The Economics of Peace and Security Journal

ISSN 1749-852X

A publication of Economists for Peace and Security (UK)

Vol. 4, No. 1 (2009)

Symposium: Defense innovation

Symposium: Economics of conflict — theory and micro-level evidence

Articles

The Economics of Peace and Security Journal

© www.epsjournal.org.uk, ISSN 1749-852X

A publication of Economists for Peace and Security (UK)

Aims and scope

This journal raises and debates all issues related to the political economy of personal, communal, national, international, and global peace and security. The scope includes implications and ramifications of conventional and nonconventional conflict for all human and nonhuman life and for our common habitat. Special attention is paid to constructive proposals for conflict resolution and peacemaking. While open to noneconomic approaches, most contributions emphasize economic analysis of causes, consequences, and possible solutions to mitigate conflict.

The journal is aimed at specialist and nonspecialist readers, including policy analysts, policy and decisionmakers, national and international civil servants, members of the armed forces and of peacekeeping services, the business community, members of nongovernmental organizations and religious institutions, and others. Contributions are scholarly or practitioner-based, but written in a general-interest style.

Articles in *The EPS Journal* are solicited by the editors and subject to peer review. Readers are, however, encouraged to submit proposals for articles or symposia (2 to 4 articles on a common theme), or to correspond with the editors over specific contributions they might wish to make. In addition, comments on published articles (<500 words) are welcome. Write to us at editors@epsjournal.org.uk or contact us via the journal's home page at www.epsjournal.org.uk.

Vol. 4, No. 1 (2009) – Contents

Abstracts

Renaud Bellais. "Defense innovation at any (out of control) cost? The stalemate of today's R&D policy, and an alternative model." Technological superiority is a key element to achieve defense effectiveness, and R&D spending is crucial to access leading-edge technologies. Nevertheless, the current defense R&D model seems to reach its limits, leading to an out-of-control burden: in spite of spending almost €70 billion yearly, defense R&D in NATO countries does not produce the expected results. This low effectiveness leads to criticism about today's model of defense R&D, especially the channeling of credits by large incumbents through the Lead System Integrator model. Spurring disruptive technologies and path-breaking innovation requires an alternative approach. It is then interesting to analyze the possible use of venture-capital mechanisms to complement the mainstream approach to defense R&D. Even though the article does not develop a full analysis of the relationship between defense R&D and the principles of venture capital, it identifies criteria for such an alternative approach. [Keywords: defense burden, defense technology, R&D, disruptive innovation, lead-system integrator, venture capital. JEL codes: G24, H61, O32]

Sylvain Daffix and **Yves Jacquin**. "Defense R&D and national R&D systems: a European outlook." Eleven years ago, Mowery studied the role of defense-related R&D in the U.S. national system of innovation. The aim of this article is similar but is based on the European countries that signed the *Letter of Intent*, with a special focus on France. We first highlight the role of defense in the national system of R&D for these countries and then detail the case for the French Ministry of Defense. The article presents information provided by the OECD science and technology database and the French statistical survey on R&D based on the OECD methodology. In terms of funding as well as in terms of performance, the organization of research efforts is quite different from one European country to another. The detailed analysis of the French case shows an approximate stability in terms of funding after the gap of the late 1990s, but also a switch from R&D that is performed in-house to contracts with the industrial sector. This comparative exercise about Europe provides insights on the specificities of national systems of innovation and their evolution. [Keywords: defense innovation, national R&D systems, France. JEL codes: H57, O32, O52]

Peter Hall and **Andrew James**. "Industry structure and innovation in the U.K. defense sector." The defense industry in the United Kingdom has experienced significant structural change since the end of the Cold War. Consolidation has occurred and overseas suppliers have entered while some domestic prime contractors have exited. The implications of consolidation for innovation have been studied in the U.S. environment but much less so in the case of Britain. In this article we report briefly on the changes to structure and review the theory and empirical evidence that suggests how such change might be related to innovation. Firms' own-financed R&D in British defense industry is proposed as a useful indicator of innovation activity and is observed to have fallen sharply. Possible explanations are sought in the recently-revived inverted-U hypothesis linking competition to innovation, and in government demand fluctuations and British defense procurement reforms. The authors caution against government taking policy action to influence industry structure until more and better data are available to analyze the potential consequences. Investigating in-contract incentives to encourage industry innovation are recommended as a possible alternative. [Keywords: defense, industry, structure, innovation, Britain. JEL code: L64, O52]

Philip Verwimp. "Symposium. The economics of conflict: theory and micro-level evidence." The brief article introduces the four articles in the symposium.

S. Mansoob Murshed. "Conflict as the absence of contract." Two phenomena have been recently utilized to explain conflict onset: greed and grievance. The former reflects elite competition over valuable natural resource rents. The latter argues that grievance fuels conflict. Central to grievance are concepts of interethnic or horizontal inequality. Identity formation is also crucial to intrastate conflict, as it overcomes the collective action problem. Conflict can rarely be explained by greed alone. The greed explanation for conflict duration and secessionist wars works best in cross-country studies but has to make way for grievance-based arguments in quantitative country-case studies. Grievances and horizontal inequalities may be better at explaining why conflicts begin, but not necessarily why they persist. Neither the presence of greed or grievance is sufficient for the outbreak of violent conflict, something which requires the breakdown of the social contract. [Keywords: civil war, greed versus grievance, social contract. JEL codes: D72, D74, O10, O40]

M. Zulfan Tadjoeddin and **Anis Chowdhury**. "Socioeconomic perspectives on violent conflict in Indonesia." Focused around the greed and/or grievance theses, a large part of the economics of conflict literature concerns itself with civil war. This article provides socioeconomic perspectives on contemporary conflict in Indonesia. Three categories of violent conflict in the country are separatist violence, ethnic/sectarian violence, and routine violence. We argue that two elements of the grievance argument, namely relative deprivation and horizontal inequality, are particularly useful for analyzing the Indonesian case. In contrast, the greed hypothesis does not appear to provide as strong an explanation of violent conflict events in Indonesia. [Keywords: Indonesia, greed, grievance, relative deprivation. JEL codes: D74, O10]

Ana María Ibáñez. "Forced displacement in Colombia: magnitude and causes." The article describes the magnitude, geographical extent, and causes of forced population

displacements in Colombia. Forced migration in Colombia is a war strategy adopted by armed groups to strengthen territorial strongholds, weaken civilian support to the enemy, seize valuable lands, and produce and transport illegal drugs with ease. Forced displacement in Colombia today affects 3.5 million people. Equivalent to 7.8 percent of Colombia's population, and second worldwide only to Sudan, this shows the magnitude of the humanitarian crisis the country is facing. The phenomenon involves all of Colombia's territory and nearly 90 percent of the country's municipalities expel or receive population. In contrast to other countries, forced migration in Colombia is largely internal. Illegal armed groups are the main responsible parties, migration does not result in massive refugee streams but occurs on an individual basis, and the displaced population is dispersed throughout the territory and not focused in refugee camps. These characteristics pose unique challenges for crafting state policy that can effectively mitigate the impact of displacement. [Keywords: conflict, civil war, forced displacement, migration. JEL codes: R23, D74]

Steven Spittaels and **Filip Hilgert**. "Are Congo's mines the main target of the armed groups on its soil?" The authors have developed a qualitative geographic research tool for case studies of conflict areas. They believe it can furnish convincing evidence on war motivation. They have used the tool to analyze concrete military actions and decisions and to trace them back to what provoked them. They have applied it to the conflict situation in the DR Congo between August 2007 and January 2008. In this article special attention is given to the role of natural resources in the armed conflict. Its importance as a war motivation factor is compared to three other conflict drivers found in the literature on the causes of war. [Keywords: Democratic Republic of Congo, Kivu, conflict, Mining, natural resources. JEL code: Q34]

Christopher E.S. Warburton. "War and exchange rate valuation." This article investigates the extent to which the dominance of the United States (U.S.) dollar as an international currency has been contingent on American diplomacy rather than the prosecution of expensive wars. Four wars are examined, the Korean War (1950-1953), the Vietnam War (1964-1975), the Persian Gulf War (1990-1991), and the Iraq War (2003-present). The historical performance of the dollar is examined in times of war and peace, and the Box-Jenkins forecasting algorithm is employed to make a short-term projection of the dollar coinciding with the Iraq war. The price of gold is used as a measure of the value of the U.S. dollar and investor confidence in the dollar during times of war and peace. The empirical evidence shows a short-term depreciation of the U.S. dollar coinciding with the Iraq War, which is not atypical of the value of the U.S. dollar in a time of war. Problems with the value of the U.S. dollar in times of war lead to the exploration of alternative forms of money, which if very successful, can erode the continued dominance of the U.S. dollar as an international currency. [Keywords: currency valuation, forecasting, optimum currency areas, war. JEL codes: F15, F51, H56]

Steve Chan. "The democratic peace proposition: an agenda for critical analysis." The proposition that democracies are more peaceful than autocracies has spawned a huge literature. Much of the relevant quantitative research has shown that democracies indeed rarely, if ever, fight each other, although they are not necessarily less bellicose than autocracies in general. This essay seeks to identify several areas of concern that offer fruitful directions for further research to extend and clarify this proposition. These concerns relate to (1) conceptual clarification, (2) methodological assumptions, (3) causal interpretations, and (4) policy relevance. [Keywords: democratic peace, republicanism, libertarianism, preventive war. JEL codes: H56]

Steve Townsend. "Friedman's First Law fail: oil prices do not predict freedom." Thomas Friedman's First Law of Petropolitics has acquired wide circulation and acceptance without being subjected to proper scrutiny. His claim that, in oil-rich, less-developed countries, there is an inverse correlation between the price of oil and the level of freedom is false. Friedman provides data for four countries to illustrate his thesis. This article shows his figures are incomplete, inappropriate, or misinterpreted. He specifies nine more countries that will conform to his prediction but when investigated, these, more often than not, also disprove his claim. [Keywords: democracy, democratization, petropolitics, resource curse. JEL code: Q34]

Ronen Bar-El, **Kobi Kagan**, and **Asher Tishler**. "Short-term versus long-term military planning." This article analyzes the allocation of the government budget to civilian and military expenditure by two rival countries that are involved in an arms race. We compare the consequences of myopic (period by period) planning versus rational (long-term) planning and show that although myopic planning is always favorable for both countries, they are likely to become locked in a prisoners' dilemma equilibrium in which they plan rationally. The prisoners' dilemma equilibrium results in higher stocks of weapon systems and lower welfare for both countries. We also find some evidence of the existence of a prisoners' dilemma equilibrium in the current Israeli-Syrian arms race. Generally, this article suggests that solving military/political conflicts that evolve into an arms race by relying only on military might is an expensive and suboptimal solution. Each country should identify this phenomenon and attempt to take it into account in its decisionmaking by using appropriate policy alternatives. There are two major policy options in this case. The first option is to reach some kind of agreement with the rival (directly or through a third party). The second option is to institute force multipliers techniques. [Keywords: arms race, intertemporal defense budget allocation, prisoners' dilemma. JEL codes: D74, D78, D90, H56, H68]

Defense innovation at any (out of control) cost? The stalemate of today's R&D policy and an alternative model

Renaud Bellais

Technological superiority is considered to be a key element to achieve defense effectiveness, and R&D plays a major role in accessing relevant leading-edge technologies. This is the reason why since World War II the United States and Western European countries have been pouring a large share of their R&D expenditures into applications for defense. However, defense R&D has become increasingly expensive while being less and less able to generate disruptive technologies. In the large arms-producing countries, spiraling R&D spending amplifies its economic burden and crowds out civilian R&D. More efficiently conducted defense R&D thus could assist civilian R&D budgets as well.

As the biggest military spender, the United States most experiences the limits of the existing model of defense R&D. In spite of doubling its budget from 2000 to 2008, with outlays growing from $37.6 billion to $74.4 billion, the Pentagon is experiencing technology shortcomings and failures. For instance, spending for developing counter-IED systems seems out of control yet with little to show for it. One may therefore wonder whether current R&D policies can achieve their goals or whether it should be transformed.

Military spending trends show that a defense technology race remains at the heart of defense procurement in the occidental world. Even though non-western countries such as Russia, India, or China have been increasing their defense R&D, they have not reached levels equivalent to the biggest NATO spenders. One must wonder how to stop the upward drift of defense R&D spending while improving its effectiveness, that is, nurturing leading-edge technologies at an affordable cost.

This article is divided into two sections. The first emphasizes the contradiction between the quest for technological superiority as the driver of defense R&D spending and the shortcomings of the dominant approach. It underlines the weight of the follow-on principle to explain spiraling R&D costs. The second section identifies the limits of the Lead System Integrator model. Relying on the innovation literature, the article proposes venture capital mechanisms as a possible alternative to both reduce the burden of defense R&D and to improve its effectiveness.

Technology races as the engine of defense procurement

Because it is considered a means to counter existing or emerging threats and because of the rising costs of defense technologies, the armies of the main arms-producing countries have been investing heavily in R&D, even after the end of the Cold War. Nevertheless, today's R&D model, set up during the Cold War, appears less and less adapted to deliver defense innovation at a sustainable cost.

Table 1: Weight of R&D in equipment spending (2006, millions of euros)

	Procurement	*R&D*	*Equipment*	*R&D share*
United States	83.0	58.0	141.0	41.1%
European Union	29.1	9.7	38.8	25.0%
United Kingdom	7.5	4.0	11.5	34.8%
France	6.3	3.8	10.1	37.4%
Germany	3.7	1.0	4.7	21.9%

Sources: European Defence Agency; European-United States defense expenditure in 2006; and 2006 National Breakdowns of European Defense Expenditure, Brussels, 21 December 2007.

Military threats and technological superiority

Since the 1940s, armed forces have fielded advanced technology through an aggressive pursuit of R&D and the development of a high-tech defense industrial base of unprecedented scale.[1] R&D represents between one-fifth and two-fifths of the equipment spending of large NATO countries (Table 1).

Once driven by the East-West arms race, technological superiority remains today at the heart of defense procurement. The prominent defense analyst Jacques Gansler writes that "although we no longer face the threat posed by a global peer competitor like the former Soviet Union, we still live in a very dangerous world. It is a world marked by uncertainty and unpredictability; a world in which multiple possible aggressors pose a wide range of potential threats and hostile actions."[2] This appraisal seems to be shared by the leading arms-producing countries.

Since the 1990s the threats have become dramatically more heterogeneous.[3] NATO countries consider advanced technology as an appropriate and adequate answer: "Potential opponents will also have access to much state-of-the-art technology, since they can purchase it on the open global market. Thus DOD must 'run faster' than others, rapidly feeding on the global base rather than relying almost exclusively on its own sponsored R&D as it did during the Cold War."[4]

Western armed forces feel that strategic surprises can be avoided by mastering innovation and disruptive technology. This explains why defense R&D remains a key dimension of defense spending. Even though technology is not the unique answer or

Table 2: Defense R&D in constant 2006 euros (millions)

	*1994**	*2006*
United States	37,200	58,000
France	4,730	3,777
United Kingdom	3,830	4,012
Germany	1,240	1,035
Italy	530	252
Sweden	360	267
Spain	270	201
European Union 26	11,500	9,700

* Available data between 1993 and 1995, 1994 being the most common year.
Sources: SIPRI Yearbook 1996, Oxford University Press, 1996, p. 381; European Defence Agency, 2006 National breakdowns of European Defence Expenditure, Brussels, November 2007; Office of the Undersecretary of Defense (Comptroller), National Defense Budget Estimates for FY2008, U.S. Department of Defense, March 2007; and author's calculations.

effective by itself, it is perceived as crucial to analyze threats and to maintain military dominance. This is the reason why merely sustaining the technological effort is not considered an option but a necessity to counter unexpected — and sometimes unpredictable — foes.

This view also explains why the United States and Europe still spend almost €70 billion a year in defense R&D. This represents a huge economic burden, and one whose legitimacy can be questioned. But maintaining technological superiority requires an effective technology policy both in terms of operational and budgetary impacts. This raises two major questions. First, are defense R&D budgets compatible with the required level of investment? Even if some countries are able to fund multiple projects (many of them in advanced technologies), none of them is rich enough to cover the whole range of potential technologies or innovations. Do defense R&D spending trends lead to an unsustainable burden? Second, is defense R&D policy effective? As the current model is not always able to produce the expected results, especially to identify and/or take advantage of emerging technologies, can an alternative model improve the effectiveness of R&D spending?

Budgetary pressures on R&D spending

In spite of technology-driven defense, R&D spending paid the price of both the end of the Cold War and strong downward pressures on public spending in the 1990s in the United States and in European countries. Cutting R&D helped preserve operational capabilities, and this choice was the easiest since it was almost without perceptible effects in the short term.[5] A bit more than a decade later, R&D spending has recovered (Table 2). The United States has steadily increased its defense R&D

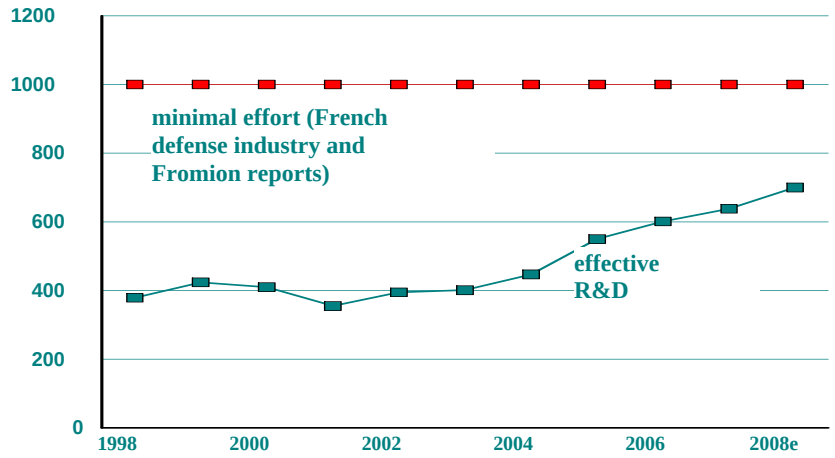

Figure 1: French R&D expenditure
Source: Directorate of Financial Affairs, French Ministry of Defense

since 2002, reaching its highest level ever. In Europe, defense R&D has increased in recent years but remains lower in real terms than in the early 1990s. This is explained, relative to the United States, by stronger budget constraints and a lower priority given to defense.

Even though defense R&D has reached an historical peak within NATO, it is insufficient to keep pace with the long-term evolution of defense R&D costs that since the 1970s have increased in real terms at a rate of 4 percent per annum.[6] Even for the Pentagon, a 63 percent budget increase in real terms from 1995 to 2008 is not sufficient to compensate for the doubling of R&D costs. The scissors effect is more important for European countries. Almost stagnant resources are, in real terms, no longer sufficient to keep pace with technology costs. For instance, despite an upward trend of its R&D effort since 2002, French minimum expenditures are still below those needed as defined by both the French defense industry association and the Fromion Report (Figure 1.[7]

Due to spiraling costs a structural gap exists between available budgets and the requested level of investment. NATO countries seem unable to sustain the technology intensiveness of defense systems. There is limited room for additional resources in France and the United Kingdom, and even in the United States projections anticipate a fall of R&D spending from $79.6 billion in 2009 to $63.4 billion in 2013 (in real terms).[8] One may therefore question whether simply to keep increasing R&D spending is the proper way ahead. Instead, one should analyze the causes of spiraling costs.

The effectiveness of defense R&D results from both the spending level in a given

technology and how such spending is managed. In fact, threshold effects result from the evolution of defense-related technologies. In particular, one cannot expect a linear relationship between R&D spending and its output: below a certain level of investment in a given technology, the output of R&D falls rapidly. Exploring investments in integrative technologies in a dynamic optimization framework, Setter and Tishler find that under nonlinear, convex development costs it is not optimal to build military forces using a myopic, short-term approach.[9] It is difficult to transform an armed force within just a few years. Consequently, early investment in technological infrastructure is required because the entry cost in technology is high and the transformation period ranges over more than a decade.

If a country's investment in a given technology is too limited, it cannot expect to keep pace with the state of the art, and it is not worth investing in that technology. But even if NATO were to carefully select the technologies in which it invests, this would only cover one part of the story. The more important concern is to understand whether the existing R&D model itself is appropriate to nurture emerging technologies.

Shortcomings of defense R&D

Defense R&D provides a low rate of social return — a lower rate than civilian R&D expenditure in equivalent fields.[10] In addition, it appears that defense projects are increasingly expensive and do not deliver on their promises.

Frank Lichtenberg shows that, holding non-defense R&D constant, defense R&D has essentially no effect on productivity growth. He underlines that "both micro and aggregate estimates of the (social) 'rate of return' to investment in government-funded (largely defense-related) R&D — or its impacts on productivity growth — are insignificantly different from zero, and are much smaller than estimates of the rate of return to privately-funded R&D."[11]

Moreover, the costs of defense R&D rise more quickly than for civilian projects. This results in part from the quest for advanced technology as well as from the obsessive drive to improve existing technologies through incremental innovations, leading to what Mary Kaldor called the Baroque arsenal. As Adelman and Augustine note, "the problem with technological sophistication is that the last 10 percent of performance sought typically adds one-third of the cost and two-thirds of problems."[12] This remains true, as demonstrated by Schinasi's critique of the major defense programs.[13] Incremental innovations are expensive, do not guarantee results, and generate limited civilian spinoffs.

In addition, defense R&D does not always fulfill its objectives: programs are delayed, costs are steadily rising, and it cannot deliver the expected results. If one takes the example of the Pentagon's major programs (Table 3), costs are drifting upward and schedules are delayed. The R&D failures explain why programs seem out of control, since programs are launched before technologies are mature enough to be incorporated. However, there is "not a strong positive association between equipment quality and ... the fraction of defense expenditure applied to R&D."[14]

Table 3: Analysis of DOD major programs

Fiscal Year	*2000*	*2005*	*2007*
Number of programs	75	91	95
Change to total RDT&E costs from first estimate	27%	33%	40%
Change to total acquisition costs from first estimate	6%	18%	26%
Estimated total acquisition cost growth	$42bn	$202bn	$295bn
Average schedule delay in delivering capabilities	16 mths	17 mths	21 mths

Source: Schinasi (2008, p. 5).

The marginal gain resulting from improving a given technology is not always the best way to achieve military dominance. It can be more relevant (operationally and financially) to develop an alternative solution. The focus on existing technology corresponds to what James Kurth has called the follow-on imperative, i.e., the evolution of systems, generation by generation, without questioning the usefulness of in-service platforms at increasing costs for decreasing marginal benefits.[15]

Because they cannot afford failure or delays with limited resources, cash-strapped armed forces favor well-known technologies rather than disruptive or emerging ones. They tend to become more risk averse and prefer limited results to unpredictable technological disruptions with unknown potential. This vicious cycle has two major consequences: defense R&D is bound to have a diminishing rate of return, and armed forces are less able to avoid technological surprises because they invest less and less outside their core technological base.

Keeping pace with emerging technology then begs for an alternative mechanism, one that helps increase the rate of return to defense R&D (and eventually reducing its economic burden) and to keep armed forces aware of technology disruptions and the related potential defense applications.

The need to transform defense R&D policy

Current defense R&D is not capable of providing troops with leading-edge, disruptive technologies. This does not mean that new technology would be perfectly effective; but one can expect a bigger bang for armed forces' buck and, eventually, a reduced defense burden.

Armed forces need to find the proper balance between incremental and disruptive technologies. Improving existing technologies through the Lead System Integrator (LSI) model is useful to some extent but not to avoid technological surprises. In this section, an alternative model, relying on venture capital, is suggested.

Table 4: Research and development in defense R&D (2006, million euros)

	R&D	*R&T*	*R&T share*
United States	58,000	13,600	23.4%
United Kingdom	4,012	733	18.3%
France	3,777	762	20.2%
European Union	9,700	2,500	25.8%

Sources: European Defence Agency; European-United States Defense Expenditure in 2006; and 2006 National Breakdowns of European Defense Expenditure, Brussels, 21 December 2007.

Overcoming the follow-on approach

The armed forces and incumbent firms are quite capable to manage the evolution of known technologies. For instance, the United Kingdom identifies required technological investments through its defense technology strategy. The French government's arms directorate (the DGA) cooperates with industry to elaborate on its thirty-year planning exercise, the "plan prospectif à 30 ans". Improvements can be anticipated and therefore organized through investment plans. But such an organization is ineffective when intending to catch unexpected technological breakthroughs which are more difficult to identify and whose results are highly uncertain.

Therefore a paradoxical trend emerges: the less money armed forces receive, the more they concentrate on D rather than R. Carter notes that "much defense R&D today goes to keep old 'legacy' systems going or to prop up faltering programs, rather than launching new leap-head military systems."[16] A large share of the R&D budget is allocated to existing technologies, even mature ones, explaining why development absorbs at least three-fourth of R&D credits (Table 4).

Some funding for nonconventional research does exist. The French DGA set up a mechanism for "non-solicited proposals." The Pentagon likewise manages an independent R&D program. Supposed to overcome the intrinsic limits of its own R&D organization and to keep industry innovative, it was designed to reimburse the defense industry for a share of its self-funded R&D so that firms were able to explore new technologies.[17]

But such mechanisms provide very limited funding, at least when compared to the overall R&D effort, and they are not as effective as expected. First, allocated credits have decreased over the last two decades (Figure 2) while the entry requirement to many technologies rose steadily. Second, a large share of credits is diverted toward conventional projects, for such mechanisms are used to achieve projects otherwise underfunded or not funded at all, in effect becoming alternative funding for incremental innovation. As Carter notes, "over time the government is tending to dictate more of the programs, making them less truly the results of the independent judgment of non-government scientists and engineers."[18]

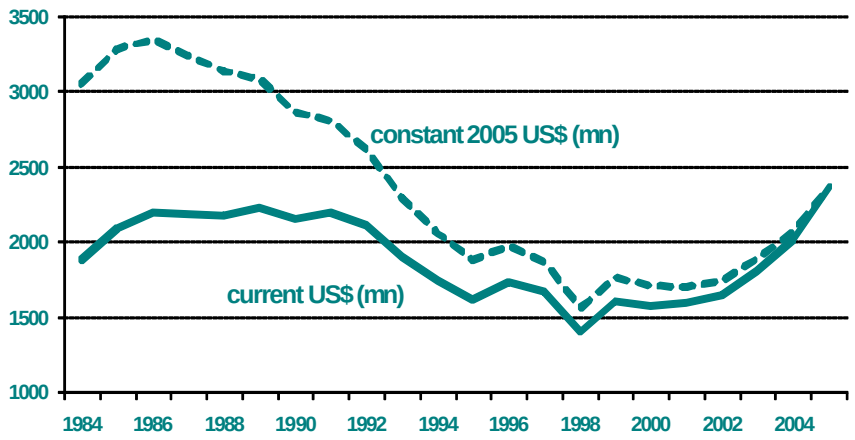

Figure 2: Independent R&D: DoD share of allowable costs
Source: Defense Procurement and Acquisition Policy, DoD, http://www.acq.osd.mil/dpap/policy/cpf/ [accessed 5 October 2007].

Moreover, since the 1990s the defense industry is less favorably disposed to disruptive innovations. The main arms-producing countries gave prime contractors a greater role, promoting a Lead System Integrator (LSI) model in which the prime contractor is the state's unique interlocutor. This channels more R&D funding toward prime contractors and large projects (especially technology demonstrators). Consequently, less money is available for innovative small firms that have limited access to the ministry of defense.

This model favors existing technologies rather than nurturing emerging or disruptive ones. Tratjenberg emphasizes: "The development of big weapon systems in the decades of the Cold War led to a high concentration of both R&D and procurement into a few large corporations, conferring on them a great deal of market and bargaining power. It is quite likely that this had detrimental effects in terms of costs and effectiveness, and it may have steered technical advance into questionable directions."[19]

Developing a growth model, Matsuyama demonstrates that a given economy can have not one but two growth regimes simultaneously: a Solow regime in which growth results from capital accumulation, and a Romer regime in which growth results from knowledge accumulation.[20] He underlines that the market rules differ from one regime to the other because of a far different weight given to innovation.

An audacious parallel could be drawn with regard to defense innovation. A monolithic R&D policy seems irrelevant to nurture simultaneously incremental improvement of existing systems and path-breaking innovations that may dramatically transform defense. This leads the United States to question whether this alternative

model for exploratory R&D corresponds to incumbent firms' abilities.

Incumbent firms and disruptive technologies

The follow-on bias is amplified by today's place given to LSIs. If incumbent firms' resistance to change and disruptive technology are widespread in many industrial sectors, this is especially true in defense:

- Incumbent firms are keen on developing incremental changes in technologies that they master (reinforcing their own position within the defense market), but they tend to resist disruptive change when it jeopardizes their technology portfolio (becoming obsolete) or requires additional investment.
- Such conservatism exists symmetrically on the demand side, as armed forces tend to promote technologies they know and for which there exists a usage doctrine. Services tend to resist new technology that requires changes in their missions and organization.

As Gansler writes: "A dialectic conflict is created by technological advancement. On the one hand, technological opportunities that result in incremental changes (…) are rapidly accepted and enthusiastically pushed forward — perhaps too unquestioningly. On the other hand, those technological opportunities that offer revolutionary changes are often outside the paradigm; thus they meet with extreme institutional and structural resistance, with every effort made to reject them."[21]

Even though large firms try to keep pace with disruptive technology, their structural functioning creates intrinsic limits. Large firms need to provide quarterly results which require both large and certain markets. Emerging technologies are the opposite: high risks, uncertain, long-term results, and nonexistent (if not entirely unpredictable) markets. Where to invest is obvious: "A project to commercialize a disruptive technology in a small, emerging market is very unlikely to be considered essential to success in a large company; small markets don't solve the growth problems of big companies."[22]

Large firms must focus on projects for which market demand and profitability are guaranteed or under control. All their organizations favor incremental innovations rather than disruptive technologies. Christensen writes that "expecting achievement-driven employees in a large organization to devote a critical mass of resources, attention, and energy to a disruptive project targeted at a small and poorly defined market is equivalent to flapping one's arms in an effort to fly: it denies an important tendency in the way organizations work."[23]

In commercial markets this dilemma is often overcome through a merger and acquisitions (M&A) strategy. Large firms invest intensively in innovations but mainly to develop or improve their knowledge capital. When they look for disruptive innovation, especially if it threatens their core business, acquiring small innovative firms is the most effective strategy.[24]

But this possibility is quite limited in defense. Absent an alternative market, defense technologies are developed only through R&D or procurement contracts; public contracts then shape the industrial base. In contrast to the civilian sector, when a firm is looking for a technology that she does not have, it is difficult to find an off-the-shelf solution. This creates a difficulty when, as emphasized previously, R&D policy does not create a window of opportunity to fund emerging technologies.

With R&D policy focused on incremental innovations and managed by LSIs, one can expect a gap to arise between the advancement of science and the knowledge mastered within the defense industrial base. If armed forces expect to keep pace with technologies and avoid surprises at an affordable cost, they need an alternative, complementary approach. In this respect, a mechanism based on venture capital, one that combines market incentives and investment strategy, may fill this gap.

Applying venture capital mechanisms

Beyond the intrinsic organizational limits of incumbent, large firms, difficulties to deal with emerging technologies arise for several reasons: technology legacies, sunk investments, the need for short-term financial results, and the influence of existing customers. One solution to overcome such limits is to set up an ad hoc structure dedicated to the development of a given technology. Here it seems of interest to explore the use of venture capital.

Venture capital (VC) is a risk-seeking, project-focused approach possibly compatible with the time horizon for emerging or disruptive technologies. Investors expect no immediate returns so long as they can forward to an attractive medium-term exit. Organized for a given project, VC can gather the critical mass of resources and is not distracted by other issues. On the financial side, risks become acceptable as investors employ a portfolio strategy: even if one project fails, they expect that the average return across all projects compensates for taking higher-than-average risks.

Empirical research underlines that VC-funding has strong positive effects on innovation. Gompers and Lerner state that "on average a dollar of venture capital appears to be three to four times more potent in stimulating patenting that a dollar in traditional corporate R&D ... venture capital, even though it averaged less than 3% of corporate R&D from 1983 to 1992, is responsible for a much greater share — perhaps 10% — of U.S. industrial innovations in this decade."[25] Venture capital can help develop technologies faster and more effectively than large firms.

