.” *The Economist* (October 29 2005): p. 60.

³⁹ Steve Ellner, “The Radical Potential of Chavismo in Venezuela: The First Year and a Half in Power,” *Latin American Perspectives* 28 (2001): pp. 5-32.

⁴⁰ Larry Rohter, “Cautiously, Argentines Warm to the ‘K Style,’” *The New York Times* June 29 2003.

⁴¹ Larry Rohter, “As Argentina’s Debt Dwindles, President’s Power Steadily Grows,” *The New York Times* January 3 2006.

Getting out of the Intellectual Trap: When Profit-Driven Humanitarian Aid Might Save More Lives

Janika Albers

Considering the increasing involvement of private, for-profit companies in humanitarian aid activities, this essay looks at the incentive structures and moral hazards of the industry. Its goal is to identify conditions under which profit-driven actors could increase the efficiency of aid delivery without compromising its humanitarian aims. The essay argues that when faced with informational asymmetries and the “temptation to cheat,” donors, for-profits and nonprofits compromise efficiency to differing degrees and in different ways. By spelling out these organizations’ respective predispositions, and their effects on performance, this essay answers when and why we should prefer for-profit to nonprofit organizations in the humanitarian aid industry.

Traditionally, humanitarian aid has been the domain of nonprofit actors. Yet in recent years, for-profit firms have increasingly begun to enter the industry, and given the direction of US policy in particular, additional involvement of private actors is expected.¹ This development clashes with a persistent skepticism toward profit-driven companies riding high on human misery, be it the result of man-made or natural disasters. The common thought seems to be “that the private sector is greedy and uncaring and that corporations cannot be trusted with the problems of poverty alleviation.”²

Surprisingly, however, research on the role of the private sector in humanitarian aid activities remains scarce. This essay tries to address this gap, at least partially. The treatment is incomplete, because due to the simple lack of data, it does not contain empirical research about the functions in emergency relief assumed by for-profit firms. Nevertheless, it tries to shed light on the issue by

Janika Albers, an M.A. candidate at the Bologna Center, holds a B.A. in philosophy and economics from the University of Bayreuth, Germany. She has worked in development in Cameroon; on human rights issues in South Korea; and, most recently, for an academic quarterly in Germany.

drawing from the existent body of economic literature about the behavior of for-profit and nonprofit firms, through an examination of empirical evidence from similar sectors.

Looking at the initial entry into the humanitarian aid enterprise by for-profit actors, furthermore, this essay's framing is normative rather than descriptive. It is valid to ask whether we want profit-driven service providers to play a role in emergency relief, and, if we do, under what conditions their involvement seems appropriate.

To be clear, however, the following discussion attempts to distance itself from the moralistic, far from enlightening sentiments mentioned above. Instead it focuses on incentive structures in face of informational asymmetries inherent to humanitarian aid provision. In other words, I will not address the fact that for-profit actors might generate competitive returns on the backs of the suffering, the starving, and the displaced, but will concentrate instead on the specific nature of the emergency relief market to determine how and when for-profit firms can make a positive contribution to the alleviation of human suffering.

The normative frame I apply can be subsumed in a simple test criterion: If saving lives is the ultimate goal of the humanitarian aid enterprise, *how* the number of lives saved is maximized ought to be subordinate to this goal. Profit-driven competition usually generates increased efficiency in markets. Given that for-profit actors improve the ability of humanitarian aid to save lives, we should simply ignore the potential bad aftertaste of profit-driven companies "helping people." If one donated US dollar in the hands of a for-profit firm generates a higher return in terms of human life saved than a dollar in the hands of a nonprofit organization, a humanitarian donor, to make a choice consistent with his values, ought to choose the for-profit firm.

First, I will elaborate on the producer-consumer structure, the resulting lack of incentives, and inefficiencies within the humanitarian aid enterprise. Second, I search for reasons why nonprofit organizations, rather than for-profit companies, traditionally inhabit the industry. The following section gets to the core of the matter, namely the different predispositions of for-profit and nonprofit organizations in the face of informational asymmetries found in most emergency relief operations. Drawing on previous findings, the essay concludes with the case for involvement of for-profit actors in the logistics sector of the aid enterprise, and more recently in the area of disaster insurance.

INCENTIVES AND EFFICIENCY IN THE HUMANITARIAN AID ENTERPRISE

"In well-functioning markets an increase in competition is almost always good news for customers and bad news for incumbents [...] [because of] its ability to push firms to innovate, cut costs, and respond to customer demand. But the aid system is at best a quasi-market."³

What Harford et al. (among many others) call attention to is the specific producer-consumer structure of the humanitarian aid enterprise. Whereas in common markets the consumer of a produced service also purchases the service, in the humanitarian aid industry the end-consumers, i.e. the aid recipients, are not the purchasers of the service. Instead, the donors purchase the service in place of the aid recipients. The aid agencies, as the producers of humanitarian relief, accordingly sell their services; the donors buy them; and the targeted population affected by an emergency finally consumes them.

From this producer-consumer structure arise informational asymmetries. "Because of this separation between the purchasers and the recipients of the service, the purchasers are in a poor position to determine whether the service they paid for was in fact performed, much less performed adequately."⁴ At one end of the transaction, the donors act without full information about the quality and cost of their purchased aid without that control that an end-consumer would normally have. At the other end of the transaction, aid recipients lack the power to "sanction" the low quantity and/or quality of the aid product given its cost. Information differs markedly between buyers and sellers. "[T]he intended beneficiaries have little ability to give feedback, and even when they do there is no mechanism by which the aid agencies suffer adverse consequences if the customers are dissatisfied."⁵

Economics has treated such situations characterized by a lack of transparency and accountability as principal-agent problems. The principal – the donor in the case of humanitarian aid – possesses incomplete information about the performance of the agent, i.e. the aid provider. Consequently, the agent, or aid agency, might compromise the donor's interest to efficiently transform his donation into aid. Easterly asserts:

With private information that is available to the agent but not to the principal, bureaucracies will ex-ante overstate the benefits of aid programs and understate the costs (adverse selection). [...] Since nobody can tell whether the aid agency efforts are really making a difference, they have only weak incentives to exert effort (moral hazard).⁶ [...] Accordingly] donor-principals face the problems of hidden action and information [...] contractor-agents [...] may conceal, withhold, or distort information harmful to their interests.⁷

Obfuscation is further enhanced by the overly complex structure of donors, donor channels, delegating and executing agencies as a study by McKinsey highlights, and with principals and agents multiplying, moral hazard problems increase.⁸

In this way the humanitarian aid industry lacks the proper incentive structure to align the interests of relief agencies and donors. The service that humanitarian aid providers deliver is beyond donors' monitoring capacities; "without adequate monitoring recipients may appropriate the contractor's resources for opportunistic gain."⁹

Another problem arises due to the fact that a large part of activities connected to the provision of aid involve complex qualitative variables that are difficult to

measure. Aid delivery is not simply a quantitative product. Often it involves human interaction between the giver and the recipient such as in emergency hospital treatment where the quality of medical care, while hard to measure, will most certainly have an impact on the number of lives saved. For example, if an aid agency deploys a particular amount of food, the effect of the food aid in terms of lives saved obviously hinges on to whom this food is distributed. Whether the aid deliverer conducts a careful needs assessment is as difficult to detect on part of the donor, as it is to prove on part of the aid agency. As the delicate relationship between output and impact of emergency relief itself would require a whole debate, an appropriate discussion goes beyond the scope of the argument presented here.¹⁰ In the following discussion I will therefore focus on two points: I define the effort to increase the impact of output (e.g. needs-based assessment) as a qualitative variable of emergency relief. The first simplification is that I assume that this effort actually translates into an increased impact (e.g. that the better needs-based assessment results in a higher number of lives saved). As such, impact becomes an attribute of output itself; output is now defined by quantity *and* quality. Second, I can thus refer to output without violating the (impact-oriented) criterion put forth in the introduction: the evaluation of relief efforts according to how many lives could be saved with a given amount of aid.

The preceding considerations have demonstrated that humanitarian aid delivery is a service neither easily monitored nor easily outlined in a contract. Most of the activities in the provision of aid involve important, immeasurable dimensions that aggravate moral hazard and contracting problems. These issues among humanitarian aid delivery agencies are not resolved by the entry of profit-driven agencies as Cooley and Ron claim.¹¹ Nevertheless, such reasoning simply states that for-profit actors are not superior to nonprofit actors in the humanitarian aid industry. Does this mean they are equally (in)efficient?

First, we have to carefully spell out what we mean by “efficiency” in the humanitarian enterprise. Commonly, economic efficiency is defined as “the maximization of aggregate consumer and producer surplus.”¹² However, as asserted above, the humanitarian aid sector is “at best a quasi-market.” With regard to humanitarian aid delivery, we do not wish the producers, or service-providers, to have a surplus. Nor do donors desire to simply maximize their “consumer” surplus. Donors will “buy” humanitarian aid services at their respective reserve prices with a belief about a specific minimum quantity and quality of the services being delivered. What donors (and I in the following) refer to as efficiency is maximizing output, namely aid products or delivery service, at a given “price.”¹³

Nonprofit and for-profit actors respond differently to the lack of incentives structures in the humanitarian aid enterprise. Given that both are susceptible to moral hazard, I contend that they do not compromise the efficiency of their aid delivery to the same degree, nor in the same way, but are guided by their respective predispositions.¹⁴ First, however, I will turn to the question of why traditionally nonprofit organizations rather than for-profit companies inhabit the humanitarian aid industry.

THE HUMANITARIAN AID INDUSTRY AS THE TRADITIONAL DOMAIN OF NONPROFIT ORGANIZATIONS

Though the question is not directly connected to the field of humanitarian aid delivery, economists have long been puzzled by the persistent growth of nonprofit organizations.¹⁵ Whereas initial investigations concentrated on historical developments and altruistic motives, studies increasingly began to focus on the economic role nonprofits play and why nonprofits concentrate in particular sectors rather than in others.¹⁶ A second research question inquired as to why nonprofit firms can compete with for-profit actors in these very industries.

At first glance, nonprofit, like for-profit organizations, constitute production units, so there is no apparent reason why output capacities should differ between them.¹⁷ Rather, the reasons why “nonprofits differ from for-profits [...] must lie in how their organizational structures affect their operations.”¹⁸ The decisive difference between nonprofit and for-profit organizations is that unlike their for-profit counterparts, nonprofits are not obliged to distribute generated returns: “Nonprofit firms operate under a non-distribution constraint. No one has a legal claim to the organization’s earnings, but such firms may, in fact, earn surpluses. These funds may be reinvested in the organization, kept as endowment, or used for other charitable purposes.”¹⁹ Furthermore, governmental statutes commonly prescribe “reasonable” operating costs for nonprofit organizations thereby limiting the margins of managerial and employee compensations and benefits.²⁰ Therefore, regulations determine the allocation of potential producer surpluses, as well as the origin of the nonprofit firms’ costs.