Even though it is not adapted to manage the whole of defense R&D, developing a defense research and technology (R&T) strategy based on venture capital could help both to identify promising technologies and to nurture their defense applications. Innovative approaches are not terra incognita for the defense world. Beyond the interesting experience led by the U.S. Defense Advanced Research Projects Agency (DARPA) that has supported emerging technologies since the 1950s, other institutions have recently been developing alternative mechanisms to improve the effectiveness of their R&D.[26] For example, the CIA launched its own fund, In-Q-Tel, in 1999; and

the U.S. Army in 2003 granted MILCOM Technologies the management of OnPoint Technologies, its VC-structure.[27]

However, venture mechanisms have limits, too. Some of the commercial VC-models are clearly designed as asset-stripping, short-run, profit-maximization vehicles. But the proposal here is not to simply channel defense funding through VC-funds but to exploit their demonstrated advantages for developing alternative means to nurture defense innovations. Defense VC-structures should have specific features corresponding to the expectations of defense. Even though the selection of projects should rely on the basic criteria of venture capital (technological and commercial potential, long-term profitability, expectable exit, and so on), its first objective would be to explore emerging or disruptive technologies and determine their potential for defense applications.

There are few experiences on which one can rely to develop an empirical analysis of the effectiveness of VC-mechanisms for defense objectives. Nevertheless, one can identify three criteria which could contribute to a successful utilization of such mechanisms to improve the social returns of defense R&D. First, a defense fund can be effective if and only if it is able to take risks and avoid limiting its potential losses ex ante. While having severe criteria of selection, it should invest in several projects and expect a big success for one investment in ten as well as an average failure rate of about 50 percent, the regular rate in the VC-industry. This is a prerequisite for truly exploring technologies with possible defense applications, and thus in avoiding technological surprises.

Second, it must avoid mimicking private venture investors. Venture capital is a cyclical industry. Cycles result from the nonlinear emergence of technological opportunities, but also from herding effects, resulting in over and underinvestment phases. A defense fund should not invest in technologies already funded by private investors, avoiding the mistake identified by Gompers and Lerner: "Government programs have frequently been concentrated during the periods when venture capital funds have been most active, and often have targeted the very same sectors that are being aggressively funded by venture investors."[28] The true aim of such a fund is to fill the gap between the advancement of science and technology and the available funding to develop innovations; then supporting risks which are not compatible with a commercial and/or profit-oriented approach. It should also counterbalance rather than exacerbate the cyclical nature of VC-investment, supporting the development of emerging technologies when private investors reduce their commitments.

Third, a defense fund should have a not-for-profit approach: it must look for a "return to technologies," rather than return to investment as In-Q-Tel does. As McCullagh notes, "the true mission of In-Q-Tel (... is) to tap the best minds in the technology sector and spur the development of products the CIA desperately needs and doesn't have the time or expertise to develop itself."[29] The aim of such a fund is to explore the defense potential of technological options so that armed forces have access to state-of-the-art technologies when needed. Forgetting this goal could jeopardize its true mission and lead the fund to behave like any other VC-fund, as underlined in the U.K. by the failure of Defence Technology Enterprise in the field of technology transfers from defense to commercial applications.[30]

Conclusion

As defense R&D faces limits in its current organization, improving its effectiveness and reducing the economic burden of defense spending requires an alternative approach, at least for emerging and disruptive technologies. The view put forward here is aimed at identifying the limits of defense R&D and analyzes its causes. Consequently, it is suggested that incremental innovations, organized through a LSI model, can deal with improvements in existing technologies but is not fully satisfying for unknown technologies.

The logic of venture capital is of interest, as its mechanism corresponds to the specific conditions of such technologies. With limited funding and adequate management rules for a specialized VC-fund, one could expect that armed forces would be more capable of keeping pace with technological change. The aim here is not to develop such an alternative approach but only to underline its intrinsic interest and suggest a research agenda on the possible relationship between defense R&D and the principle of venture capital.

The stakes are not limited to the effectiveness of public spending and access to military dominance. Improving the effectiveness of defense R&D is a good means both to reduce taxpayers' bills and to reallocate R&D funding toward civilian, socially-sound projects. Indeed, even though defense R&D has sometimes been used as a substitute for civilian R&D policy, one cannot consider defense as the best catalyst for R&D. It would be more relevant to invest directly in civilian projects. One could then expect that by improving the utilization of defense R&D, it would be easier to limit its growth and eventually push armed forces to rely on civilian R&D rather than supporting a specific agenda.

Notes

Renaud Bellais is Manager for European Defense and Security at Astrium, Paris, and Associate Professor of economics at the École Spéciale Militaire St Cyr Coëtquidan. He may be reached at renaud-bellais@voila.fr. The author is grateful to Geoffroy Rémond for research assistance, Jean-Michel Oudot, Prof. Asher Tishler, and the anonymous referees for useful comments. The usual disclaimers apply.

1. Technology change is the engine of defense procurement since the Renaissance, but it is only after World War II that the alliance between armed forces and science became effective. Thereafter, defense R&D became really significant as underlined by the works of Charles Wright Mills, Seymour Melman, or Claude Serfati (Bellais,

1999).

2. Gansler (1998, p. 1).

3. See, for instance, Bellais and Le Blanc (2002) on asymmetrical threats or Tratjenberg (2006) on terrorism.

4. Carter (2001, p. 131).

5. Middleton, *et al*. (2006) demonstrate that there is a relationship between defense R&D and equipment quality about 10 to 25 years later. Reciprocally, impacts of budget cuts appear only about a decade later.

6. Hartley, *et al*. (2008).

7. CIDEF, L'industrie de défense française: Une dynamique à soutenir, des enjeux européens, Paris, 2004. Yves Fromion, MP, Les exportations de défense et de sécurité de la France, Report to the French Prime Minister, Paris, July 2006.

8. OUSDC (2008, p. 76).

9. Setter and Tishler (2006).

10. Bellais (2004).

11. Lichtenberg (1995, p. 456).

12. Adelman and Augustine (1990, p. 142).

13. Schinasi (2008).

14. Middleton, *et al*. (2006, p. 117).

15. See Kurth (1993).

16. Carter (2001, p. 135).

17. “Contractors shall be encouraged to undertake IR&D activities that may further national security in a broad sense, may lead to a superior military capability, or may lower the cost and time required for providing that capability” (DoD, Directive 3204.1, 10 May 1999).

18. Carter (2001, p. 135).

19. Tratjenberg (2006, p. 195).

20. Matsuyama (1999).

21. Gansler (1989, p. 218).

22. Christensen (1997, p. 138).

23. Christensen (1997 p. 139).

24. This is the strategy followed by big R&D spenders such as Cisco, Microsoft, Novartis, and GSK.

25. Gompers and Lerner (2003, p. 19).

26. For a presentation of DARPA’s history, see e.g., Edwards (2005).

27. OnPoint Technologies is a $25 million, not-for-profit, strategic private equity organization aimed at investing in promising power-supply technologies.

28. Gompers and Lerner (2003, p. 24).

29. McCullagh (2000).

30. A technology broker created by financial and investment institutions to exploit MOD’s technology portfolio, DTE failed as it only tried to make profit in a short-term perspective — assuming that once a blueprint, patent, or idea had been located, it could be easily transferable — rather than investing to develop commercial applications of defense-funded applications. For further analysis on the misleading mix between commercial targets and defense issues, see Bellais and Guichard (2006).

References

Bellais, R. 1999. Production d'armes et puissance des nations. Paris: L'Harmattan.

Bellais, R., ed. 2004. Économie et Défense: Nouvelles frontières entre États et marchés. Paris: Éditions Descartes et Cie.

Bellais, R. and R. Guichard. 2006. “Defense innovation, technology transfers and public policy.” *Defence and Peace Economics*, Vol. 17, No. 3 (June), pp. 273-286.

Bellais, R. and G. Le Blanc. 2002. “How is Terrorism Disruptive for Conflict Analysis.” Sixth Annual Middlesex Conference on Economics and Security.

London. June.
Carter, A. 2001. “Keeping the technological edge,” pp. 129-163 in A. Carter and J. White, eds. *Keeping the Edge: Managing Defense for the Future*. Cambridge, MA: The MIT Press.
Christensen, C. 1997. *The Innovator’s Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press.
Edwards, J. 2005. *The Geeks of War: The Secretive Labs and Brilliant Minds Behind Tomorrow's Warfare Technologies*. New York: Amacom Books.
Gansler, J. 1989. *Affording Defense*. Cambridge, MA: The MIT Press.
Gansler, J. 1998. “Defense Science and Technology: Looking to the Future.” DARPA 40th Anniversary Symposium. Arlington, 8 April.
Gompers, P. and J. Lerner. 2003. “Short-term America revisited? Boom and Bust in the Venture Capital Industry and the Impact on Innovation,” pp. 1-27 in A. Jaffe *et al*., eds. *Innovation Policy and the Economy*, Vol. 3. Cambridge, MA: The MIT Press.
Hartley, K., R. Bellais, and J.P. Hébert. Forthcoming. “The Evolution and Future of European Defense Firms” in M. Chatterji and J. Fontanel, eds. *Economics of International Security*. New York: North Holland/Elsevier.
Kurth, J. 1993. “The Follow-on Imperative in American Weapons Procurement, 1960-90,” pp. 304-321 in J. Brauer and M. Chatterji, eds. *Economic Issues of Disarmament*. London: McMillan.
Lichtenberg, F. 1995. “Economics of Defense R&D” in T. Sandler and K. Hartley, eds. *Handbook of Defense Economics*, Vol. 1. Amsterdam: Elsevier Science.
Matsuyama, K. 1999. “Growing through Cycles.” *Econometrica*, Vol. 67, No. 2, pp. 335-347.
McCullagh, D. 2000. “CIA Looks to Vnture Capital.” *Wired*. 19 October.
Middleton, A., S. Browns, K. Hartley, and J. Reid. 2006. “The Effect of Defence R&D on Military equipment Quality.” *Defence and Peace Economics*, Vol. 17, No. 3 (June), pp. 117-139.
[OUSDC] Office of the Under Secretary of Defense — Comptroller, 2008. National Defense Budget Estimates for FY2009. Washington, DC: U.S. Department of Defense, March.
Schinasi, K.V. 2008. “Better weapon program outcomes require discipline, accountability, and fundamental changes in the acquisition environment." Testimony before the Committee on Armed Services, U.S. Senate, Washington, DC: U.S. Government Accountability Office, 3 June.
Setter, O. and A. Tishler. 2006. “A Brave Leap or a Gradual Climb? The Dynamics of Investment in R&D of Integrative Technologies.” *Defence and Peace Economics*, Vol. 17, No. 3 (June), pp. 201-222.
Trajtenberg, M. 2006. “Defense R&D in the Anti-terrorist Era.” *Defence and Peace Economics*, Vol. 17, No. 3 (June), pp. 177-199.

Defense R&D and national R&D systems: a European outlook

Sylvain Daffix and *Yves Jacquin*

Defense R&D addresses a highly specific purpose: equipping the armed forces with up-to-date technologies. This R&D responds to the requirements of a monopsony, the government, which acts as the market maker for these products by direct procurement and export authorizations. But defense R&D is also embedded in national R&D systems. The question of the place and role of the defense R&D effort within a national R&D system is of particularly high interest if we take into consideration the possible links and externalities between civil and defense R&D.

A 1998 study of the U.S. national innovation system arrived at two stylized facts on the evolution of R&D efforts there during the 1990s. First, the share of federal funding in R&D was decreasing due to the reduction of defense-related R&D. Second, industry funded a growing part of the R&D effort, thus inducing a high priority given to development rather than basic research.[1] But instead of considering national systems of innovation (NSI) in general, in this article we concentrate our attention on a more specific item and, to some extent, on the core of NSI and technological innovation, namely, the organization of research and development. In particular, we consider "a set of institutions that (jointly and individually) contribute to the development and diffusion of new technologies. These institutions provide the framework within which governments form and implement policies to influence the innovation process." [2] This definition applies both to innovation in general and to R&D in particular.

This article reviews differences in European national R&D systems especially in terms of the place of defense R&D. It is based on annually collected data by the OECD, consistent with the conventions of the so-called Frascati manual.[3] The data contain information on who *finances* and who *performs* R&D. Elements of defense R&D are included, and this allows us to illustrate the place of defense R&D in national R&D systems.

The next section discusses the data, and previous work based on it, and focuses on the evolution of defense R&D in the six European countries who signed the so-called Letter of Intent (or LOI) and who together account for more than 90 percent of defense R&D in the EU.[4] The section thereafter details the role of defense R&D in the French national system of innovation in particular.

European comparisons

We first discuss the role of government in the respective national R&D efforts, then we detail the split of public funding into civil and defense objectives. All monetary amounts are expressed in US$ purchasing power parities (exchange rates are defined by reference to a representative basket of goods and services). But methodological constraints on data gathering are examined first to highlight that and why one cannot directly compare defense R&D across different funding sources for gross domestic expenditure on research and development (GERD).[5]

The U.K. has little public R&D funding, yet a large part of this is dedicated to defense; France has a larger public R&D funding but a relatively small part of this is dedicated to defense. In contrast, Germany has a low level of defense R&D despite a large overall R&D effort, and despite having twice Spain's population, it spends less on defense R&D than does Spain.

Methodology

The OECD collects annual data on R&D from its member states. To be comparable these data must be built on the same definition and data collection method. An OECD meeting with national experts took place in Frascati, Italy, and the resulting manual is therefore called the Frascati manual.[6] But because two types of data gathering are used, namely survey data and budgetary data, the data are not directly comparable.

Several studies using these data to examine defense R&D have already been published.[7] One author uses the data to show the growing defense R&D gap between the United States and Europe. For the 1991 to 2003 period, three results in particular are consistent with our own findings. First, "defense R&D spending in Europe is highly concentrated"; second, "defense R&D budgets fell significantly during the post-Cold War period"; and third, "patterns of defense R&D spending vary amongst the LOI Six countries."[8] (The six countries are France, Germany, Italy, Spain, Sweden, and the United Kingdom.) Our article goes further in that it presents defense R&D not merely on its own terms, but as embedded within the respective national R&D systems. Nevertheless, there are limits to the OECD data, for example in the fuzzy definition of defense R&D and especially the inadequate consideration of dual-use technologies.[9] But no other data source is available for our purposes.

Before presenting the European data, we take a brief look at the U.S. case. In absolute terms, the U.S. defense R&D effort is six times larger than the sum of the efforts of the six members of the European Letter of Intent. The U.S. figure represents 56.9 percent of total public R&D funding; the corresponding figures in the LOI countries range from 3.6 to 31 percent.[10] The U.S. number was even larger in the late 1950s, approximately 80 percent.[11] "Defense-related R&D and procurement programs provided a powerful impetus to the development and commercialization of new civilian technologies in commercial aerospace, semiconductors, computers, and

Table 1: Gross domestic expenditure on research and development (GERD), 2005, by funding source and performance

Panel (a): GERD, 2005, by source of funds in billions $ ppp (and in % of total)

	Business sector	*Government sector*	*Other nat. sources*	*Non-national sources*	*Total*
France	21.4 (52.6)	15.5 (38.2)	0.8 (1.9)	3.0 (7.3)	40.6
U.K.	14.8 (42.1)	11.5 (32.8)	2.1 (5.9)	6.8 (19.2)	35.2
Germany	42.7 (67.6)	17.9 (28.4)	0.2 (0.3)	2.3 (3.7)	63.1
Italy	7.2 (39.7)	9.2 (50.6)	0.3 (1.7)	1.4 (8.0)	18.1
Spain	6.2 (46.3)	5.8 (43.0)	0.7 (5.0)	0.8 (5.7)	13.4
Sweden	7.4 (65.7)	2.6 (23.5)	0.3 (3.1)	0.9 (7.7)	11.3

Panel (b): GERD, 2005, by sector of performance (billions $ ppp)

	Business sector	*Government sector*	*Higher ed. sector*	*Private, non-profit sector*	*Total*
France	25.5 (62.5)	7.2 (17.6)	7.6 (18.6)	0.5 (1.3)	40.7
U.K.	21.7 (61.6)	3.7 (10.6)	9.0 (25.6)	0.8 (2.2)	35.2
Germany	43.7 (69.4)	8.9 (14.1)	10.4 (16.5)	0.0 (0.0)*	63.1
Italy	9.1 (50.4)	3.1 (17.3)	5.5 (30.2)	0.4 (2.1)	18.1
Spain	7.2 (53.9)	2.3 (17.0)	3.9 (29.0)	0.0 (0.1)*	13.4
Sweden	8.4 (73.8)	0.5 (4.7)	2.4 (21.2)	0.0 (0.3)*	11.3

Note: * $ ppp are not zero but rounded to one decimal place, hence small percentages do result in the parenthetical expression.
Source: OECD (2007/2).

computer software."[12] If so, it should be worthwhile studying the role of European defense R&D efforts in their own national R&D system.

National R&D systems in Europe

Research capabilities, both private and public, are a key element of the future success of the economy in industrialized countries. In 2000, the European countries decided in Lisbon to undertake an unprecedented R&D effort to make Europe the most productive and competitive knowledge economy in the world. The quantitative target is for R&D expenditure to reach three percent of GDP in 2010. For the six LOI countries, only Sweden has met the target (in fact, it has exceeded it). Worryingly, some of the other countries are showing a decreasing R&D spending trend for the past decade: by 2005 France was down to 2.13 percent and the U.K. to 1.78 percent.[13]

These differences in national R&D efforts are even more striking when partitioned by source of funds as well as by sectors of performance, showing different institutional arrangements to deal with R&D.

Table 1, panel (a) shows significant differences from one country to another. Due to its size, Germany spends the most, followed by France and the U.K. The German business sector alone spends as much on R&D as all of France does. The business sector provides two-thirds of R&D funding for Sweden and Germany but only one-half in France and Spain, and a little less than that in the U.K. Government funding is particularly important for France (more than one third) and for Italy and Spain. Britain shows an unusual pattern in that an important share of R&D funding comes from abroad. Table 1, panel (b) shows the distribution of GERD by the sector in which R&D is carried out (the sector of performance). In this regard, the LOI Six are more comparable than the distributions of funding but two groups of countries nonetheless stand out: Italy and Spain show a relatively low share of R&D carried out in the business sector and with a comparatively high share in higher education; in contrast, Germany, France, and the U.K. show that business performs about two-thirds of R&D performance and the government and higher education shares are relatively balanced. Sweden stands apart with about three-quarters of GERD performed by the business sector and very little by the government sector.

Some initiatives are taken to enhance both the quantitative effort of R&D at the European level and in terms of international collaboration. The most well-known and, in budget terms, important initiative are the Framework Programs for Research and Technological Development which helped to structure the European Research Area, but in the past the defense field was explicitly excluded from this funding. The seventh Framework Program, starting in 2007, includes a cooperative program on "security" for the first time. But the budget is only €1.4 billion for the period 2007-2013. Another initiative, specific to the defense sector, should also be mentioned: the relatively new European Defense Agency (EDA) which already has launched multinational programs for Research and Technology (R&T).[14] None of these initiatives apply automatically to all participating member states.

Despite efforts to develop European collaboration within the Research Areas, the organization of national R&D efforts shows large differences, and it is noticeable that countries with stronger efforts in terms of the R&D share in GDP (Germany and Sweden) are also those where private funding and private sector R&D work are prominent. At the opposite end of the scale, Italy and Spain depend heavily on public efforts.[15]

Defense R&D

As compared to other studies, the definition of defense R&D used in this article is quite restrictive, but this allows us to give a clearer view of its publicly funded share and also of its evolution throughout the last decade. Table 2 presents data on Government Budget Appropriations or Outlays for R&D (GBAORD) dedicated to defense, i.e., public funding of defense R&D wherever it is performed.

In relative terms, the main defense R&D efforts are made by the United Kingdom and France, both of whom also have the largest defense budgets in Europe.[16] Sweden and Spain have show a significant share of GBAORD dedicated to defense. The numbers for Germany and Italy are surprisingly low. For the latter, this may be due to the fact that defense R&D is supported by the Ministry of Industry. Comparing the numbers for 1995 with 2005, the tendency is decreasing for all except Spain (the explanation being that in Spain defense R&D was switched from the Ministry of Industry to the Ministry of Defense with the advent of the Eurofighter Typhoon program).[17]

Comparing the numbers in Tables 1 and 2, defense R&D budgets are particularly important in the national R&D efforts of France and Spain, due to the high shares of public funding in GERD associated with the significant shares of defense R&D in GBAORD. The United Kingdom is in the same situation especially for its important defense share in GBAORD. In contrast, Germany has both comparatively low public funding of GERD and a low share of defense R&D in GBAORD. Its defense R&D effort is comparatively smaller than those of its European counterparts.

In monetary terms, defense R&D is highest in the United Kingdom and France, Spain being third and significantly above Germany (although the Spanish numbers are not consistent with the data published by EDA in 2006). Italy and Sweden lag far behind in monetary terms, the former because it has a small share of defense R&D in GBAORD, and the latter due to the small size of its economy. The Swedish GBAORD is only about one quarter of Spain's and a mere 13 percent of Germany's. Generally speaking, the Letter of Intent effort is increasing in the later years of the period studied, reflecting the end of the post-Cold War peace dividend era.

Comparing 1995 and 2005, the level of defense R&D has been roughly stable for France, Italy, and Sweden. But Italy and Sweden show a sharp drop during the late 1990s and a similar recovery since then. The United Kingdom and especially Spain have shown continuous defense R&D growth, and Germany a slight decrease. As indicated, some of these trends can be explained by external factors such as changes in ministerial responsibilities and boundaries. Except for the Spanish case, the results appear to be globally consistent with the statistical survey of the EDA.[18] The EDA figures also show that only ten percent of the R&T of its participating member states is conducted under European collaboration. While Europe's main industrial groups are well-integrated, they still are organized along national lines so as to benefit from national funding. Even if a program is nominally conducted under European collaboration, a nation-state based prime contractor has difficulties to transfer knowledge and components from one country to another. Thus, room remains for a higher level of integration in the defense R&D area.

Table 2: Defense budget R&D as a percentage of total government budget appropriations or outlays on research and development (GBAORD) and defense budget R&D (in billion $ ppp)

	1995 (%)	*2000* (%)	*2005* (%)	*1995* ($)	*2000* ($)	*2005* ($)
France	30.0	21.4	20.8	4.1	3.2	3.9
United Kingdom	36.5	36.2	28.3	3.3	3.8	4.1
Germany	9.1	7.8	5.8	1.4	1.3	1.1
Italy	4.7	0.8	3.6	0.3	0.1	0.4
Spain	10.4	26.2	16.4	0.3	1.3	1.6
Sweden	20.9	7.1	17.4	0.4	0.1	0.4

Source: OECD MSTI 2006/1 and 2007/2.

The French case

The role of defense in the national R&D system

The French Ministry of Defense performs, by itself or through contracts with other organizations, both public and private, the R&D needed for equipping its armed forces. (French defense R&D means as funded by the French Ministry of Defense.) Claude Serfati showed that since the late 1950s, defense *grand programs* played an essential role in building and developing the French national system of innovation.[19] Figure 1 shows the links between public and private actors of the national effort of R&D. A significant amount of R&D is paid by the public sector but performed by the private sector and vice versa.

For 2004, the whole of public sector R&D funding represents nearly half of the total R&D funding (Figure 1). But in terms of where R&D is performed, its share is significantly lower (at €13.3 billion, it is just above one-third of the total). Two-thirds of public funding of R&D performed by the business sector is coming from defense (€1.7 billion defense R&D vs. €0.9 billion civilian R&D). Of R&D performed in-house by the public sector, defense represents approximately 10 percent (€1.2 billion vs €10.7 billion). And of the total government-funded defense R&D, €1.7 billion flows into the private sector, and only €1.2 billion stays within the government

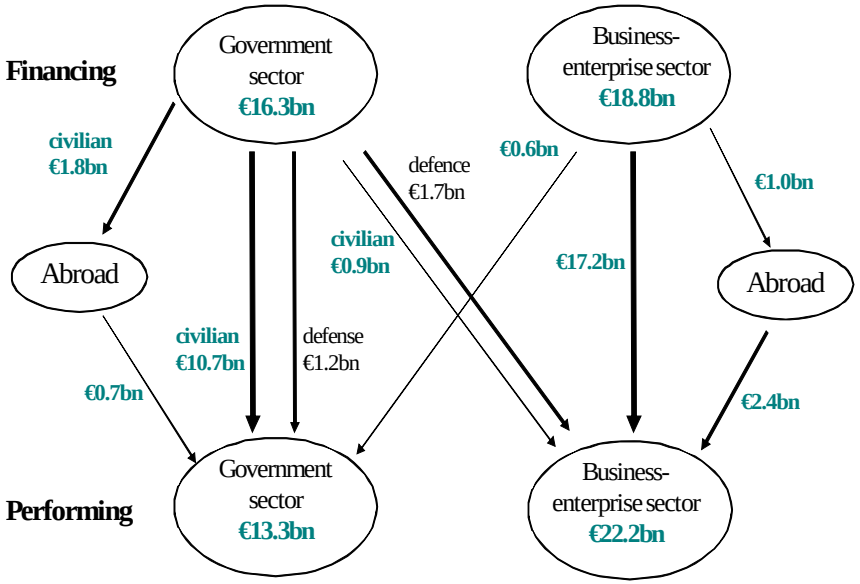

Figure 1: The French system of research and development
Source: French Ministry of Research data, 2004

sector. Thus, defense R&D constitutes a key link between public and private R&D efforts.[20]

Despite the recent defense R&D increase due to the 2003-2008 Military Programming Law that gave priority to equipment procurement, as compared to the early 1990s, the share of defense R&D in national R&D has declined (see Table A1 in the appendix). The share of defense R&D is U-shaped for the period 1992-2005, the bottom phase occurring in 2001. The level attained in 2005 is still far lower than the level attained in the mid-1990s, even when measured in nominal euro terms.

This quantitative change was accompanied by a structural change in the objectives of public funding. Traditionally, the French national system of innovation was both highly concentrated and publicly managed.[21] Nowadays, this is less true.[22] The strategy of the *grand programs*, both civil and defense, is no longer a priority and funding of bottom-up initiatives from the industrial sector and especially from small and medium enterprises (SMEs) are increasing. The growing role of complementary funding for private initiatives like the OSEO group, the Agence Nationale pour la Recherche and other institutions dedicated to the new *pôles de compétitivité* have changed the basic role of the public sector; its orientation and management activity is decreasing and it increasingly performs strictly a funding role.[23]

These new types of interventions carry strategic implications for defense R&D, even if not directly funded by the state. Since 2005, the French government identified 71 research clusters, each accompanied by a ministerial committee: out of the 7 "global competitiveness clusters," the Ministry of Defense (MoD) participates in the 5 committees dedicated to software, aeronautics, and nano and biotechnology. Among the 10 "globally-oriented competitiveness clusters," the MoD is present in 3 committees, and it also participates in 5 of the 54 other "competitiveness clusters."[24] This permits the MoD to keep in touch with the leading edge of technology, even if it is not (yet) specific to the defense sector. It is of major importance to identify possible disruptive technologies and other solutions are being explored, such as the use of venture-capital in strategic sectors.[25]

The major part of useful defense technology still comes from direct investments by the MoD. Nevertheless, since the early 1990s not only did the amount of defense R&D change significantly but so did its distribution. Table A2 shows two major trends. First, total R&D spending by the French Ministry of Defense decreased from 1992 to 1998 and then remained stable until 2001. From this date onward, an increased effort has taken place, but still not enough to bridge the gap. Second, the distribution of defense R&D funding between in-house performance and contracts given to firms and other research institutes also changed.[26] While both types of spending were reduced, the cut was 44 percent for internal research, but only 20 percent for research contracts performed outside the Ministry. Private firms therefore carried out a growing share of defense R&D. This finding is consistent with the new French arms procurement strategy, the *délégation générale pour l'Armement*, to give more responsibility to the business sector even in terms of research and development.[27]

The changing role of the defense industrial base in the French R&D effort

R&D funded by the Ministry of Defense is increasingly carried out by private firms. However, this funding is concentrated on a small number of companies, about 110 to 120 firms each year. Even if the number of SMEs financed directly by the Ministry remains stable over the period, the funding they received has been reduced. This can be explained in part by the transfer of responsibilities to large industrial prime contractors. Certain R&D contracts for SMEs are still maintained by the Ministry to secure access to disruptive technologies and to develop specific areas of interest.

Table A3 shows that defense R&D funding is a significant part of firms' total R&D expenditure, and its contribution is roughly stable (around 7 percent) since the late 1990s. Even if highly concentrated, the defense sector represents a significant source of funding for the business sector.

About one-quarter of firms' R&D in France is carried out by the few firms whose research is directly financed by the Ministry of Defense. These firms also account for about one-third of R&D outsourcing. This includes all industrial prime contractors and some very innovative SMEs. This is likely to have an important effect on driving the whole defense industrial base, as shown by its share in R&D and in particular in

Table 3: Distribution of defense funding toward firms by industrial sector

Sector	*Share of funding (%)*
Aerospace	48.8
Medical, precision, and optical instruments	17.4
Telecommunications	14.3
Weapons and ammunition	13.1
Parachemistry	3.3
Services	0.7
Other	2.4
Total	100.0

Source: Ministry of Research data, 2004.

the research outsourced by firms. These one hundred firms represent 10 percent of the sales of innovative firms, 23 percent of their R&D performed, and 33 percent of their outsourced R&D. The comparisons show that these firms are leaders of innovative networks. They are also highly concentrated in terms of the industrial sectors in which they operate (Table 3). The aerospace sector obtains nearly half of defense R&D funding. Together, the four main sectors (aerospace, instruments, telecommunications, and weapons) absorb more than 90 percent of the total.

Despite the significant overall funding decrease during the 1990s, defense funding steered toward industrial firms is still significant. The findings presented in this article demonstrate the growing role of industrial firms in the French defense R&D system as well as in the technical definition of defense equipment. Industrial skills and technical competencies seem to move from the Ministry to the main prime contractors. This evolution is due to internal as well as external factors, such as shrinking budgets, defense equipment cost escalation, and international cooperation led by European prime contractors.

The distribution of R&T credits (*études amonts*) among prime contractors illustrates the major role taken by Lead Systems Integrators such as EADS and Thales that capture almost half of the funding (Table 4). This is due both to the recent acquisition and merger activity within the French defense industrial base and to the new acquisition strategy of the French arms procurement agency, DGA, that tends to globalize contracts. It should also be noted that small and medium enterprises receive a steady share (10-11%) of the *études amonts* funding.

Conclusion

While limited in scope, the data used for this article permit us to extract stylized facts regarding patterns of national R&D systems in selected European countries. One finding is to show up the diversity of national R&D systems and the even greater diversity in defense R&D. There is no general link between the role of public effort in total R&D spending and priorities given to defense concerns.

The two European leaders in defense R&D are France and the United Kingdom, consistent with their overall defense effort and their strategic and international objectives. But their national R&D systems are quite different. The U.K. has little public R&D funding yet a large part of this is dedicated to defense (comparable to the United States); France has a larger public R&D funding but a relatively small part of this is dedicated to defense. In contrast, Germany has a low level of defense R&D despite a large overall R&D effort; despite having twice Spain's population, it spends less on defense R&D than does Spain.

As regards France, the trend in defense R&D is rising since the late 1990s but it has not yet closed the gap to the early 1990s, even in nominal terms. The data show the effect of the new strategy of the French defense procurement agency favoring industrial firms as from the late 1990s. Increasingly, France relies on the private sector to perform R&D, even for the defense sector. In 2004, two-thirds of publicly funded but privately performed R&D came from the Ministry of Defense; it plays therefore a major role in the relation between public and private R&D efforts.

Table 4: Main contractors in terms of R&T (études amonts)

	1998	*2003*
Alcatel	2	3
Dassault Aviation	6	3
Aérospatiale/Matra	8	-
Eurocopter	2	-
EADS	-	22
Giat industries	3	3
Sagem	3	3
Snecma	5	4
SNPE	3	1
Thomson	25	-
Thales	-	27
SME	10	11
others	33	23
Total	100	100

Source: Fromion (2005).

Notes

Sylvain Daffix, the corresponding author, is with the Economic Observatory of Defense, French ministry of Defense, and with GREDEG, UMR CNRS 6227, Université de Nice-Sophia-Antipolis. He can be reached at daffix@laposte.net. **Yves Jacquin** also works at the Economic Observatory of Defense, French ministry of Defense and at INSEE, the Institut National de la Statistique et des Etudes Economiques. The views and opinions presented here are solely those of the authors and do not represent those of the French Ministry of Defense.

1. Mowery (1998).

2. Following Metcalfe (1995), as cited by Sharif (2006).

3. OECD (2002).

4. EDA (2006).

5. OECD (2002).

6. The latest version is OECD (2002).

7. James (2006) for the United Kingdom; Molas-Gallart for Spain (1999); and Sachwald (1999) and Guichard (2004) for France. The quotes that follow in the text are taken from James (2006).

8. James (2006, p. 226, Table II).

9. See Molas-Gallart (1999).

10. OECD (2007).

11. Mowery (1998).

12. Mowery (1998, p. 640). In a recent book on defense technology, Ruttan also underlined the development of general purpose technologies induced by World War II and the Cold War (Ruttan, 2006, p.185).

13. On the U.K., see Hall and James in this issue.

14. The R&T aggregate is specifically used in the defense area and corresponds, roughly, to both basic and applied research.

15. In the 1980s the share of private R&D funding was increasing to the detriment of government funding in every OECD country studied (France, Germany, Japan, the United Kingdom, and the United States). But no clear-cut evidence in terms of R&D execution was found. See Mowery (1998).

16. EDA (2006).

17. Molas-Gallart (1999).

18. EDA (2006).

19. Serfati (1998 p. 21).

20. Also see Daffix and Jacquin (2007).

21. Mustar and Laredo (2002).

22. Foray (2001).

23. OSEO is a holding company with public status. Its mission is to provide assistance and financial support to French SMEs and VSEs in the most decisive phases of their life cycle: start up, innovation, development, business transfer, or buy out. By sharing the risk, it facilitates the access of SMEs to financing by banking partners and equity capital investors.

24. Ministry of Defense, PLF2008 (2007).

25. See Bellais, in this issue.

26. Bellais and Daffix (2005).

27. See Bellais, in this issue.

References

Bellais, R. 2009. “Defense innovation at any (out of control) cost? The stalemate of today’s R&D policy and an alternative model.” *The Economics of Peace and Security Journal*, Vol. 4, No. 2.

Bellais, R. and S. Daffix. 2005. “Evolutions de l’effort public et privé de recherche-développement et capacités d’innovation des firmes,” pp. 87-110 in R. Bellais, ed., *Economie et Défense*. Paris: Editions Descartes.

Daffix, S. and Y. Jacquin. 2007. “Quelques repères statistiques sur la R&D de Défense en France.” *Ecodef*. No. 48, (October) pp.1-4.

[EDA] European Defense Agency. 2006. *National Defense Expenditure in 2005*. http://www.eda.europa.eu/genericitem.aspx?area=Facts&id=179 [accessed 15 December 2007].

Foray, D. 2001. “On the French system of innovation: between institutional inertia and rapid changes.” Working paper 01/03. Paris: IMRI, Dauphine University.

Fromion, Y. 2005. “Rapport d'information n˚2150 sur la recherche de défense et de sécurité.” Paris: Assemblée nationale.

Guichard, R. 2004. “Eléments pour un repositionnement de la R&D de Défense au sein du système d'innovation français.” *Revue d'économie industrielle*, No. 108, 4ième trimestre, pp.7-20.

James, A. 2006. "The Transatlantic Defense R&D Gap: Causes, Consequences and Controversies." *Defense and Peace Economics*, Vol. 17, No. 3 (June), pp. 223-238.

[MESR] Ministry of Research. Various annual editions. *La recherche et développement en France*. Paris: MESR.

Ministère Français de la Défense. 2006. *Annuaire statistique de la Défense*. Paris: Ministry of Defense.

Molas-Gallart, J. 1999. "Measuring Defense R&D: A Note on Problems and Shortcomings." *Scientometrics*, Vol. 45, No. 1, pp. 3-16.

Mowery, D. 1998. "The changing structure of the U.S. national innovation system: implications for international conflict and cooperation in R&D policy." *Research Policy*, Vol. 27, pp. 639-654.

Mustar, P. And P. Laredo. 2002. "Innovation and research policy in France (1980-2000) or the disappearance of the Colbertist state." *Research Policy*, Vol. 31, pp. 55-72.

[OECD] Organization of Economic Development and Cooperation. Various years. *Main Science and Technology Indicators*. Paris: OECD.

[OECD] Organization of Economic Development and Cooperation. 2002. *The Frascati Manual*. Paris: OECD.

[OECD] Organization of Economic Development and Cooperation. 2007. *Main Science and Technology Indicators*. Paris: OECD.

Ruttan, V. 2006. *Is War Necessary for Economic Growth? Military Procurement and Technology Development*. New York: Oxford University Press.

Sachwald, F. 1999. "Banalisation et restructuration des industries de défense." *Les notes de l'Ifri*, No.15. Paris: IFRI.

Serfati, C. 1998. *Les industries européennes d'armement, de la coopération à l'intégration?* Paris: La documentation française.

Sharif, N. 2006. "Emergence and development of the National Innovation Systems concept." *Research Policy*, Vol. 35, pp. 745-766.

Appendix

Table A1: Distribution of national R&D funding

	National expenditure on R&D	*Business sector*	*Gov't sector (civilian)*	*Gov't national (defense)*	*Defense/ national expenditure*	*Defense/ total gov't sector*
1992	26,229	12,769	9,136	4,324	16.5%	32.1%
1993	27,002	13,307	9,620	4,075	15.1%	29.8%
1994	26,995	13,468	9,529	3,998	14.8%	29.6%
1995	27,563	13,916	10,184	3,463	12.6%	25.4%
1996	28,091	14,373	10,337	3,381	12.0%	24.6%
1997	28,005	15,025	10,320	2,660	9.5%	20.5%
1998	28,724	15,865	10,423	2,436	8.5%	18.9%
1999	29,885	16,618	10,760	2,507	8.4%	18.9%
2000	31,438	17,166	11,738	2,534	8.1%	17.8%
2001	33,570	18,897	12,163	2,510	7.5%	17.1%
2002	34,759	19,082	12,897	2,780	8.0%	17.7%
2003	34,395	18,505	13,061	2,830	8.2%	17.8%
2004	35,136	18,831	13,395	2,910	8.3%	17.8%
2005	37,125	20,156	13,861	3,108	8.4%	18.3%

Note: Data in € millions
Source: [MESR] Ministry of Research, several years

Table A2: Distribution of the defense R&D budget by R&D-performing sector

	1992	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*
Defense expenditure on R&D	4,323	4,075	3,999	3,464	3,381	2,660	2,436	2,507	2,533	2,509	2,781	2,829	2,910	3,108
intramural R&D	1,973	2,005	1,945	1,589	1,585	1,013	774	777	777	848	874	993	1,075	1,175
Extramural R&D	2,350	2,070	2,054	1,875	1,796	1,647	1,662	1,730	1,756	1,661	1,907	1,836	1,835	1,933
of which performed by:														
- the government sector	163	158	222	206	242	230	362	299	234	206	278	205	130	119
- the higher education sector	17	16	19	26	25	14	11	8	7	6	7	8	7	6
- the nonprofit private sector	4	5	8	8	8	0	0	0	0	0	0	0	1	1
- the business sector	2,149	1,874	1,788	1,618	1,500	1,386	1,273	1,407	1,497	1,432	1,604	1,608	1,680	1,790
- the foreign (abroad) sector	17	17	17	17	21	17	16	16	18	17	18	15	18	17

Note: Data in € millions
Source: [MESR] Ministry of Research, several years

Table A3: Total expenditure of firms' R&D by sector of funding

	*Total expenditure**	*| financed by: | (a) firms*	*(b) government*	*of which: (a) defense*	*(b) abroad*
1991	16,567	11,395	3,423	n/a	1,750
1992	17,664	12,588	3,136	2,133	1,940
1993	17,957	13,221	2,897	1,874	1,840
1994	17,890	13,352	2,681	1,788	1,858
1995	17,979	13,778	2,351	1,626	1,850
1996	18,471	14,195	2,331	1,500	1,945
1997	18,612	14,785	1,994	1,386	1,834
1998	18,972	15,497	1,824	1,273	1,651
1999	20,004	16,183	2,174	1,407	1,646
2000	20,971	16,962	2,259	1,497	1,749
2001	22,591	18,680	2,110	1,432	1,800
2002	23,605	18,871	2,498	1,604	2,236
2003	23,021	18,318	2,444	1,608	2,258
2004	23,562	18,545	2,620	1,680	2,396

Note: Data in € millions
Source: [MESR] Ministry of Research, several years
* Total expenditure includes R&D performed by firms plus R&D outsourced by firms toward government sector and abroad

The Economics of Peace and Security Journal, ISSN 1749-852X
© www.epsjournal.org.uk – Vol. 4, No. 1 (2009)

Industry structure and innovation in the U.K. defense sector

Peter Hall and *Andrew James*

In the United Kingdom, as elsewhere, the defense industry has undergone substantial structural change since the end of the Cold War. This is reflected in consolidation at the overall sector level, although with wide variations among subsectors in the potential for competition. While the nature and strength of causal connections remains a subject of debate, economic theory suggests that such consolidation might be expected to have implications for industry innovation. In this article, we observe that the defense industry's own-funded research and development (R&D), a key metric for innovation, fell sharply as a proportion of total defense industry R&D between 1989 and 2005. Statistics available in the public domain constrain exhaustive analysis of the connections, if any, between structural change in the sector and innovation, but we draw on recent advances in the theoretical debate to argue that observed changes are consistent with a nonlinear (inverted U-shape) relationship.[1] We suggest that Ministry of Defence (MoD) policies may also have played a role, and that innovation may have fed back into structural change.

The changing structure of the U.K. defense industry

The notion of "defense industry" is ambiguous for reasons that are well known. Inputs ultimately used to form military capability, for example, are also used in the non-defense economy and produced by firms that predominantly serve civil markets. Many firms that could, in principle, produce for defense, and may in the past have done so, do not currently sell to the military and do not seek to. And if one considers the defense industry of a specific country, confounding issues surround the nationality of firms' ownership and the location of production to meet the demands of the domestic defense purchaser. Which firms may therefore be said to comprise the defense industry, generally and in specific nations, ultimately becomes a matter of judgment. In this study, we focus on firms which were awarded prime contracts by the U.K.'s MoD. Some of these firms were foreign-owned, a matter taken up again below. Some component manufacture and production work to meet contractual obligations has undoubtedly been performed offshore (and may well have increased) but information on the public record is insufficient to allow us to quantify the proportions involved.

Our analysis of change is based on observations of the firms in receipt of major contract awards first in 1990/1 and then in 2005/6. We note first that the number of companies undertaking the largest defense contracts shrank during our period of study. At the level of the largest contracts (over £375 million in 1990/1 and over £500 million in 2005/6), five firms were at work at the start of the 1990s but by the mid-2000s, the number had dropped to four. In relation to the next lower-valued range of contracts, £151-375 million (1990/1) and £250-£500 (2005/6), the number of firms fell from 11 to 9.[2] Second, membership of the lists of prime contractors changed substantially. Of the total of 16 firms in the two contract-value brackets in 1990/1, only four remained among the 13 in the corresponding groups in 2005/6. British Aerospace had merged with GEC Marconi to form BAE Systems, by 2005/6 the single largest business in the U.K. industry, and Babcock International and Rolls Royce were also still players. Among firms to disappear from the list of largest prime contractors were VSEL and while a number of medium-sized contractors (e.g., Racal, Ferranti, and Alvis) also exited. Third, the supplier base became more international as U.S., Italian, French, and European transnational companies entered the U.K. industry to replace the mostly local firms that departed. New overseas players undertaking contracts in the two value-brackets included EADS (Europe), EDS, General Dynamics, Halliburton, Lockheed Martin, and Raytheon (all U.S.), Finmeccanica (Italy), and Thales Defense (France).

As regards platform systems level capability, three subsectors in the British defense industry are now dominated by a single firm: BAE Systems for fast jet combat aircraft and for armored fighting vehicles, and Marshall for strategic airlift capacity. In the helicopter subsector there are two companies, five in complex surface warships, and as many as 15 in non-embedded C4ISTAR. How may these different industry structures affect innovation within the industry?

The outcome of the changes noted was consolidation throughout U.K. defense industry, on the one hand, and an increase in the role of foreign-owned companies on the other. While the overall effect has been to reduce the number of firms competing with each other, the extent of competition within subsectors of the industry (judged by number of firms) remains highly variable. MoD information on companies possessing what it calls Platform Systems Level Capability (PSLC) in the U.K. captures some of the differences.[3] At one extreme, three subsectors are now dominated by a single player: BAE Systems has for many years been the single source of PSLC in the case of fast jet combat aircraft. Thanks to consolidation over the last decade or so, it now also occupies a similar position in relation to armored fighting vehicles. And only Marshall of Cambridge offers PSLC in the strategic airlift (C-130) subsector. At the other end of the spectrum, as many as 15 firms appear to offer PSLC in non-embedded C4ISTAR. Between the extremes, the remaining five sectors are populated by as few as two firms with PSLC (helicopters) and as many as five (complex surface warships and royal fleet auxiliary).

The consequences of industry consolidation for innovation in the defense sector have attracted considerable academic and policy attention outside the U.K., especially in the United States. Some have argued that the main danger posed by mergers is that technological competition will decline and that any remaining supplier(s) will stand fast against new design approaches that may threaten their position.[4] Others have observed that increased innovation in combat aircraft occurs at times of increased demand and the emergence of new component technologies and significant changes in military threat perceptions. Generally, larger numbers of experienced and credible prime contractors are more likely than lower numbers to promote the greater competition to innovate that leads to significant new technology development. In many cases, the key innovations that have led to radical change have come from firms who were not the dominant players at the time but members of a larger group surrounding them.[5] In a study of competition and innovation in the U.S. fixed wing military aircraft industry, it has also been argued that there were potentially serious questions about the level of competition and that innovation in a future environment might be dominated by one or two credible prime contractors.[6] Until now, there has been little or no academic discussion of the impact on industry level innovation of changes in the structure of the U.K. defense industry, and this is a gap to which we wish to draw attention. But the lack of data comparable to that available in the United States inevitably leads us to be more speculative than definitive in our analysis. In such cases, guidance from theory can be useful.

Industry competition and innovation: theory

Consolidation implies a reduction in the number of firms competing in an industry. As well as the number of firms currently active within it, however, the level of competition in an industry is affected by actual and perceived threats of potential entry. Furthermore, the relationship between the effect on competition of increasing or decreasing the number of firms may well be nonlinear: increasing the number from one (monopoly) to two (duopoly) may well yield a greater increase in competition than increasing the number from, say, five to six. Economists have often argued that when change of this kind occurs, the level and intensity of innovation effort may be affected. While the most commonly used measure of the level of competition is some version of the average price-cost margin within the industry (see the studies reported below), such information is impossible to construct for defense industry from the publicly available data. (And while many firms involved in defense contracts also supply civil markets, their published accounting data do not permit separation of the profits of one type of work from the other.) We therefore focus solely on the number of firms. We also follow the common practice in many economic studies of taking R&D as an acceptable proxy for innovation activity.[7]

Stated succinctly, the theory linking innovation to competition rests on the operation of forces that work in opposite directions. On the one hand, a so-called Schumpeter effect predicts that an increase in competition will whittle away profits that would otherwise have been available for investments in innovation. This implies a negative relationship between the level of competition and innovation. On the other hand, some scholars have argued that increasing competition prompts previously slack management teams now to invest more in innovation, with a view to avoiding the threat of takeover or bankruptcy by enhancing their firm's competitiveness.[8] Taken at face value, the first effect suggests consolidation should yield increasing innovation and the second decreasing innovation. The theory offers little guidance on (1) the relative sizes of the opposing effects, (2) whether the relative sizes of the effects might vary with the initial level of competition, and (3) whether we should be entitled to assume symmetry in the size and duration of the effects when competition decreases as well as when it increases. Empirical attempts to resolve such issues also need to acknowledge the implications of an important additional theoretical claim: that the levels of competition and innovation in an industry are simultaneously co-determined by a deeper force, "technological opportunity," that varies in its impact from one sector to another.[9] Technological opportunity is reflected in the (downward) responsiveness of production costs to a given percentage increase in innovation investment such as R&D. Over a cross-section of industries with similar demand responsiveness to price (price elasticity), higher levels of technological opportunity are predicted to yield both greater concentration and more innovative activity. It is thus necessary to control for technological opportunity when seeking to identify the relationship between industry structure and competitiveness on the one hand and industry innovation on the other.

Focusing on industries of similar technological opportunity (and also facing similar price elasticities), the predictions of theory turn out to be ambiguous. A given variation in the level of competition could be associated with either higher or lower levels of innovation. From any given starting point, an increase in competition could occur alongside either an increase or decrease in innovation, depending on the relative strengths of the two forces described above. The same is also true of a decrease in competition. When the Schumpeter effect dominates, more (less) competition is associated with less (more) innovation. When the Schumpeter effect is relatively less important, more (less) competition is associated with more (less) innovation.

Industry competition and innovation: recent empirical evidence

The relationship between industry structure and innovation has been of longstanding interest to economists, and much empirical analysis has been undertaken to clarify the nature and strength of the links involved. As Nickell points out, however, almost all of the earlier work in the area can be regarded as unenlightening because it fails to control for industry-specific variations in technological opportunity.[10] Studies that do allow for such variations come to opposite conclusions on the innovation impact of increasing competition.[11] But more recently researchers have started to converge on

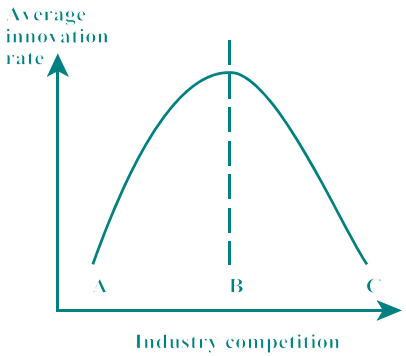

Figure 1: The inverted U-shape relationship between industry competition and innovation.

the finding that there is a nonlinear, inverted U-shape relationship between levels of competition and innovation.[12] When the number of competitors in an industry increases from the lowest possible base of a single firm (monopoly), industry-wide innovation intensity at first rises and then declines again as competition increases further. Representing such findings schematically in Figure 1, the segment AB is the range in which a positive relationship is found between industry competition and innovation, an increase (decrease) in one being associated with an increase (decrease) in the other. Along the segment BC, the relationship becomes negative, an increase (decrease) in one being associated with a decrease (increase) in the other.

If the theory we have outlined has explanatory power, then (controlling for technological opportunity and price elasticity) the inverted U-shape implies the Schumpeter effect dominates at low levels of competition but ceases to be dominant when the number of firms increases above a critical level. In their empirical work, Tingvall and Poldahl find the AB segment encompasses variations between monopoly and duopoly and possibly a three-firm oligopoly, with the peak of the inverted U occurring at a point between two and three firms in the industry.

We now turn to the innovation experience of the U.K. defense industry, taking R&D as a proxy, to discover how that has varied with the structural change discussed above.

U.K. defense industry innovation

Economic explanations for variations in business R&D spending are most potent when it can be assumed that firms fund the investment themselves. Figure 2 compares changes in the value of defense R&D conducted by industry with that conducted by government and universities.[13] As may be seen, total real defense R&D spending has fallen sharply since the Cold War: by 2005 it was 40 percent lower than in 1989. Defense R&D performed by industry now represents by far the largest share of defense R&D conducted in the United Kingdom. Even before the formation of QinetiQ, industry performed around two-thirds of U.K. defense R&D. The value of defense R&D performed by industry fell dramatically, by 39.5 percent, between 1989 and 1996, the spending low-point. Despite an increase in later years, defense R&D conducted by industry in 2005 was still almost 20 percent below its 1989 level.

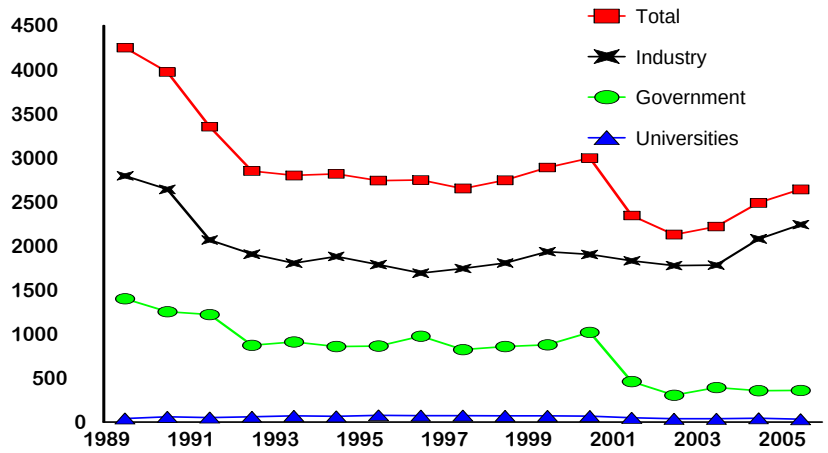

Figure 2: Defense R&D conducted in the U.K., by performer, 1989-2005 (£ mn in constant 2005 prices). *Source*: Office of National Statistics, http://www.statistics.gov.uk/StatBase/tsdataset.asp?vlnk=584&More=N&All=Y.

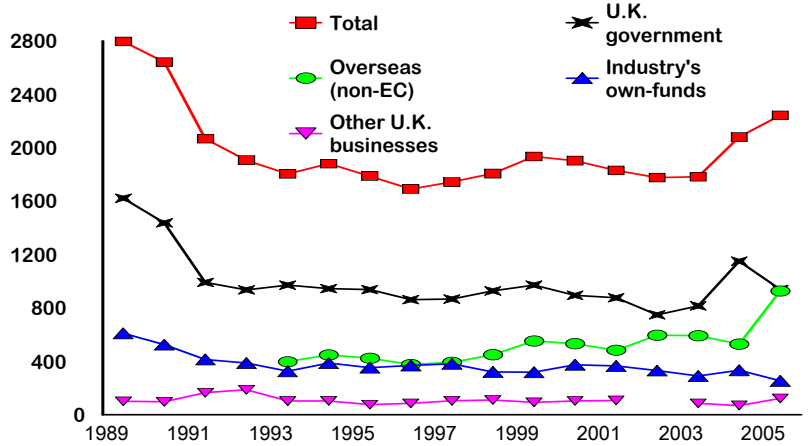

Figure 3: Sources of funds for defense R&D in UK businesses, 1989-2005 (£ mn in constant 2005 prices). *Source*: Office of National Statistics, http://www.statistics.gov.uk/StatBase/tsdataset.asp?vlnk=571&More=N&All=Y
.

Figure 3 shows changes in the sources of funds for defense R&D in U.K.

While the U.K. defense industry has been funding substantially less R&D from its own sources there has been a striking increase in the importance of funding from overseas.

businesses between 1989 and 2005. The figure reveals that industry's own-funded R&D saw a remarkable decline, constituting an increasingly modest share of the defense R&D conducted by U.K. businesses. In 1989, industry's own-funded defense R&D accounted for 20 percent of its R&D, but by 2005 this had fallen to a little over 11 percent. In absolute terms, in 1989 the industry funded defense R&D worth £610 million (in 2005 prices). By 2005 this had fallen to a little over £250 million. This represents a dramatic decline of almost 60 percent in real terms between 1989 and 2005.[14]

While the U.K. defense industry has been funding substantially less R&D from its own sources there has been a striking increase in the importance of funding from overseas. Overseas sources of funding for defense R&D conducted by U.K. companies has more than doubled since 1993 to reach £927 million in 2005. In percentage terms, overseas sources have increased from 22 percent to 41 percent of funding of defense R&D in U.K. businesses. (We speculate that this reflects the growing importance of international collaborative projects, particularly the Joint Strike Fighter, and increasing efforts by U.K. companies to obtain research contracts from foreign governments, particularly the United States. At the same time, U.K. government funding for defense R&D conducted by U.K. businesses has declined.)

Discussion

The number of companies awarded major MoD contracts (as defined earlier) fell over our period of analysis from 16 to 13. At the same time, industry own-funded R&D fell sharply and overseas-funded R&D rose. Are these observations connected?

Structural explanation

Taking the industry overall, the simultaneous fall in own-funded R&D and rise in consolidation could be accounted for by arguing either that the Schumpeter effect was the only causal relationship at work or that it became increasingly dominant as competition declined. If the Schumpeter effect were the only one at work, we would have to conclude that the U.K. defense industry was an exceptional case that did not fit with the findings of the empirical studies suggesting the existence of an inverted U-shape relationship, which implies the operation of opposing forces. If both effects were at work, we would have to say that the peak of the inverted U occurred with a structure of at least 15 firms, which seems implausible and is very much at odds with the findings of Tingvall and Poldahl.

We believe, however, that it may not make sense to view the "defense industry," with all its diversity and heterogeneity, as a construct comparable with, say, "tire manufacture" or "electricity generation." It may make more sense to disaggregate the industry into its component parts.[15] We noted before that in three subsectors (fast jet combat aircraft and maritime patrol fixed wing; strategic airlift; armored fighting vehicles) conditions had become less competitive in our terms or, where there was monopoly to start with, no more competitive. There are also low numbers of competitors in helicopters (2) and submarines (3), and in the remaining sectors, two with 4, one with 5 and one (non-embedded C4ISTAR) with 15. Of these, helicopters and submarines had become no more competitive in terms of firms performing MoD work. But this is a mixed bag, and it is not impossible that some subsectors (in particular non-embedded C4ISTAR) may have become more competitive rather than less.[16]

Clearly, it would be dangerous to put down observed changes in R&D to changes in industry structure alone. But a possible explanation for the observed fall in own-funded R&D comprises three parts. First, for subsectors starting out with a small number of competitors, consolidation took those elements of the overall industry down the AB segment of the inverted U-shape in Figure 1, leading to a reduction in R&D along with a reduction in competition. Second, for the remaining subsectors, competition either increased, decreased, or remained unchanged, with ambiguous outcomes for R&D. And third, the overall impact on industry R&D would be negative if (a) the effects of parts 1 and 2 were both negative or if (b) any positive effects of part 2 were outweighed by negative effects of part 1. The previous paragraph indicates that five out of nine subsectors had three firms or fewer in 2005, and we know that consolidation played a major role in producing a monopoly in at least the case of armored fighting vehicles. One of the remaining four, non-embedded C4ISTAR clearly experienced new entry during the period which, given the relatively large number of players in the subsector, may have taken it down the BC segment of the inverted U-shape in Figure 1. We could not identify any cases in which a change in subsector player numbers might be expected to yield an increase in R&D, given the operation of an inverted U-shape relationship. In such cases, initial one-firm or two-firm subsectors would have experienced a rise in the number of players, or initial many-player subsectors would have seen a modest fall in the number of firms. On the basis of these observations, we would argue that changes in the industry structure alone were more likely to suppress business innovation than encourage it.

We would not wish, however, to look to consolidation per se as the only explanation of declining own-funded R&D. As Figure 3 indicates, at the same time as industry's own funds have declined as a proportion of the total, non-EC overseas sources of funding have increased. As stated before, between 1993 and 2005, these sources of funding for defense R&D conducted by U.K. companies increased from 22 percent of the total to 41 percent. The Office of National Statistics (ONS) offers no further information on the exact source of these funds, but the rising inflow might have arisen from the increasing significance of international collaborative projects and

the fruits of research contracts won with non-EC overseas governments.

Data are not available in the public domain to test these propositions further against the evidence but, in any case, we believe that other factors may well also have been at work. We refer to these as nonstructural explanations.

Nonstructural explanations

Three types of nonstructural explanation are considered here. First, total government demand for U.K. defense industry output has fallen. As in many other parts of the western world, military equipment spending declined in real terms between 1990 and 2000 in the United Kingdom, offering fewer opportunities and incentives for industry to invest in innovation. But this also looks like a partial explanation, at best, since MoD equipment spending fell by 21 percent while industry own-funded R&D fell by an even larger proportion, 28 percent. Second, procurement reforms may also have had an effect. The so-called Levene defense procurement reforms of the 1980s led to a shift from cost-plus to fixed-price contracts with a view to reallocating program risk from the MoD to prime contractors. This may have had the effect of reducing business incentives to undertake high-risk technological innovation since contractors may seek to maximize profit (or minimize loss) under fixed-price contracts by cutting back on quality or adhering rigidly to specifications rather than experimenting with alternatives. Third, another sort of explanation for the relationship that we have observed may lie in the direction of causation flowing from innovation to industry structure rather than structure to innovation. We know that the cost, complexity, and risk of defense innovation have much increased in recent decades and this, in itself, may have contributed to consolidation. If large multidivisional and multiproduct defense companies are better positioned to spread the costs and risks of innovation, we may have been observing innovation driving consolidation in the U.K. defense industry. Many medium-sized companies concluded in the late 1980s and early 1990s that rising R&D costs and risks, along with shrinking market opportunities, meant that they should either exit through divestment or seek to grow through acquisition. Some chose to exit and were replaced by larger operators, often with international reach.

Policy implications

On the basis of the discussion in this article, we believe that U.K. defense industry policymakers have a number of policy options available to them. First, if further research were able to confirm the inverted U-shape relationship found in many other industry environments, policies aimed at increasing the number of competitors in one or two-player subsectors could stimulate an increase in industry innovation. MoD has already made the U.K. defense market contestable by threatening or actually opening up its national markets to competition from foreign firms. Equally, MoD might enhance competition through procurement reforms that make it easier for nondefense suppliers to act as prime contractors. For subsectors already well-populated with firms, however, such action could be counterproductive. At most, therefore, policymakers should consider influencing the industry structure (i.e., number of players) only in some circumstances. They would need to have better data than we have been able to access on the parameters of the inverted U-shape relationship or, better, about the relative strength of the forces shaping the relationship. And they would need to have a reasonable degree of confidence on the stability of the relationship. Given the defense industry environment globally, we also caution that increasing competition might reinforce incumbents in their belief that the U.K. is an unattractive market and cause them to redouble their efforts to shift the focus of their investments from the United Kingdom to the United States.

Second, to the extent that declining demand may have stifled innovation investments, increasing demand might restore the strength of incentives. But the potential for increased defense-related innovation should not be used as an argument for expanding defense equipment purchases. Increased defense spending or changes in its composition can only be justified logically in terms of changed strategic requirements. Third, if fixed-price contracts inhibit industry innovation, current efforts to introduce risk-sharing into defense contracts should assist, and MoD should persist, with further variations on the contractual incentives theme to uncover what achieves best results. Optimal contracts in this area appear more a matter of trial and error than prior knowledge.[17]

Fourth, policy should encompass the broader defense industry supply base. The U.K. MoD's Defence Industrial Strategy (2005) and Defence Technology Strategy (2006) both recognize that procurement policies need to ensure that suppliers receive proper returns for innovation and that value flows through to the layers of the supply chain that are innovating rather than being retained by prime contractors.

Conclusion

We have speculated in this article that structural issues, found to be relevant to explaining industry innovation generally, may also have played a part in U.K. defense industry. Given the constraints on publicly available data, however, we cannot at this stage take the argument further and readily concede that other factors such as demand and contractual incentives need to be taken into consideration. Future research might also seek to obtain better measures of competitive pressure (related to price-cost margins) and of innovation. Our analysis has been based on an input measure of innovation (defense R&D spending) and it would be desirable to develop suitable output measures. A starting point could be a time-series analysis of patenting data for U.K. defense companies. There is also a need for case studies of defense equipment innovation in the U.K. as a means of understanding the dynamics of the innovation process and the relative importance of competition, industry R&D, government conducted R&D, and linkages with overseas sources of technology.

Notes

Peter Hall is Emeritus Professor in the School of Business, University of New South Wales at the Australian Defense Force Academy, Canberra, ACT, Australia. He may be reached at p.hall@adfa.edu.au. **Andrew James** is Senior Lecturer at the Manchester Business School, University of Manchester, Oxford Road, Manchester, M13 9PL, U.K. He may be reached at andrew.james@mbs.ac.uk. The research assistance of Chris Redpath is gratefully acknowledged.

1. See Aghion, *et al*. (2005); Tingvall and Poldahl (2006).

2. U.K. Defence Statistics (1991; 2006).

3. MoD (2005).

4. See Kovacic and Smallwood (1994).

5. See Lorell (2003).

6. Birkler, *et al*. (2003).

7. We fully accept that R&D as an input measure for innovation is an incomplete and imperfect metric for innovation performance more generally. That said, aerospace and defense, together with pharmaceuticals, account for over half of the R&D performed by the U.K.'s top 750 companies (Harris, Nightingale, and Acha, 2006, p. 20), suggesting that R&D may be more important as a driver of innovation in defense than in many other sectors of industry.

8. For example, Scherer (1967) and Hart (1983).

9. Dasgupta and Stiglitz (1980).

10. Nickell (1996).

11. For example, Geroski (1990); Nickell, (1996).

12. Aghion, *et al*. (2005); Tingvall and Poldahl (2006).

13. Universities conduct only a modest amount of defense R&D in the U.K. and are not considered here. On the role of British universities in defense work, see Langley (2008).

14. To clarify, Fig. 2 shows how all of the defense R&D conducted in the U.K. is divided up among the various performers of the R&D. The "total" shown in Fig. 2 is thus the result of adding together the R&D done in any given year by all of the defense R&D performers. R&D performed by industry in 1989 (at 2005 prices) was 2,795 million pounds, within an overall total of R&D performed of 4,248 million pounds. Fig. 3 relates only to defense R&D carried out in U.K. business and shows the sources of funds for that R&D work. The "total" shown in Fig. 3 is thus the result of adding together the amounts of funds from all sources of funding for U.K. business defense R&D and is the same (actually, 2,796 million rather than 2,795) as the amount shown in Fig. 2 as the amount spent by industry on defense R&D. In brief, "total" in the two figures means different things, but everything adds up.