Now a valid reason why nonprofits traditionally inhabit the humanitarian aid sector can be discovered by looking at the high dependency of emergency relief on private donations. Who is better at attracting donations? It is not difficult to see that many donors will have ideas about the best use of funds they provide. The ideological foundation of most nonprofits, as well as the diversity of their respective philosophies, make this type of organization capable of satisfying such wishes: “Nonprofits are better positioned than either governmental or for-profit firms to serve this reification purpose. Government agencies are usually too constrained by legislative mandates and demands for uniformity, and for-profits must satisfy a market survival test.”²¹ Arguably the diversity in philosophies among agencies serves a positive function for accommodating donors’ equally diverse views.²²

Secondly, as Rose-Ackerman suggests, their non-distribution constraint makes nonprofit organizations preferable to profit-driven organizations when donors are faced with a choice between the two.²³ Easley and O’Hara depict quite vividly:

[I]n buying food for our own consumption we can observe the products we are purchasing. In donating money to feed Ethiopian famine victims, however, we are not likely to travel to Ethiopia to observe the delivery of the food. The nonprofit’s nondistribution

constraint assures us that our entire donation will not end up in the manager's bank account. If we donate to a for-profit firm, however, what assurance do we have that the money provides food to Ethiopians rather than profits to the owners?²⁴

Even though for-profits could imitate the non-distribution constraint through specification of limits to profit-distribution in private contracts, high transaction costs of establishment and enforcement still render the for-profit disadvantageous vis-à-vis the nonprofit.²⁵

Nonprofit organizations are better at accommodating donors' interests insofar as ideological criteria direct the use of financial contributions. People's common, intuitive uneasiness to put for-profit actors on a par with nonprofit actors in the industry is further caused by the fear that their voluntary donations are turned into profit rather than used for the intended charitable purposes. The legal statutes regulating the behavior of nonprofit firms, particularly the "non-distribution constraint," assure that at least there will be sanctioning of such behavior, which, in addition, does not involve such high transaction costs as private contractual devices may. Both arguments not only illustrate that, through their superior ability to attract funding, nonprofits traditionally inhabit the humanitarian aid industry, but also that opaqueness is the biggest obstacle to humanitarian aid financing, and the humanitarian aid enterprise as such.

Nonetheless, when humanitarian ends are taken into account, it would be insufficient to look exclusively at the donors' (justified or unjustified) preference for nonprofits, even though this seemingly explains why nonprofits rather than for-profits have come to dominate the aid sector.

Recalling that the ultimate goal of humanitarian aid is to maximize the number of lives saved, financial means should be allocated to the organizations through which they are most efficiently translated into aid output. In concrete terms, this means that if one donated US dollar in the hands of a for-profit firm generates a higher return in terms of lives saved than that dollar in the hands of a nonprofit firm, even if it was not used according to our (i.e. the donors') ideological conception of how to employ it, and/or profits were deducted from the donation, we still should choose the for-profit firm.

In light of the preceding considerations, our central concern thus becomes the effect of the organizational cultures of for-profit versus nonprofit firms on their performance in the face of the ever-present informational asymmetries in the humanitarian aid sector.

NONPROFIT VS. FOR-PROFIT ACTORS: PREDISPOSITIONS IN THE FACE OF INFORMATIONAL ASYMMETRIES

The market structure of the humanitarian aid enterprise, along with its inherent informational asymmetries, yield the question of whether nonprofit and for-profit actors respond differently to their often misaligned interests with donors and the resulting threats of moral hazard and adverse selection. Both for-profit

and nonprofit actors are subject to the “temptation to cheat” donors who cannot control the end product that they finance.

As I argue below, each type of organization indeed shows distinctive behavior due to the characteristics of its culture and accompanying incentive structures. Given that neither resists moral hazard, I assert that they do not compromise the efficiency of their aid delivery to the same degree, nor in the same way, but are guided by their respective organizational predispositions.

First, it seems reasonable to expect that nonprofit actors have “good intentions,” i.e. a moral or ideological commitment to delivering aid to those in need. Employees of nonprofit firms are likely to be “ideologues,” i.e. persons “with strong beliefs about the proper way to provide a particular service.”²⁶ The renunciation of profits, and the resulting lower pay among the employees, suggests that there must be a moral reward making up the difference between the for-profit and the nonprofit salary. Otherwise, employees in the nonprofit sector would simply move to the private sector.

This view is confirmed by recruitment theories.²⁷ Equally, empirical evidence exists for the claim that “high level professional employees may accept lower levels of pay in return for greater certainty that their efforts are actually helping to achieve their altruistic goals.”²⁸ Whereas “entrepreneurs preoccupied with profits apply their talents elsewhere,”²⁹ ideologically or morally committed employees are more likely attracted by nonprofit firms, since “the lack of equity holders is a signal to employees that their selflessness is not enriching someone else.”³⁰ It becomes apparent that the nonprofit employees’ interests are more closely aligned with those of the donors than the for-profit employees’ interests are.

Accordingly, in the absence of donor control, we may expect the nonprofit employees’ moral and ideological commitment to counteract efficiency-compromising behavior induced by informational asymmetries. On the other hand, we have no grounds to assume such a behavioral corrective among employees of for-profit organizations.³¹ Nevertheless, this claim does not posit that employees of nonprofit organizations are resistant to prestige or do not engage in rent-seeking at all. Whereas ideological commitment counteracts efficiency-compromising behavior, it does not eliminate it. What it says, though, is that ideological commitment mitigates such behavioral tendencies.

Intuitively, “the nonprofit form of institution [...] produces an incentive structure that leaves managers more neutral as between decisions that would increase profit and decisions that would do more to meet social needs, such as increasing aid to the poor.”³² The nonprofits’ legal constraint of non-distribution acts as a “broader contract” supplementing the discipline of the market, and improving the incentive structure.³³ Thus, “[f]rom the viewpoint of modeling economic behavior [...] under conditions of informational asymmetry, nonprofits may be valuable precisely because they face the constraint on distribution of profit.[...This constraint] is generally counterproductive when information is widely available, but it can play a useful social role when informational asymmetries are large.”³⁴

However, this is not to say that the nondistribution constraint could simply substitute for the function of the ideological or moral commitment we find among nonprofit employees. Indeed, in the United States for instance, the state devotes few resources to regulation and enforcement of nonprofit corporation law.³⁵ That abuses still remain the exception confirms that “ethical constraints may be far more important than legal sanctions in causing the managers of nonprofits to adhere to their fiduciary responsibilities.”³⁶

Given that, due to informational asymmetries, both nonprofit and for-profit actors encounter weak incentives to exert effort in actually delivering their services efficiently, nonprofit actors will compromise efficiency to a lesser degree than for-profit firms. Therefore, nonprofits not only are an “institutional form potentially able to deal with asymmetric information,”³⁷ but may evolve precisely in response to “a market niche when information asymmetry and trust are important.”³⁸

Significantly, there is quite probably another difference between nonprofit and for-profit actors in their response to informational asymmetries. Apart from whether they compromise the efficiency of aid delivery to the same degree, it is a particularly interesting question to ask whether nonprofit and for-profit actors compromise efficiency in the same way. As mentioned before, in speaking of “efficiency,” we refer to maximizing output, i.e. the delivery of aid, at a given cost. To hold “costs” constant for comparative reasons,³⁹ should not obfuscate that there are two variables involved in efficiency considerations regarding humanitarian aid: the output, i.e. quality and quantity of aid delivered, as well as the input, i.e. the cost of providing aid. In other words, efficiency comprises the provision of high quantities of quality aid at a low cost.

I argue that due to their predisposition, nonprofit actors are more likely to compromise on the cost side. By virtue of their ideological or moral commitment, they will avoid compromising the quality and quantity of their aid delivery as much as possible. But, they will produce it at a higher cost than necessary, and they are able to do so given the inherent lack of transparency in the humanitarian aid enterprise.⁴⁰ Nonprofit firms, by virtue of their organizational structure and the employees they attract, thus have a “quality control advantage.”⁴¹ Empirical findings suggest that “higher quality comes at a cost:”⁴² “[W]hen nonprofits provide higher quality services, they may also charge higher prices to compensate, reducing their advantage.”⁴³ Accordingly, while nonprofits are less prone than for-profits to give in to potential moral hazards and to compromise the quality of their output, they still display a higher tendency to engage in “adverse selection” of aid programs, as recognized by Easterly.

I contend that, by contrast, for-profit actors, not constrained by any sort of ideological or moral commitment, and that are by definition, profit-seeking, are more likely to operate at a comparatively lower cost. Nevertheless, they will tend to compromise their produced output, particularly in terms of qualitative characteristics that involve variables that are difficult to measure and difficult to define in a contract.⁴⁴ Indeed, studies provide evidence that on average “for-profits [sic] firms provide lower quality and less expensive services.”⁴⁵ If “purchasers of a firm’s output cannot ascertain whether the output is actually produced,”⁴⁶ output

quantity likewise may belong to this hard-to-measure category.

To summarize, we can state that the impact is dependent on the organizational form, i.e. nonprofit versus for-profit firms, in response to informational asymmetries inherent to the emergency relief sector. While it is reasonable to believe that in the face of moral hazards, nonprofit firms compromise the efficiency of their services to a lesser degree than for-profits, they are more likely to engage in adverse selection by overstating costs of particular programs they promote. At the same time, for-profit actors appear to display a more careful attitude toward the use of resources meaning they tend to be more cost-conscious. However, whereas nonprofits have a quality control advantage, for-profit organizations are more prone to compromise output-variables that are hard to measure, as is most probably the case concerning quality.⁴⁷ As defined, lower quality lessens the impact of the aid delivered.

In light of these findings, what conclusions follow? Are for-profit providers able to provide a remedy for persistent inefficiencies of the emergency relief industry? I suggest that despite the negative aspects of for-profit actors' performance spelled out in the preceding discussion, these very considerations should help us to identify the conditions under which for-profit actors *can* contribute to the ultimate goal of maximizing the number of lives saved. Moreover, they should highlight with *what kind* of inefficiency a "true" humanitarian will be more concerned.

WHAT FOLLOWS? MAKING THE CASE FOR FOR-PROFIT ACTORS

In the case of informational asymmetries stated above, it appears that nonprofit organizations compromise efficiency at the input stage, while for-profit organizations compromise it at the output stage. Vis-à-vis the goals of humanitarian intervention, if we take it as the declared aim of the humanitarian enterprise to save the maximum number of lives *and* we cannot eliminate the inefficiencies inherent to the industry, we might do better to accept higher costs to achieve that aim rather than compromise the aim itself. That is to say, given the alternatives, we care more about the output than the input of humanitarian aid.

However, referring to our initial considerations, we have in fact identified the cases in which we must neither accept higher costs, nor compromise the humanitarian aim itself. By looking at economic incentive structures, the preceding argument has revealed valid reasoning, why and when we should actually prefer for-profit actors, and why and when not to do so. "[P]rivate firms are the preferred institutional mechanism when all of the desired characteristics of a commodity or service can be stated in a contract and monitored easily."⁴⁸ Therefore, "[t]he more important the unmeasurable dimensions, [...] and the more complex and ambiguous the proposed measures, the stronger is the social efficiency case for the nonprofit alternative."⁴⁹ Moreover, "there may be a case for [...] [the nonprofit] that explicitly divorces rewards from the easily measured aspects of performance when those measures do not accurately reflect the quality of the output."⁵⁰

In other words, we can in fact now estimate *when* our donated US dollar in the

hands of a for-profit firm is likely to be more efficient than the same dollar in the hands of a nonprofit. The common, deeply moralized discussion about the involvement of for-profit firms in the humanitarian enterprise depicts a “self-imposed intellectual trap”⁵¹ that hinders the achievement of such important insights, and contributes little. In the words of Hansmann, “[it] is commonly suggested that certain services [...] are organized on a nonprofit basis because it would be unthinkable to have people profiting from the misery and ill health of others or to leave the availability of such vital services in the hands of profit-seeking entrepreneurs. Such reasoning is unpersuasive. [...] The central factor, rather, seems to be contract failure.”⁵²

In line with these conclusions, I would see a limited yet distinct role for for-profit service providers where it is possible to state quality dimensions and the requisite accountability in contracts, and/or where qualitative dimensions may be dominated by cost considerations. For instance, there is no human dimension to the service provided during the logistical stage of humanitarian intervention, and few non-accountable qualitative variables in general. Consequently, for-profit actors might well be more efficient in this area.

Another area where for-profit actors can contribute to capacity is disaster risk management. This specifically illustrates how moral preconceptions could lead the sector to miss out on innovative approaches with potentially high benefits. As there are no informational asymmetries at all, according to the argument developed, involvement of for-profit actors is completely unproblematic. Disaster risk management includes predicting the scale and consequences of disasters, like earthquakes and droughts, and the financial means needed to mitigate their impact on livelihoods. Insurance instruments such as index-based weather derivatives or catastrophe bonds are used to transfer disaster risk to private capital markets.⁵³ Essentially, these novel insurance instruments are based on physical triggers that can easily be stated in a contract, and allow quasi-perfect monitoring.⁵⁴ Hard-to-measure variables are not involved. Once specified in a contract, the for-profit derivative or bond investor cannot influence the output defined in pure monetary terms. At the same time, the capital market, through the respective prices of the insurance instruments, grants transparency about “inputs,” or costs at given benefits. The very absence of informational asymmetries between donor and for-profit contractor, as well as potential market demand, make the donor-subsidized provision of risk insurance on the micro and macro level a promisingly efficient future supplement to traditional aid programs.⁵⁵

In conclusion, moral hazards and contractual complexities among humanitarian aid delivery agencies are not resolved by the entry of profit-driven agencies. In areas of the industry such as logistics or insurance-based disaster risk management, it is possible, and probable, that the existence of for-profit firms may spur the efficiency of the enterprise, and therefore contribute to humanitarian aims.

NOTES

¹ “...traditional Republican scepticism about the utility of international aid remains undiminished and [...] other delivery channels – for example those provided by the private sector [...] might take a market share from “traditional” aid deliverers if they can demonstrate greater impact and efficiency (and presumably malleability) than the conventional aid system.” Nicholas Stockton, “The Changing Nature of Humanitarian Crises,” *The Humanitarian Decade* (United Nations Office for the Coordination of Humanitarian Affairs: United Nations Publications, 2004) p. 34.

² C.K. Prahalad, *The Fortune at the Bottom of the Pyramid* (Upper Saddle River, NJ: Wharton School Publishing, 2005) p. 9.

³ Michael Harford, Bitá Hadjimichael, and Michael Klein, „Aid Agency Competition,” *Public Policy for the Private Sector*, Note No. 277, The World Bank Group Private Sector Development Vice Presidency (October 2004): p. 3.

⁴ Henry B. Hansmann, “The Role of Nonprofit Enterprise,” *Yale Law Journal* Vol. 89 (April 1980): p. 847.

⁵ William Easterly, “The Cartel of Good Intentions: Bureaucracy vs. Markets in Foreign Aid,” *Policy Reform* Vol. 00, Center for Global Development (Washington 2002): p. 22.

⁶ *Ibid.*, p. 22.

⁷ (Cooley and Ron, 2002:15) Alexander Cooley and James Ron, “The NGO Scramble: Organizational Insecurity and the Political Economy of Transnational Action,” *International Security* 27, No.1 (Summer 2002): p. 15.

⁸ McKinsey & Co: *Delivering the Promise of Humanitarian Assistance: The Role and Performance of the UN – Appendix A (Exhibits)*, Inter-Agency Working Group on Humanitarian Aid Coordination (London: October 2002). As Easterly specifies, “[p]rincipal-agent theory says having multiple principals dramatically weakens the incentives for the agent (the international agency). “[...] Having many agents monitored by the many principals makes things worse, as the joint product of the many agents makes it hard to evaluate the efforts of any one agent.” (Easterly, p. 23)

⁹ Cooley and Ron, p. 16.

¹⁰ Needs-based assessment, e.g., is by far not an indisputable procedure to increase the impact of humanitarian aid. It may place incentives for aid-recipients in their struggle of survival to demonstrate non-existent needs and “fool” the scheme applied by the aid agency.

¹¹ *Ibid.*, p. 6.

¹² Robert S. Pindyck and Daniel L. Rubinfeld, *Microeconomics*, 6th edition (New Jersey: Pearson Prentice Hall 2005) p. 306.

¹³ Of course, one could argue, that an increased price of aid-/service-units, i.e. less quantity provided at the same price, might result in decreasing demand on the

donors' side to purchase aid. Yet as this does not contribute to the subject of my following investigation, I will not further discuss this point.

¹⁴ As Rose-Ackerman puts it: "The impact of organizational form on firm performance and survival depends upon the interaction between institutional structure and other features of the environment such as entrepreneurial motivations [...] and the overall competitive environment. No organizational form is invariably superior, but one can isolate factors that will favor one form over another." Confer Susan Rose-Ackerman, "Altruism, Nonprofits, and Economic Theory," *Journal of Economic Literature* 34, No.2 (June 1996), pp. 701/2.

¹⁵ See, for instance, Paul J. DiMaggio and Helmut K. Anheier, "The Sociology of Nonprofit Organizations and Sectors," *Annual Review of Sociology* 16 (1990), pp.137-159.

¹⁶ David Easley and Maureen O'Hara, "The Economic Role of the Nonprofit Firm," *The Bell Journal of Economics* 14, No.2 (Autumn 1983), pp.531. Nonprofit firms "are encapsulated in a relatively small, albeit important, set of industries [...] in the United States, hospitals, daycare, museums, universities, home health care, social services, broadcasting, and a few others" (Dimaggio and Anheier, p. 139) and "virtually all nonprofits [...] are producers of services" (Hansmann, p. 872).

¹⁷ "Nonprofits, like their profit-oriented counterparts, use inputs – labor and capital – to produce outputs – goods and services. Viewed this way, it is apparent that nonprofits are firms and that whether a firm is for-profit or nonprofit does not determine the output it is capable of producing." (Easley and O'Hara, pp. 531/2).

¹⁸ Easley and O'Hara, p. 532.

¹⁹ Rose-Ackerman, pp. 715/6). For a more detailed account, confer Hansman, pp. 838-40: "What makes an organization nonprofit?"

²⁰ Easley and O'Hara, e.g., in order to distinguish nonprofit from for-profit firms in their contracting model, simply differentiate according to whether the managerial, employees' compensation respectively is specified (nonprofit), or unspecified (for-profit). (Easley and O'Hara, p. 532.)

²¹ Rose-Ackerman, p. 717. The same advantage of nonprofits holds vis-a-vis UN agencies. As Minear and Smillie report: "National NGOs, in their advocacy as well as their delivery modes, are often more attractive to donors than more remote, bureaucratic, and sometimes unpopular UN organizations "The Quality of Money: Donor Behaviour in Humanitarian Financing," *Humanitarianism and War Project*, ." Larry Minnear and Ian Smillie (April 2003) Feinstein International Famine Centre, Tufts University, p. 11. <http://hwproject.tufts.edu/new/pdf/donor_behav.pdf>

²² See, for example, Caroline Moorehead, "Speak No Evil" *Financial Times* 19 June 2005; Mary Anderson, *Do no Harm: How Aid Can Support Peace – Or War*, (Boulder: Lynne-Reinner 1999); or David Rieff, *A Bed For the Night* (New York: Vintage 2002).

²³ "Given the difficulty of monitoring charitable work, donors may fear that for-profit firms will convert gifts into profits for their owners. Although nonprofit managers can, of course, embezzle or misuse funds, at least using a nonprofit for

private gain is illegal.” (Rose-Ackerman, p. 716.)

²⁴ Easley and O’Hara, p. 532. See also Hansmann’s account led by the question: “If you rely on a for-profit distributor to provide the food to feed your own children, why should you not also turn to a for-profit firm to provide the food you purchase for children overseas?” (Hansmann, pp. 846/7.)

²⁵ “Under the contractual approach, the patrons would need to insist upon receiving regular, audited financial statements from the firm containing sufficient detail to permit the patrons to determine whether the owners are adhering to the contractual constraints. Should it appear to a given patron at any point that the owners are not keeping their promises, then the patron would need to bring suit to enforce the contract. In contrast, under the nonprofit form, the state is empowered to bring suit if the organization’s management compensates itself too generously.” (Hansmann, p. 853.)

²⁶ Rose-Ackerman, p. 719.

²⁷ “Recruitment theories hold that NPO [nonprofit organization] employees are willing to work for less because their values differ systematically from those of employees of FP [for-profit] firms: [...] NP [nonprofit] employees value job quality more and wages less than FP staff.” (Dimaggio and Anheier, p. 148.)

²⁸ Rose-Ackerman, p. 720.

²⁹ Dimaggio and Anheier, p. 140.

³⁰ Ibid.

³¹ For proof, Rose-Ackerman for instance asserts that empirical evidence collected by Weisbrod shows that, e.g., “nonprofits providing long-term care generally take less advantage of their informational superiority over customers than do for-profits” (Rose-Ackerman, p. 722). While long-term care and post-disaster assistance are quite different situations, decisively they carry the same characteristics of informational asymmetry on which my argumentation rests.

³² Burton A. Weisbrod, “Rewarding Performance that is Hard to Measure: The Private Nonprofit Sector,” *Science*, New Series 244, No.4904 (May 1989): p.543. “Unlike private enterprise, such an institution [the nonprofit organization] does not reward sellers who take advantage of consumers’ informational handicaps.” (Ibid., pp. 541/2.)

³³ “The nonprofit producer, like its for-profit counterpart, has the capacity to raise prices and cut quality in such cases [of informational asymmetries] without much fear of customer reprisal; however, it lacks the incentive to do so because those in charge are barred from taking home any resulting profits. In other words, the advantage of a nonprofit producer is that the discipline of the market is supplemented by the additional protection given the consumer by another, broader ‘contract’, the organization’s legal commitment to devote its entire earning to the production of services.” (Hansmann, p. 844.)

³⁴ Weisbrod, p. 542.

³⁵ Hansmann, pp. 873-4.

³⁶ Ibid., p. 875.

³⁷ Ibid.

³⁸ Rose-Ackerman, pp. 722. Concerning the humanitarian aid enterprise, one should, however, not overstate the case for nonprofit organizations. A vast literature refers to the negative effects and inefficiencies of the industry due to the lack of transparency and accountability. Following the subject of the essay, the reader should not forget that the present argument specifically aims at taking a comparative perspective of nonprofit *versus for-profit actors* in emergency relief.

³⁹ As explained, this seemed appropriate in the case of humanitarian aid, as donors will make contributions at their reserve price.

⁴⁰ Weisbrod confirms this view: "If a seller is able to benefit from added profit, the incentive exists to "chisel" – to provide lower quality than was promised. If, however the seller is effectively prevented, by the nondistribution constraint, from reaping benefits from any profit generated through underproviding quality, the incentive to chisel is weakened. At the same time that this adverse incentive is reduced, however, so, too, there is a weakening of the manager's incentive to be efficient." (Weisbrod, p. 543.)

⁴¹ Rose-Ackerman, p. 719.

⁴² Ibid., p. 722.

⁴³ Ibid.

⁴⁴ Hansmann calls this case a "contract failure" and finds nonprofit organizations to be the market response to "police producers by ordinary contractual devices" (Hansmann, 1980: p. 845.)

⁴⁵ Rose-Ackerman, p. 722 (in this specific example, referring to child care). In terms of economic theory, "[w]hen quality is costly for buyers to monitor, provision of higher quality is financially unrewarded and, hence, is unprofitable, however socially valuable it might be." (Weisbrod, p. 544.)

⁴⁶ Easley and O'Hara, p. 532.

⁴⁷ For a very interesting, economic model of explanatory strength see Hansmann, pp. 899-901 (Appendix). The model treats the screening of ideologically committed employees, managers respectively, as well as the quality signaling of the nonprofit organizational form alongside for-profit actors and implications for the market.

⁴⁸ Easley and O'Hara, p. 531.

⁴⁹ Ibid.

⁵⁰ Weisbrod, p. 542.

⁵¹ Prahalad, p. 9.

⁵² Hansmann, pp. 880-1. Italics mine.