15. Kovacic and Smallwood (1994) do something similar in their analysis of competitiveness in U.S. defense industry.

16. The picture is complicated by the fact that subsectors can be defined in different ways. Whichever way subsector boundaries are drawn, it remains quite possible for competition to have decreased in some cases and increased in others and to yield the same overall result.

17. Rogerson (1995).

References

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *The Quarterly Journal of Economics*, Vol. 120, No. 2, pp. 701-728.

Birkler, J., A. Bower, J. Drezner, G. Lee, M. Lorell, G. Smith, F. Timson, W. Trimble, and O. Younossi. 2003. *Competition and Innovation in the U.S. Fixed-Wing Military Aircraft Industry*. Santa Monica, CA: RAND.

Dasgupta, P. and J.E. Stiglitz. 1980. "Industrial Structure and the Nature of Innovative Activity." *The Economic Journal*, Vol. 90, No. 358, pp. 266-293.

Geroski, P.A. 1990. "Innovation, Technological Opportunity and Market Structure." *Oxford Economic Papers*, Vol. 42, pp. 586-602.

Harris, M., P. Nightingale, and V. Acha. 2006. *The Innovation Gap: Why Policy Needs to Reflect the Reality of Innovation in the UK*. London, NESTA (National Endowment for Science, Technology and the Arts).

Hart, O. 1983. "The Market Mechanism as an Incentive Scheme." *Bell Journal of Economics*, Vol. 14. No. 2, pp. 366-382.

Kovacic, W.E. and D.E. Smallwood. 1994. "Competition Policy, Rivalries, and Defense Industry." *Journal of Economic Perspectives*, Vol. 8, No. 4, pp. 91-110.

Langley, Chris. 2008. "Universities, the military, and the means of destruction in the

United Kingdom." *Economics of Peace and Security Journal*, Vol. 3, No. 1, pp. 49-55.

Lorell, M. 2003. *The U.S. Combat Aircraft Industry, 1909-2000*. Santa Monica, CA: RAND.

[MoD] Ministry of Defence. 1991. *UK Defence Statistics 1991*. London: Ministry of Defence.

[MoD] Ministry of Defence. 2005. *Defence Industrial Strategy*. London: Ministry of Defence.

[MoD] Ministry of Defence. 2006. *Defence Technology Strategy*. London: Ministry of Defence.

[MoD] Ministry of Defence. 2006. *UK Defence Statistics 2006*. London: Ministry of Defence.

Nickell, S. 1996. "Competition and Corporate Performance." *Journal of Political Economy*, Vol. 104, No. 4, pp. 724-746.

Rogerson, W.P. 1995. "Incentive Models of the Defense Procurement Process." Chapter 12 in K. Hartley and T. Sandler, eds. *Handbook of Defense Economics*, Vol. 1. Amsterdam: Elsevier.

Scherer, F.M. 1967. "Market Structure and the Employment of Scientists and Engineers." *American Economic Review*, Vol. 57, pp. 524-531.

Tingvall, P.G. and A. Poldahl. 2006. "Is There Really an Inverted U-shape Relation Between Competition and Innovation?" *Economics of Innovation and New Technology*, Vol. 15, No. 2, pp. 101-118.

Symposium. The economics of conflict: theory and micro-level evidence

Philip Verwimp (Symposium guest editor)

For far too long, certain bad outcomes such as war, famine, and genocide where not studied by economists. Based on a core assumption of economics, voluntary exchange in peaceful markets, economists saw violent conflict as an aberration of the main model that did not merit further consideration, theory building, or analysis. Economists neglected the strategic use of violence and force by those who had a comparative advantage in using it.

Two groups of economist, largely independent from each other, changed this unsatisfactory situation. First are scholars in the field of defense and peace economics. Their contributions concentrated on the analysis of the arms race, the military-industrial complex, arms sales, military expenditure, nuclear arms control, and related topics. These topics came to the forefront during the Cold War, a period in which academics and policymakers were very much occupied with interstate war, alliances, and nuclear arms. Today, this problematic is still with us, witness the invasions of Afghanistan and Iraq, the stand-off between Russia and Georgia, the mounting tension between India and Pakistan, and the nuclear issue in Iran, to name but some interstate conflicts.

After the end of the Cold War, the world has observed an upsurge in another type of conflict, intrastate war. Governments and rebel movements previously held in check by the Cold War superpowers broke out of the straightjacket and sized the opportunity to voice their own long-standing grievances and/or claim their own stake in the affairs of the state. That upsurge has become the center of attention for academic economists and political scientists since the second part of the 1990s.

In the first contribution to this symposium, Mansoob Murshed reviews merits and shortcomings of the intrastate war literature. He acknowledges that the greed versus grievance debate has had an influential impact on researchers and policymakers alike. Murshed argues that the greed versus grievance dichotomy is a useful entry point into the debate about the causes of conflict. In certain instances, where there are substantial quantities of capturable natural resource wealth such as diamonds, oil, or drugs, greed may be the dominant factor prolonging conflict. Without group formation, however, for which some historical grievances are important, violent collective action cannot take place. It follows that conflict analysts should consider greed as well as grievance-based motives with an eye for their potentially salient but different role in the initiation versus the duration of conflict.

Murshed argues that violent conflict is unlikely to take hold if a country has a framework of widely-agreed rules, both formal and informal, that govern the allocation of resources, including resource rents, and the peaceful settlement of grievances. Such a viable social contract can be sufficient to restrain, if not eliminate, opportunistic behavior such as large-scale theft of resource rents, and the violent expression of grievances. The occurrence of large-scale violence thus constitutes a breakdown of the social contract.

Murshed concludes that the competing greed versus grievance hypotheses offer complementary explanations for conflict. Insofar as they do provide alternative views, a fair test for their relative explanatory powers is best conducted at the level of a quantitative country-case study. That is exactly what the three other contributions to this symposium have to offer. The debate in the economics of conflict has been based for a very large part based on cross-country regressions. A third generation of scholars now shifts the debate to the micro-level, to within-country analysis, guided by a thorough understanding of any one country's history and institutions. The focus of this new generation of studies is on the behavior of groups, households, and individuals as drivers, bystanders, or victims of conflict. The three articles each offer such contribution from Latin America, Africa, and Asia.

The article by Ana María Ibáñez argues that the main reasons for forced migration in Colombia are the direct and indiscriminate attacks of which the civilian population is the victim. The population is deliberately attacked. Aggressions against the civilian population allow insurgents to take over valuable assets in order to finance their war operations, but also for their members' personal gain. These attacks also allow the insurgents to strengthen territorial supremacy by driving possible opponents from the region, avoiding civilian resistance, weakening social networks, or separating the civil population from the rebel groups. Kidnappings, massacres, selective murders, forced recruitment, and land mines, among others, are some of the aggressions the civil population face and the strategies used by the rebel groups.

While forced migration is one of the direct consequences of armed conflict worldwide, Ibáñez argues that migration strategies are country-specific. African conflicts are often characterized by a massive flux of people who settle in a refugee camp. In Colombia, most displacement occurs on an individual basis, most municipalities are affected by a certain degree of displacement, displaced people do not migrate out of Colombia and do not settle in refugee camps. This underlines the need for within-country, micro-level analysis. Additionally, in some regions of Colombia there is a close relationship between drug-dealing and displacement when the land acquisition by the former undermines the state's power to mediate conflicts between peasants and landowners.

A second article, by Steven Spittaels and Filip Hilgert, employs a geographic method to study war motivations in the Democratic Republic of the Congo. They argue that greed as well as grievance motivations are present, but that they are not directly related to the control of the mining sector. Warring factions such as the rebellion by Tutsi general Laurent Nkunda and the Democratic Forces for the Liberation of Rwanda (FDLR) are involved in lucrative business ventures but not in

direct control of the mines themselves. The authors argue that the war strategy of these parties, in particular the control over space, seems to be driven by grievance and security reasons. Tutsi as well as Hutu-affiliated groups feel that their very existence is threatened and seek a territory were they feel safe. Congolese Tutsi claim a stake in the Congolese nation and state whereas Hutu rebels claim a stake in Rwandan politics. A solution to the Congolese situation must thus include a country-specific institutional arrangement that addresses both the politically-based grievances that are at the root of the conflict as well as the economic interests responsible for the duration of the conflict.

The third article, written by Mohammad Zulfan Tadjoeddin and Anis Chowdhury offers a socioeconomic perspective of different forms of conflict in Indonesia. Investigating mainly two kinds of civil strife, to wit, separatist and ethnic/sectarian violence, they argue that the grievance over the loss of a previously superior socioeconomic position is an important driver of these conflicts. To make their claim for separatist violence, the authors focus on the troubled relationship between the central government and the four regions rich in natural resources. They attribute a large chunk of separatist violence to a drive by the central government toward equalization of welfare between the rich and the poor regions, thereby nurturing a feeling of relative loss in the minds of the rich. Tadjoeddin and Chowdhury argue that ethnic/sectarian violence in Indonesia is not explained by widening intergroup horizontal inequality but by three other factors: first, the grievance due to the loss of a group's relative (superior) position vis-à-vis a previously downtrodden group, an explanation thus akin to separatist violence but this time at the intergroup level and not at the interregional level; second, the discrepancy between relatively high levels of education (aspirations) and relatively low levels of welfare (reality); and, third, the increase in local budgets after decentralization making local government posts a prize worth fighting for.

As in the Colombian and Congolese cases, the Indonesia article shows a great need to offer within-country, institution-based analysis of the conflict. Only then can we begin to understand the greed and grievance-based motives of the warring factions.

Note

Philip Verwimp is Research Fellow at the Fund for Scientific Research (Flanders, Belgium) based at the University of Antwerp and Visiting Researcher at Ecares, Université Libre de Bruxelles. He codirects the Households in Conflict Network (www.hicn.org) and can be reached at philip.verwimp@gmail.com.

Conflict as the absence of contract

S. Mansoob Murshed

Civil war is a multi-faceted problem. Not only does it produce human tragedies on a colossal scale, but it creates humanitarian crises that are of concern to the international community, as well as contributing to global and regional insecurity. Civil war is also a major cause of underdevelopment, and perpetuates poverty.[1] According to the rational choice paradigm, conflict is a result of choice. The rationality may be of a myopic nature, as negotiated settlements avoiding the losses that ensue from war are usually superior from a universal point of view.

In recent years, two phenomena have been utilized to explain conflict onset among academic economists: greed and grievance. The former is due to the influential work of Paul Collier, and is popular among economists. According to this view, conflict reflects elite competition over valuable natural resource rents, concealed with the fig leaf of collective grievance. Above all, there was the assertion that inequality played no part in adding to the risk of civil war. It also provides intellectual excuses for direct, colonial-style intervention to prevent failing states from collapsing. But in many ways, these views go against the grain. There is a long-standing position in political science that relative deprivation and the grievance that it produces fuels internal violence. Identity is also crucial to intrastate conflict. This is due to the collective action problem. It is difficult to mobilize large groups to undertake collective action, because of mutual mistrust, monitoring difficulties, and the free-rider problem. Ethnic identities, whether based on race, language, religion, tribal affiliation, or regional differences may serve as the most effective amalgam for the purposes of group formation, compared to other forms of more transient differences that are traditionally stressed such as socioeconomic class. More recently, Frances Stewart has introduced the notion of horizontal inequality, the inequality between groups, rather than the inequality that may exist among an ethnically homogenous population (vertical inequality).[2]

The purpose of this article is to emphasize that violent conflict is essentially a manifestation of the breakdown of the social contract. The three sections of this article examine the greed hypothesis, issues related to grievance, and a synthesis of the greed and grievance views related to malfunctioning institutions that may be described as the failure of the social contract.

The greed explanation for conflict

The greed motivation behind civil war has been popularized by empirical work on the causes of civil war where a cross-section of conflicts in different nations is analyzed econometrically, and greed is proxied by the availability or abundance of capturable natural resource rents. For Collier and Hoeffler, civil wars stem from the greedy behavior of a rebel group in organizing an insurgency against the government. Greed is about opportunities faced by the rebel group. The opportunities can be divided into three components: financing, recruitment, and geography.[3]

Violent conflict is essentially a manifestation of the breakdown of the social contract. A synthesis of the greed and grievance views related to malfunctioning institutions may be described as the failure of the social contract.

The most common sources of rebel finance are the appropriation of natural resources, donations from sympathetic diasporas residing abroad, and contributions from foreign states or multinational companies interested in the region. Natural resource wealth is chief among the three in terms of its relative importance. Recruitment is about the opportunity to induct fighting manpower; something made easier when there is a high proportion of young unemployed males in the population, in a setting of endemic poverty and poor education. Geographical situations favorable to rebel groups are mountainous terrain and other safe havens for insurgents. In short, greed simply means the economic opportunity to fight, and should be distinguished from sociopolitical grievances. Collier and Hoeffler's empirical findings conclude that the set of variables representing rebel opportunity or greed akin to loot-seeking are the main reasons for civil war. By implication, the alternative hypothesis of grievance (justice-seeking) focusing on ethnic or religious divisions, political repression, and horizontal inequality is dismissed, although its invalidity is not formally tested for. Natural resource rents constitute booty and this fact has been used to emphasize the greed or criminal motivation for civil war. Central to Collier and Hoeffler's empirical testing for the greed hypothesis is the role of primary commodities in the economic structure. They measure the dependence on natural resources by the share of primary commodity exports in GDP, and the validity of this metric as well as the statistical robustness of the relationship between resource rents and the risk of conflict, has been called into question. The empirical controversy over the link between natural resource wealth and greed hypothesis are about the saliency of mechanisms in-between natural resource rents and conflict, as well as measurement issues and estimation techniques.

Humphreys for example argues that other mechanisms may be present.[4] First is the greedy outsider mechanism: the existence of natural resources may be an incentive for third parties — states and corporations — to engage in or foster civil conflict. Second is the grievance mechanism: natural resource dependence could in fact be associated with grievances rather than greed. Third is the weak-state mechanism: natural resource dependent economies may have weaker states, which stems from the nature of state revenue that is dependent on resource rents. On the one hand, untaxed citizens have less ability or incentive to monitor state activity. On the other hand, governments relying more on natural resource revenues than on taxation have weaker

incentives to create accountable bureaucratic structures.

Any measure of natural resource dependence may also be endogenous to conflict, which has two implications: first, reverse causality, in which civil wars might cause resource dependence by reducing the size of a country's non-resource sector, and second, spurious correlation, where both civil war and resource dependence might be independently caused by an unmeasured third variable such as poor institutional quality.

Observe that in Collier and Hoeffler the term "primary commodity" includes both agricultural commodities and minerals and fuels, but crucially excludes illegal substances (coca and heroin) as well as illegal alluvial diamonds. Certain varieties of resources are more easily captured: they may be lootable such as alluvial diamonds (in Sierra Leone, Angola), available along river beds using artisanal techniques, illicit drugs such as coca in Colombia, or obstructable like an oil pipeline.[5] Illicit gemstones and drugs are arguably more crucial to financing rogue conflict entrepreneurs in a greed-based conflict; their omission is a serious flaw. Collier and Hoeffler do not differentiate among different types of natural resources such as between lootable and nonlootable natural resources or between point-source and diffuse natural resources.[6] Lootable point source natural resources in particular are prone to be illegally exploited and traded. Collier and Hoeffler are only concerned with past natural production, neglecting future prospects for extraction.[7] They also only focus on exports, even though production might be a better measure of the availability of these resources.

On estimation techniques, Fearon provides the strongest challenge to Collier and Hoeffler's empirical finding on the link between primary commodity exports and civil war. He re-estimates Collier and Hoeffler's model using country-year observations, as opposed to country-five year observations employed by Collier and Hoeffler, and finds that the significance of statistical associations between primary commodity exports and civil war onset vanish in the country-year regression, meaning that the previous claim of such a relationship is simply not robust. In other words, this cross-country result will not withstand variation in sample and data coverage. A similar view is shared by Ross who reviews 14 cross-country empirical studies on natural resources and civil war. He concludes that the claim that primary commodities are associated with the onset of civil war does not appear to be robust, oil dependence appears to be linked to the initiation of conflict, but not its duration, and illicit gemstones and drugs seem to lengthen pre-existing wars.[8]

There are additional reasons to be skeptical of the greed theories. First, they generally neglect the destructiveness of war, and its capacity to ravage productive capacity, additional to direct military expenditure. Second, there is no growth in these models, something which would raise the opportunity costs of war. Third, the possibilities of peaceful exchange need to be limited in order to rationalize conflict. In traditional economics the gains from trade arise mainly from differences in tastes, technology, and endowments, and these gains from trade need to be minimized in order to make conflict an optimal choice. Violent means are attractive when the intention is to extract resources (as in the case of colonial plantations and mines) or accumulate surpluses at the expense of others (mercantilism). Fourth, these models imply full information. In the presence of asymmetric information, misperceptions about contest success, the opposition's intentions, and so on, wars that do not maximize expected utility under full information may break out, akin to problems associated with moral hazard and adverse selection. Fifth, such theorizing is broadly blind to institutions and the presence of transactions costs that breed mutual mistrust. Wars can also reflect the absence of institutions facilitating negotiation and peaceful exchange.

Despite these limitations, there is much in these models that can explain the greedy behavior as analyzed by the empirical exponents of the greed hypothesis. First, the presence of readily capturable natural resource based rents may make conflict more attractive when compared to peaceful production, as can a shortage of intermediate inputs due to population pressure. These resources are best regarded as a non-produced "prize" such as oil or diamonds (which apart from extraction costs are like manna from heaven), whose ownership is violently contested. Second, contributions from a sympathetic diaspora (or aid from a superpower in the Cold War era) can raise the probability of victory of a potential rebel group against the state. Third, the inability of the state to act as a preemptive leader in a potentially divided nation may raise the chances of war between groups in a manner similar to the weak-state capacity mechanism discussed above.

Grievances and horizontal inequality

Central to grievances is identity and group formation. An individual's utility may be related to his identity, specifically the relative position of the group he identifies with in the social pecking order.[9] An individual may derive utility from certain normative forms of behavior appropriate to his identity but considered deviant by other groups and may even face sanctions from like-minded group members if he deviates from them. This type of behavioral paradigm may be related to solving the collective action problems without which organized large-scale violence is impossible, even if we believe conflict is primarily motivated by greed. As noted, some appropriate definition of ethnicity may be a superior basis for group formation compared to social class in an ethnically homogenous society. We may subdivide theories of grievance into relative deprivation, polarization, and horizontal inequality.

Relative deprivation

The notion of relative deprivation dates back to the work of Ted Gurr who defines it as the discrepancy between what people think they deserve and what they believe they can get, in short the disparity between aspirations and achievements.[10] Thus, educational achievements may raise the aspirations of young people, but they will

become frustrated if unemployed, occasionally venting their feelings in mass political violence.

Polarization

A related notion is that of polarization. It occurs when two groups exhibit great intergroup heterogeneity combined with intragroup homogeneity. Economic polarization (along with high vertical income inequality) can occur in societies that are culturally homogenous. Ethnic polarization could, in principle, exist along with a degree of economic equality. What is useful is a hybrid concept that combines identity and economic polarities, as in Østby. In their original and seminal concept of polarization, Esteban and Ray focus on the identification and alienation framework. Their idea is that polarization is related to the alienation that groups of people feel from one another and that such alienation is fueled by the feeling of within-group identity. Furthermore, they argue that traditional measures of inequality are only concerned with interpersonal alienation, but fail to capture the dimension of group identity. It is important to note that ethnic polarization requires two or a few ethnicities. When a society has a very large number of identities, then the term ethnic fractionalization is more appropriate. Therefore, polarization is what may matter for conflict, rather than fractionalization, and/or overall vertical (inter-individual) inequality. Montalvo and Reynal-Querol find that ethnic polarization is a significant explanatory variable for civil war onset while ethnic fractionalization is not.[11]

Horizontal inequality

The notion of horizontal inequalities among groups, classified by ethnicity, religion, linguistic differences, tribal affiliations, and so on, is thought to be an important cause of contemporary civil war and sectarian strife, but not routine violence. The idea of horizontal inequality may overlap with the notion of relative deprivation and polarization as will be indicated by alternative measures discussed below. The horizontal inequality expression originates from the work of Frances Stewart and is distinguished from vertical inequality, which is the inequality within an otherwise homogenous population.[12] Four sources of horizontal inequality may be highlighted:

- Discrimination in public spending and taxation: discrimination in the allocation of public spending and unfair tax burdens can cause unrest.
- High asset inequality: agrarian societies with high inequality, for example El Salvador, Guatemala, Nepal, the Philippines, and Zimbabwe, have high asset inequality and are very prone to conflict.
- Economic mismanagement and recession: conflict-ridden countries have usually suffered prolonged economic mismanagement and growth collapse. Economic mismanagement is often associated with an uneven and unfair distribution of the burdens of subsequent adjustment; public spending benefitting the elite and the military is protected, often favoring particular ethnic groups, with the burden of adjustment placed on expenditure of value to the poor and disadvantaged groups.
- Grievances related to resource rents: natural resource rents can by themselves become a source of grievance if local populations feel that they are not getting a fair share of these, as in the Niger Delta region of Nigeria.

Relative deprivation and horizontal inequalities have been found to significantly affect conflict in Nepal and Indonesia, to cite two examples of its application to individual nations. Østby manages to construct polarization indices and horizontal inequalities across 36 developing countries for 1986-2004 based on ownership of consumer durables and educational attainment from household surveys. In her panel and cross-sectional analysis, she finds that social and economic polarization and social horizontal inequality based on education significantly contribute to conflict, whereas vertical inequality and purely ethnic or socioeconomic polarization do not. A priori one would expect more enduring horizontal inequalities based on health, education, political exclusion, and asset holdings to be more significant as compared to transient income differences.[13]

Three further points are worth emphasizing at this juncture. First, horizontal inequality has to be measured at the level of the nation-state. In a sense it refers to cross-sectional variation within a specific country. The data in different countries on horizontal inequality is still embryonic. Indices for horizontal inequality can be used for cross-country comparisons, whereas for a single conflict onset, gap measures may be sufficient. Second, most nation-states do not keep detailed or systematic data on group inequalities (say between Catholics and Protestants, Hutus and Tutsis, Muslims and Christians) because of obvious political sensitivities. However, ethnic questions in future household surveys across the developing world will go a long way in helping us to enumerate data on inter-group differences in socioeconomic achievement. Third, horizontal inequality as a cause of conflict can work in two directions: the rich may initiate conflict to extricate themselves from the relatively poor (the rage of the rich), or the poor may rise up in revolt against the rich (the rage of the poor). The former may be more likely in cases where a region suddenly discovers it can exist viably on its own resources, thus wishing to secede and not hand over revenues to the rest of the country. The latter is likely to be manifested in rebellions and revolutions.

Synthesis and social contract

The greed versus grievance dichotomy is a useful entry point into the debate about the causes of conflict. In certain instances, where there are substantial quantities of capturable natural resource wealth present such as alluvial diamonds, oil, or drugs, greed may be the dominant factor prolonging conflict, but without group formation (for which some historical grievances are important) violent collective action cannot

take place. Although greed and grievance are regarded as competing views, yet they may be complementary. Greed may produce grievances after war breaks out as people are affected by conflict, and the converse is also true as movements that begin by espousing a grievance-based cause may become used to exacting rents from the populations that they control or the diasporas and other support groups that back them. The greed or grievance explanations (or some hybrid form of both) may be necessary for the outbreak of civil war, but arguably they are not sufficient. This is because the causes enumerated in the preceding two sections contribute to the risk of civil war, yet some societies despite having conditions predisposing them to civil war, such as horizontal inequality, polarization, and natural resource rents, do not descend into conflict. It can be argued that for the forces behind either greed or grievance to take the form of large-scale violence there must be other factors at work, specifically a weakening of what may be described as the social contract following classical thinkers such as Hobbes, Locke, and Rousseau among others.[14] This is similar to the weak-state capacity, and by implication poor institutional quality, arguments made before. Therefore, even if rents from capturable resources do constitute a sizeable prize, violent conflict is unlikely to take hold if a country has a framework of widely-agreed rules, both formal and informal, that govern the allocation of resources, including resource rents, and the peaceful settlement of grievances. A viable social contract can be sufficient to restrain, if not eliminate, opportunistic behavior such as large-scale theft of resource rents, and the violent expression of grievance.

Civil war is a reflection of the breakdown or degeneration of a contract governing interactions between various parties. Hirshleifer draws our attention to the fact that within a society, social contracts can be vertical if they are authoritarian in the sense of Thomas Hobbes, or they may be horizontal if fashioned with popular consent, as advocated by John Locke. The former may be described as dictatorial, and the latter as democratic.[15]

The breakdown of social contract

What factors lead to the breakdown of the social contract within a nation-state? What circumstances create incentives for groups within societies to choose war rather than resolve disputes peacefully? Clearly these seem to occur in failing states. Yet, the term "failed state" may be too vague and unhelpful in this regard. Among the various factors, three reasons may be highlighted. The first refers to the fiscal and revenue sharing agreements the state (or those in power) have with various stakeholders, and the breakdown of these arrangements can produce greed and/or grievance. Second, there is the political system. In the face of an unstable polity where the separation of powers and the sources of (legitimate or illegitimate) power are inherently unstable, it is important to focus on individual incentives faced by rulers that may or may not cause them to promote development and modernization. Third, the famous Lipset modernization hypothesis states that demands for democracy surely follow economic development and the attainment of a high standard of living; once a particular (high) level of average income is achieved violence becomes a very costly means of settling disputes. The road to peace and democracy is therefore along sustained economic growth, and, thus, the real culprit as far as the breakdown of the social contract is considered could be growth failure in low-income developing countries because it creates conditions where violence is more attractive. Low growth also implies a more undiversified economic structure, susceptible to terms of trade shocks and dependence on external aid.[16]

The greed versus grievance dichotomy is a useful entry point into the debate about the causes of conflict. Regarded as competing views, they may be complementary. Perhaps more important, violent conflict is a reflection of the breakdown or degeneration of a contract governing interactions between parties.

Within nation-states, the fiscal system will secure a workable social contract if the allocation of public expenditure and the apportionment of taxes are judged to be fair, or at least not so unfair that some groups judge taking resources by force the better option. There are many examples of conflicts emerging out of fiscal disputes. Disputes over the apportionment of revenues from natural resources are especially common and, as in Indonesia and Nigeria, these take on ethnic and regional dimensions. Contemporary civil wars are more often related to the breakdown of explicit or implicit arrangements to share resources, rather than their complete absence (although in certain instances such as with the discovery of new minerals or fuels a revenue sharing agreement may not have been devised). One reason that a contract to share revenues encounters difficulties is the imperfect credibility with which the side that controls the "pot" honors its commitment, due mainly to malfunctioning institutions and political malfeasance.

Conflict-affected nations have histories of weak social contracts (or a once strong social contract that has degraded). This weakness is in many instances a legacy of colonialism that institutionalized mechanisms favoring settlers over indigenous peoples (Guatemala, Zimbabwe, South Africa); divide-and-rule favoring one ethnic group over another (as in Rwanda); market controls to create rents for settlers to the cost of locals (Zimbabwe); and the expropriation of land and resource rents (Angola, and the Belgian Congo).

Anocracies

Hegre *et al*. point out that the risk of conflict is lower in both well-established democracies and autocracies. They suggest that conflict risk is at its highest during transitions to and away from democracy when state capacity is weak, and also in fledgling and imperfect democracies (anocracies). Most developing countries are

imperfect democracies, or at any early stage of the democratic transition. A final complexity in fatally weakening social contracts was the interaction of these "domestic" factors with external events, notably the Cold War that provided finance and ideological succor to ruling elites and rebels. The net result of these processes is the accumulation of grievances within the context of a disintegrating social contract that would otherwise have provided the rules of the game to govern distribution and to achieve peaceful conflict resolution. These circumstances can also promote greed-based motivations aimed at controlling natural resources.[17]

Pure versions of the greed hypothesis are, on their own, unsatisfactory explanations for the causes of conflict. Addison, Le Billon, and Murshed construct a game-theoretic model of contemporary conflict involving the competition for resources combined with historical grievances.[18] In addition to resource rents, grievances also play their part in fueling conflict by explaining intergroup noncooperation and serving to lower the cost of participation in conflict. Conflict can increase because of heightened intrinsic grievances, or because there are more lootable resources. In reality, the competing greed versus grievance hypotheses may, after all, be complementary explanations for conflict. Insofar as they do provide alternative views, a fair test for their relative explanatory powers is best conducted at the level of a quantitative country-case study, because cross-country comparisons of horizontal inequality are still at very early stages of development due to the lack of data.

Indonesia's resource rich regions that have had separatist conflicts with the federal government offer us a striking contrast in trying to gauge the relative explanatory power of the greed versus grievance explanations for conflict. When viewed via the lens of a detailed quantitative case study, the grievance and horizontal inequality explanations dominate any greed motivation.[19] Yet, when looked at as one observation among many through the prism of a cross-country study, Indonesia's resource-rich regions are examples of a modified form of the greed explanation (resources helping to prolong the duration of conflict and encouraging secession.[20] It would appear, therefore, that the greed explanation for conflict duration and secessionist wars works in cross-country studies, but has to make way for grievance-based arguments in quantitative country-case studies. Grievances and horizontal inequalities may, after all, be better at explaining why conflicts begin, but not necessarily why they persist. Neither the presence of greed or grievance is sufficient for the outbreak of violent conflict, something which requires institutional breakdown and the failure of the social contract.

As yet, no empirical models at the level of cross-national analysis exist to properly test for the relative power of greed vis-à-vis horizontal inequality type grievances in explaining conflict onset. This is not just a result of constraints posed by insufficient data. Greed and grievance can and do coexist. Because one breeds the other, a model of their simultaneous determination is required, along with the contribution of poverty (which is chiefly about the lack of growth) and institutional quality. Furthermore, the existing econometric literature regarding the causes of conflict allows us to infer little about the true nature of the causal links between the phenomena examined. Tests for causality require sufficiently long time-series data; unless techniques of dynamic panel data analysis are employed, inferences about causality will remain limited in nature.

Despite these shortcomings, a review of the existing empirical literature on the causes of conflict informs us that the most robustly significant predictor of conflict risk and its duration is some indicator of economic prosperity (or lack of poverty) such as income per capita within a cross-section where average income does vary. This is because at a higher income people have more to lose from the destructiveness of conflict;[21] and higher per capita income implies a better functioning social contract, institutions, and state capacity. Yet economic development, even if it eventually diminishes motives for conflict, may at first increase violence in poor institutional settings, if growth is not pro-poor and disadvantages some groups, as Tadjoeddin and Murshed have demonstrated for Indonesia.[22] Above all, there is less poverty; masses of impoverished individuals provide the best recruitment grounds for rebel fighters. Ultimately, the political economy of growth failure and institutional degradation must inform us about the causes of conflict, along with theories of deprivation and alienation.

Notes

Syed Mansoob Murshed is at the Birmingham Business School, University of Birmingham. He also holds positions at the Institute of Social Studies,The Hague, and at the Center for the Study of Civil War (CSCW), Peace Research Institute, Oslo. He may be reached at murshed@iss.nl. Financial support from the European Union-funded MICROCON project and the suggestions of a referee are acknowledged.

1. See Collier, *et al*. (2003); Collier (2007).

2. Greed: e.g., Collier and Hoeffler (2004); relative deprivation: Gurr (1970); collective action: Olson (1965); horizontal inequality: Stewart (2000).

3. Collier and Hoeffler (2004).

4. Humphreys (2005).

5. See, e.g., Ross (2003) on these issues.

6. Lujala, Gleditsch, and Gilmore (2005); Murshed (2004).

7. Humphreys (2005).

8. Fearon (2005); Ross (2004).

9. Akerlof and Kranton (2000); Sen (2006).

10. Gurr (1970).

11. Esteban and Ray (1994); Østby (2008); Montalvo and Reynal-Querol (2005).

12. Stewart (2000).

13. Nepal/Indonesia: Murshed and Tadjoeddin (2008); indices: Østby (2008).

14. Addison and Murshed (2006); Murshed (2002).

15. Hirshleifer (1995).

16. Lipset (1960).

17. Hegre, *et al*. (2001).

18. Addison, Le Billon, and Murshed (2002).

19. Tadjoeddin (2003).

20. Tadjoeddin, Suharyo, and Mishra (2003).

21. Lipset (1960).

22. Tadjoeddin and Murshed (2007).

References

Addison, T., P. Le Billon, and S.M. Murshed. 2002. "Conflict in Africa: The Cost of Peaceful Behaviour." *Journal of African Economies*, Vol. 11, No. 3, pp. 365-386.