⁵³ Index-based weather derivatives are "contingent contracts with a payoff determined by future weather events, such as a specified lack of precipitation measured at a weather station" (Joanne Linnerooth-Bayer, Reinhard Mechler and Georg Pflug, "Refocusing Disaster Aid," *Science*, Vol.309 (August 2005) pp.1045). A catastrophe bond can be explained as "an instrument whereby the investor receives an above-market return when a specific catastrophe does not occur within

a specified time [...] but sacrifices interest or part of the principal after the event.” (Ibid.)

⁵⁴ Linnerooth-Bayer, Mechler and Pflug, p. 1045. Given investment in new technologies that electronically convey weather data, and prevention of bribery of reporters of weather station readings, respectively. (Jerry Skees, et al., “Can Financial Markets be Tapped to Help Poor People Cope with Weather Risks?” in Stefan Dercon, (ed.) *Insurance Against Poverty* (Oxford: Oxford University Press, 2005) pp. 432-3.

⁵⁵ Skees, et al., p. 422.

The Case for Cash: Delivering Aid Through Direct Payments to the Poor

Fatima B. Kassam

While it is quite common in the West to provide monetary assistance to people in emergency situations, the practice is highly underutilized in developing countries. Under certain conditions, giving cash is one of the most cost-efficient methods of delivering assistance to a high number of people in a short amount of time while engendering future economic growth. This paper counters the arguments of skeptics and calls on key figures in the industry to reflect upon the paternalism and inefficiency that they may continue to foster through their institutions.

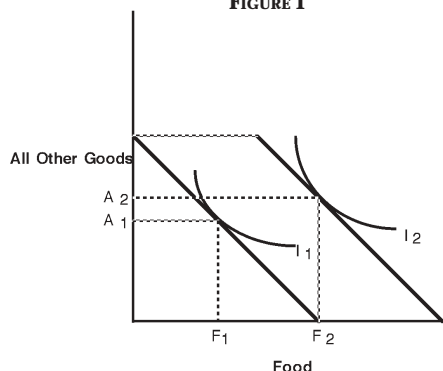
While it is quite common in the West as a part of relief action or through insurance companies to provide monetary assistance to people in emergency situations, the practice is highly underutilized in developing countries. When appropriate, a cash-based approach allows beneficiaries of aid to exercise more choice; is usually more cost-efficient for both providers of assistance and recipients; and can have positive short- and long-term effects on the economy of the affected area. However, donor agencies list a myriad of reasons why this approach cannot work in the developing world. Inflationary and security risks, while they have not been thoroughly explored, pose real concerns for both providers of aid and affected populations. Critics also cite the lack of control over where the money is spent and reservations about how the social structure of a community will affect the distribution and control of currency. These challenges affect the plausibility of expanding the scale of cash-based interventions – but they are surmountable. Though not perfect, the cash-based approach is becoming a popular alternative to providing not only food assistance but also other forms of relief, particularly in places where NGOs overlap in the services they offer. Under certain conditions, giving cash is one of the most cost-efficient methods of delivering assistance to a

Fatima B. Kassam is an M.A. candidate at the Bologna Center concentrating in international development. Her previous work includes economic development in Mozambique and South Africa, legislative analysis and advocacy on trade and security in the United States, and work with nonprofit organizations around civic engagement and political strategy.

high number of people in places that have experienced either man-made conflicts or natural disasters while engendering future economic growth. This paper will examine the effects of the infusion of cash by taking a closer look at the cases of Afghanistan¹ and Mozambique and the conditions necessary for successful cash-based approaches to relief elsewhere in the world.

THE ECONOMIST'S BASIS FOR CASH

FIGURE 1



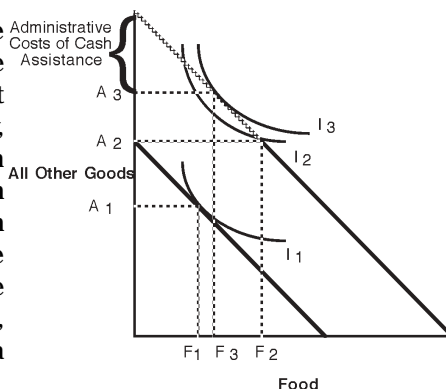
With only a simple graph, economists have long helped aid workers to justify the shipment at a moment's notice of millions of tons of provisions and other materials into post-disaster or post-conflict areas. The argument follows that food aid results in a better-fed population that enjoys a higher level of utility.² As shown in Figure 1, given that an individual has a limited amount of disposable income to spend, one can construct a line exhibiting this budget constraint. If he spends all of it on food, he arrives at the intercept of the x-axis.

Likewise, if he spends all of it on other goods, he arrives at the intercept of the y-axis. Food relief shifts the budget line to the right, parallel to the original. After the shift, the individual can obtain more food on the same budget and redirect savings to consumption of other goods. With food aid, an individual consumes more of everything and can move to a higher indifference curve, and therefore a higher utility.

If the goal of delivering assistance is a higher nutritional intake of the affected populations, this approach makes sense. However, despite the improvement in the situation, this is not always the most effective use of aid monies. If the goal is to efficiently spend resources available in a way that will maximize the utility of the recipient – that is, in a way that satisfies as many of his needs as possible – then cash is a better alternative.

FIGURE 2

The fungibility of cash allows the new budget line to extend upwards (see Figure 2). With food aid alone, the highest indifference curve achievable is I_2 . Clearly, the recipient is better off with aid than without it, consuming higher levels of both food and other goods. However, with cash, utility can be maximized on the curve I_3 . As exhibited in Figure 2, at the intersection of I_3 and the new budget line, the optimal point of consumption for a



beneficiary, it is possible that he consumes less food (point F_3) than with food aid (point F_2), but more of other goods.

With what goods might beneficiaries replace food? This is a point of controversy. Those in the humanitarian aid industry who are unfamiliar with evaluations of cash-based interventions often, in an effort to err on the side of caution, err instead on the side of paternalism. Interviews with aid workers have revealed that many of them worry that adequate dietary intake for a family may be sacrificed by male heads of household for alcohol consumption or dowries for new wives.³ Later sections of the paper will show that such notions have been negated with empirical evidence from both Afghanistan and Mozambique.

THE PRACTITIONER'S BASIS FOR CASH

Thoughts on food insecurity have evolved throughout the past few decades. In the 1970s, the focus was on global and national supply shortfalls, while in the 1980s it became clear that the fundamental problem was not food scarcity, but food distribution. Availability at the national level did not necessarily translate to access at the household level. Amartya Sen explained this disconnect by introducing the concept of food "entitlement" in a 1981 publication. "Entitlement refers to the set of income and resource bundles available (e.g. assets, commodities) over which households can establish control and secure their livelihoods."⁴ This new understanding of food-security allowed for consideration of a number of socio-economic factors not previously given as much attention.

Research throughout the 1990s focused on the relationship between malnutrition and poverty. More specifically, practitioners observed how the poor absorb various risks and found that food is only one of a whole range of factors that determine why people in desperate situations make certain choices. Balancing competing interests in order to prolong subsistence in both the short- and long-term was found to be more complex than previously thought. Many leading organizations have adopted this latest view that food security is a fundamental need but not independent of a broader number of factors. CARE USA calls this "household livelihood security" and defines it as:

... Adequate and sustainable access to income and resources to meet basic needs (including adequate access to food, potable water, health facilities, educational opportunities, housing, time for community participation and social integration) ... therefore livelihoods are secure when households have secure ownership of, or access to, resources and income earning activities, including reserves and assets, to offset risks, ease shocks and meet contingencies.⁵

Both food aid and cash can prevent or delay risk-absorption methods that might be detrimental to livelihoods, like the sale of assets to cope with hunger. However,

cash allows recipients more flexibility, and in turn a higher level of household livelihood security.

A CASH FAMINE IN AFGHANISTAN

“On behalf of the people of Afghanistan, we would like to make an appeal ... The humanitarian transition must be marked by an emphatic shift from kindness “in-kind” to opportunities “in cash.”

-Minister of rural reconstruction and development in Afghanistan

In post-Taliban Afghanistan, severe drought over a three- to four-year period in some regions has exacerbated the economic effects of the US-led invasion of the country and the subsequent change in government. In spite of the security risks posed by protracted warfare in some regions, cash-based approaches have not only worked well, but are preferred by many of the international agencies currently operating in the country and the beneficiaries of their programs. The findings of a Tufts University report prepared in 2002 for USAID found what the authors described as a “cash famine” in Afghanistan. A number of factors contributed to this phenomenon: loss of crops and livestock after the drought; depression of wages in a labor market flooded with new social groups, such as women, now being allowed to work and refugees returning from neighboring countries; and the stress on credit institutions from non-performing loans and failing currency markets. The report states:

More people need access to markets to achieve food security than ever before yet fewer people have the cash resources necessary to buy goods in the market... Throughout Afghanistan, there are crises of purchasing power, production and credit that continue to directly threaten household food security.⁶

Though the combination of events leading to the widespread loss of disposable income in Afghanistan may be particular to that country, seizure of land and livestock, forced migration, and stress on traditional financial institutions is not uncommon to conflict zones more generally and can bring about the same negative effect.

Though well-intentioned, agencies delivering food aid can actually hurt local producers and traders. It is true that food aid has saved lives, deterred families from migrating out of desperation, and even protected them from falling further into debt. This is the case for particularly vulnerable households with no surplus labor to contribute to Cash-for-Work programs and other such forms of assistance. However, sales of commodities distributed by relief organizations are widespread.⁷ This monetization not only indicates a preference for cash or goods other than those distributed, but also depresses prices of goods that can be produced and supplied locally. Wheat prices in Nangahar, for example, fell by 15 to 20 percent as a direct result of the World Food Program's presence in the area. It was confirmed through focus groups that this price depression served as a disincentive for farmers who would normally try to cultivate surplus wheat for sale.⁸ The non-distortionary

nature of cash would guard against this change in both producer and consumer behavior. Inefficiently targeted food aid, in this case, could potentially set back replenishment of household livelihood security.

BEYOND BASIC NEEDS IN MOZAMBIQUE

One year prior to the bombing campaign in Afghanistan, devastating floods hit the southern African country of Mozambique, causing 700 deaths and displacing approximately half a million people in the Limpopo, Save, and Zambezi valleys. Having examined the market conditions and the availability of food supply, USAID Mozambique created the Resettlement Grant Program (RGP), which provided 1,500,000 meticaís (about \$92) in cash to each of a little more than 106,000 rural families. Due to Congressional inertia, approval and appropriation of funds to USAID Mozambique was delayed until well after the flooding took place.⁹ By the time monies were made available, other NGOs had already been able to meet most people's basic needs for everyday survival. In conjunction with millions of dollars allotted to other projects aimed at increasing local production, USAID also launched the Resettlement Grant Program (RGP). The scale of the RGP is negligible compared to that of any other programs in the country. Nevertheless, the thorough evaluation of the endeavor provides practitioners with excellent insight into what recipients consumed in place of food. As program designers envisioned, grants were spent predominantly in local villages on items not available through other programs. Though purchases of food were high, most of the money was spent on household goods, clothes, and livestock. Expenditures on seeds and construction materials were also substantial, indicating an investment in future livelihoods. Interviews revealed that some of the in-kind donations from other programs, such as zinc roofing, was often re-sold in order to purchase different building materials. Again, this monetization suggests that cash would have been more efficient for both providers and recipients of aid.¹⁰ One of the most important findings of an impact evaluation performed by Abt Associates is that grant monies almost exactly reflect the normal spending patterns of households, proving the theory that cash-based interventions are non-distortionary.¹¹

THE SECRET TO SUCCESS

The economic, social, geographic, and political landscapes of a post-conflict or post-disaster region are all important elements to consider when deciding whether a cash-based approach is appropriate and optimal. Though there has not been much published on what conditions are necessary for a successful venture, this paper pulls from current working papers on the subject and the realities of relief efforts in Afghanistan and Mozambique.