Addison, T. and S.M. Murshed. 2006. "The Social Contract and Violent Conflict," pp. 137-163 in H. Yanacopoulos and J. Hanlon, eds. *Civil War, Civil Peace*. Oxford: Currey.

Akerlof, G. and R.E. Kranton. 2000. "Economics and Identity," *Quarterly Journal of Economics*, Vol. 115, No. 3, pp. 715-753.

Collier, P. 2007. *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It*. Oxford: Oxford University Press.

Collier, P. and A. Hoeffler. 2004. "Greed and Grievance in Civil Wars." *Oxford Economic Papers*, Vol. 56, No. 4, pp. 563-595.

Collier, P., L. Elliot, H. Hegre, A. Hoeffler, M. Reynal-Querol, and N. Sambanis. 2003. *Breaking the Conflict Trap: Civil War and Development Policy*. Washington, DC: World Bank.

Esteban, J.-M. and D. Ray. 1994. "On the Measurement of Polarization." *Econometrica*, Vol. 62, No. 4, pp. 819-851.

Fearon, J. 2005. "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution*, Vol. 49, No. 4, pp. 483-507.

Gurr, T.R. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Hegre, H., T. Ellingsen, S. Gates, and N.P. Gleditsch. 2001. "Towards a Democratic Civil Peace? Democracy, Civil Change, and Civil War 1816-1992." *American Political Science Review*, Vol. 95, No. 1, pp. 17-33.

Hirshleifer, J. 1995. "Anarchy and its Breakdown." *Journal of Political Economy*, Vol. 103, No. 1, pp. 26-52.

Humphreys, M. 2005. "Natural Resources, Conflict, and Conflict Resolution. Uncovering the Mechanisms." *Journal of Conflict Resolution*, Vol. 49, No. 4, pp. 508-537.

Lipset, S. 1960. *Political Man: The Social Bases of Politics*. New York: Doubleday.

Lujala, P., N.P. Gleditsch, and E. Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Conflict Resolution*, Vol. 49, No. 4, pp. 538-562.

Montalvo, J.G. and M. Reynal-Querol. 2005. "Ethnic Diversity and Economic Development." *Journal of Development Economics*, Vol. 76, No. 2, pp. 293-323.

Murshed, S.M. 2002. "Civil War, Conflict and Underdevelopment." *Journal of Peace Research*, Vol. 39, No. 4, pp. 387-393.

Murshed, S.M. 2004. "When Does Natural Resource Abundance Lead to a Resource Curse." *IIED-EEP Working Paper* 04-01.

Murshed, S.M. and M.Z. Tadjoeddin. 2008. "Revisiting the Greed and Grievance Explanations for Violent Conflict." *Journal of International Development*, forthcoming.

Olson, M. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Østby, G. 2008. "Polarization, Horizontal Inequalities and Civil Conflict." *Journal of Peace Research*, Vol. 45, No. 2, pp. 143-162.

Ross, M.L. 2003. "Oil, Drugs and Diamonds: The Varying Role of Natural Resources in Civil Wars," pp. 47-70 in K. Ballentine and J. Sherman, eds. *The Political Economy of Armed Conflict: Beyond Greed and Grievance*. Boulder, CO: Lynne Rienner.

Ross, M.L. 2004. "What Do We Know About Natural Resources and Civil Wars?" *Journal of Peace Research*, Vol. 41, No. 3, pp. 337-356.

Sen, A.K. 2006. *Identity and Violence: The Illusion of Destiny*. New York: Norton.

Stewart, F. 2000. "Crisis Prevention: Tackling Horizontal Inequalities." *Oxford Development Studies*, Vol. 28, No. 3, pp. 245-262.

Tadjoeddin, M.Z. 2003. "Communal Conflicts and Horizontal Inequalities in Indonesia: Dynamics and Consequences." Paper presented at the IPSK-CNRS Group Meeting on Conflict, LIPI Jakarta, 15 January 2003.

Tadjoeddin, M.Z. and S.M. Murshed. 2007. "Socioeconomic Determinants of Everyday Violence in Indonesia: An Empirical Investigation of Javanese Districts, 1994-2003." *Journal of Peace Research*, Vol. 44, No. 6, pp. 689-707.

Tadjoeddin, M.Z., W.I. Suharyo, and S. Mishra. 2003. "Aspiration to Inequality: Regional Disparity and Centre-Regional Conflicts in Indonesia." Paper presented at the UNU/WIDER Conference on Spatial Inequality in Asia, Tokyo, Japan, 28-29 March 2003.

Socioeconomic perspectives on violent conflict in Indonesia

Mohammad Zulfan Tadjoeddin and *Anis Chowdhury*

Indonesia's post-independence history has been punctuated by internal conflicts of various kinds with varying intensity. Soon after its war of independence, a series of regional rebellions took place in the 1950s. They included armed struggles in several parts of Indonesia, aiming at the establishment of an Islamic state. The 1960s were marked by communist rebellion and anticommunist violence. President Suharto, who came to power in 1967 through a military coup that toppled President Sukarno, was able to impose a measure of peace and stability for the next three decades, although violent conflicts continued for example in Aceh, Papua, and East Timor. But the image of relative peace and stability disappeared with the fall of Suharto in 1998. More than 1,000 people were killed in a two-day anti-Chinese riot in Jakarta only a week before Suharto was removed from office. After his fall, armed separatist struggles in Aceh, Papua, and East Timor intensified. Close to 10,000 people died in interethnic violence that flared up mainly in the eastern part of the country. Routine, everyday violence also increased markedly.[1] Latent conflict became open and manifested itself in civil riots among ethnic groups as well as in regional armed revolts, partly due to the legacy of the authoritarian regime.[2]

The rise of interethnic and center-regional conflicts as well as of routine, everyday violence coincided with Indonesia's transition to democracy. In the literature, a reasonably strong argument has been made that transition to democracy is accompanied by sociopolitical turbulence (internal conflicts), with widespread violence. In addition, the modernization hypothesis causally links income levels to democratic transition, arguing in particular that prosperity breeds democracy and not the other way around. At a low level of income, it is thought, democracy is more likely to generate regression to repression rather than to accountability (which is more likely to occur in a more affluent society and is expected to bring peace).[3]

The hypothesized link between democratic transition in low or middle-income countries and violence takes on an added dimension when the political transition takes place amidst market liberalization. They can be a deadly mix breeding ethnic hatred and violence.[4] This occurs when the poor majority, suddenly feeling empowered due to democratization, attacks the wealthy minority that benefits from market liberalism. Thus, democratic transition may be problematic not just on its own account. Casual observations suggest examples ranging from the former Soviet Union to Rwanda, the Philippines, Venezuela, and Indonesia.

Following a synopsis of the two main current variants of conflict theories (greed and grievance), a basic categorization of violent conflict in Indonesia is presented. The thrust of the article then examines, for the case of Indonesia, the basic conflict categories in light of the greed and grievance debate. The final section concludes.

Indonesia experienced an outburst of violent conflict in the wake of the economic crisis of 1997-98. This was followed by the collapse of the autocratic regime of Suharto. Between 1996 and 2002, violence claimed an estimated 19,000 lives. About 9,000 of these died in secessionist violence in East Timor and Aceh, and 10,000 in ethnic strife and "routine" violence.

Theoretical perspectives: greed and grievance

Much of the literature of the past 10 years has debated two possible causal attributes of violent within-state conflict: greed and/or grievance. According to the greed perspective, conflict reflects elite competition over valuable point-source natural resource rents, concealed with the fig leaf of collective grievance. Rebellions and civil wars are more likely to start up and endure when they are carried by natural resource-based rents or financed by sympathetic diasporas. In contrast, the grievance explanation stresses the perception of unjust treatment of people and groups that share similar identities. It can be referred to as justice-seeking motivation. Central to grievance is identity and group formation. Theories of grievance are divided into emphasizing relative deprivation, polarization, or horizontal inequality. Of course, the greed and grievance themes may be complementary rather than competitive explanations of violent intra-state conflict.[5]

Some categories of violent conflict in Indonesia

Indonesia experienced an outburst of violent conflict in the wake of the economic crisis of 1997-98. This was followed by the collapse of the autocratic regime of Suharto. Between 1996 and 2002, violence claimed an estimated 19,000 lives. About 9,000 of these died in secessionist violence in East Timor and Aceh, and 10,000 in ethnic strife and "routine" violence. Violent internal conflict is not new to Indonesia. Around 40,000 people were killed in the country between 1950-61 and around 500,000 to 600,000 alleged communist were murdered in the mid 1960s.[6]

The economic and social costs of these violent conflicts also were substantial. For example, while between 1999 and 2002 Indonesia experienced average annual GDP growth of four percent, the economy of Aceh province, where a secessionist rebellion was active, contracted by 2.3 percent annually and its poverty rate doubled from 14.7 to 29.8 percent. The within-Indonesian Human Development Index (HDI) ranking for ethnic strife-ridden Ambon declined from third in 1999 to twenty-ninth in 2002.[7] From 1999 to 2000, the ethnic conflict-torn districts of Central Maluku and Southeast Maluku saw their regional GDPs drop by 22 and 40 percent, respectively. The only plausible explanation for these declines, while the country as a whole was recovering,

is religion-based violent conflict between Christians and Muslims.

In addition to high-profile violence such as separatist, ethnic, and religious affiliation-based conflict, Indonesia also experienced "low intensity" violence, centered around sporadic group brawls and vigilantism. These kinds of violence tend to generate low numbers of casualties per event and happen in more "routine, everyday" fashion (hereafter referred to as routine violence). They usually cause little economic damage.

Another kind of conflict occurs when, without a distinct separatist agenda, a community is unhappy with the State, for whatever reason.[8] This may become manifest for example in initially peaceful protests and strikes. But a number of factors, such as unruly protest behavior, overzealous law-enforcement authorities, or counter-protests by government supporters, can trigger violence. Conflict also involves community against business enterprises over the use of local resources and opportunities. This happened mostly in extractive industries; one much publicized incident involved a multinational mining company (Freeport) in the province of Papua.[9]

This article examines three categories of conflict: first, separatist violence; second, ethnic/sectarian violence; and third, routine violence. The other two (community versus state, and community versus companies) are left out either because of insufficient data or because they do not show a clear enough pattern to extract reliable generalizations.

For clarity, we define our three categories of conflict as follows. First, secessionist conflict refers to conflict between the central government and regions wanting to secede from the federation or at least having some degree of secessionist aspirations. This is also labeled as center-regional conflict and is "vertical" in nature. Second, ethnic violence is the conflict between different ethnic groups without any formal state involvement and is not directed against the state; it is "horizontal" in nature. Third, routine violence is the residual of the first two and centers on group brawls and vigilantism. These three are examined in the next section.

The case of Indonesia

Center-regional conflict

The focus here is on the troubled relation between the central government and four regions rich in natural resources: Aceh, Papua, Riau, and East Kalimantan.[10] To varying degrees, each has posed a secessionist challenge to the central government. Levels of conflict in these regions differ considerably from each other. Aceh has had the highest level of conflict, in part because both political and military wings of the rebel organization have been active for about three decades (since 1976). According to the internationally agreed definition, it is a clear case of a civil war.[11] It ended in 2005 when the Helsinki accord between the Government of Indonesia and the Free Aceh Movement (GAM) was agreed. Since then, Aceh has been recovering and rebuilding its socioeconomic and political life under the so-called self-government provision.

Table 1: Characteristics of four resource-rich Indonesian regions

Province	*Main resources*	*Conflict level*	*Conflict manifestation*
Aceh	natural gas, timber	high	well-articulated, secessionist movement; significant, violent insurgency by an organized rebel group (GAM)
Papua	oil, copper, natural gas, gold, timber	medium	fragmented, poorly-articulated secessionist, political movement; minor, violent insurgency by a less-well organized group (OPM)
Riau	oil, timber natural gas, minerals	low	minor political secessionist sentiment
East Kalimantan	oil, timber natural gas, minerals	low	minor political secessionist sentiment

The Papuan conflict could be considered an intermediate case. Although both political and military wings of rebel groups are active there, no civil war, according to the commonly agreed definition, has occurred. Compared with Aceh, Papua's secessionist movement has been significantly weaker. It is fragmented and less organized, partly because it has to deal with animosities among hundreds of tribal groups, and because it neither has charismatic leaders to unite the community (like Hasan Tiro or Daud Beureuh in Aceh) nor a strong diaspora community to provide support. Its military wing Organisasi Papua Merdeka (OPM, or Free Papua Organization) has only been able to launch sporadic violence directed against the Indonesian army or police, foreign companies, and migrant groups. The 1998 democratic transition in Indonesia provided momentum for renewed secessionist demands from the province and culminated in the second Papuan People Congress in 2000. The Congress openly demanded an independent Papua, separate from Indonesia, through peaceful and democratic means. The province was granted special

autonomy status in 2001 and since then the once popular call for independence, although not extinguished, has diminished significantly.

The center-regional conflicts in Riau and East Kalimantan can be considered as minor because no rebel groups exist and only relatively minor secessionist sentiments have been put forward. The second Riau People's Congress (KKR, or Kongres Rakyat Riau) held in Pekan Baru, January 2000, issued a decree calling for Riau independence, separate from Indonesia. In a roughly similar tone, in November 1999, the provincial parliament of East Kalimantan officially issued a decree (No. 28/1999) demanding federal state status from Indonesia. The demands from these two regions were the climax of growing secessionist sentiments in the regions following the fall of Suharto in mid-1998. Conflict and other characteristics of the four provinces are summarized in Table 1.

Subnational entities reacted to the centralistic nature of the New Order government in three ways. First, the industrialized resource-poor Java region enjoyed benefits during the New Order authoritarian past as in Suharto's New Order regime, the idea of power was very much derived from Javanese culture of paternalistic relationships. Second, non-Javanese resource-poor regions felt unhappy with the centralistic and autocratic style of the regime but acknowledged that they benefitted from central government subsidies. Third, non-Javanese resource-rich regions felt unjustly treated because they were subsidizing the New Order's equalization policy scheme without having much political say in the center.

To face grievances, the center had available a range of policy options and instruments. They included the implementation of the equalization policy with suppression, offering transfers or subsidies, sharing resource rents, or a combination of these three. Overt armed rebellion was an extreme form of noncooperative bargaining and a clear case of social contract failure, akin to what Hirshleifer called the vertical social contract. In strategic game-theoretic terms, one could say that the lack of credibility of the central government policies or its new initiatives toward the regions contributed to this social contract failure and fostered the continuation of secessionist conflicts, particularly in Aceh.[12] But explaining the continuation of center-region conflicts in Indonesia during the Suharto era and thereafter is different from explaining their onset. The separatist movement in the resource-rich provinces was triggered by socioeconomic grievances. In particular, socioeconomic grievances arise from the phenomenon of "aspiration to inequality."[13] This refers to demands by the rich regions for a degree of community welfare that would correspond to their relatively high regional prosperity due to natural resource endowments, a situation that could be phrased as "the rage of the potentially rich." It reflects regional protests against the central government's equalization scheme by distributing resource-rents across the whole of the nation.

Regional prosperity may or may not result in community welfare. The difference between the two therefore has to be highlighted. The former refers to regional output measured solely by GDP. This calculates value-added in monetary terms derived from a region or geographical unit (district, province, or country) in some accounting period, usually one calendar year. The latter refers to indicators of physical well-being of the people living in the region. Measures of community welfare include consumption expenditure and indicators such as of health, education, and poverty, in short, human development measures.

Table 2: Regional prosperity and community welfare (Indonesia = 100)

Provinces (resource-rich districts)	*GDP per cap. 1996*	*Purchasing power 1996*	*Poverty headcount 1996*	*1999*	*Human development index 1996*
Aceh	142	98	72	63	102
(Aceh Utara)	350	96	87	75	103
Papua	170	97	241	234	89
(Fak-Fak)	1,616	92	278	242	97
Riau	241	99	72	60	104
(Bengkalis)	435	97	77	74	103
(Kepulauan Riau)	283	96	54	42	101
East Kalimantan	404	100	55	86	105
Resource-poor non-Javanese provinces	71	99	118	104	99
Java provinces (excluding Jakarta)	80	102	104	111	101

Note: The first numeric column is a measure of regional prosperity; the others reflect community welfare. Purchasing power is based on household expenditure data (National Socioeconomic Survey, SUSENAS).
Source: Authors' calculation based on BPS-Statistic Indonesia data.

Levels of interregional income (regional GDP) inequality have been substantial, mainly because oil, gas, and key minerals are concentrated in only a few regions of Indonesia. But in terms of community welfare, interregional inequality has been extremely low, and this is due to the success of the New Order equalization policies.[14] According to a measure of interdistrict inequality (the L-index), regional output inequality (based on GDP) is almost four times higher than is regional expenditure

inequality (based on a consumption survey measure). This contrasts, for example, with Indonesia's neighbor Thailand, where the L-index figure suggests virtually no correlation across regions between community welfare indicators and the regional prosperity measure.[15]

Table 2 suggests the relative success of the equalization campaign. In GDP per capita terms, the resource-rich regions stack up very well relative to the country as a whole. Yet welfare measures such as purchasing power and human development are not significantly different from the national average or from the non-resource Java and non-Javanese provinces. Only Papua lagged behind the nation as well as the other provinces.

"Aspirations to inequality" are a response to people's first-hand experience of their community welfare being reduced to, or even below, the national average, even though their regions are rich in natural resources. While this shows the success of the equalization scheme, the resource-rich provinces felt that they had been robbed. They expected their welfare level to match their wealth (resources), hence their rage against the center or their aspiration to inequality.[16] This was especially so as the equalization scheme was designed and implemented unilaterally by a Javanese ruler and they had no say in it.

This "rage" was aggravated by visible relative deprivation of the native population, often engineered by the central government. For example, Javanese migrants outnumbered Acehnese in top-strata jobs (government officials, professionals, and technicians) as well as in land holding compared with native Bataks vis-à-vis Javanese migrants in the neighboring province of North Sumatra. In fact, migrant Javanese have become the second-largest ethnic group in both provinces. The government's transmigration policy played a large role in this shift in the ethnic mix. Relative deprivation of the native population vis-à-vis migrants is also evident in Papua. In 1971, native Papuans accounted for 96 percent of the province's population; by 2005 this had been reduced to 59 percent. The remainder are migrants originating from government-sponsored transmigration and voluntary in-migration. Migrants dominate the urban economic sectors and the more productive agricultural activities, while the majority of native Papuans are stuck in the traditional extractive sectors.[17]

Decentralization policies have been effective in dealing with secessionist demands in these resource-rich provinces. The 2001 Special Autonomy Law for Papua has calmed the separatist movement. The provinces of Riau and East Kalimantan that posed strong demands for autonomy have been happy with the two decentralization laws passed in 1999 and their subsequent 2004 revisions. Aceh has been an exception. Having experienced violence for a longer time period, and seeing the central government failing to deliver on promises, earlier decentralization laws did not curb the separatist conflict. For Aceh, the credibility of the central government's policy was at issue. But it seems that the Helsinki agreement between the government and GAM is viewed as credible; it needed a popularly elected President to generate sufficient good-will. Since the signing of the Helsinki agreement, the peace has held.[18]

Table 3: Major episodes of ethnic violence in Indonesia

Main cleavage	*Provinces*	*Affected districts*	*Time-span*	*Estimated casualties*
Dayak-Madurese	West Kalimantan	Bengkayang, Pontianak, Landak, Sambas, Sanggau	30 Dec 1996 - 28 Feb 1997	1,006
Malay-Madurese	West-Kalimantan	Bengkayang, Sambas	19 Jan to 26 Apr 1999	481
Dayak-Madurese	Central Kalimantan	Kotawaringin Timur, Kotawaringin Barat, Kapuas, Palangkaraya	2 Dec 2000 - 6 July 2001*	1,255
Christian-Muslim	Maluku	Ambon, Maluku, Tengah, Maluku Tenggara, Buru	15 Jan 1999 - 2 Nov 2002	2,023
Christian-Muslim	North-Maluku	Ternate, Tirode, all Halmahera	19 Aug 1999 - 7 Dec 2000	2,782
Christian-Muslim	Central Sulawesi	Poso	Apr - July 2000	613
Anti-Chinese	Jakarta		May 1998	1,206
Anti-Chinese	Central Java	Solo	May 1998	33
Total number of deaths				9,399

Note: * The peak of the violence took place the third week of February 2001 in Kotawaringin Timur and claimed 1,200 lives.
Source: Tadjoeddin (2008), summarized from UNSFIR conflict database.

Ethnic strife

From 1996 to 2000, various parts of Indonesia experienced a surge of ethnic violence (see Table 3), especially between the collapse of the New Order in mid-1998 and the full implementation of regional autonomy (decentralization) in 2001.[19]

Excluding the episodes of anti-Chinese violence that took place primarily in Jakarta, the national capital, and in Solo in Central Java (the hometown of Suharto's wife) a few days before Suharto's fall in May 1998, the other major episodes of ethnic violence have been referred to as "small town wars." The ethnic, small-town warfare episodes outside Java, and in eastern Indonesia in particular, were led by urban middle-class elements. These small towns and regions were particularly dependent on central government subsidies The ratio of civil servant to nonagricultural workers has been interpreted as measuring the degree of local reliance on state resources. Indeed, high ratios are found for four provinces (West and Central Kalimantan, Central Sulawesi, and Maluku) where major ethnic strife occurred. Thus, limited economic options in the nonstate-sector may render a region more vulnerable to interethnic competition.[20]

The explosion of ethnic violence coincided with the decentralization reforms of 1999-2000. This may be viewed as a moment of opportunity for local actors to take control of the political and financial resources expected soon to be transferred to local entities under the decentralization scheme. Therefore, the move toward decentralization would significantly inflate the expected benefits of controlling local power, over which conflict and violence would be highly possible. The expected gains from conflict arise either from increasing returns to fighting efforts or from the elimination or conquest of the opponents.[21] This analysis corresponds closely to the greed hypothesis of conflict.[22]

In some cases, interethnic violence goes beyond spontaneous intergroup clashes or ethnic riots, taking the form of serious interethnic warfare. It can spread to a wider geographic area and penetrate deep into the interior, as in the case of the episodes of Dayak-Madurese violence in Kalimantan and Christian-Muslim clashes in Maluku and Sulawesi. Some authors have labeled these as communal war, not simply ethnic violence. For various reasons, in communal war, formal state authority (the police and the military) often has no role to play. On occasion, elements of state authority, motivated by ethnic loyalty, may take sides with the warring parties, leaving behind their formal duty to maintain law and order. Such a situation can be described as anarchy.[23]

An alternative, and perhaps superior, explanation for ethnic or communal violent conflict during the post-Suharto transition period may be offered by the grievance rather than the greed hypothesis.[24] Yet the origin of violence may not lie in the widening of opportunities (horizontal inequality) but in the narrowing of inequality. A previously privileged ethnic group may feel aggrieved when it loses position relative to another group; groups may tend to fight when they become more equal.[25] As we have seen, three decades of development during the Suharto era brought economic and social dislocations for certain groups and resulted in significantly diminished relative positions vis-à-vis others, such as in the case of Dayaks in Kalimantan.[26] Power loss-related grievance by the relatively rich is a new interpretation of horizontal inequality: it is "the rage of the previously rich."[27]

Due to the convergence of socioeconomic conditions, econometric work cannot find convincing horizontal inequality-related evidence for deadly ethnic violence in Indonesia. In one work, for instance, cross-district logistic regressions suggest only child mortality as a statistically significant factor associated with deadly ethnic strife. Other measures of group horizontal inequality such as income, education, land holding, young male unemployment, and government employment did not result in statistically significant effects on ethnic violence.[28] These results have little intuitive appeal; it is hard to believe that child mortality differences between groups can be instrumentally used by ethnic elites to mobilize coethnics.

In contrast, in a two-stage regression approach using district-level data, another study finds that ethnic strife is rooted in relative deprivation-related grievance of local educated, but still poor communities. However, this then mutates into local elites' greed as they compete for political power at the local level, due to high local dependency on state resources. This suggests that greed and grievance are mutual, and inseparable, explanations of conflict.[29]

In the first-stage regression, it is found that districts experience more severe ethnic strife with higher poverty, both in terms of consumption poverty (headcount ratio) and capability poverty (human poverty index), even as they achieve higher levels of education (years of schooling). The finding is robust when taking differences between rural and urban districts and between Java and non-Java, into account. It points to the existence of an expectations gap between actual welfare and higher future welfare expected on account of higher education. The second-stage regression finds a positive and statistically significant relationship between local-level budget allocation by the central government and the severity of violence. After decentralization, budget allocations were increased (in relative terms) to regions that experienced more severe ethnic strife. Regions with severe ethnic strife became more reliant on state resources at the local level, and the state at the local level became a "prize" to compete for. This can be qualitatively linked to the local level political setting in post-ethnic strife periods as the relative bargaining power of ethnic groups changed. In some cases groups have succeeded in turning the clock back, for example Christians in Maluku and Dayaks in Kalimantan. The relative importance of the size of the state at the local level would therefore seem to matter. Rent-seeking reflects greed and can take on violent forms.

Routine violence

Routine group violence is differentiated from separatist and ethnic violence in that it

Table 4: Routine violence in four Javanese provinces* (1994-2003)

	Incidents		*Deaths*		*Deaths per*
Categories	*Total*	*%*	*Total*	*%*	*10 incidents*
a. Popular justice/vigilante	592	30.3	373	59.6	6.3
b. Group brawls	516	26.4	174	27.8	3.4
c. State-Community	244	12.5	27	4.3	1.1
d. Political party & factions	227	11.6	19	3.0	0.8
e. Economic	323	16.5	19	3.0	0.6
f. Others	52	2.7	14	2.2	2.7
Total	1,954	100.0	626	100.0	3.2

Notes: * The provinces are Banten, and West, Central, and East Java. They account for 90 percent of Java's population. (The two Javanese provinces remaining provinces are the Jakarta and Yogyakarta.)
Source: Calculated from the UNSFIR conflict database.

usually occurs episodically.[30] It is treated as the residual, after the other forms of violence have been accounted for. Routine violence centers on vigilante violence or popular justice, and intergroup or neighborhood brawls (see Table 4).[31] Our focus so far has been directed at higher-profile and lethal episodic kinds of incidents that exclusively take place in the form of ethno-communal and separatist violence in the outer islands. In contrast, the low-profile and low-intensity routine kind of violence is most common in Java. Rarely does it produce headlines. Incidents are not reported beyond the local newspapers, unless they turn very ugly. This sort of violence has attracted minimal attention from either researchers or policymakers.

Routine violence is closely linked to widespread social frustration, for example resulting from weak law and order as in the case of popular justice. It might be caused by socioeconomic declines and lower levels of human development and both are likely to create a situation where the opportunity cost of engaging in violence is extremely low. This encourages people — especially unemployed youth — to participate in violence. Routine violence can also be linked to relative deprivation as in the case of the rising segment of an inverted U-shaped relation between education and routine: at relatively low levels of education, violence tends to rise, and with higher levels of education, violence tends to fall.[32] (The increase in income levels might have lagged behind expectations commensurate with increased education levels.) Another possibility is that the increase in education from a very low level might change society to become more dynamic and increase its ability to express dissatisfaction. This, in turn, might result in higher levels of violence until it reaches the turning point. Falling violence at higher levels of income reflects that the opportunity cost of engaging in violence could be high; therefore violence declines.

In the long run, routine violence will decline as society becomes more developed with higher income, education, and overall levels of human development. Such a situation would transform into a stronger state, advance law and order, and in general lead to better institutions.

Conclusion

Early conflict studies in Indonesia were dominated by detailed ethnographic accounts of conflict episodes and sociopolitical and historical perspectives thereon. They have not emphasized socioeconomic perspectives. Several studies have emerged lately that point to the important role of socioeconomic factors. In particular, grievances due to relative deprivation, horizontal inequality, and marginalization can act as conflict drivers. The greed hypothesis of conflict in natural resource-rich regions does not appear to be as strong an explanation.

Notes

Mohammad Zulfan Tadjoeddin, the corresponding author, is a doctoral candidate at the University of Western Sydney, Australia, and visiting researcher at the Institute of Social Studies (ISS), the Netherlands. He may be reached at z.tadjoeddin@uws.edu.au. **Anis Chowdhury** is professor of economics at the University of Western Sydney, Australia. His address is a.chowdhury@uws.edu.

1. See Tadjoeddin (2002); Varshney, Tadjoeddin, and Panggabean (2008).

2. Bertrand (2004).

3. Literature: See, e.g., Huntington (1968); Snyder (2000); Hegre, Ellingsen, Gates, and Gleditsch (2001). Also see Collier and Rohner (2008) for theoretical and empirical findings on democratic transition and violence in poor and lower middle income countries. Modernization: due to Lipset (1959).

4. Chua (2002).

5. Greed: Collier, *et. al*. (2003); Collier and Hoeffler (2004; Collier (2007). Grievance: Stewart (2000; 2008); Gurr (1970); Montalvo and Reynal-Querol (2005). Complementary: e.g., Murshed and Tadjoeddin (2008) and Murshed (this issue).

6. 1996-2002: Klinken (2007); Varshney, *et al*. (2008). 40,000: PRIO-Uppsala conflict database); communists: Sulistyo (2000). The post-independence conflicts in Indonesia between 1945 and the mid-1960s were mostly revolutionary affairs when conflicting groups were negotiating what kind of nation-state they were going to have. In contrast, conflicts during Suharto's military-backed "New Order" authoritarian regime (mid-1960s until the late 1990s) were dominated by state-sponsored violence as a means of repression (Anderson, 2001).

7. Human Development Index: see UNDP, BPS, and BAPPENAS (2004).

8. The reasons may vary from a local issue, such as a development project or the disapproval of a local official to a national issue such as a new regulation or law perceived to be unfair.

9. For a comprehensive discussion of different kinds of conflict in Indonesia, see Tadjoeddin (2002).

10. East Timor is not included, as it has not been part of Indonesia since the 1999 U.N.-organized referendum. It also has a distinctly different history as it was colonized by Portugal, while the rest of Indonesia was under the Dutch rule.

11. Civil war, according to the definition commonly agreed by scholars in the field, takes place when an identifiable rebel organization challenges the government militarily and the resulting violence results in more than 1,000 combat-related deaths, with at least 5 percent on each side (Collier, *et al*., 2003).

12. Vertical social contract: Hishleifer (1995); game-theory model: see Addison and Murshed's (2001).

13. See Tadjoeddin, *et al*. (2001).

14. Under a centralized system, the central government collected all revenues from natural resources and distributed them across all regions, mainly in the form of agricultural development and social expenditure such as basic education and health. The guiding principle for fiscal distribution was the welfare of all Indonesians; the regions which were behind in social development received more, especially through discretionary Presidential allocations, known as Inpres.

15. See Tadjoeddin, *et al*. (2001).

16. In the case of Aceh, Dawood and Sjafrizal (1989) had put forward a similar concern nearly two decades ago. Although by 1985 Aceh, together with Riau and East Kalimantan, had been the largest three provinces contributing to Indonesian exports. Aceh's "benefits to [the] local economy have been much smaller and [the] cost-benefit calculus more problematic, perhaps ironically one of the most staunchly independent regions, long in conflict with the central government, is now subsidizing that government and the rest of the country" (p. 115).

17. North Sumatra: see Brown (2005); Papuan ethnic mix: Elmslie (2008); Papuan economic activity: McGibbon (2004).

18. Effective: Tadjoeddin (2007). Law: Furthermore, the 2006 direct elections for the Papuan governor opened intra-provincial divisions along political and subregional factional affiliations and weakened separatist sentiments (Mietzner, 2006). The separatist sentiment is, however, still there, although only in muted form. Aceh: After the 2004 devastation caused by the tsunami, separatist leaders realized the difficulties of continuing the conflict. The relief and recovery work jointly done by the rebels and the military also helped built trust.

19. The UNSFIR conflict database for Indonesia is constructed based on leading provincial newspapers and complemented with any available information, case studies, and consultation with knowledgeable sources at local the level (see Varshney, *et al*., 2008).

20. See Klinken (2007). In 1990, the North Maluku province was part of Maluku.

21. Return to fighting: Hirshleifer (1995); conquering opponents: Skaperdas (2002).

22. See Klinken (2007).

23. Communal war: Tomagola (2000); Klinken (2007); anarchy: Hirshleifer (1995). Interestingly, since the introduction of decentralization some kind of con-sociational political arrangements at the local has emerged in regions that previously experienced severe ethnic violence. For example, in direct elections it has became a common practice for candidates from previously warring ethnic groups to team up on the same ticket to campaign for head and deputy-head of local executive political office. Thus, after anarchy, previously warring parties have been able to cooperate. This is an indication of some sort of horizontal social contract emerging.