First, a thorough understanding of formal and informal market systems as well as geographical distances are key to determining whether cash will be useful. Local markets must be able to provide essentials at reasonable prices, accommodate additional demand for certain goods, and be competitive and accessible. These facts often relate to how much time has passed since the height of the emergency. Markets are often disrupted in the very short term (four to six weeks) but can be

functioning well not long thereafter. An assessment of market indicators is not as complex as it may sound. Price monitoring for key commodities is already a part of most early-warning systems; interviews with traders and shopkeepers are simple and can be conducted quickly. With the increase between 2002 and 2004 of cash-based projects in Afghanistan, some organizations there have developed their own list of indicators. Knowing the following has helped even the government make more informed choices regarding how to deliver assistance: whether bartering is an important form of exchange, the time and cost of accessing the nearest market center clinic or school, the number of months during the winter when market access is restricted, community views on cash versus food programming as a response, and so on. In a recent report published by the Humanitarian Policy Group, Harvey urges that:

... Better analysis of markets is arguably something that aid agencies should be doing anyway, in order to better understand how people are surviving and how best they can be supported. As Lautze argues, village markets are a critical arena [sic] for information, political exchange and socialising; monitoring markets should be one of the fundamentals of a livelihoods strategy.¹²

Despite recent progress, cash-based approaches have yet to be considered seriously in the range of options available to organizations. It is imperative that the industry change the current focus on resource-driven needs assessments and fold indicators, like those mentioned above, into their routine vulnerability mapping.

Social patterns in a community have bearing on targeting decisions for an organization, who will make spending decisions for a household and how the money will be spent. In Mozambique, program managers targeted families that had not experienced a second harvest, indicating that their assets were more depleted than others. Beyond this, and based solely on intuition, grants were dispersed only to female heads-of-household, under the assumption that they would have more incentive to spend on children and to buy items that would benefit the entire household.¹³ In evaluating the program, Abt Associates found this hypothesis to be true – for the most part. There were a small number of cases in which a wife simply gave the money to her husband and thus was not able to report on what it was spent. At least one case of domestic abuse was reported, indicating that the cash placed the women at a disadvantage, as opposed to food, which their husbands perceived gave them less power. Perhaps to the dismay of veterans in the industry, almost none of the money was spent on alcohol.¹⁴ In Afghanistan, much of the money went towards paying off debts, a grave social burden which many families bore and which had resulted in forced, premature marriage of many young girls in efforts to pay back lenders. By 2002, the practice was widespread.¹⁵ Cash in Afghanistan then restored a balance of social power that had been lost in the medium term post-conflict.

The type and place of an emergency will always influence the decision to administer a cash program. Natural disasters in relatively secure places with developed financial institutions will be far easier for implementation than remote

war zones punctuated with treacherous mountain ranges, hidden mines, and sniper fire. Despite the risks, the case of Afghanistan can be seen as proof that cash can work well, and is often preferred, in addressing complex emergencies:

... Wars distort economies. Patterns of production, employment, trade, and services shift to war-related activities and patterns... Aid can reinforce market distortions by feeding the war economy and undermining peacetime production and productivity... When aid agencies import goods that can be produced locally and distribute them at no cost, they can undermine peacetime economic incentives.¹⁶

Political power structures and relationships are routinely explored by aid agencies for the benefit of staff safety and logistics concerns. However, the implications of delivering aid in traditional ways is not always as rigorously examined. "Aid can also affect trading patterns that link people. If aid organizations import goods that were previously supplied by one group to another and those two groups are in conflict, aid reinforces the new division."¹⁷ Finding the necessary information and information that is reliable, particularly in emergency situations, in order to assess the appropriateness of a cash-based program can be risky and difficult to accomplish in a timely manner. Innovations in the area of needs assessments, however, are the next step that international aid agencies should engage in to better administer their existing programs and better serve their clients in the future.

SECURITY AND CORRUPTION

The idea that cash is necessarily more vulnerable to insecurity and corruption than in-kind donations requires further evidence. International aid agencies inevitably find a way to pay staff stationed in remote areas; no one seems to suggest that their salaries be paid to them in-kind. The fact that transport of cash is possible on a smaller scale provides some hope for larger projects. To date, studies of numerous cash-based initiatives have, in fact, shown that the opposite may be true. Although proponents of food aid often argue that the "bulky, and highly visible" nature of shipments or grain stocks make them difficult to steal, many counter this argument with empirical evidence of attacks of these assets in complex emergencies by armed groups. Mobile assets in smaller amounts can provide far more security to staff especially.¹⁸

In order to implement a cash program effectively, it is important for agencies to understand the informal financial institutions that exist. Inhabitants of an area always find ways to transfer currency somewhat safely. In Afghanistan, agencies have made use of the local *hawala* system.¹⁹ Staff from Save the Children's Cash-for-Work programs in the country stated that these programs were easier to manage than in-kind distribution projects, and that they faced less pressures to redirect funds to local commanders because the bundles of currency were in such small amounts.²⁰ Comparatively, distribution of cash in Mozambique was less straightforward and efforts toward increasing security presented enormous

logistical challenges. After the floods, program managers contracted different companies to manage security, transport, communications and financial accounting respectively. Oxfam has generated a list of simple techniques to reduce security risks for agencies, including: limiting knowledge of cash movements, limiting access to bank transactions, decentralizing disbursement responsibility, not scheduling cash transfers to keep them unpredictable, using staff who have been with the organization for a while and are trusted.²¹ From a high-level survey of evaluations (which should not be viewed as comprehensive), the only major sources of corruption seemed to be *within* the organizations administering the programs. This should not be surprising given the attraction of cash to everyone, not only those in need. Although precautions can be taken in targeting certain beneficiaries of the program, deciding who handles cash internally can be far more challenging.

FOR AGENCIES, CASH IS COST-EFFECTIVE

The cost-effectiveness of cash is usually stated relative to in-kind distributions. The comparison can be a tricky exercise due to the fact that agencies calculate overhead costs in different ways. Additionally, there are certain costs, post-delivery, that the beneficiary might incur, transport costs for example, which are not counted in overall expense assessments. These variations in accounting methods can account for the wide range in cost-savings reported by different agencies. For projects studied through the Humanitarian Policy Group, cash-based programs were estimated to be between 20 percent and 50 percent less expensive for agencies to operate. A significant example can be pulled from a study of food aid in Afghanistan which reported that wheat aid was provided at a cost of \$332 per metric ton (MT), whereas local markets could provide it at \$112/MT to \$167/MT.²² Cash was not only more cost-efficient for the agency, but probably contributed to the strengthening of local markets. It has been suggested that a portion of such cost savings be directed to enhanced security and financial monitoring to ensure proper disbursement and expenditure.

FOR RECIPIENTS, CASH PRESERVES DIGNITY THROUGH CHOICE

While economists laud cash for its ability to deliver a greater level of utility and practitioners point to a higher level of household livelihood security and cost-efficiency, beneficiaries cite dignity, self-esteem and choice as the most beneficial aspects of cash-based programs. Sen summarizes the significance of this “agency aspect” of the individual:

The success of a society is to be evaluated, in this view, primarily by the substantive freedoms that the members of that society enjoy... Greater freedom enhances the ability of people to help themselves and also to influence the world, and these matters are central to the process of development.²³

Through the optic of a relief-development continuum, “the notion that relief, rehabilitation/ mitigation and development interventions are a continuum of

related activities, not separate and discrete initiatives,”²⁴ the idea of preserving self-worth and individual agency has profound implications for how quickly an area will emerge from crisis. Understanding what might inhibit the cultivation of individual agency should be examined. The philosopher Margalit argues for the “decent society” in which institutions exhibit respect, or “non-humiliation” through their actions, not only in their theoretical origins:

Social institutions can be described in two ways: abstractly, by their rules or laws, or concretely by their actual behavior. Analogously, one can think of institutional humiliation by law, as manifested by the Nuremberg laws or those of apartheid, in contrast to concrete actions of institutional humiliation, such as the Los Angeles police officers’ treatment of the black motorist, Rodney King. In the concrete description of institutions the distinction between a noncivilized and a nondecent society is blurred.”²⁵

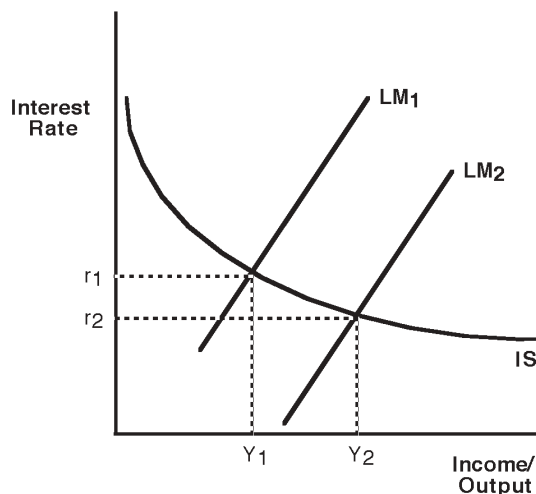
He further posits that NGOs operating in what he calls a welfare society can be compared to public institutions in a welfare state.²⁶ Paternalistic views that recipients do not know what they need, that they will squander money on alcohol or drugs such as qat and hashish, have been institutionalized through the concrete behavior of an industry that gives little credence to cash-based approaches. Harvey states that “these attitudes of paternalism and superiority remain an important factor in humanitarian response despite the professed commitment to greater participation on the part of affected populations, and that some of the reluctance to use cash is linked to these prejudices.”²⁷ It is imperative that leading agencies engage in introspection and examine the internal structures preventing a necessary paradigm shift in both the aid and development industry.

SHORT-TERM AND LONG-TERM EFFECTS ON THE ECONOMY

In the short-term, recipients of aid monies tend to spend the most in local areas, on locally produced goods that will be consumed in the near future and will bring benefit to most members of the household. Evaluations of projects in both Afghanistan and Mozambique illuminate the expenditure of a significant portion of funds on debt repayment, livestock, or other investments in future livelihood security. This stimulates local and regional markets and can strengthen the credibility of currency transactions and associated formal and informal institutions. In the long-term, the self-worth gained through an ability to exercise choice after sustained desperation and to make investments in future productivity can only have positive effects.

Critics commonly cite inflation as a significant risk. To date, there has been little evidence collected of major price inflation as a direct result of cash-based programs. Evaluations that cite an abnormal increase in prices also note very few cases of supply shortages or minor price gouging. This type of price fluctuation can be attributed to normal factors affecting market equilibrium or asymmetric information on the part of shopkeepers. To date, only smaller cash-based projects have been implemented, so inflation on a macroeconomic level has not been a truly significant risk to consider. Given the large amount of cash-based

FIGURE 3



programming in Afghanistan, it will be interesting to examine the impact on national exchange rates, inflation rates and wage rates in the coming years. The following graphs are an attempt to model the economic impact of a large-scale cash infusion, for which there is currently no empirical evidence or thorough economic analysis.

Using a Keynesian derived model of aggregate demand, one can model the effects of an increase in the money supply with the IS-LM model. According

to the model, if the money supply is increased, people have more money than they want to hold at the prevailing interest rate. As they begin to deposit extra money in banks or to buy bonds,²⁸ the interest rate falls until people are willing to hold all of the money “created” by the injection of cash into the market. This brings the money market to a new equilibrium (shift of curve from LM_1 to LM_2 as exhibited in Figure 3), lowering the interest rate and raising the level of income. In the short-run, the assumption of this model is that prices are sticky and need more time to adapt. An increase in the money supply then necessarily translates to an increase in real money balances. Thus, there is no price inflation, only an increase in income and a fall in interest rates.

IS-LM curve analysis is most useful when studying industrialized countries with strong central banks and financial sectors. Thus, for developing countries, the more relevant aspect may be in the effects of an expansionary fiscal policy. Cash disbursement on a very large scale might be compared to an increase in government purchases or a decrease in taxes. Both of these policies shift the IS curve upward, generating the opposite effect on interest rates (raising them), while still raising income (see Figure 4). This is a result of an actual shift in aggregate demand at any given price level. So, again, using the Keynesian model, there would be no changes in price level; only output would change. Assuming that the effect on interest rates would apply to both formal and

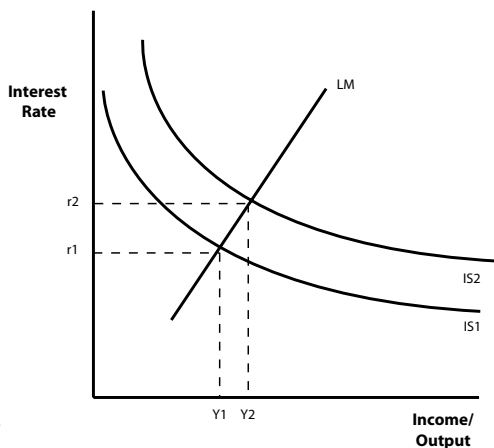


FIGURE 4

informal money markets, this may pose a significant risk to already indebted recipients. In the collection of baseline data, particularly for organizations that adopt a livelihoods-assessment strategy, level of debt security could be a crucial factor to monitor, and perhaps even more important than price fluctuation.

Both the Keynesian and classical models of short-run and long run equilibria assume that there is a natural rate of income. In the classical theory, however, prices may fluctuate, and adjust to ensure that national income/ output is always at a natural rate. Thus, according to classical theory there may be some increases in prices, but they are a reflection of increased output and income in the economy and not necessarily negative effects.

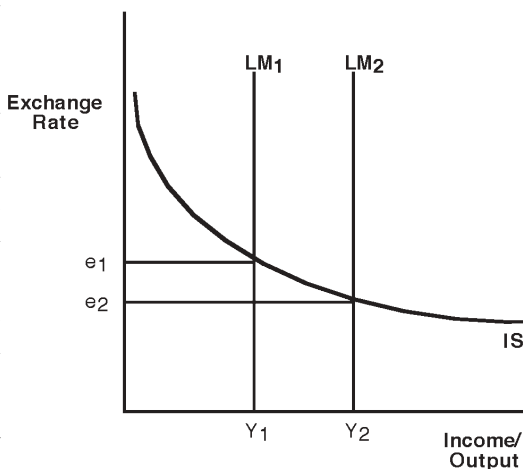
Finally, with an eye toward later development, one should also examine the IS-LM model for a small open economy (Figure 5). In this case, an increase in the money supply shifts the LM curve to the right, increasing income and lowering the exchange rate. A decrease in the exchange rate makes goods produced domestically cheaper for foreign consumption and can fuel export sectors, in particular. For cash responses aimed at fueling local production, it may be beneficial in the long run to target export sectors, after local needs are met, in order to cultivate surplus for sales abroad.

Contrary to critics who cite inflationary risks as a real concern, the theoretical modeling suggests that any significant increase in price would be temporary or a result of normal factors such as scarcity of supply. Moreover, they will adjust as income- and output-generating activity increase. What practitioners should note is the fluctuation of the interest rate, which could potentially be detrimental to recipients. For development strategists, if the country is open to trade, cash could positively affect export-oriented industries. This subject has been explored at length and there is still debate regarding which methods work best.

WFP AND THE POWERS THAT BE

Are donors and NGOs responsible for providing the most appropriate resources for meeting needs of affected populations? Judging from the current structure of the industry, it would not seem so. There is virtually no representation of cash responses in the UN's consolidated appeals process. According to researchers, "Much of this is to do with the dominant operational role of the World Food Program (WFP). WFP's mandate restricts it to using food as a resource. WFP also plays an important and often leading role in vulnerability assessments,

FIGURE 5



which often lead, with little consideration of alternatives, to food aid.”²⁹ Though not explored in this paper, there is valid reason to believe that the continued tying of food aid to donor country surpluses, particularly the United States, accounts for the dominance of food aid; it is simply what is available. Revision of the Emergency Food Assessment Handbook and recent publications by Edward Clay seem to suggest some willingness to change within the UN. Outside of the UN, NGOs that have adopted livelihoods-assessment methodologies appear to be much less threatened by cash responses.

Perhaps it is too simplistic an approach to characterize agencies as threatened or nonthreatened. While there are certain barriers in the architecture of the humanitarian aid industry that inhibit change, it is important to note that this new approach also requires skill sets that are not readily available among staff in the industry. Logisticians might lack an understanding of financial accounting, social workers do not necessarily know how to evaluate weather risks, development economists do not always have a strong security background, etc. In order to fully realize the benefits of cash-based approaches, the internal make-up of the industry must change significantly through a re-evaluation of staffing needs and innovations in development strategy. Given the benefits of this form of assistance, key figures in the industry might find it worth such an investment to curb the inefficiency and paternalism they may unknowingly continue to foster by not choosing an approach that, for beneficiaries of humanitarian assistance, preserves dignity through choice.

NOTES

¹ There is much dispute as to whether Afghanistan is “in conflict” or in a “post-conflict” transitional phase. The political implications associated with both labels can change how aid is administered and which models of assistance delivery that have been implemented there are relevant to this paper. Although this discussion of labels is outside the scope of this piece, it is important to note that for this paper Afghanistan is considered a post-conflict region.

² This paper will contrast food aid with cash, however, the analogy can be extended to other types of in-kind donations like building materials, clothing, blankets etc. The x-axis in the example would measure the level of the donation being considered and food would be folded into the “All Other Goods” category on the y-axis.

³ Paul Harvey, “Cash and Vouchers in Emergencies,” *Humanitarian Policy Group*, 10 October 2005 <www.odi.org.uk>.

⁴ “The Household Livelihood Security Concept,” T.R. Frankenberger and M.K. McCaston (1998), FAO Corporate Document Repository, p. 6, 4 January 2006 <www.fao.org>.

⁵ *Ibid.*, pp. 1-3.

⁶ “Qaht-E-Pool: A Cash Famine,” Sue Lautze, et. al., (2002), United States Agency for International Development, Washington, p. 10, 7 January 2006 <http://nutrition.tufts.edu/pdf/research/famine/cash_famine.pdf>.

⁷ *Ibid.*, p. 39.

⁸ Ibid., p. 39.

⁹ Information garnered from personal interviews with RGP staff.

¹⁰ "Mozambique 1999-2000 Floods Impact Evaluation: Resettlement Grant Activity," Maputo: Abt Associates, (2002), United States Agency for International Development, pp. 30-35, 27 October 2005 <http://www.sarpn.org.za/documents/d0000811/P907-Mozambique_floods_1999-2000_USAID_072002.pdf>.

¹¹ Though vouchers and cash programming are often studied together, vouchers are by nature distortionary and not advocated in this paper. This important difference is not emphasized enough.

¹² Harvey, p.26.

¹³ Personal communication with RGP staff.

¹⁴ Abt Associates, pp. 22-24.

¹⁵ Lautze, pp. 45-48.

¹⁶ Mary B. Anderson, *Do No Harm: How Aid Can Support Peace – or War* (London: Lynne Reinner Publishers, Inc., 1999) pp. 42-43.

¹⁷ Ibid., p.44.

¹⁸ Harvey, pp. 39-42.

¹⁹ "Cash Transfer Programmes in Afghanistan: A Desk Review of Current Policy and Practice," Charles-Antoine Hofmann (2004), Overseas Development Institute, p. 6, 5 January 2006 <www.odi.org.uk>.

²⁰ Harvey, p. 38.

²¹ Ibid., p. 39.

²² Ibid., p. 35.

²³ Amartya Kumar Sen, *Development as Freedom* (New York: Alfred A. Knopf, Inc., 1999) p.18.

²⁴ Frankenberger and McCaston, p.3.

²⁵ Avishai Margalit, *The Decent Society* (London: Harvard University Press. 1996) p. 1.

²⁶ Ibid., p. 235.

²⁷ Harvey, p. 46.

²⁸ Of course, at this point, it is difficult to believe that anyone will buy bonds in Afghanistan. However, for the purpose of predicting the effects of holding money, the model is still useful.

²⁹ Harvey, p. 44.

Human Rights and Sexual Orientation: An Expansion of the Equality Doctrine

Elaine Hensle

When the seminal documents of human rights were written, no thought was given to the inclusion of sexual orientation minorities. As the movement for equality of orientation expands and the existing human rights paradigm becomes increasingly challenged by feminist critique, the question is growing of where sexual orientation belongs within human rights. Attempts at including sexual orientation have largely been through the right to privacy. Through an examination of American jurisprudence and changes in American and European legislation, this paper argues that the equality doctrine can and should be extended to include sexual orientation.

INTRODUCTION

The question of sexual orientation presents one of the greatest challenges to the human rights doctrine of equality. While the majority is able to envision itself as the minority across other dimensions of discrimination, this has not generally been the case with sexual orientation; sexual orientation minorities, meaning gay men, lesbians, bisexuals, asexuals, trans- and intersex individuals and the other categories emerging under the more encompassing category of queer are not in the majority anywhere.² There is no society where heterosexuals receive treatment as a sexual minority. This has prevented the heteronormative population from recognizing that its orientation is only one of many. Further, this refusal has contributed to a trend toward hate crimes that in 1988 prompted Mario Cuomo, then mayor of New York City, to comment that based on statistics, sexual orientation minorities made a stronger case for hate crime legislation in New York City than any other group.¹ Due to the rise in overall hate crimes, the US Congress passed

Elaine M. Hensle is an M.A. candidate at the Bologna Center, pursuing a degree in conflict management and international economics with a specialization in quantitative methods and economic theory.

the Hate Crimes Statistics Act in 1990, including sexual orientation as one of the categories of hate bias. In 2004 alone there were 1,482 victims of hate crimes motivated by sexual orientation.² However, it should be remembered that these statistics only account for reported incidents; the majority is believed to go unreported because of fear of further victimization. Many other sources have detailed the long, depressing history of violence against sexual orientation minorities, and that is not the goal of this paper. Instead, these brief statements are offered only to create a context in which to frame the debate surrounding sexual orientation and equality.

As statistics continued to be collected and the queer movement grew in strength, nations were forced to recognize the plight of sexual orientation minorities both inside and outside their borders. The first recognition of foreign oppression of the queer community abroad came in 1993 when Canada established the right of homosexuals to claim asylum.³ The United States followed suit the next year. Today, the two countries are joined by the United Kingdom, Israel and Australia. This step merely serves to recognize an existing international dilemma but does nothing to correct it.

The United Nations has long been silent. The first openly homosexual individual did not speak in the UN until 1992. He indicted the organization for its lack of representation and acknowledgement of the homosexual community.⁴ It took eleven more years for the first resolution explicitly mentioning sexual orientation to be brought forward when Brazil brought a resolution on "Human Rights and Sexual Orientation" to the 59th meeting of the Commission on Human Rights.⁵ Unfortunately, the Commission continues to table the issue, refusing to act against the coalition of the Vatican and many Third World nations. It appears that the UN's doctrine of equality as entrenched in the jurisprudence of human rights has reached its stumbling block. Sexual orientation minorities do not qualify.

Over the course of the following pages, this paper challenges this belief. Beginning with the concepts of equality and minority status, it examines the historical evolution of equality to include new minority groups and traces developments in American jurisprudence and legislation as well as the regional legislation of the European Union. Through this examination, it establishes grounds for the expansion of the equality doctrine to sexual orientation minorities.