24. Grievance: Steward (2000); greed: Collier and Hoeffler (2004).

25. Horizontal inequality: see Stewart (2000). For such an explanation of Christian-Muslim violence in the Malukus and Sulawesi, see Klinken (2007), Bertrand (2002; 2004), and Aragon (2001). Also see Besancon (2005).

26. See Klinken (2007); Davidson (2008); Peluso and Harwell (2001). Interestingly, changes in the relative position of various ethno-religious groups did not result in the widening of inequality among them. Instead, in most cases the development policies of Suharto's New Order government caused convergence of socioeconomic conditions of different ethno-religious groups; see Tadjoeddin (2003).

27. The logic of "the rage of the rich" is also applicable to the case of secessionist sentiments posed by several resource-rich provinces in Indonesia (see the previous subsection). In a way, the aspiration to inequality can be regarded as the rage of the rich.

28. Mancini (2008).

29. Tadjoeddin (2008).

30. For the Indonesian context, the phrase was introduced by Varshney, *et al*. (2008); also see Tadjoeddin and Murshed (2007).

31. The classes in Table 4 are based on Varshney, *et al*. (2008) and are mutually exclusive, but they can be criticized for using inconsistent bases. Categories (a) and (b) concern forms of violence along intercommunal lines but do not involve ethnic dimensions (otherwise we would call them ethnic violence). Neither do they occur along the lines of categories (c), (d), or (e). Category (c) is based on the state-community cleavage; categories (d) and (e) are similar to (a) and (b). However, if we just universally apply the form-based categorization, all categories take the form of either popular justice/vigilante violence or group brawls without involving ethnic dimensions.

32. Frustration: Welsh (2008); education: Tadjoeddin and Murshed (2007).

References

Addison, T and S.M. Murshed. 2001. "From Conflict to Reconstruction: Reviving the Social Contract." UNU/WIDER Discussion Paper 48. Helsinki: UNU/WIDER.

Anderson, B., ed. 2001. *Violence and the State in Suharto's Indonesia*. Southeast Asia Program. Ithaca, NY: Cornell University.

Bertrand, J. 2002. "Legacies of the authoritarian past: religious violence in Indonesia's Moluccan islands." *Pacific Affairs*, Vol. 75, pp. 57-86.

Bertrand, J. 2004. *Nationalism and Ethnic Conflict in Indonesia*. Cambridge, UK: Cambridge University Press.

Chua, A. 2002. *World on Fire: How Exporting Free Market Democracy Breeds Ethnic Hatred and Global Instability*. New York: Doubleday.

Collier, P. 2007. *The Bottom Billion*. Oxford, UK: Oxford University Press.

Collier, P. and A. Hoeffler. 2004. "Greed and Grievance in Civil Wars." *Oxford Economic Papers*, Vol. 56, No. 4, pp. 563-595.

Collier, P. and D. Rohner. 2008. "Democracy, Development and Conflict." *Journal of the European Economic Association* (forthcoming).

Collier, P., L. Elliot, H. Hegre, A., Hoeffler, M. Reynal-Querol, and N. Sambanis. 2003. *Breaking the Conflict Trap: Civil War and Development Policy*. Oxford, UK: Oxford University Press.

Dawood, Dayan and Sjafrizal. 1989. "Aceh: The LNG Boom and Enclave Development," in H. Hill, ed. *Unity and Diversity: Regional Economic Development in Indonesia Since 1970*. Oxford, UK: Oxford University Press.

Davidson, J.S. 2008. *From Rebellion to Riots: Collective Violence on Indonesian Borneo*. Madison, WI: University of Wisconsin Press.

Elmslie, J. 2008. "Demographic Transition in West Papua and Claims of Genocide." West Papua Project, University of Sydney.

Hegre, H., T. Ellingsen, S. Gates, and N.P. Gleditsch. 2001. "Towards a Democratic Civil Peace? Democracy, Civil Change, and Civil War 1816-1992." *American Political Science Review*, Vol. 95, No., 1, pp. 17-33.

Hirshleifer, J. 1995. "Anarchy and its Breakdown." *Journal of Political Economy*, Vol. 103, No. 1, pp. 26-52.

Huntington, S.P. 1968. *Political Order in Changing Societies*. New Haven, CT: Yale University Press.

Klinken, G.v. 2007. *Communal Violence and Democratization in Indonesia: Small Town Wars*. London: Routledge.

Lipset, S.M. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review*, Vol. 53, No. 1, pp. 69-105.

Mancini, L. 2008. "Horizontal Inequality and Communal Violence: Evidence from Indonesian Districts," in F. Stewart, ed. *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. London: Palgrave Macmillan.

McGibbon, R. 2004. "Plural Society in Peril: Migration, Economic Change, and the Papua Conflict." Policy Studies 13, Washington, D.C.: East-West Center.

Mietzner, M. 2006. "Autonomy, Democracy and Internal Conflict: The 2006 Gubernatorial Elections in Papua." Seminar presentation, RSPAS, Australian National University, 15 May.

Montalvo, J.G. and M. Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review*, Vol. 95, No. 3, pp. 796-816.

Peluso, N.L. and E. Harwell. 2001. "Territory, Custom, and the Cultural Politics of Ethnic War in West Kalimantan Indonesia," in N.L. Peluso and M. Watts, eds. *Violent Environments*. Ithaca, NY: Cornell University Press.

Skaperdas, S. 2002. "Warlord Competition." *Journal of Peace Research*, Vol. 39, No. 4, pp. 435-446.

Stewart, F. 2000. "Crisis Prevention: Tackling Horizontal Inequalities." *Oxford Development Studies*, Vol. 28, No. 3, pp. 245-262.

Stewart, F., ed. 2008. *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. London: Palgrave Macmillan.

Sulistyo, H. 2000. "Palu Arit di Ladang Tebu: Sejarah pembantaian massal yang terlupakan, Jombang-Kediri 1965-66." Jakarta: Gramedia.

Tadjoeddin, M.Z. 2002. "Anatomy of Social Violence in the Context of Transition: The Case of Indonesia 1990-2001." *Politics Administration and Change*, Vol. 38, pp. 1-35.

Tadjoeddin, M.Z. 2003. "Communal Conflicts and Horizontal Inequalities in Indonesia: Dynamics and Consequences." Presentation at the IPSK-CNRS Group Meeting on Conflict, LIPI Jakarta, 15 January.

Tadjoeddin, M.Z. 2007. "A Future Resource Curse in Indonesia: The Political Economy of Natural Resources, Conflict and Development." CRISE Working Paper 35, University of Oxford.

Tadjoeddin, M.Z. 2008. "Educated but Poor: Localized Ethnic Violence during Transition in Indonesia." Paper accepted for presentation at the International Studies Association (ISA) Annual Conference, New York, 15-19 February 2009.

Tadjoeddin, M.Z. and S.M. Murshed. 2007. "Socio-Economic Determinants of Everyday Violence in Indonesia: An Empirical Investigation of Javanese Districts, 1994-2003." *Journal of Peace Research*, Vol. 44, No. 6, pp. 689-709.

Tadjoeddin, M.Z., W.I. Suharyo, and S. Mishra. 2001. "Regional Disparity and Vertical Conflicts in Indonesia." *Journal of the Asia Pacific Economy*, Vol. 6, No. 3, pp. 283-304.

Tomagola, T.A. 2000. "The Bleeding Halmahera of North Moluccas," in O. Törnquist, ed. *Political Violence: Indonesia and India in Comparative Perspective*. Center for Development and the Environment. Oslo: University of Oslo.

UNDP, BPS, and BAPPENAS. 2004. "Indonesia National Human Development Report 2004: The Economics of Democracy, Financing Human Development in Indonesia." Jakarta.

Varshney, A., M.Z. Tadjoeddin, and R. Panggabean. 2008. "Creating Datasets in Information-Poor Environments: Patterns of Collective Violence in Indonesia (1990-2003)." *Journal of East Asian Studies*, Vol. 8, No. 3, pp. 361-394.

Welsh, B. 2008. "Local and National: Keroyokan Mobbing in Indonesia." *Journal of East Asian Studies*, Vol. 8, No. 3, pp. 473-504.

Forced displacement in Colombia: magnitude and causes

Ana María Ibáñez

The intensification of civil conflicts during the 1990s resulted in a substantial increase in the number of civilian victims. Reasons for the rise of these wars include unresolved conflicts, on-going conflicts since the end of World War II, the dissolution of the Soviet Union, the subsequent proliferation of armaments, and the strengthening of the illegal drug trade. All contributed to an increase in civil conflicts and intrastate wars during the 1990s. This was accompanied by an increase in the average duration of conflict: in 1947, the average conflict lasted approximately two years; by 1999 this had gone up to 15 years.[1]

In contrast to interstate conflicts, whose victims mostly are members of the armed forces, victims in intrastate conflicts are mainly civilians. The civilian population is not only recruited by the rebel forces, but also attacked in order to eliminate civilian leaders, frighten civilians, widen rebels' area of influence, evacuate their opponents from a specific area, and take over the population's assets. Thus, violence against civilians is not a random and unfortunate consequence of war. Instead, it is a deliberate strategy of war. Consequently, the percentage of civilian casualties of war has increased, from ten percent at the beginning of the twentieth century, to 90 percent by 1990. Increased threats and attacks on civilians have resulted in a constantly rising tendency in the number of refugees and displaced peoples since the end of the Second World War. Since 1990, the number of displaced people has fluctuated around 25 million and the number of refugees around 13 million.[2]

The objective of this article is to analyze population displacement in Colombia. The first section examines the economic literature on the dynamics of forced migration and the close relation it holds with the war strategies adopted by armed groups, whether government or rebel forces. The second section deals with the evolution and causes of forced displacement in Colombia. The third section concludes.

The civilian population in civil conflicts

In intrastate violent conflict armed groups attack civilians as part of their war-fighting strategy. Attacks come from insurgent violence, state violence, and the interaction between the two. State violence against civilian populations has been frequent in some African countries, creating high refugee flows, while attacks on civilians by insurgent groups are mainly found in Latin America, generating forced migrations within a country's borders.[3]

Violence against civilians is not a random and unfortunate consequence of war. Instead, it is a deliberate strategy of war.

Armed groups attack civilians for a variety of reasons. The exercise of territorial control is among them. Some areas are attractive for their economic wealth, for example for the cases of illicit crops, the existence of valuable natural resource reserves, such as mining or oil resources, or if they are considered potential routes for transporting illegal drugs or weapons with ease. Other regions are important for rebel forces as part of their war strategy and the need to expand territorial strongholds. Aggression against civilians allows insurgents to take over valuable assets to finance their war operations or for their members' personal gain. Furthermore, they allow insurgents to strengthen territorial supremacy by driving possible opponents from the region, avoiding civilian resistance, weakening social networks, or separating civilians from rebel groups. Rebel groups frighten the population by murdering civic leaders and threatening the population at large. This prevents the creation of opposition groups to gain influence outside the battlefield, all at a low cost. For all these reasons, violence toward the population is not a byproduct of war. It is a deliberate strategy of war.[4]

To achieve their objectives, selective murders, massacres, death threats, forced recruitment, forced disappearances, and the destruction of property are some forms of attacks carried out by rebel forces to frighten civilians. Such attacks force the population to migrate to protect itself, and this creates flows of refugees and displaced people. The close relation between civil war and forced migration either within a country's borders (displaced population) or outside its borders (refugees) has been widely documented in a series of important studies. They have focused primarily on establishing the relation between internal conflict and the flow of refugees to other countries. Studies that evaluate migration data within a country, for example for Colombia, Guatemala, and Indonesia, have examined the relation between actions carried out by armed rebels and forced displacement. All confirm that the main reason for forced migration lies in the direct and indiscriminate attacks of which civilians are the victims.[5]

Forced displacement in Colombia

Colombia faced surges of violence for most of the twentieth century. The first half of that century was characterized by the struggle for control of political institutions, land, and resources, while the emergence of drug-trafficking and the weakening of the judicial system were important additional conflict determinants during the second half. These conflicts have exerted a progressively heavier toll on civilians. In particular, two conflicts that arose during the second half of the twentieth century targeted and inflicted significant losses on civilians. The first emerged from the

political struggle between the two traditional and major political parties in Colombia, the Liberals and the Conservatives. The national conflict extended to Colombia's regions and fueled traditional confrontations for land and economic resources. After the assassination of a major political leader of the Liberal Party in 1948, the period known as *La Violencia* intensified. Homicide rates soared throughout this period. To end the violence, a power-sharing arrangement between the parties was negotiated in 1958. This paved the way for a peace deal and eased violence significantly, but many of the deep causes of the conflict persisted.[6]

Second, during the early 1960s, three left-wing guerrilla groups — the FARC, ELN, and ERP — emerged in the southern regions of the country. Although the objective of these groups was to seize power, their initial activities were restricted to isolated regions and consisted of sporadic attacks on government forces and occupation of rural towns and lands.[7] Guerrilla violence intensified and expanded throughout Colombia in the late 1970 and early 1980s with the appearance of illegal drug cultivation and trade. The drug trade provided massive monetary resources to finance rebel groups. Besides funding these groups, the drug-trade instigated the creation of right-wing paramilitary groups by drug-barons and landowners. Their main purpose was to defend themselves from guerrilla aggression and to contest their territorial control over certain regions. The emergence of paramilitary groups, coupled with resources from the illegal drug trade, intensified the Colombian conflict. This resulted in increasing attacks on the civilian population.[8]

The worsening conditions led to a tripling of homicide between 1970 and 1991. The origins of today's violence are manifold and complex: the illicit drug trade, a weakening of the legal system, the appearance of armed rebels, poverty, inequality, and decades of armed conflict are among the causes.[9] The impact of violence differs between urban and rural areas. While urban areas face high homicide rates, rural areas are subject to armed confrontations, massacres, attacks by armed rebels, and forced displacement. Kidnappings, massacres, selective murders, forced recruitment, and land mines, among others, are among the aggressions civilians face. In 2001, the number of kidnappings reached three thousand people, one of the highest in the world. By 2006, this had significantly decreased to 687 reported cases from 28 of Colombia's 32 administrative departments. The number of massacres also decreased: in 2006 only 37 cases were reported with 193 associated deaths. Active or abandoned land mines, the use of which has yet to end in Colombia, caused 1,100 victims in 1,964 incidents, a 14 percent increase from 2005.[10]

Forced displacement in Colombia: magnitude and geographic expansion

Forced displacement, the most dramatic victimization of civilians has reached high levels. According to the Registry for Displaced Populations (RUPD), by 31 March 2008, the displaced population numbered 2,452,152 people.[11] The objective of RUPD is to legally recognize displaced households and to quantify the demand for aid. The registry is a demand-driven instrument, whereby displaced households must approach government offices to declare, under oath, the circumstances of their displacement. After a victim makes a declaration, government officials validate whether it is truthful and, if so, the legal status to be granted the members of the displaced household. About 30 percent of the displaced population is believed not to be registered, so that the total number of displaced people may number over 3.5 million, or 7.8 percent of the national population.[12]

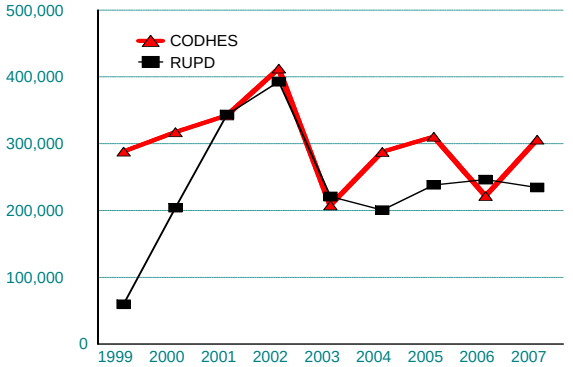

Figure 1: Number of displaced people in Colombia. *Source*: Acción Social and CODHES.

Figure 1 shows the displacement trend, beginning in 1999, the year in which the RUPD was adopted. The figure illustrates displacements, both according to official records and according to CODHES, a nongovernmental organization. Both trend lines indicate a strong increase in forced migration between 1999 and 2002 when Colombia lived through a period of heavy conflict. After 2002, forced displacement decreased, yet an average of 266,635 people still migrated involuntarily each year, the equivalent of a medium-sized city in Colombia. Moreover, even when peace negotiations with paramilitary groups moved forward and the intensity of the conflict diminished, more than half of all forced migrations between 1999 and 2007 fell in the period of between 2003 and 2007. This indicates that a decrease in conflict does not necessarily imply an immediate reduction of forced displacement.

The magnitude of forced migration is revealed when compared with international indices. For 2008, the Internal Displacement Monitoring Center for example estimates total worldwide internal population displacement at approximately 24.5 million people. Of those 14.3 percent are found in Colombia, second only to Sudan.[13] This ranking persists independently of the data source used.

The number of people displaced is not only large; it also has steadily expanded across the country. By 2002, the number of municipalities affected by out-placement (expelled residents), had risen to 949 to affect more than 90 percent of the country´s municipalities. The expansion of displacement is shown in the two panels of Figure 2 that indicate which municipalities either expelled or received displaced populations between 1999 and 2008. (Municipalities left blank neither expelled nor received displaced populations. Only a few municipalities did not expel or receive displaced

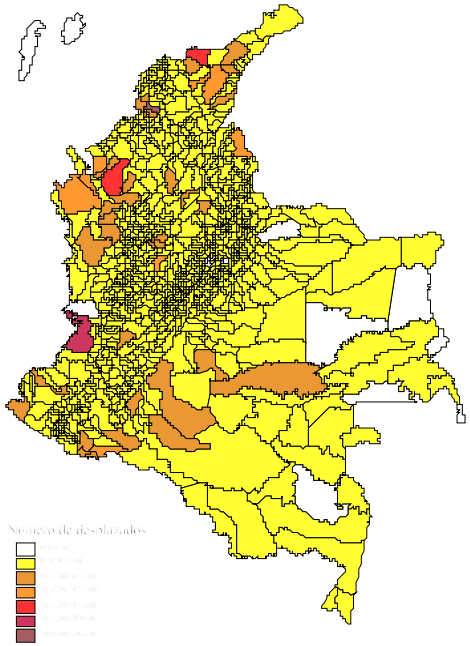

Figure 2a: Geographic distribution of displacement. Municipalities that expelled populations: 1999-2008. *Source*: Ibáñez and Velásques (2008).

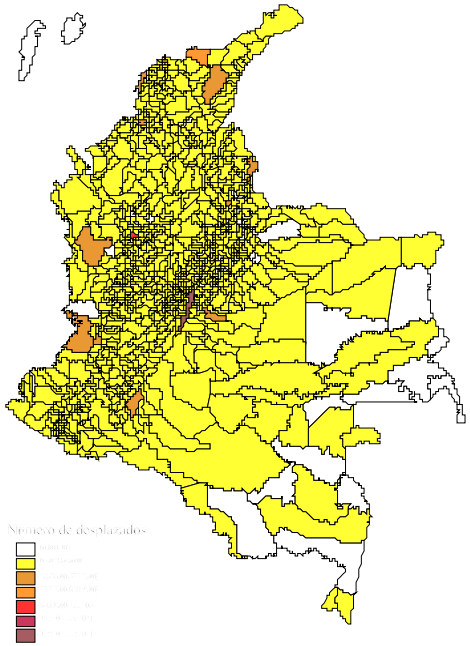

Figure 2b: Geographic distribution of displacement. Municipalities that received populations: 1999-2008. *Source*: Ibáñez and Velásquez (2008).

populations.) As may be seen, the phenomenon is not confined to remote or isolated municipalities; to the contrary, it extends throughout the whole of Colombia.

Although displacement seems to affect all of the country's municipalities, impact is more substantial in some regions than in others. When the intensity of displacement is calculated, that is, the percentage of the displaced population in relation to total population, one finds that some municipalities have lost more than half of their population. Some of the most dramatic examples are Bojayá (Chocó) with an intensity of 94.7 percent, Cocorná (Antioquia) with 93.8 percent, El Tarra (Norte de Santander) with 82.4 percent, Peque (Antioquia) with 77.5 percent, and Riosucio (Chocó) with 76.1 percent. Moreover, ten percent of Colombia´s municipalities have lost close to one-fourth of their population as a direct consequence of forced displacement. Pressure indicators, which calculate the arrival of displaced populations in relation to the native population, likewise indicate the difficulties faced by mid-sized cities in absorbing large migration flows. The pressure indicators for selected department capitals demonstrate this: Mocoa (Putumayo) 33.3 percent, Quibdó (Chocó), 26.4 percent, Sincelejo (Sucre) 24.6 percent, Florencia (Caquetá) 20.3 percent, and San José (Guaviare) 18.5 percent. For the eight years to 2007, all of these municipalities, which already have difficulties attending to their own population's needs, have received about one person per year for every five locals. Even though municipalities of course deal with ordinary, voluntary population movements, those that move involuntarily face extremely difficult conditions due to the victimization process described earlier, such as large asset losses, poor conditions in the receiving municipalities, and the sheer difficulty of participating in the labor markets. This constitutes a humanitarian emergency that cannot be solved using the assigned resources.

Immediate and underlying causes of displacement[14]

Given their constant attacks on civilians, the main parties responsible for the population displacement in Colombia are guerilla and paramilitary groups. The worsening conflict between them has led both to adopt ever harsher measures of attack. Aside from facilitating territory control, asset appropriation, and weakening of social networks, this also forces the population to seek refuge in other municipalities. Data by Social Action reveals that when the perpetrator is reported, 45.8 percent of the time guerilla are responsible for the displacement; in 21.8 percent of the cases it is paramilitary groups, and 1.1 percent for state forces (with the remainder for other groups).[15]

Population displacement is provoked by any of the following situations: attacks by armed groups, indiscriminate violence, or the mere presence of armed groups. After a direct attack, civilians are displaced, but to avoid possible victimization, preventive displacement occurs as well. Table 1 lists various types of attacks civilians face: the multiplicity of attacks on the population is chilling. The main migration triggers are threats (54.5 percent), homicide or homicide attempts (53.4), indiscriminate violence (39.1), confrontations between armed groups (36), eviction notices (29.6), and massacres (21.1). Displaced populations suffer from double

Table 1: Motive for displacement*

Variables	*Percentage of households*
Direct threats	54.5
Homicide attempt	18.9
Homicide	34.5
Disappearance or torture	14.4
Forced recruitment	17.3
Massacres	21.1
Kidnapping	7.6
Armed confrontations	36.0
Indiscriminate violence	39.1
Eviction notice	29.6

Source: Ibáñez and Velásquez (2008).
Note: * Percentages do not add up to 100 percent as one family may be the victim of several events.

victimization: they are subject to crimes against humanity and they are forced into a condition of displacement.

Migration strategies for displaced populations in Colombia are specific to the Colombian conflict. Ordinarily, in countries where forced migration occurs, populations generally move in massive numbers and settle in refugee camps. This happens after confrontations between rebel groups and armed forces, or after massive and direct attacks on local communities. Lack of protection by official armed forces, or their participation in the events, generates this mass migration toward other countries in search of protection.[16] The situation in Colombia is very different. For the most part, people migrate individually, and few move outside the country's borders. Figure 3 illustrates the evolution of individual and mass displacement between 1999 and 2007.[17] Close to 80 percent of displacement in Colombia is individual. The highest levels of mass displacement were reported from 2000 to 2002, when attacks on and combat in municipalities were common. Individual displacement has been predominant and rising since then.

Regardless of the clear relation between the presence of armed groups, violence, and forced displacement, the underlying causes for forced migration in Colombia are difficult to identify. The last incident in a chain of events is generally the immediate trigger that forces a decision to finally flee in search of refuge. However, the origins of forced displacement can be found in the dynamics of the Colombian conflict. The following paragraphs describe some of the hypotheses on displacement in Colombia that recurrently appear in the literature.

Land disputes and the illegal takeover of lands are considered underlying causes of forced displacement. Illegal occupation of land is one of the crucial strategies of war used to clear territories of possible opponents, increase territorial control, and take over valuable land. Displacement is more intense in municipalities with high degree of informality regarding land ownership. This is particularly so in regions that are contested by armed rebels and where these groups are trying to consolidate their supremacy. These conditions generate a high displacement index among small rural landowners, and abandoned lands are calculated to range from 1.7 to 4 million acres.[18]

Forced displacement also occurs where illicit crops and drug trafficking are found. The growth of illicit crops adds pressure on land and displacement because of the acquisition of lands for cocaine and poppy crops and also because of the importance of transport routes for drugs. Illegal crops' aerial fumigation leads to asset destruction for farmers, directly affecting their income, and increases confrontations as well. The increase in aerial fumigation in the last few years has been the cause of an important migration wave. According to CODHES, 36,000 people have migrated for this reason since 1998. An estimated 13,153 people migrated during 1999 from the cocaine regions and war zones in Guaviare, Meta, Caquetá, and Putumayo. This number rises to 20,000 if interdepartmental migration is included, and to 30,000 or 40,000 if cocaine-leaf farmers and settlers are added.[19]

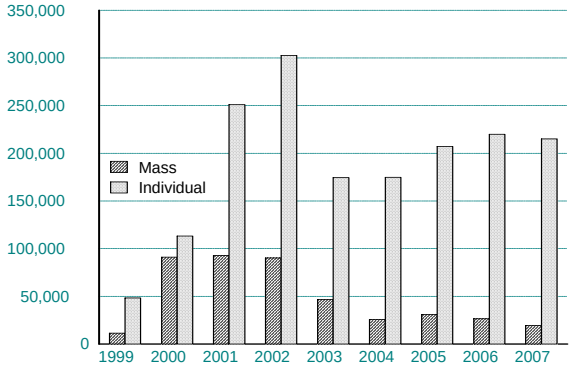

Figure 3: Mass vs. individual displacement.
Source: Acción Social (RUPD).

In some regions, a close relation between drug dealing and displacement is observed. Land acquisition by drug traffickers as a means to launder money generates land speculation, reduces the state's purchasing power, and impedes possibilities of negotiation between peasants and landowners. In addition, when purchasing land, drug dealers sometimes inherit social conflicts that are then dealt with by creating self-defense groups and lead to an increase in conflict, and of course displacement.[20]

Population expulsion as a strategy of war is also a way to impede civil resistance movements, weaken social networks, and intimidate the population in order to exercise more control. Attacks on the population weaken support by the opposition and avoid civil uprisings: 65 percent of displaced people were active members of community organizations, and slightly over 11 percent participated in union or political organization in their home towns. People with strong social networks are more likely to suffer an attack by armed rebels.[21]

Finally, families living in rural areas may involuntarily migrate in order to avoid forced recruitment of their children by armed rebels. Children as young as eight have been recruited as soldiers for armed rebel groups in Colombia. For example, after an armed confrontation in October 2001, Colombian military forces discovered that close to 43 percent of guerrilla members killed in combat and 41 percent of captured members were younger than 18 years of age. UNICEF estimates indicate that close

to 14,000 minors have been recruited by rebel groups as soldiers.[22]

Conclusions

The data presented in this article indicate that forced displacement in Colombia is a humanitarian emergency. Levels of forced displacement in Colombia are high, even in comparison to other countries that suffer from a similar situation. The profile of the displaced population indicates their condition both as victims of displacement and victims of violence. However, contrary to the situation in many other countries, forced displacement is not generated by state organs but by nonstate armed rebel groups. In addition, displacement in Colombia is generally individual and victims disperse within the country's borders, whereas in other countries victims participate in mass migrations and settle in refugee camps. The causes of forced displacement lie in the nature of armed conflict in Colombia. Institutional weakness, the weakening of social networks, the takeover of lands, and a desire to profit from natural resources in some areas are among the causes of forced displacement in Colombia. To achieve their war objectives, the targeting of civilians is part of the strategy of war adopted by armed rebel groups.

Notes

Ana María Ibáñez is Associate Professor, Department of Economics, Universidad de los Andes, Bogotá, Colombia. She may be reached at aibanez@uniandes.edu.co. This article is based on two chapters of her 2008 book *Desplazamiento forzoso en Colombia: ¿un camino sin retorno a la pobreza*.

1. Fearon and Laitin (2003).

2. Strategy of war: Azam and Hoeffler (2002); casualties: Cairns (1997); displaced persons: Internal Displacement Monitoring Centre; see www.internal-displacement.org [accessed 28 March 2008]; also see Weiner (1993); Wood (1994); Davenport, *et al.* (2003); Lubkemann (2005); refugees: U.S. Committee for Refugees and Immigrants; see www.refugees.org [accessed 28 March 2008].

3. Sources of attacks: Melander and Öberg (2006); Africa: Azam and Hoeffler (2002).

4. Separation: Hedmanm (2005); low cost: Lubkemann (2005); strategy of war: Ibáñez and Vélez (2008); Azam and Hoeffler (2002); Stanley (1987).

5. Important studies: Melander and Öberg (2006); Davenport, *et al*. (2003); Schmeidl (1997); Stanley (1987); Zolberg, *et al*. (1986); country studies: see Czaika and Kis-Katos (2007); Morrison and May (1994); Schultz (1971).

6. Regions: Oquist (1976); homicide rates: Brauer and Gomez-Sorzano (2004); Brauer, *et al*. (2004).

7. Southern regions: Brauer, *et al*. (2004); isolated activities: Echeverry, *et al*. (2001).

8. Gaviria (2000); Thoumi (2002).

9. Vélez (2002); Sánchez and Núñez (2001); Gaviria (2000).

10. Kidnappings, 2001: see www.policia.gov.co [accessed 16 March 2008]; kidnappings, massacres, land mines, 2006: see www.derechoshumanos.gov.co [accessed 22 April 2008].

11. Law 387 of 1997 defines an internally displaced person as "anyone who has been forced to migrate within the national boundaries, living outside her residence or her habitual economic activities because either her life, her physical integrity, or her freedom have been either violated or threatened by situations such as armed conflict, generalized violence, violations to human rights, and any other situation that may alter public order."

12. Ibáñez and Velásquez (2006).

13. See www.internal-displacement.org [accessed 22 April 2008].

14. This section is based on the work by Ibáñez and Vélez (2008) "Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia" and Ibáñez and Querubín (2004) "Access to land and forced displacement in Colombia"

15. See www.acciónsocial.gov.co [accessed 21 April 2008].

16. Moore and Shellman (2006).

17. According to the official definition, a massive displacement occurs when a group of 50 or more people or 10 or more homes migrate simultaneously.

18. Underlying causes: Reyes and Bejarano (1998); Kay (2007); informal land ownership: Ibañez (2008); displacement index and abandoned acreage: PMA (2001).

19. Puyana (1999). A more recent quantification is not available since forced displacement caused by the eradication of illegal crops is not included in Law 387 of 1997.

20. Pérez (2002); Reyes (1997). Pérez (2002) indicates that pressure on land and subsequent conflicts occur in areas where strategic resources are found, not only illicit crops, but also coal, oil, and emeralds, creating situations of violence and speculation.

21. Avoid civil uprising: Henao (1998); active in community organizations: Lozano and Osorio (1999); strong social networks: Engel and Ibáñez (2007).

22. Child recruitment: Salazar (2001); killed/captured: USCR (2003); UNICEF: www.unicef.es [accessed 6 May 2008].

References

Azam, J.P. and A. Hoeffler. 2002. "Violence against Civilians in Civil Wars: Looting or Terror?" *Journal of Peace Research*, Vol. 39, No. 4, pp. 461-485.

Brauer J. and A.G. Gómez-Sorzano. 2004. "Homicide Cycles in Colombia, 1950-1999." *International Journal of Applied Econometrics and Quantitative Studies*, Vol. 1, No. 1, pp. 1-22.

Brauer J., A.G. Gómez-Sorzano, and S. Sathuraman. 2004. "Decomposing Violence: Political Murder in Colombia, 1946-1999." *European Journal of Political Economy*, Vol. 20, pp. 447-461.

Cairns, E. 1997. *A Safer Future: Reducing the Human Cost of War*. Oxford, UK: Oxfam Publications.

Czaika, M. and K. Kis-Katos. 2007. "Civil Conflict and Displacement: Village-Level Determinants of Forced Migration in Aceh. HICN Working Paper No. 32.

Davenport, C.A., W.H. Moore, and S.C. Poe. 2003. "Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964-1989." *International Interactions*, Vol. 29, No. 1, pp. 27-55.

Echeverry, J.C., N. Salazar, and V. Navas. 2001. "¿Nos parecemos al resto del mundo? El conflicto colombiano en el contexto internacional." *Archivos de Macroeconomía*. Documento 143. Departamento Nacional de Planeación. Bogotá, Colombia.

Engel, S. and A.M. Ibáñez. 2007. "Displacement Due to Violence in Colombia: A Household Level Analysis." *Economic Development and Cultural Change*, Vol. 55, No. 2, pp. 335-365.

Fearon, J.D. and D.D. Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review*, Vol. 97, No. 1, pp. 75-90.

Gaviria, A. 2000. "Increasing Returns and the Evolution of Violent Crime: The Case of Colombia." *Journal of Development Economics*, Vol. 61, No. 1, pp. 1-25.

Hedman, E.L.E, 2005. "A State of Emergency, A Strategy of War: Internal Displacement, Forced Relocation, and Involuntary Return in Aceh," in E.L.E. Hedmann, ed. *Aceh under Martial Law: Conflict, Violence and Displacement*. Refugee Studies Centre, University of Oxford.

Henao, H. 1998. *Desarraigo y futuro. Vida cotidiana de familias desplazadas de Urabá*. Cruz Roja Colombiana. Medellín, Colombia: Cruz Roja.

Ibáñez, A.M. and P. Querubin. 2004. "Acceso a tierras y desplazamiento forzado en Colombia." Documento CEDE 2004-23, Bogotá, Colombia.

Ibáñez, A.M, nd A. Velásquez. 2006. "El Proceso de Identificación de Víctimas de los Conflictos Civiles: Una Evaluación para la Población Desplazada en Colombia." Documento Cede 36, Bogotá, Colombia.

Ibáñez, A.M. and C.E. Vélez. 2008. "Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia." *World Development*, Vol. 36, No. 4, pp. 659-676.

Ibáñez, A.M. 2008. *El desplazamiento forzoso en Colombia: un camino sin retorno hacia la pobreza*. Bogotá, Colombia: Editorial Uniandes.

Kay, C. 2007. "Reflections on Rural Violence in Latin America." *Third World Quarterly*, Vol. 22, No. 5, pp. 741-775.

Lozano, F.A. and F.E. Osorio. 1999. "Horizontes de comprensión y acción sobre el desplazamiento de población rural en Colombia (1995-1997)." CODHES, Bogotá.

Lubkemann, S.C. 2005. "Migratory Coping in Wartime Mozambique: An Anthropology of Violence and Displacement in 'Fragmented Wars'." *Journal of Peace Research*, Vol. 42, No. 4, pp. 493-508.

Melander, E. and M. Öberg. 2006. "Time to Go? Duration Dependence in Forced Migration." *International Interactions*, Vol. 32, pp. 129-152.

Moore, W. and S. Shellman. 2006. "Refugee or Internally Displaced Person? To Where Should One Flee?" *Comparative Political Studies*, Vol. 39, No. 5, pp. 599-622.

Morrison, A.R. and R.A. May. 1994. "Escape from Terror: Violence and Migration in Post Revolutionary Guatemala." *Latin American Research Review*, Vol. 29, No. 2, pp. 111-132.

Oquist, P. 1980. *Violence, Conflict, and Politics in Colombia. Series in Social Discontinuity*. New York: Academic Press.

Pérez, L.E. 2002. "Desplazamiento forzado en Colombia 1995-1999: Una aproximación empírica a las relaciones entre desplazamiento, conflicto armado y desarrollo," in *El desplazamiento forzado en Colombia: compromisos desde la universidad*. Bogotá, Colombia.

PMA. Programa Mundial de Alimentos. 2001. "Estudio de caso de las necesidades alimentarias de la población desplazada en Colombia." Bogotá, Colombia.

Puyana, A.M. 1999. "Cultivos Ilícitos, fumigación y desplazamiento en la Amazonía y la Orinoquía," in F. Cubides and C. Domínguez, eds. *Desplazados, Migraciones Internas y Reestructuraciones Territoriales*, Bogotá, Colombia: Centro de Estudios Sociales - Universidad Nacional y Ministerio del Interior.

Reyes, A. 1997. "La compra de tierras por narcotraficantes en Colombia," in F. Thoumi, ed. *Drogas Ilícitas en Colombia*. Bogotá, Colombia: Editorial Planeta.

Reyes, A. and A.M. Bejarano. 1998. "Conflictos agrarios y luchas armadas en la

Colombia contemporánea." *Análisis Político*, Vol. 5, pp. 6-27.

Salazar, M.C. 2001. "Consequences of armed conflict and internal displacement for children in Colombia." Winnipeg Conference on War Affected Children, Report of the United Nations High Commissioner for Human Rights on the Human Rights Situation in Colombia.

Sánchez, F. and J. Núñez. 2001. "Determinantes del Crimen Violento en un país altamente violento: el caso de Colombia," in A. Martinez, ed. *Economía, Crimen y Conflicto*. Bogotá, Colombia: Universidad Nacional de Colombia.

Schmeidl, S. 1997. "Exploring the Causes of Forced Migration: A Pooled Time-Series Analysis, 1971-1990." *Social Science Quarterly*, Vol. 78, No. 2, pp. 284-308.

Schultz, T.P. 1971. "Rural-Urban Migration in Colombia." *The Review of Economic and Statistics*, Vol. 53, pp. 51-58.

Stanley, W.D. 1987. "Economic Migrants or Refugees from Violence? A Time-Series Analysis of Salvadoran Migration of the United States." *Latin American Research Review*, Vol. 22, No. 1, pp. 132-154.

Thoumi, F.E. 2002. "Illegal Drugs in Colombia: From Illegal Economic Boom to Social Crisis." *Annals of the American Academy of Political and Social Science* Vol. 582, pp. 102-116.

USCR. 2003. *World Refugee Survey 2003*, Washington. D.C.

Vélez, C.E. 2002. *Colombia Poverty Report*. Washington, D.C.: World Bank.

Weiner, M. 1993. "Security, Stability, and International Migration." *International Security*, Vol. 17, No. 3, pp. 91-126.

Wood, W. 1994. "Forced Migration: Local Conflicts and International Dilemmas." *Annals of the Association of American Geographers*, Vol. 84, No. 4, pp. 607-634.

Zolberg, A.R., A. Suhrke, and S. Aguayo. 1986. "International Factors in the Formation of Refugee Movements." *International Migration Review*, Vol. 20, No. 2, pp. 151-169.

Are Congo's mines the main target of the armed groups on its soil?

Steven Spittaels and *Filip Hilgert*

The Democratic Republic of Congo (DRC) has a soil rich in minerals and a recent history characterized by an enduring armed conflict with enormous humanitarian costs. It is often assumed that there is a direct relation between these two characteristics.

Throughout the various stages of the conflict, the DRC's minerals have been used by the different armed groups to finance their war effort. After the presidential elections of 2006, natural resources all over the DRC have continued to attract armed men but it would be wrong to assert that minerals constitute the core issue of the current conflict dynamics. Our geographical research on four different war motivations for the period August 2007 to January 2008 shows how in the Kivu provinces of the DRC other factors, such as personal security issues, ethno-economic grievances, and the protection of individual business interests, play a role that is more important than the direct struggle over natural resources.[1]

Since the surge of econometric conflict analyses at the end of the 1990s, in the wake of the greed versus grievance debate, less attention has been given to in-depth studies of the motivations behind specific conflicts. For little less than a decade, quantitative econometric methodologies have dominated much of the research activity on the issue. Initially this research had a great appeal to policymakers and the general public. As opposed to the existing qualitative research, the statistical analyses, based on rational-choice theories, resulted in simple and clear-cut economic explanations of the phenomenon of civil war. However, the findings of the statistical analyses have proven to be highly sensitive to coding rules (and errors) in the quantitative data sets used, sometimes leading to contradictory correlations, which has made it clear that there is a continuing need for detailed and case-specific research on the drivers of war.[2]

For this reason, we have developed a qualitative geographical research tool that we have used for our DRC analysis and that we believe can furnish convincing evidence on war motivation. The geographical research model is intended to complement the existing qualitative and quantitative research on the drivers of (civil) wars. Whereas classic qualitative research on conflict motivation always risks being suspected of insufficient objectivity, the existing quantitative research, only draws conclusions on a general level. We felt that a scientific method assessing the behavior of individual conflict parties in a concrete war situation and making use of evidence open to objectification, was needed.

In the first, introductory, section of this text, we explain the geographical research methodology we apply in our analyses on war motivation. In the second section, we turn to the analysis we made of the situation in the Kivu provinces at the end of 2007 using that research tool. We discuss the role of the mining sector in the armed conflict while comparing its importance as a war motivation to other conflict drivers found in literature on the causes of war. Although all warring parties receive income from the mineral sector, their behavior and tenacity are motivated by other factors. In the final section, we broaden our scope by arguing that the Eastern DRC is only one of Congo's important mining areas. Many of the other mineral deposits are located in post-conflict zones, where they continue to experience serious security problems of their own.

Our geographical research on four different war motivations for the period August 2007 to January 2008 shows how in the Kivu provinces of the Democratic Republic of the Congo other factors, such as personal security issues, ethno-economic grievances, and the protection of individual business interests, play a role that is more important than the direct struggle over natural resources.

The geographic research tool on war motivation

The theoretical model we use to analyze motivations of armed groups in the DRC, schematized in Figure 1, starts from the assumption that war is a means used to attain a certain goal or objective, for example the secession from an existing state. Such objectives are driven by a certain motivation: for example, people want to secede because they are being discriminated against and oppressed. The *objectives* of a war are an answer to the question to what end it is fought. War *motivation*, in contrast, explains why warring parties want to attain a certain objective. Motivation precedes objective. Why some motives lead to the outbreak of war and others do not, depends on opportunities and other situational factors.

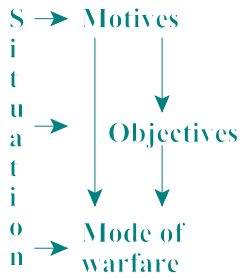

Figure 1: Model for the relation among motives, objectives, and the way in which a war is being waged.

Warring parties may claim to be driven by certain motives and toward certain objectives but this could be a mere misrepresentation. In order to reveal the real motivations of warring parties, we need to start from the facts on the ground. The objective one wants to attain and the motivation behind it will influence the way in which a war is being waged. If an armed group seeks power and self-rule for a certain region and fights to secede the territory from the motherland, this will show in its military actions. We can assume the group will

attempt to secure the internal borders and to conquer the regional capital.

In short, our tool analyses concrete military actions and decisions and traces them back to what provoked them. A full description of the methodology is available online at www.ipisresearch.be/mapping under the heading "Handbook." Because the tool is still being developed, the handbook is a draft that we revise every eight months.

In our analysis of the second half of 2007 phase of the Eastern DRC war, we have compared the relative importance of four basic motivations in explaining the behavior of armed groups on the ground:

- Profit is the central motivation in the "greed" theory of conflicts. War can create huge personal gain, which may be enough reason to wage one. Profit-motivated conflicts entail different phenomena, like trafficking of natural resources, pillaging, and illegal taxation.
- Grievance: In every society there are groups and individuals that oppose the existing political and/or social situation. Their dissatisfaction is caused by feelings of inequality, oppression, discrimination, hatred, and injustice. When these feelings remain unaddressed, they can become a driver of violent conflict.
- Survival: People or peoples who feel threatened in their survival quite often resort to violence to safeguard their future. Essential elements for survival are: access to food, access to water, physical security, shelter, living space, and outlet possibilities.
- Power: History has seen many examples of politicians using war to win more political or territorial power. Wars driven by the search for power are wars for conquest. The relationship between power and geography has always been the central topic in geopolitics.[3]

For each of these motivations we have identified geographical targets on a digital map of the Eastern DRC. We have produced, for example, a map with mining areas that represent targets for profit and a map that indicates dense forest areas, which can be a target for armed groups that want to hide and survive. The motivational maps are static maps. The geographical features on these maps have a fixed location. In our analysis, we have compared these sets of static motivational maps with a second collection of maps composed of dynamic elements. The dynamic maps are a series of snapshots of the positions of the armed groups in the area and the military incidents or human rights abuses in which they have been involved. All the online maps make reference to the sources used to build them. In general, the static maps have been constructed by combining existing cartographical material from geographic and other scientific or research institutes. Some maps, such as the natural resources and trade routes maps, also contain data acquired by the authors during interviews when interviewees were asked to pinpoint elements of their story on a map. The dynamic maps are based chiefly on data from interviews, with some elements from written sources with geographical references.[4]

We have compared the static and dynamic maps by making use of GIS software. A GIS (Geographic Information System) is software capable of integrating, editing, analyzing, and presenting geographical information. Characteristic of a GIS is that the different features of the digital map are organized in separate layers (e.g., places, rivers, territorial boundaries, land use, etc.) that can be combined at will (switched on or off) and that to each of these layers tables with additional information can be linked. Because it combines spatial with nonspatial data, it is a powerful analytic tool. Within the GIS environment we have superimposed the dynamic maps upon the maps of the static set. In this manner we have been able to check which motivational targets are present in those territories where armed groups have significant troop concentrations or where important military confrontations took place. We have worked under the assumption that armed groups concentrate their troops and efforts in those areas where their interests lie.

The armed conflict in the East

The following findings are based on our reading of a set of digital maps produced at the International Peace Information Service (IPIS). A web version of these maps is published on our web server. It is advisable to consult these maps while reading the text in order to be able to fully follow the matters that are being discussed. The level of detail on the maps can be changed by zooming in or out. The maps are available at three different scales: 1:2,500,000 (initial view), 1:1,000,000, and 1:500,000. An advanced geographical search function is available and one can also search thematically for data by clicking the "Lists" button. The map will center on the requested map element and automatically a table will appear with additional information on the map feature. The same additional information on map features can be retrieved by clicking on the item directly on the map itself. The complete written conflict analysis and all conclusions are also available online. All facts mentioned below were taken from the maps and the analysis, where there are references to their sources.[5]

Despite relatively successful elections in 2006, the North Kivu province of the Democratic Republic of Congo and to a lesser extent South Kivu, have been occupied by four warring parties involved in a protracted civil war. At the center of the conflict are the territories of Masisi and Rutshuru in North Kivu where the rebel movement of General Laurent Nkunda fights the government army and two other armed groups. The politico-military movement of Nkunda is called the National Congress for the Defense of the People (CNDP). It is a recent movement that was launched by Nkunda during the 2006 elections. The bulk of the CNDP forces are Rwandophones. Both Hutu and Tutsi are represented in significant numbers but most of the higher cadres are Tutsi. It is generally recognized that Nkunda and his movement serve the interests of (at least a part of) the Tutsi minority in the Kivus.[6]

One of the main reasons of existence of the CNDP is to fight the Forces

Démocratiques de Libération du Rwanda (FDLR), a movement that originates from ex-FAR (Forces Armées Rwandaises) soldiers, ex-Interahamwe militiamen, and Hutu civilians that fled the offensive of the Rwandan Patriotic Front (RPF) in neighbouring Rwanda in 1994. The FDLR are the largest armed group in the Kivu provinces. The FDLR still comprise a few high ranking members that were implicated in the 1994 genocide but the large majority, among whom are many children, did not play any role in that event.[7]

Besides the FDLR, two Congolese forces are fighting the CNDP: the official Congolese army (FARDC) and the people's Mayi-Mayi militias. Since the beginning of 2007, the main bands of Mayi-Mayi warriors have forged an alliance called the Coalition of Congolese Resistant Patriots, "PARECO. In this text we will not deal with the Mayi-Mayi motivations.

Conflict motivations

When analyzing the positions and behavior of the CNDP between August 2007 and January 2008, we find only limited evidence of direct involvement in illegal mining. Nkunda and his militia occupy only two important mining sites, one of which is not even operational. However, apart from the mining story, greed should not be ruled out as an important war motivation for them. Several of the main CNDP financers have considerable business interests in the region, especially in cattle farming. A considerable percentage of the areas occupied by the CNDP are hilly grazing lands very well suited for cattle raising.

In the past Nkunda declared an independent Republic of the Volcanoes with its own flag and anthem. However, looking at our dynamic maps, it does not seem as if the CNDP is really interested in gaining power. It does not control any administrative capital and it has never seriously tried to take one since its inception.

With regard to its grievances, to a certain extent reality matches the discourse. The CNDP has set itself up as the protector of the Tutsi population in the region. On the one hand, it has frequently engaged the FDLR, heirs of the 1994 *génocidaires* that it considers to be a threat to the Tutsi population. On the other hand, it occupies many of the villages from which more than 40,000 Tutsi inhabitants fled to the neighboring countries and to which they want to return. Arguably, however, the CNDP is more concerned with safeguarding its territory than its people. In December 2007, for example, Nkunda seriously endangered the lives of "his people" when he used them as a human shield against a large-scale FARDC offensive.

Besides the larger towns and the Nkunda zones, the FDLR are virtually omnipresent in the Eastern DRC at the end of 2007. Some of the rare areas where they have no presence are the main mining sites. They control many smaller mining sites but the money they can make at these sites is very limited compared to the revenues from the larger ones. Moreover, the FDLR have deployed several of their 11 battalions in areas that are almost devoid of mines. In these areas, they use other methods to generate an income, for example growing and trading marihuana. As in the case of the CNDP, the FDLR use their weapons as leverage to make a profit. Some FDLR members may prefer to stay on Congolese soil out of greed-related reasons. But it can hardly be claimed that the FDLR members lead a luxury life in the jungles and villages where they have been staying for 14 years. The FDLR do not control any major town center. They hold positions along several roads but in general their forces are concentrated in heavily forested areas. It seems that the large majority of them are in fact hiding themselves and surviving. This is especially true for the small minority of FDLR with a genocidal past.

Concerning their grievances, the FDLR claim their fight is not against the DRC government but against Rwanda. They say they seek a peaceful solution through an inter-Rwandan dialogue when security conditions in their home country are met. FDLR grievances seem a plausible explanation for their behavior in the field. They have taken a defensive stance, and since 2005 they have stopped carrying out armed attacks on Rwandan soil. It is true that they have never used their military power to try to gain political control in the Congo. However, there have been some reports of armed confrontations with the FARDC, and the FDLR have used force in a series of plundering raids on Congolese villages.

In general, the FARDC positions in the Kivu provinces during the same period are concentrated in areas neighboring the zones controlled by Nkunda and the CNDP. Most of the FARDC brigades are involved in military operations against the CNDP. They are trying to reestablish the power of the Congolese state, which is their legitimate task as the official protectors of the Congolese territory and its population. But some of the FARDC units are deployed in remote areas where they hold positions with little strategic value and far away from the frontlines. In these areas there is only a small presence of rebel groups but there are many mines. The main motive of these troops is to enrich themselves and their superiors by preying on the mining activity or by digging themselves. Besides mining, FARDC soldiers and officers are often also involved in other types of illegal economic activities such as timber, coffee, and marihuana trafficking and taxation.

The war in the East and the mines: some examples

If it is not the main driving factor, then what role does the mining sector play in the war in the Kivu provinces? In this section, we provide additional insight into the involvement of armed groups in illegal mining and trafficking of minerals by giving some concrete examples from the August 2007 to January 2008 time period.

First, there are those militiamen who work as diggers on account of their superiors. For example in the Bibatama mine, the largest operational mine controlled by Nkunda, CNDP soldiers are mining tin ore, coltan, and wolframite. The CNDP is not the only conflict party that is directly involved in mining. FDLR and FARDC soldiers also take part. For instance, one of the incidents we have included on our

maps describes how FARDC soldiers of the 11th Integrated Brigade, deployed in the territory of Shabunda in South Kivu, have been digging for cassiterite at the airstrip of Tshonka causing serious damage to the runway. The minerals gathered by the soldiers are sold to trusted traders or sometimes the military commanders run a trading company of their own, most often through civilian intermediaries.

Second, each of the armed groups levies all kinds of taxes throughout the mining process. Artisanal miners often have to pay a fee to enter the mine or they have to hand over a part of their daily earnings. Other common types of taxation are the *péage de route* (toll) and other transport taxes. The FDLR control a stretch of road between Bukavu and Shabunda and have installed six barriers along it. At each barrier passers-by have to pay a dollar to be accompanied to the next checkpoint. In the same territory, the FARDC levy a tax on every cargo flight that leaves for the provincial capital of Bukavu. The CNDP controls the strategic crossings along which minerals are transported to Goma. In Mushaki and Kitchanga it has erected roadblocks where all vehicles passing by are stopped. It is estimated that the CNDP earns $10,000 each week from taxes at the barrier of Mushaki alone.

Third, there is protection money. In Bisie, the biggest tin mine in the whole of the DRC, an FARDC brigade "secures" the site.[8] They are paid by the administrator of the territory who receives a share of 10 percent of the production for organizing the security there. It is, however, the FARDC themselves that cause the biggest security problems. In the past, different units of the 85th brigade currently deployed in the area have even fought among themselves.

Sometimes the control of certain mines by these armed groups is total. The militias determine the prices at which the minerals are traded, they preside at business meetings, control markets, and hire personnel. When several armed groups are present in the same mining area, it is mostly the FARDC that control the town center and the central pits and the rebel groups that control the outlying deposits. This is, for example, the case for the Numbi mine at the border of North and South Kivu. The FDLR and the FARDC each have their own pits and do not engage each other militarily.

The large majority of the minerals mined in the Kivu provinces arrives in Bukavu and most importantly in Goma before being exported. Often, they have gone through the hands of several middlemen. It is therefore nearly impossible to control the origin of the cassiterite or coltan that is being sold. Minerals sold by the militias end up in Europe and Asia.

Mining and conflict in the rest of the DRC

The facts mentioned in the following paragraphs are also based on the maps and reports published at our web site.[9]

There are more than 2000 mining concessions in the DRC. The majority of them lie in post-conflict areas or in areas the Congo wars did not reach. But also in these regions the mining sector is often characterised by conflicts and violence.[10] A first fault line runs between the artisanal miners on the one hand and international mining companies on the other hand. Congolese state-owned enterprises have signed a series of joint ventures with private companies to restart industrial production. However, the concessions these companies should operate are often occupied by large groups of artisanal miners who earn their living in the mines. The forced expulsion of artisanal workers from mining sites in the past has led to mass protests with violent exchanges between workers and the police. During a violent protest in Likasi in March 2008, one worker was killed.

The joint ventures between the Congolese state-owned enterprises (SOEs) and the international companies are a problem in themselves. Most of the contracts have been signed during the war years or the transition period. The negotiation process has often been conducted in a very nontransparent way and the resulting deals are highly unfavorable to the Congolese state. A commission was appointed by the Congolese government to revise 61 mining contracts. Its final report was published in November 2007 and concluded that none of the existing contracts is acceptable: 39 should be amended and the other 22 cancelled.[11] The contracts the commission has selected for renegotiation contain some very unfavorable deals for the Congolese SOEs.[12] In a contract with the Australian mining company Anvil Mining for the Dikulushi mine in Katanga, for example, it has been agreed that Anvil should not pay any taxes for 20 years. Another example is the Tenke Fungurume Mining joint venture, in which the participation of the Congolese SOE Gécamines has been reduced to a meagre 17.5 percent. The renegotiation promises to be an outright legal battle with huge costs but also enormous potential benefits.

The joint ventures only concern a small percentage of all the Congolese mines. There are many other mines where only artisanal miners are active. Often these mines lie in remote areas and, although they are located outside the war provinces, they are being controlled by armed men. In most cases it is the FARDC who control these mines, in some it is other state officials or, in rare cases, remnants of people's self-defense groups. The soldiers present at the mining sites commit serious human rights violations. They extort money and minerals from the population, they force people to mine for them, they commit sexual crimes, they allow diggers in closed mining sites with radioactive material, they organize traffic of minerals, and commit many other crimes. Sometimes even a small-scale war erupts at such sites between units of the FARDC. For example, at the gold mine of Lunga in 2007 and the coltan mine of Kisengo in 2008, both in Katanga, gunshots were exchanged between army units. In Kisengo, which lies more than 200 km away from the Tanganyika Lake, even a unit of the Congolese Navy showed up. The perpetrators of these crimes go unpunished. In several instances in Northern Katanga, soldiers who had been arrested by the Military Justice were liberated by force by their comrades and commanders.

In the DRC new mines are being discovered regularly. Especially at such new sites the situation is comparable to the Wild West: population booms, epidemics, land

conflicts, exclusion of indigenous people, widespread crime, and the almost total absence of the state of law.

Conclusion

We carried out in-depth research on a well-defined time period of the protracted war in the Eastern DRC. In our research we have compared the importance of four war motivations that could be driving the conflict and its armed groups. In this article we have summarized the conclusions drawn from that research while focusing specifically on the role of the mining sector which is often cited as the main source of the conflict.[13]

Our analysis reveals that, as opposed to persistent popular belief, Congo's mines are not the main target of the armed groups on its soil. The government army and the rebel groups control several mining sites in Eastern Congo but the fighting is not concentrated around them. Moreover, the sites controlled by the rebel groups are of relatively minor importance, especially when compared with mines from outside the war provinces. Other motives, such as grievances and security, deserve at least as much attention as natural resources.

However, Congo's mines remain places of insecurity even when they are far away from the front. If there are no rebel groups harassing the local population, it is most likely the FARDC who control the area and behave in a similar way as the rebel groups.

The significance of our research lies in the fact that the relative importance of four types of war motivations — greed, grievance, power, and security — was compared for each of the belligerent parties. Our findings acknowledge the importance of the greed versus grievance debate, but they also demonstrate the need to take the other two explanations for war motivation, power and security, into account. Moreover, they provide a detailed insight in how different motivations of warfare can interact.

More added value to existing research stems from the fact that a qualitative geographical analysis, as opposed to the more typical econometric studies, adds detailed information on the behavior of conflict parties and the war events on the ground. The insights generated from such geographical case studies can credibly help to explain some contradictory findings from quantitative research. Additionally, the geographic tool provides a solid amount of logical arguments, open to objectification, to sustain its conclusions (an aspect that is often lacking in nongeographic qualitative analyses).

Finally, the authors are convinced that should they have studied the further development of the same conflict, they could have clearly shown how motivations can change throughout a conflict. At the time of writing this article (October 2008), Laurent Nkunda has transformed his CNDP into a movement that will pursue the "total liberation" of the DRC. He has taken over the administration of the regional capital of Rutshuru and he would have captured the provincial capital of Goma had not the United Nations intervened. Clearly, his change of discourse has also generated a change in his battlefield behavior, which could be convincingly illustrated on an accurate map.

Notes

Steven Spittaels and **Filip Hilgert** are researchers at the International Peace Information Service, a research institute in Antwerp, Belgium. The Institute specializes in conflict, arms trade, natural resources, and Corporate Social Responsibility studies in sub-Saharan Africa. The authors may be reached at mapping@ipisresearch.be.

1. Spittals and Hilgert (2008).

2. Dominated: The most prominent researcher on the topic is Paul Collier. As an illustration of the enduring influence of his work it suffices to mention that in 2008 Collier and Hoeffler (2004) was cited in 13 different articles published in the *Journal of Peace Research*; explanations: Woodward (2005); clear-cut explanations: Cramer (2002); case-specific: Woodward (2005).

3. Greed: see, e.g., Collier (2000); Collier, *et al*. (2003); Hirshleifer (1995); Keen (1998); grievance: see, e.g., Azar's theory on protracted social conflict in Miall, Ramsbotham, and Woodhouse (1999); Harff and Gurr (2004); Horowitz (2000); Sambanis (2001); survival: see, e.g., Homer-Dixon (1996); Peluso and Watts (2001); power: see, e.g., Agnew (1998); Chauprade (1999).

4. Elements taken from reliable written sources have been directly included on the maps. Those gathered through interviews have been presented for confirmation to at least two other local experts. Furthermore, because all geographical information is publicly available, permanent control by everyone involved is possible.

5. Maps: http://www.ipisresearch.be/maps/Oost-Congo/web/index.html; written analysis: http://www.ipisresearch.be/mapping_kivu.php [both accessed 11 December 2008].

6. International Crisis Group (2007).

7. Romkema (2007).

8. The case of Bisie has been studied extensively by the U.N. Group of Experts and it has been the subject of many research reports. Some of the most important are Miller (2005); FinnWatch (2007); Global Witness (2005); Tegera and Johnson (2007);

Garret (2007).

9. See http://www.ipisresearch.be/mapping_katanga.php; http://www.ipisresearch.be/updates-katanga.php [accessed 11 December 2008].

10. Concessions: The report of the U.N. Group of Experts on the Democratic Republic of the Congo of 18 July 2006 speaks of 2,144 concessions listed at that time. The examples used in this section all are based on research carried out in 2007 and 2008.

11. See http://www.miningcongo.cd [accessed 4 December 2008)].

12. Custers (2008).

13. Some recent examples include Global Witness (2008); Hari (2008); Fédération des comités de Solidarité avec l'Afrique (2008).

References

Agnew, J. 1998. *Geopolitics. Re-visioning World Politics*. London: Routledge.

Chauprade, A. 1999. *Introduction à l'analyse géopolitique*. Paris: Ellipses.

Collier, P. 2000. "Economic Causes of Civil Conflict and their Implications for Policy." Washington, D.C.: World Bank.

Collier, P., *et al*. 2003. *Breaking the Conflict Trap. Civil War and Development Policy*. Washington, D.C.: World Bank.

Collier, P. and A. Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers*, Vol. 56, No. 4, pp. 563-595.

Cramer, C. 2002. "Homo Economicus Goes to War: Methodological Individualism, Rational Choice and the Political Economy of War." *World Development*, Vol. 30, No. 11, pp. 1845-1864.

Criekemans, D. 2007 *Geopolitiek. 'Geografisch geweten' van de buitenlandse politiek?* Antwerp: Garant.

Custers, R. 2008. "Democratic Republic of the Congo. Mining Contracts - State of Affairs. The Congo Wants to Raise the Profits of its Mining Sector." *IPIS article*. Antwerp, Belgium: IPIS.

Fédération des comités de Solidarité avec l'Afrique. 2008. "À qui profite la nouvelle guerre au Congo?" Press release. Spain: Umoya. 29 October 2008.

FinnWatch. 2007. "Connecting Components, Dividing Communities: Tin Production for Consumer Electronics in the DR Congo and Indonesia." FinnWatch Report. Helsinki: FinnWatch. http://www.finnwatch.org/pdf/DRCwebversio.pdf [accessed 5 December 2008].

Garret, N. 2007. "The Extractive Industries Transparency Initiative (EITI) and Artisanal and Small-Scale Mining (ASM)." Draft report. EITI. http://eitransparency.org/files/publication_file/NG_EITI_Report_22_10_Final.pdf [accessed 5 December 2008].

Global Witness. 2005. "Under-Mining Peace. Tin: The Explosive Trade in Cassiterite in Eastern DRC." Report. London: Global Witness. http://www.globalwitness.org/media_library_get.php/238/Under-Mining%20Peace.pdf [accessed 5 December].

Global Witness. 2008. "Resource Plunder still Driving Eastern Congo Conflict." Press release. London: Global Witness.

Harff, B. and T.B. Gurr. 2004. *Ethnic Conflict in World Politics*. 2nd ed. Oxford: Westview Press.

Hari J. 2008. "How we Fuel Africa's Bloodiest War. What is Rarely Mentioned is the Great Global Heist of Congo's Resources." *The Independent*. 1 November 2008.

Hirshleifer, J. 1995. "Theorizing about Conflict." Los Angeles, CA: Department of Economics, University of California, Los Angeles.

Homer-Dixon, T.F. 1996. "nvironmental Scarcity, Mass Violence, and the Limits to Ingenuity." *Current History*, Vol. 95, No. 604, pp. 359-365.

Horowitz, D.L. 2000. *Ethnic Groups in Conflict*. London: University of California Press.

International Crisis Group. 2007. "Congo: Bringing Peace to North Kivu." Crisis Group Africa Report, No. 133. Nairobi/Brussels: International Crisis Group. http://www.crisisgroup.org/home/index.cfm?id=5134&l=1 [accessed 5 December 2008].

Keen, D. 1998. "The Economic Functions of Violence in Civil Wars." Adelphi Paper. London: Institute for International Strategic Studies.

Miall, H., O. Ramsbotham, and T. Woodhouse. 1999. *Contemporary Conflict Resolution*. Cambridge: Polity Press.

Miller, J. 2005. "Tin Soldiers." TV report. London: Channel 4. http://www.channel4.com/news/articles/world/congos%20tin%20soldiers/108670 [accessed 5 December 2008].

Peluso, N.L. and M. Watts, eds. 2001. *Violent Environments*. Ithaca, NY: Cornell University Press.

Romkema H. 2007. "Opportunities and Constraints for the Disarmament and Repatriation of Foreign Armed Groups in the Democratic Republic of Congo." MDRP commissioned report. De Veenhoop, The Netherlands: Conflict & Transition Consultancies.

Sambanis, N. 2001. "Do Ethnic and Nonethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution*, Vol. 45, No. 3, pp. 259-282.

Spittaels, S. and F. Hilgert. 2008. "Mapping Conflict Motives: Eastern DRC." IPIS Report. Antwerp: IPIS, http://www.ipisresearch.be/mapping_kivu.php [accessed 11 December 2008].

Tegera, A. and D. Johnson. 2007. "Rules for Sale: Formal and Informal Crossborder

Trade in Eastern DRC." Report. Goma: Pole Institute. http://www.pole-institute.org/documents/regard19_anglais.pdf [accessed 5 December 2008].