** The terms sexual orientation minorities, homosexuals and queers are used interchangeably throughout the paper as to avoid the overuse of the longwinded, but politically correct, terminology "sexual orientation minority." While the use of the term "queer" may seem unusual to those not familiar with the movement, the community or the literature, it is widely accepted and used to indicate any person whose sexual orientation is outside the dominant heterosexual paradigm and is challenging the oppressive social/political structures of heteronormative society. In the academic realm, Andrew Parker has defined queer as "a non-gender-specific rubric that defines itself diacritically not against heterosexuality but against the normative," while Michael Warner uses the broader definition of "resistance to regimes of the normal." Parker and Warner quoted from LaborLawTalk.com <<http://encyclopedia.laborlawtalk.com/Queer>>.*

EQUALITY

Although a complete review of the history of the philosophical debate surrounding the equality of man is outside the scope of this paper, the philosophical origins of equality and their influence over the legal progression are essential to an argument based on the idea of equality itself. To this end, this section recognizes two of the prominent philosophers of equality, their opinions and their influence in the beginnings of human rights legislation before moving on to a discussion of the equality doctrine in the international human rights legal norms that emerged in the 20th century and the definition of equality itself. This is by no means to be an authoritative review of the debate, but merely a glance at the works of two men who greatly shaped the literature and the legal texts that have followed.

Primary among philosophers of the rights of man is John Locke, who developed the theory of the natural rights of man, including life, liberty and property, in the 17th century. In his *Second Treaties of Government*, he argued that all men are born "...with a title to perfect freedom and uncontrolled enjoyment of all rights and privileges of the law of nature equally with any other man or number of men in the world..."⁶ From this conclusion he followed with an argument that all men are therefore equal before all legal and government institutions, and that the institutions are established to further the substantive appreciation of man's natural rights by all men without discrimination. This view greatly influenced the drafting of the Virginia Declaration of Rights, considered by many to be the first bill of rights of the modern era. Its first article stated:

That all men are by nature equally free and independent, and have certain inherent rights, of which, when they enter into a state of society, they cannot, by any compact, deprive or divest their posterity; namely, the enjoyment of life and liberty, with the means of acquiring and possessing property, and pursuing and obtaining happiness and safety.⁷

Written in 1776, the Virginia Declaration then had a great impact on the Declaration of Independence drafted later that year, as well as the *Declaration of the Rights of Man and the Citizen*. Even much later the influence of the wording of this preliminary declaration – and through it that of John Locke – is seen in the Universal Declaration of Human Rights, which will be discussed in more detail shortly.

Following on the influence of Locke and the events of the 18th century that colored his own life, the American Thomas Paine took a more legalistic approach to the topic of human rights. In his 1791 work, *The Rights of Man*, Paine came to three conclusions focusing on the equality of man and the nature of political associations. Like Locke, he argued that all men were born equal, but he further stated that they always continue to be equal.⁸ The legal nature of his work focuses on the direct links of man's rights and equality and the functions of political associations. First among these, he argues that the end goal of any such association is the preservation of these rights and resistance of oppression. Later works even went so far as to proscribe schemes of reform for the British government. Paine's

influence was found mainly on the personal level of such individuals as Abraham Lincoln and George Washington. The wide dispersion of Paine's works incited many contemporary elites, like the latter, to actions that Paine perceived to be grave injustice. Paine viewed government as only a necessary evil to preserve the natural freedoms of man.

From this philosophical basis, equality has entered the human rights dialogue as a cornerstone to the substantive appreciation of all other rights, for if a person is not treated equally as others, then it must be assumed that at some level one of that individual's rights is being infringed. As such, virtually every international, regional and national legal document (be they charters, conventions, constitutions or named in some other manner) begins with a statement of the premise of equality of all persons. In the Universal Declaration of Human Rights (UDHR), which began the proliferation of international human rights in the modern era, this finds its expression in first clause of the preamble:

Whereas recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world...⁹

And then in Article 1:

All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.¹⁰

Article 2, which outlines the basis for antidiscrimination legislation, will be discussed in the next section. Through the auspices of the UDHR, the same belief in the equality of men has been incorporated in the other UN Conventions and the founding documents of regional human rights bodies.

This recognition of equality as the key to the enjoyment and enforcement of human rights creates an important question. What is equality? Before it is possible to discuss the extension of equality to a new community or examine its appreciation by any community, the essence of equality must be understood. At the very basic level, equality can be understood as treating like situations in a similar manner and conversely, not treating different situations in a similar manner.¹¹ The extensive jurisprudence on the subject illustrates that in practice this is not as simple as it sounds. Equality requires that individuals step away from their own personal prejudices and assess a situation in its pure form. For example, an employer should assess applicants based solely on their qualifications, as an ideal meritocracy dictates, and without regard to differentiating characteristics such as gender, race, sexual orientation or national origin. This must be done to the extent that the reason an applicant is chosen has nothing to do with anything other than they were best qualified for the position. Likewise, employers should offer equal salaries and benefits to employees who are employed in similar situations. Conversely, this does not mean treating different situations similarly. The doctrine of equality is not a promise of every individual receiving equal rewards in life; instead, it is a guarantee that every individual will be treated the same when he is in similar situations, such that no one is privileged over another person for any reason other

than incomparable status. This qualified status would exist in a comparison such as that of an individual holding a doctorate to one with no college education; the individuals involved are not in similar situations due to the divergence of their education background and as such should not be treated similarly.

MINORITY STATUS

The difficulties in implementation of the equality doctrine have led to the designation of certain groups within global society as minorities and as such enjoying certain protections of antidiscrimination. This has manifested itself in what are deemed antidiscrimination clauses included in international, regional and national human rights documents. Under the UDHR, this takes the guise of the above-mentioned article 2, which reads:

Everyone is entitled to the rights and freedoms set forth in this Declaration without distinction of any kind, such as race, color, sex, language, religion, political or other opinion, national or social origin, property, birth or other status...¹²

This has provided the traditional format for antidiscrimination clauses and contains two elements illustrative of the forward-thinking mentality that can be attributed to the drafters. Though the clause contains a list of categories, it is prefaced by the “such as” and concluded with “other status,” which indicate that the list is not meant to be definitive but can adapt as the political culture evolves or new categories subject to discrimination arise. A simple glance at the list also shows that the categories are not ones that would by themselves give rise to an incomparable situation. The purpose of these clauses and the recognition of minority groups in general is to heighten the scrutiny of judicial proceedings, as the terms are defined in American legislation. Further insight into the purpose behind the granting of minority status – or suspect classification – can be found in the criteria used to term whether a group qualifies for the designation.

For simplicity of discussion, the American mechanism for suspect classification is the only one discussed here, but the general mechanism is similar across regions. Under the US Supreme Court there are seven requirements to be satisfied for designation as a suspect classification:

- 1) they have suffered a history of intentional unequal treatment;
- 2) the classification imposes on them a stigma that brands them as inferior;
- 3) they have been the object of widespread prejudice and hostility;
- 4) the unequal treatment they have suffered has often resulted from stereotyped assumptions about their abilities;
- 5) they constitute a ‘discrete and insular minority’ whose political participation has been seriously curtailed because of prejudice;
- 6) the basis of the classification is an immutable (and often high visible) personal characteristic that each such individual possesses;

7) the characteristic is irrelevant to their ability to perform in or contribute to society (and to any legitimate public purpose).¹³

From these stipulations it can be deduced that suspect classification is meant to provide further protection against discrimination for members of groups that have historically suffered due to a shared, immutable characteristic. The characteristic in question is one that is part of the very definition of the group but has no bearing on their abilities otherwise; hence, it would not be possible to base an argument for incomparable situations on the said characteristic alone. As such, distinctions made between similar situations involving members of this group are clearly discriminatory and likely based upon a prejudice of the offending party.

With the above definition in mind, it is now possible to ask if sexual orientation minorities, i.e., any individuals identifying outside the dominant heterosexual paradigm, qualify as a suspect classification. Two cases of American jurisprudence will be discussed in detailed below, so here the discussion is only to form general support for the case that sexual orientation is indeed a suspect classification. The first four criteria can easily be satisfied by a cursory examination of the historical treatment of members of this minority group. Such an examination would serve to elaborate on the general pattern of discrimination and violence directed toward persons known to belong (or suspected of belonging to) a non-dominant category of sexual orientation.

Though queers are not without political power, as evidenced by the proliferation of interest groups with a sexual orientation focus or branch, their ability to exert political power remains severely limited. Stigmas attached to identification with this category have long prevented politicians from openly identifying as members of this minority group and has inhibited the efforts of openly queer candidates seeking election. Internationally, this can be seen in the number of national level representatives in four large democracies: the United States, three members of Congress; Australia, two senators; the United Kingdom, 15 members of Parliament (including 2 members of the Scottish Parliament and 1 member of the European Parliament); and Canada, six in the two legislative houses.¹⁴ Queers are more underrepresented and exert less political influence than either racial minorities or women, the two groups that have served as the basis for the suspect classification doctrine.

Having satisfied criterion five, numbers six and seven focus on the quality of the shared characteristic itself. The immutability of homosexuality is debated in scientific circles, but this paper follows from the belief that it has widely, if not popularly, become accepted as a fact of nature through observation of homosexual activity in other species if nothing else. Criteria seven is the same as that under the equality doctrine: that differences in individual characteristics only matter if they establish an incomparable situation or status. As sexual orientation by itself does impact the other aspects of comparison, such as intellect or physical strength, it satisfies the seventh criterion and fulfills the requirements for suspect classification. Legally, the mere satisfaction of these requirements is not enough to guarantee the group receives this designation; this depends on the prevailing culture of jurisprudence.

EXTENDING THE EQUALITY DOCTRINE

The basic statements of the equality doctrine were written long before the more inclusive statements of nondiscrimination. The evolution of that brought the world from the first to the second is well-documented. A comprehensive discussion of this process is outside the scope of this paper. Instead, a superficial reminder of the dynamic nature of the equality doctrine will serve to reinforce the changing ethos of global jurisprudence.

The pattern of the equality doctrine followed one akin to recognition of individual citizenship, and with it the right to vote, as the earliest statements of equality among men had a very specific group of men in mind: white men of property. As times changed through economic development patterns and men began to make a living on activities that were not based on land, this was extended to all white men of a certain age. Still, the distinctions of white and male remained. Through political upheaval and wars of liberation, this was extended to both men of other races and eventually, women. However, these groups did not, and arguably still have not, achieved full equality. The process of reaching this status has been more difficult than the passage of a law or amendment recognizing members of these categories as individuals of equal worth as all others. Recognition of these difficulties is manifested in the United Nations International Conventions on the Elimination of All Forms of Racial Discrimination and of All Forms of Discrimination against Women, respectively. Over the years, other groups suffering from discrimination due to a shared characteristic have also been added to the list of internationally recognized suspect categories such as religious minorities, indigenous people, children, persons with disabilities, the aged and migrant workers, among others. This illustrates recognition within the international community that the equality doctrine and the concept of suspect classification under it are not static concepts. Instead, they evolve with the political culture as new categories emerge and show the need for recognition under law. One of the necessary steps to become an internationally recognized suspect category is recognition at the national level by several states; to this end, two cases from the US Federal Court system will now be examined.

In 1985, a non-tenured high school guidance counselor named Marjorie Rowland, who was suspended from her position and did not have her contract renewed, filed suit in US federal court on the basis of her Fourteenth Amendment right to equality. In the initial decision by a trial jury it was found that “[s]he was discharged merely because she is bisexual and revealed this fact to acquaintances at her workplace.”¹⁵ This decision was then overturned by the Sixth Circuit and appealed to the Supreme Court, where it was denied certiorari. However, the dissent from this decision written by Justice William Brennan, joined by Justice Thurgood Marshall, offers broader implications for the jurisprudence of homosexuality. Throughout his decision, Brennan relies on the doctrine of equality and equates the treatment of the petitioner as necessitating a heightened level of scrutiny due to suspect classifications. In particular, he takes issue with the denial of certiorari:

[b]ecause determination of the appropriate constitutional analysis to apply in such a case continues to puzzle lower courts and because this Court has never addressed the issues presented...¹⁶

He further argues for extension of suspect classification to queers on the basis of Justice Burger's dissent in *Plyler v. Doe* in which Burger wrote: "The Equal Protection Clause protects against arbitrary and irrational classifications, and against invidious discrimination stemming from prejudice and hostility."¹⁷ From this Brennan follows with his own judgment of the prevailing situation of queers:

Under this rubric, discrimination against homosexuals or bisexuals based solely on their sexual preference raises significant constitutional questions under both prongs of our settled equal protection analysis.¹⁸

Though he writes in dissent, Brennan starts a precedent for other judges to follow as they extend the doctrine of equality to encompass homosexuality. His reliance on the equality doctrine more than the First Amendment protection of free speech signaled the beginning of a transition in judicial thinking, creating a basis for sexual orientation as a classification analogous to race or sex.

In a 1989 case, *Watkins v. the US Army*, the US Court of Appeals for the Ninth Circuit continued the transition begun by Brennan, solidifying the designation of sexual orientation as a suspect classification. Watkins had been discharged and refused re-enlistment despite exemplary service due to his homosexual status, which he had stated at the time of his enlistment. The court found that "[t]his is a case where equity cries out and demands that the Army be estopped from refusing to reenlist Watkins on the basis of his homosexuality."¹⁹ The concurring opinion of Judge Norris goes further stating plainly that "Watkins is entitled to relief because the Army denied him the equal protection of the laws by discharging and refusing to re-enlist him solely on the basis of his homosexuality."

Norris continues through his opinion to make distinctions between homosexual acts and homosexual orientation. These distinctions are important, as under US law at the time, homosexual acts such as sodomy remained largely illegal, but a person's homosexual orientation provided no basis for discrimination or assumption of participation in illegal acts. The portion of the Army regulations in question were found to target homosexuality itself, taking statements of orientation as *prima facie* evidence that the individual had engaged in illegal activity and vice versa that statements of having engaged in such activities necessitated the holding of a particular sexual orientation.²¹ This reading of the regulations in question enables Norris to find that

[i]n sum, the discrimination against homosexual orientation under these regulations is about as complete as one could imagine. The regulations make any act or statement that might conceivably indicate a homosexual orientation evidence of homosexuality; that evidence is in turn weighed against any evidence of a heterosexual

orientation. It is thus clear in answer to my threshold equal protection inquiry that the regulations directly burden the class consisting of persons of homosexual orientation.²²

This firmly places the question of sexual orientation in the category of suspect classification. In order to reaffirm his argument for heightened scrutiny with regard to this category, Norris provides a point by point analysis of the seven aforementioned criteria, finding each of them satisfied and designating sexual orientation as a suspect classification worthy of heightened scrutiny.²³ In doing so, he refutes a previously held belief that the Supreme Court decision in the 1986 *Bowers v. Hardwick* challenge to an anti-sodomy law precluded the designation of sexual orientation as a suspect classification; as he indicated, the basis of the *Hardwick* claim was not equal protection, leaving the question of suspect classification outwith the scope of the decision. This decision clarified the basis for suspect classification judgment and level of scrutiny due to cases of sexual orientation and equality. The ramifications of the case continue as Judge Norris' opinion is still cited for its thorough examination of the requirements for suspect classification.

NATIONAL AND REGIONAL LEGISLATION: THE UNITED STATES AND THE EUROPEAN UNION

As much as changes in jurisprudence influence the formal thinking about sexual orientation, they can do little for the instances of discrimination that go unrecognized and are not brought before a court for adjudication. These instances must be protected against by legislation at the local, national, regional, or international level. Since the day-to-day activities of companies and individuals outside a specific issue-public often take place without information about courtroom proceedings, it is necessary to examine where nations stand with respect to the possibilities for equal protection as outlined above.

The United States has a very mixed record with regard to the rights of sexual orientation minorities. In some respects the record is quite positive while in others some would call it appalling. Twenty-two states maintain sodomy laws on the books and in five of these the law applies only to homosexual couples, but in June 2003, the Supreme Court ruled that such laws are unconstitutional. Still, the remaining states have not deemed it worthwhile to repeal their constitution-violating laws. While only four states prevent homosexuals from adopting children, all except for one have laws against same-sex marriage.^{**} Of the remaining 49 states, 39 do not recognize out-of-state same-sex marriage licenses, a recognition they happily extend to opposite-sex marriage licenses. And 19 states amended their constitutions to ban gay marriage as part of the 2004 election cycle.²⁴ This in many ways reflects an overriding homophobia among the general American populace; however, this should not be seen as forbidding any forward movement by the national government. As education increases and people become more aware of the plight of sexual orientation minorities, attitudes are slowly beginning

to change. This is evidenced by the national outrage at the brutal murder of Wyoming college student Matthew Shepard in 1998, which facilitated the process of passing comprehensive hate crime bills.²⁵ To this end, September 2005 saw the passage of a comprehensive hate crimes bill by both the House and the Senate.²⁶ The version passed by the House even included explicit language inclusive of transgender individuals, a category often ignored in discussions of sexual orientation. Though the US still lacks Constitutional-level legislation for sexual orientation minorities, changes within national legislation and jurisprudence over the past five years show a growing political culture moving in that direction. However, any momentum will have to contend with large pockets of homophobia remaining in the country.

Lessons for the United States may come from the other side of the Atlantic, where the European Union (EU) has been working on expanding its doctrine of equality through antidiscrimination legislation for the past ten years. While the original European Convention provided very minimal protection for sexual orientation minorities under the concept of fundamental freedom, more recent legislation has steadily improved upon this beginning.²⁷ In the 1997 Treaty of Amsterdam the discrimination clause reads: "discrimination is prohibited on the grounds of sex, racial or ethnic origin, religion or belief, disability, age, or sexual orientation."²⁸ This was further strengthened with the passage of the Charter of Fundamental Rights of the European Union in 2000. Article 21 of the Charter contains the most all-encompassing antidiscrimination clauses prohibiting

discrimination based on any ground such as sex, race, color, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation.²⁹

The inclusion of sexual orientation in this clause is an example of top-down legislation at the regional level. At the time the clause was written, the majority of EU member-states did not have similar legislations protecting sexual orientation minorities from discrimination. However, the Charter was widely accepted and helped member-states with reluctant publics change their laws. An analogous strategy may be applied by the United States in prompting change from reluctant states and citizens. Passage of legislation at the national level while requiring a lot of political capital may spur needed reforms at lower levels of government in a

***These statistics should be understood to reflect only the status of same-sex marriage in the United State, not civil unions, which exist as a legal mechanism outside of marriage. Civil unions create a status of not single, yet not married, and confirm selected benefits of partnership equal to that of married spouses in some states. However, due to the nationwide Defense of Marriage Act [DOMA], the federal government is under no obligation to recognize civil unions. Currently, Massachusetts is the only state which recognizes same-sex marriages, while civil unions are possible in six states and Washington, D.C. Civil union statistics from Lambda Legal, <www.lambdalegal.org>.*

shorter period of time than would be likely if the reforms had to come from the bottom.

CONCLUSION

As the above analysis shows, there is room in the equality doctrine for sexual orientation minorities. They do not have to rely solely on fundamental freedoms, rights to expression or privacy, or whatever names the corresponding national, regional or international guarantees take. Further, the precedents set by the American judicial system indicate that sex discrimination clauses do not need to be broadly interpreted to include sexual orientation. Quite to the contrary, sexual orientation can and should be afforded a separate designation of suspect classification distinct from that afforded to sex.

The path that activists have been taking to overturn laws to legalize homosexual activity should continue, as it is a key element in changing existing prejudice, but it is by no means the only avenue available. Passage of a comprehensive hate crimes bill in the United States that was even inclusive of the transgender community shows that the masses may be compelled into reform by the elite. European experience supports this at a broader level with the inclusion of sexual orientation in the Charter of Fundamental Rights of the EU despite the lack of inclusive legislation in many member states. A similar process could be pursued by likeminded central governments to quicken the pace of reform in their jurisdictions. Another path that must be continued is education. Once people are able to identify with an individual member of a minority group, it helps to mitigate existing prejudice that even the passage of law is unable to achieve.

An increase in action at the national and regional levels should push the international community in the form of the UN to take action. As stated at the beginning of this paper, the UN has been resisting making a statement on the rights of sexual orientation minorities since the topic was introduced by Brazil in 2003. This situation cannot be ignored. With its goal of promoting equality and peace for all citizens, the UN cannot afford to let the discrimination of sexual orientation minorities pass without comment. Lessons from both the limited measures of the United States and the more extensive ones of the European Union should guide the future efforts of the UN. The resistance of a minority of nations should be overlooked in favor of the citizens suffering within those nations and the UN should take a more progressive stance than it has been willing to do thus far. In an extension of President Harry S Truman's statement, the UN should state that "When we say human, we mean all humans."³⁰ In a time of embracing diversity and democracy, no group should be excluded from the doctrine of equality simply as the consequence of an immutable characteristic. The possibility of extension exists; the door has been opened at both the national and regional levels. The only thing remaining is international action.

NOTES

¹ *Lambda*, 10 Jan. 2006 <<http://www.lambda.org>>.

² "Crime in the United States 2004," Hate Crime, 7 Jan. 2006 <<http://www.fbi.gov/>

ucr/cius_04/offenses_reported/hate_crime/index.html>.

³ "Australia sets world precedent for gay asylum? Not at all," *Fathers for Life*, 17 January 2006 <<http://www.fathersforlife.org/letters/letter8.htm>>.

⁴ Douglas Sanders, "Getting Lesbian and Gay Issues on the International Human Rights Agenda," *Human Rights Quarterly* 18, no. 1 (1996): p. 67.

⁵ *Human Rights Campaign*, 12 Jan. 2006 <<http://www.hrc.org>>.

⁶ John Locke, *The Second Treatise of Government* (Indianapolis, IN: The Bobbs-Merrill Company, Inc., 1952) p. 48.

⁷ "The Virginia Declaration of Rights," Gunston Hall, 13 January 2006 <<http://www.gunstonhall.org/documents/vdr.html>>.

⁸ Thomas Paine, *The Rights of Man* (London: Everyman Press, 1969) p. 43.

⁹ Walter Laquer and Barry Rubin, eds., *The Human Rights Reader* (New York: Meridian Books, 1990) p. 197.

¹⁰ *Ibid.*, p. 198.

¹¹ Rhona K. M. Smith, *Textbook on International Human Rights* (Oxford: Oxford UP, 2005) p. 185.

¹² Laquer and Rubin, p. 198.

¹³ Robert Wintemute, *Sexual Orientation and Human Rights: The United States Constitution, the European Convention and the Canadian Charter* (Oxford: Clarendon Press, 1997) p. 62-3.

¹⁴ "Homosexual Rights around the World," *International Gay & Lesbian Human Rights Commission* (11 December 2004), Brussels, Belgium, 12 December 2005 <<http://www.actwin.com/eatonohio/gay/world.htm>>.

¹⁵ *Rowland v. Mad River Local School District*, No. 470 U.S. 1009 (US Supreme Court, 22 April 1985) p. 2.

¹⁶ *Ibid.*, p. 1.

¹⁷ *Ibid.*, p. 3.

¹⁸ *Ibid.*, p. 3.

¹⁹ *Watkins v. US Army*, No. 875 F.2d 699 (U.S. Court of Appeals for the Ninth Circuit, 3 May 1989) p. 11.

²⁰ *Ibid.*, p. 12.

²¹ *Ibid.*, p. 15.

²² *Ibid.*, p. 16.

²³ *Ibid.*, p. 17-21.

²⁴ All cited statistics from "Homosexual Rights Around the World" unless otherwise indicated.

²⁵ Lambda.

²⁶ Human Rights Campaign.

²⁷ Wintemute, p. 142.

²⁸ Quoted in Smith, p. 185.

²⁹ Quoted in Smith, p. 186.

³⁰ Lambda.