Woodward, S.L. 2005. "The Inequality of Violence: On the Discovery of Civil War as a Threat to 'The North' in the 1990s and the Debate over Causes and Solutions." New York: The Graduate Center, City University of New York.

War and exchange rate valuation

Christopher E.S. Warburton

This article investigates the extent to which the dominance of the United States dollar as an international currency has been contingent on American diplomacy rather than the prosecution of expensive wars. Four wars are of interest here, the Korean War (1950-1953), the Vietnam War (1964-1975), the Persian Gulf War (1990-1991), and the Iraq War (2003-present). The historical performance of the dollar in times of war and peace is examined. The Box-Jenkins forecasting algorithm is employed to make a short-term projection of the dollar coinciding with the Iraq war. The price of gold is used as a measure of the value of the U.S. dollar and investor confidence in the dollar during times of war and peace. The empirical evidence shows a short-term depreciation of the U.S. dollar coinciding with the Iraq War, a finding not atypical of the value of the dollar in times of war. Problems with the value of the U.S. dollar in times of war lead to the exploration of alternative forms of money, which if successful, can erode the continued dominance of the U.S. dollar as an international currency.

Historically, the basis of the dollar's dominance as an international currency has been the result of diplomacy before and after World War II and of geopolitical diplomatic arrangements in the 1970s and 1980s that enabled the dollar to be closely linked with gold and, subsequently, to become a numeraire for international currencies in international commodity markets. The dollar has been very influential in trading arrangements and commodity markets, one result of which prevented its collapse after the Vietnam War.

The empirical evidence shows a short-term depreciation of the U.S. dollar coinciding with the Iraq War, a finding not atypical of the value of the dollar in times of war.

In order to determine the sustainability of the dollar as a dominant international currency, this study on economic performance and currency valuation, unlike others, pointedly focuses on the historical value of the dollar in times of war and the reaction of economic agents to its value during such periods. To evaluate responses to the dollar's value in war time, Bank of England data on the monthly dollar price of gold for the period of January 2000 to March 2008 is employed, i.e., a time period mostly coincident with the ongoing Afghan and Iraq wars. The Box-Jenkins method is then used to forecast the dollar's direction.

The article is structured as follows. The first section contains a brief overview of the exchange rate literature. This is followed by an analysis of the emergence of the dollar as a dominant international currency. The penultimate section discusses the value of the dollar in the context of the relevant wars, followed by a discussion of the forecasting method. The final section presents the empirical results and conclusions.

The exchange rate literature

The exchange rate literature can be classified into three broad categories: exchange rate determination and forecasting, currency crashes (causes and consequences), and monetary nationalism or optimum currency areas. Apart from Krueger's early summary of the literature in terms of the current account, the capital account, or the interactions of both, research on exchange rate determination has focused on key economic variables such as interest rate, international trade (including capital flows), and the money supply or inflation.[1] Some of the important fundamentals such as errors in money demand, foreign-exchange risk premiums, and the equilibrium real exchange rate are not observed by econometricians.

Exchange rate modeling has been controversial and inconclusive. For many years exchange rate forecasting models were judged on their performance against the random-walk model for forecasting exchange rates. This performance criterion was popularized by Meese and Rogoff who discovered that empirical exchange rate models of the 1970s that seemed to fit well within-sample did not have good out-of-sample fit, even though they use realized explanatory variables to predict exchange rates.[2]

The work of Meese and Rogoff encouraged review of three models of analyses — overshooting, monetary, and portfolio models — that were prevalent in the 1970s and early 1980s largely as a result of the work of Dornbusch, and Frenkel and Mussa.[3] There appears to be some agreement in the literature that standard models that relate exchange rates to monetary variables like prices and interest rates perform poorly.[4]

In 1995 interest in empirical exchange rate model was revived, partly as a result of the finding by Mark that models could predict exchange rates over long horizons.[5] However, some researchers are uncomfortable with long-term projections, although empirical evidence suggests that in the long-run exchanges rate can be explained by the variables predicted by the monetary model. It has been argued elsewhere that current account balances have an impact on the short-run dynamics of exchange rate movements in some cases, but that they do not affect the long-run dynamics of exchange rates. This evidence is consistent with the view that in a world of high capital mobility, exchange rates are a monetary phenomenon in the long-run.[6]

Although many researchers thought that standard models of exchange rates based on macroeconomic variables such as prices, interest rates, and output provide weak empirical results, Engel and others present evidence to the contrary while conceding that beating the random walk in forecasting is too stringent a criterion for accepting an exchange rate model.[7] Comparatively recent studies highlight the role of expectations in determining exchange rate movements. As it stands, the literature shows a lack of consensus whether models are helpful in projecting exchange rate

movements or that the models can be effectively evaluated.[8]

In forecasting the movement of the dollar during the planning and execution of the Afghan and Iraq wars, this study applies a univariate model for a variety of reasons. First, monetary variables — money supply, prices, and income — tend to exhibit endogeneity (interdependence); second, modeling expectations is beyond the scope of this work; third, the primary forecast motive of this article is to project the direction of the dollar during a period of war by relying on the data itself; and fourth, econometric theory shows that if theoretical speculations about economic structure are well-founded, then it can be shown that one manifestation of that structure generates an autoregressive moving average (ARMA) process for each of the endogenous variables in the structure.[9]

The literature on exchange rates has been diversified further as a result of the East Asian crises of the 1990s. Close scrutiny has been given to the causes and consequences of currency crises and crashes and were largely seen in terms of capital reversibility, destabilizing speculation, and moral hazard.[10] As the literature evolved to its present form, war and currency preeminence has not formed a significant component of its content.

As a consequence of war, currency dominance is also contingent on the competition of successful optimum currency areas.[11] In international monetary arrangements, optimum currency areas in competition with the dollar entail consequences for the dollar in times of peace and war.

The theory of the optimum currency area was developed by Mundell in the 1960s to show that fixed long-run exchange rates and common monetary policy could be beneficial for a region of nations.[12] In the 1980s he identified the possibilities of competition that the U.S. dollar could face. It was noted then that European Monetary Union (EMU) countries will eventually comprise a transaction domain that is considerably larger than the dollar area and that the euro will become an international currency on the same scale as the dollar.[13] Anxiety over the value of the dollar and its war time depreciation makes gold and the euro more competitive and attractive as an alternative store of value.

Much more recent studies on the European currency area have focused on the evolution of the union and the conditions under which it might be beneficial to participate in an optimum currency area.[14] There is a special interest in efficiency and openness in these recent studies. This article contributes to the literature in the area of war and questionable currency preeminence by maintaining that strong optimum currency areas, lack of dollar diplomacy, geopolitical anxiety, and the availability of competing forms of money challenge the dominance of the U.S. dollar as an international currency.

Dollar diplomacy, Pax Americana, and the dominance of the U.S. dollar

Three significant events promoted the U.S. dollar as an international currency by peaceful means between 1900 and 1975: first, dollar diplomacy; second, the Bretton Woods arrangement after the Second World War; and third, the successful arrangement with the Organization of Petroleum Exporting Countries (OPEC) to denominate the price of oil in international markets in U.S. dollars.

Dollar diplomacy (1900-1928) was a foreign policy phase designed to extend American economic interests abroad by peaceful means. Although American financial and commercial interests were extended to various areas of the world (including China) during this period, the policy was strategically developed to extend American influence in Latin America and the Caribbean after the Spanish American War of 1898.

The overriding objective of dollar diplomacy as it was articulated in 1911 by the then-Assistant Secretary of State, Francis Wilson, was to substitute dollars for bullets. Notwithstanding its setbacks in Nicaragua and China under President Taft, dollar diplomacy created amazing opportunities for American commerce and national income that ultimately contributed to the stability of the dollar and its dominance in Latin America. In 1904, Panama dollarized fully and a substantial number of Latin American countries subsequently embraced de facto or full dollarization in the twentieth and twenty-first centuries. The problem of "original sin" — when a country cannot borrow in its own currency in international financial markets — hastened the pace of dollarization.[15]

American trade and investment with Latin America more than doubled between 1900 and 1910, and the U.S. State Department established a separate Latin American Bureau in 1909. American investment in Cuba rose to $200 million in 1911, and it increased its economic power in the Dominican Republic.[16]

At the height of dollar diplomacy (1909-1913), before Woodrow Wilson repudiated it in theory although not necessarily in practice, the closing price of an ounce of gold on the New York Mercantile Exchange remained constant at $20.67 — an indicator of relative stability and confidence in the dollar for twelve years. The dollar price of an ounce of gold as quoted by the Exchange actually remained at this price from 1878 until the Great Depression of 1932.

Relative peace and booming commerce and industry placed the United States in an enviable position after World War II. As the Europeans decimated their infrastructure and ruined their economies during the World Wars, America emerged with a stronger economy, accumulated a balance of trade surplus, and extended the Marshall Plan to the Europeans. The World Wars created a monetary system in disarray, one that was characterized by a chaotic system of competitive devaluation and plummeting volume of trade. In the subsequent arrangement of the international monetary system (i.e., Bretton Woods), the dollar rose to prominence over the British pound and other currencies, and the new period of prosperity coincided with what became known as Pax Americana — the post-World War II period of American peace and economic and military superiority. Major conflicts among the great powers were avoided until the Korean War of the 1950s, a war which expedited the subsequent

After the Bretton Woods fiasco, and the abandonment of the gold standard, the dollar's fortunes were restored in the international oil markets. The OPEC arrangement was propitious because it tied the dollar to the demand for a good with negligible price sensitivity; demand for oil and dollars became inextricably linked. It should be recalled, however, that the new financing of American debt with recycled petrodollars was largely the result of two wars (Korea and Vietnam) and expansionary monetary policy.

decline of the dollar.

The Bretton Woods system was a fixed exchange rate system in which nations determined the value of their currencies in relation to gold or the U.S. dollar equivalent of $35 an ounce (called mint parity). As such, the system is sometimes referred to as the Gold Exchange Standard in which the dollar became a numeraire. The Bretton Woods system replaced this standard, and with the new system, the dominance of the British pound was dislodged by the dollar. Bretton Woods also brought the International Monetary Fund (IMF) into existence, the institution that oversees the smooth running of the international financial system. The new leadership role of the dollar created responsibilities for the U.S. government and its monetary authority.

The dominance of the dollar in the international monetary system meant that the United States was to maintain the price of gold at $35 an ounce and be prepared to redeem dollars for gold at that price without restrictions. In effect, the United States was to guarantee the principle of convertibility. A challenging proposition, it meant fiscal and monetary discipline. As a result of two major wars (Korea and Vietnam) and unsustainable budget deficits, U.S. President Richard Nixon suspended convertibility in 1971 and the dollar was devalued by about 9 percent, from $35 to $38 per ounce of gold.[17]

The wars created loss of confidence in the dollar, symptomatic to a run on the Federal Reserve Bank system. In response, American businesses and policymakers resorted to strategic geopolitical diplomacy to sustain the dollar's dominance. For example, then-Secretary of State, Henry Kissinger, also famous for his shuttle diplomacy in the Middle East during the 1970s, was instrumental in establishing the United States-Saudi Arabian Joint Commission on Economic Cooperation. This was a forum for Saudi and American financial officials to discuss matters of common concern, and it provided a trajectory for the Saudi government to help U.S. companies increase their exports to Saudi Arabia, while the Saudis recycled petrodollars into long-term U.S. securities to finance U.S. debt. This economic cooperation soon included military protection for Persian Gulf states after the Soviet invasion of Afghanistan in 1979.[18]

As the specter of geopolitical instability loomed over the Middle East, the Gulf Cooperation Council (GCC) of Saudi Arabia, Bahrain, Kuwait, Oman, Qatar, and the United Arab Emirates was formed in 1981 to resist outside intervention in the Gulf and to promote regional economic integration. Under the leadership of Saudi Arabia, the GCC supported oil pricing in U.S. dollars, purchased U.S. debt securities, and accumulated U.S. dollars as foreign reserves. After the Bretton Woods fiasco, the dollar's fortunes were restored in international oil markets. The OPEC arrangement was propitious because it tied the dollar to the demand for a good with negligible price elasticity of demand; demand for oil and dollars became inextricably linked. It should be recalled, however, that the new financing of American debt was largely the result of two wars and expansionary monetary policy.

War and the decline of the U.S. dollar

There are substantive reasons why war diminishes the value of a currency. To prosecute wars, particularly over extended periods of time, nations lose foreign reserves and borrow or print money. For example, at the end of World War I Germany experienced hyperinflation mainly because it printed money to meet its reparation obligations. The price of a daily newspaper increased from 0.30 marks in 1921 to 70 million marks on 17 November 1923.[19] After the Korean and Vietnam Wars, the United States lost a tremendous amount of gold reserves, which made it unrealistic to defend the dollar-gold exchange rate. Inability to maintain convertibility ultimately resulted in a devaluation of the dollar in the 1970s. The road to devaluation actually started in Korea in the 1950s. The Congressional Research Service (CRS) estimates the cost of the Korean War to have been $295 billion (in 2007 dollars).[20]

Although the price of gold was held steady at $35 per ounce because of the fixed exchange rate regime, it was exceedingly costly to maintain the mint parity or defend the dollar at that rate. The U.S. capital account (the sum of all long-run, private economic transactions by U.S. residents with other nations for a given time period) was in deficit in almost every year since 1950. From $1billion per year between 1950 and 1957, it rose to $3 billion per year from 1958 to 1970, and the U.S. financed its deficits from 1950 to 1970 with a $13 billion loss of its gold reserves, equivalent to a decline in gold reserves of about 54 percent.[21] For the duration of the Korean War, the price of gold fluctuated between $35.5 and $40.25 on the New York Mercantile Exchange.

The Vietnam War was twice as costly as the Korean War ($670 billion, in FY2007 dollars),[22] and by 1968 the signs of pending inconvertibility were already evident. The surplus in the current account (the account which records transactions in goods, services, remittances, and income) deteriorated from $8.5 billion in 1964 to $4.8 billion in 1967. This coincided with the rapid expansion of U.S. military commitments in Southeast Asia and U.S. government imposed capital controls that restricted foreign investment by U.S. residents and reduced asset earnings abroad.[23]

The resulting outflow of dollars and gold increased destabilizing speculation and doubts about the ability of the United States to sustain convertibility. On 8 August

1971 newspapers reported that the French were about to present $191 million of reserves to the United States in exchange for gold to make loan repayment to the IMF.[24] Speculation against the dollar intensified and gold flowed out of the country on a daily basis. On 15 August 1971 President Nixon made it known that the United States was no longer committed to exchange dollars for gold; he in effect suspended convertibility. The Smithsonian Agreement of December 1971 was an attempt to remodel the beleaguered international monetary system of the late 1960s and early 1970s, and it formalized the devaluation of the dollar by 9 percent.

Beyond Vietnam, the United States has been involved in two major wars, the Persian Gulf War of 1990-1991, which can be classified as a war of collective self-defense in international law, and the ongoing Afghan and Iraq wars which started in 2001 and 2003, respectively. The Gulf War showed no detrimental impact on the U.S. currency, although it had of course economic and human costs. Data from the Bank of England show that the annual average price per ounce of gold for 1990 and 1991 was $383.57 and $362.10, respectively. In fact, gold depreciated against the dollar. The reason for this is that the United States did not have to print or borrow a significant amount of money to finance the cost of that brief war. The CRS indicates that the Gulf War cost $94 billion (in 2007 dollars), of which the United States paid only $7 billion. The rest of the money was paid by Saudi Arabia, Kuwait, and other countries.[25] Unlike the Gulf War or earlier wars, financing the Iraq War has been a very expensive endeavor, with consequences for the value of the dollar. As of this writing, it is estimated that the cost of the war, excluding Afghanistan, is likely to reach $2.7 trillion.[26]

Salient lessons show that the cost of war is realistically inestimable. Much of the financial outcomes of wars depend on tactical strategies of belligerents (e.g., guerilla versus conventional warfare), post-war welfare remediation, and contingent economic shocks. Escalating oil prices and U.S. budget deficits resulted in stagflation in the 1970s, and of a milder form of it in 2008, partly because war-time planning has not generally considered the prospects of economic shocks adequately — for example, inflation, unemployment, or mortgage crises. For all the wars under consideration there has been a corresponding downturn in economic activity: July 1953 through May 1954, December 1969 through November 1970, July 1990 through March 1991, and arguably October 2007 to the time of writing.

The Box-Jenkins method

The forecasting method employed in this article provides a projection of the dollar's value in terms of gold during the Iraq War. It thus assesses the sustainability of the dollar as a dominant international currency. The dollar price per ounce of gold from January 2000 to March 2008 was obtained from the Bank of England, providing a total of ninety-nine observations.

Gold is a good measure of the value of the dollar because it is an alternative form of money that is universally acceptable, and it is generally used to hedge against the adverse effects of inflation or the depreciating value of a currency. It captures speculative behavior and it is also a good numeraire to calculate cross-exchange rates.

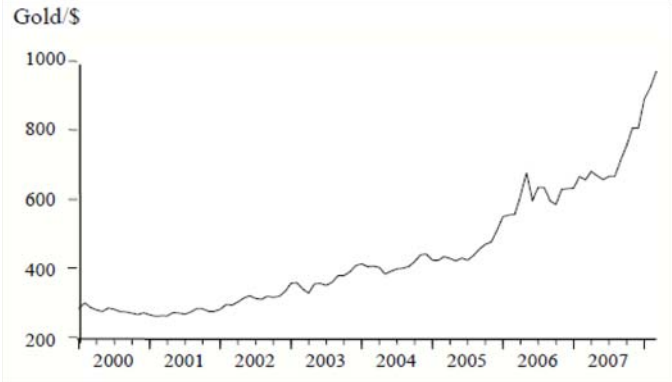

Figure 1: Monthly dollar price per ounce of gold (January 2000-March 2008)
Source: Bank of England

The Box-Jenkins (BJ) method is the preferred short-term forecasting tool used here. It enables one to evaluate the future direction of the dollar through stationarity and avoids the endogeneity problem associated with monetary exchange rate models.[27] There are four important procedures to implementing the BJ algorithm. First, a visual analysis of the data to determine if there is a deterministic or seasonal trend; second, ensuring stationarity; third, diagnostic testing; and fourth, forecasting.

A visual representation of the dollar per ounce of gold exchange rate is provided in Figure 1. The illustration shows an upward deterministic trend, suggesting dollar depreciation (more dollars per ounce of gold) for the period under consideration. To speculate on the future of this trend it is tempting to mentally extend the curve upwards. But this is misleading unless knowledge of the statistical or random (stochastic) mechanism responsible for the trend is available (the data generating process, or DGP). To obtain knowledge of the DGP, it is essential to difference the data to remove the trend and make the series stationary. To achieve stationarity, we difference the log of the exchange rate to eliminate larger variances over time (heteroskedasticity). This also permits analysis of dollar depreciation in terms of percentage changes.

Figure 2 provides visual evidence that the deterministic trend shown in Figure 1 has been removed to obtain a constant mean of zero and unit variance. This second requirement of the BJ algorithm must be evaluated to ascertain successful stationarity because of the potential problem of underdifferencing or overdifferencing the data. Statistically, stationarity is ascertained by diagnostic testing or visual analysis of the autocorrelation (AC) and partial autocorrelation functions (PACF), the correlogram and the Q-statistic.[28]

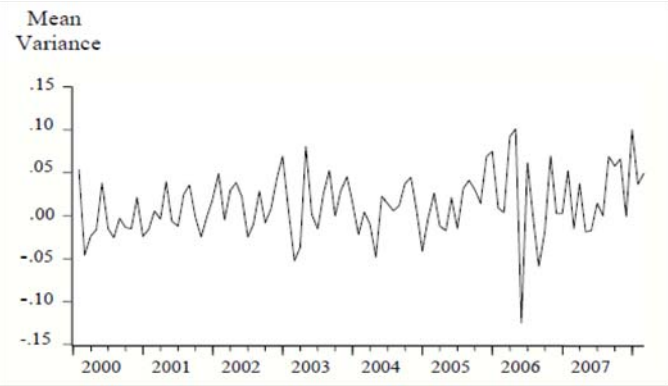

Figure 2: Mean and variance of the differenced/stationary gold/dollar series

Information from the diagnostic test is used to model the data generating process to make a reasonable projection (in this case, a static forecast) of the direction of the dollar's value in terms of the price of gold. The correlogram is a guide which indicates when autoregressive (AR) or moving average (MA) terms must be added or reduced, depending on the significance of the lags. The PACF suggests the required number of AR-terms because it acts like a partial difference, while the ACF suggests the required number of MA-terms to dampen the amount of unnecessary differencing.

After differencing the data and evaluating the correlogram and the Q-statistics, 7.44 and 12.27 at lags three and four, respectively, the corresponding AC-terms are found to be significant with p-values of 0.059 and 0.015. The following forecasting model is specified:

$$(1) \qquad r_t = \mu + \theta e_t + \theta e_{t-3} + \theta e_{t-4}$$

where r_t, is the gold-U.S. dollar exchange rate at the end of March 2008, μ captures the average percentage monthly depreciation of the dollar, θ is the parameter of random shocks to be estimated, and *e* refers to the random shocks associated with the exchange rate in war time. Shocks in the past three and four months prove to be very significant. These coincide with the fuel and mortgage crises. In ARIMA notation, the model can be expressed as (0, 1, 4) which means, first, a zero autoregressive term, second, that the data is differenced once, and third, that the model is a moving average of the fourth order. The probable autoregressive terms are not found to be significant. The expression

$$(2) \qquad \theta e_t = r_t - \mu - \theta e_{t-3} - \theta e_{t-4}$$

reflects the continuing assessment of past information on current decisions and its solution is normally referred to as the invertibility condition.[29] An ARIMA model generates a solution for analysis if it is considered to be stable or invertible; otherwise it becomes explosive and unreliable.

Table 1: Gold-U.S. dollar exchange rate January 2000-March 2008, Ordinary Least Squares Estimates (t-stat in parenthesis)*

Diagnostics	*Value*	*Variables*	*Coefficient*
Akaike criterion	-3.83	Mean	0.012 (2.41)
Bias proportion	0.002	MA(3)	0.23 (2.43)
Variance proportion	0.04	MA(4)	0.28 (2.87)
Covariance proportion	0.96	...	...
Inverted MA roots	0.52 + 62i	...	...
	0.52 - 62i	...	...
	-0.52 - 0.40i	...	...
	-0.52 + 0.40i	...	...

* Estimates done with E-Views v5.1

Empirical findings

The overall results of the model and diagnostics are provided in Table 1. They indicate that for the period under consideration, the value of the dollar, measured in terms of an ounce of gold, depreciated at a monthly average rate of 1.2 percent during the Iraq War. The moving average coefficients (0.23 and 0.28) are significant at the 95 percent level of confidence and the null hypotheses that they are zero are rejected.

The Akaike information criterion, or AIC, of -3.83 is used to ascertain the required number of regressors (parsimony) and the performance of competing models. It penalizes for excessive regressors, and smaller values are generally preferred.

The bias, variance, and covariance proportions are used to evaluate the forecasting model because the root mean square error is sensitive to the units in which the relevant variable (i.e., the exchange rate) is estimated. The bias proportion of 0.002 indicates how far the mean of the forecast is from the mean of the actual series, the variance proportion of 0.04 indicates how far the variation of the forecast is from the variation of the actual series, and the covariance proportion of 0.96 is a measure of the remaining unsystematic forecasting errors. The forecast results shows that the bias and variance proportions are very low. These conditions are preferred for a better fit. The inverted MA roots show that the model is stable (can be solved) and not explosive.

The forecast of the dollar value is provided in Figure 3. The static forecast shows short-term depreciation of the dollar (more dollars per ounce of gold) with the commencement of the Afghan and then Iraq wars. The general trend of this

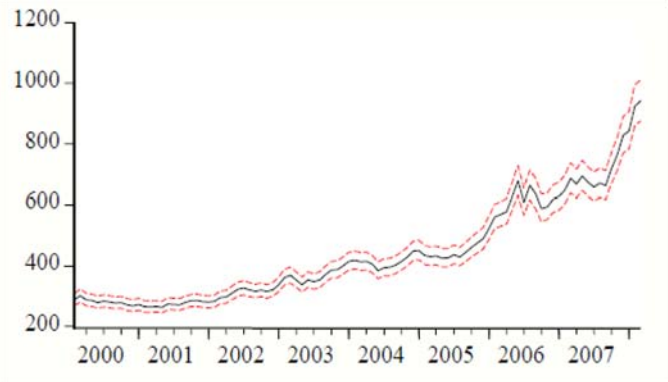

Figure 3: Static (short-term) forecast of gold-U.S. dollar exchange rate

depreciation shows no difference whether or not actual values are inserted for out of sample observations. The dynamic forecast (not shown) also results in a short-term depreciation of the dollar but is prone to a higher margin of bias and error.

Conclusion and policy implications

The empirical results give credence to the argument that the currencies of nations at war come under pressure during times of war. In the case of the United States this has been the case since the Korean War when the United States lost a significant amount of its gold reserves. The continued loss of reserves as a result of the Vietnam War led to destabilizing speculation, suspension of convertibility, and depreciation of the dollar.

The cost of war has historically been miscalculated not only in terms of treasure, but also in terms of strategy and contingent economic shocks like inflation, oil prices, or mortgage crises. Vietnam and Iraq support the notion that the costs of war are inestimable and unpredictable when enemies choose guerilla warfare as their preferred strategy of warfare. When economic shocks conflate with war-time expenditure, geopolitical anxiety, and competition from optimum currency areas, the dominance of the dollar is endangered. It should be recalled that after Vietnam the dominance of the dollar was insured only by geopolitical alliances of the 1970s and 1980s.

All the major wars of the twentieth and twenty-first centuries have been accompanied by recessions, and at times stagflation. These, and bloated budget deficits, cause long-term economic disruptions that encourage the search for alternative forms of money and call into question the continued dominance of the dollar.

War-time animosity against the United States and anxiety over the value of the dollar encourages nations to consider dislodging the dollar from its position of preeminence, as evidenced for example by Russia and Iran in 2007: Russia's attempt to replace the dollar for rubles in its financial markets, and Iran trying to persuade OPEC to accept euros rather than dollars for oil to undo the arrangements of the 1970s and 1980s. The Asian Clearing Union, headquartered in Tehran, is also considering the inclusion of the euro to compete with the dollar at a time of weakness. There is no evidence that these measures can be successful in the short-run, but they challenge the once dominant position of the dollar and portend long-run destabilization associated with protracted war.

All the major wars of the twentieth and twenty-first centuries have been accompanied by recessions, and at times stagflation. These, and bloated budget deficits, cause long-term economic disruptions that encourage the search for alternative forms of money and call into question the continued dominance of the dollar. War as an instrument of foreign policy endangers the vitality of the dollar, and its continued dominance under such conditions is highly precarious.

Notes

Christopher E.S. Warburton is Assistant Professor of Economics at the John Jay College of Criminal Justice, New York, U.S.A. He may be reached at Cwarburton@jjay.cuny.edu.

1. Krueger (1983).

2. Meese and Rogoff (1983, pp. 3-24).

3. Krugman (1993).

4. Bacchetta and Wincoop (2006).

5. Mark (1995).

6. Johnston and Sun (1997).

7. They examine in-sample fit, but emphasize the importance of the monetary policy rule and its effects on expectations in determining exchange rates. They then present evidence that exchange rates incorporate news about future macroeconomic fundamentals and demonstrate that the models might well be able to account for observed exchange-rate volatility. See Engel, Mark, and West (2007).

8. Engel and West (2005, pp. 485-517).

9. Johnston and DiNardo (1997, pp. 204-206). ARMA models, unlike single equation or simultaneous equation models, analyze the probabilistic properties of economic time series. In this case the dollar value of gold is explained in terms of its past values and shocks. See Gujarati (2003).

10. Mishkin (1992); Obstfeld (1994); Dornbusch, Goldfajn, and Valdes (1995); Sachs, Tornell, and Velasco (1995); Eichengreen, Rose, and Wyplosz (1995); Frankel and Rose (1996); Krugman (1998); Cooper (1999); Salvatore (1999).

11. Optimum currency areas allude to monetary integration of nations and the use of a single currency against others, for example the euro. With permanently fixed exchange rates, an optimum currency area is likely to generate price stability, unlike with flexible rates among the member nations. Greater price stability is attributable to the cancellation or minimization of random shocks within the economies of member states, which make the currency a relatively reliable store of value. See Salvatore (2006).

12. Mundell (1961).

13. Mundell (1998).

14. Ricci (2007); Mongelli (2008).

15. A country is fully dollarized when it abandons its local currency as an official policy to adopt a much more stable international currency. The concept was originally explained in terms of the adoption of the U.S. dollar. Today, it is loosely applied to the adoption of other international currencies. De facto dollarization occurs when residents of a nation prefer to hold international currencies in order to conduct domestic and international business, although the local government has not repudiated the local currency.

16. Smith (2005, pp. 66-71).

17. Salvatore (2001).

18. Spiro (1999); Essayad and Marx (2001).

19. Mussa (1981).

20. Stiglitz and Bilmes (2008, p. 240).

21. Salvatore (2001, p. 755).

22. Stiglitz and Bilmes (2008, p. 240).

23. Dudley and Passell (1968, p. 437).

24. Daniels and Vanhoose (2005, p.81).

25. Stiglitz and Bilmes (2008, p.236).

26. Stiglitz and Bilmes (2008, p. 33).

27. Stationarity says that the mean and variance of the series are constant through time; the mean and variance of intervening periods will be identical. This facilitates the projection of the series in the short-run.

28. Autocorrelation is the correlation between members of a time series. The sample autocorrelation is the ratio of the sample covariance (γk) to sample variance (γ) and the sample autocorrelation function at a particular lag (ρ) is (γk)/(γ). The value of the autocorrelation function always lies between -1 and +1 and the correlogram is the visual (graphical) representation of the autocorrelation and partial autocorrelation functions. The partial autocorrelation is the correlation between two lags at different periods when intervening lags are not considered; that is the amount of correlation between a variable and its lag that is not explained by correlations at all lower-order lags. For a time series to be considered stationary, its autocorrelations at various lags must hover around zero and the null hypothesis that the autocorrelation values are not jointly zero must be rejected. The Q-statistic, developed by Box and Pierce, is the test statistic used here. It tests for a pure random or white noise process.

29. An infinite MA model is invertible if for $A(L)= 1- \alpha_1 L - \alpha_2 L^2 \ldots \alpha_p L^p$ a solution of the form $1/(1- \alpha_1 L) + 1/(1-\alpha_2 L^2) \ldots + 1/(1-\alpha_p L^p)$ can be obtained; where L is a lag operator indicating the number of lags and α is a parameter to be estimated and presumed to be less than 1.

References

Bacchetta, P. and E. Wincoop. 2006. “Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?” *American Economic Review*, Vol. 96, pp. 552-576.

Cooper, R. 1999. “The Asian Crisis: Causes and Consequences,” pp.17-28 in A. Harwood, R. Litan, and M. Pomerleano, eds. *Financial Markets and Development: The Crisis in Emerging Markets*. Washington, DC: Brookings Institution Press.

Daniels, J.P. and D.D. Vanhoose. 2005. *International Monetary and Financial*

Economics. Mason, OH: Thomson South-Western.
Dornbusch, R., I. Goldfajn, and R.O. Valdes. 1995. "Currency Crises and Collapses." *Brookings Papers on Economic Activity*, Vol. 2, pp. 219-270.
Dudley, L. and P. Passell. 1968. "The War in Vietnam and the United States Balance of Payments." *The Review of Economics and Statistics*, Vol. 50, No. 4, pp. 437-442.
Eichengreen, B., A. Rose, and C. Wyplosz. 1995. "Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks." *Economic Policy*, Vol. 21, pp. 249-312.
Engel, C. and K.D. West. 2004. "Accounting for Exchange Rate Variability in Present Value Models when the Discount Factor is Near One." NBER Working Paper No. 10267. Cambridge, MA: National Bureau of Economic Research.
Engel, C. and K.D. West. 2005. "Exchange Rates and Fundamentals." *Journal of Political Economy*, Vol. 113, pp. 485-517.
Engel, C., N. Mark, and K.D. West. 2007. "Exchange Rate Models Are Not as Bad as You Think." NBER Working Paper No. 13318. Cambridge, MA: National Bureau of Economic Research.
Essayad, M. and D. Marx. 2001. "OPEC and Optimal Currency Portfolios." *Oil, Gas, and Energy Quarterly*, Vol. 49, No. 2, pp. 363-384.
Frankel, J.A. and A.K. Rose. 1996. "Currency Crashes in Emerging Markets: An Empirical Treatment." *Journal of International Economics*, Vol. 41, pp. 351-366.
Gujarati, D. 2003. *Basic Econometrics*. New York: McGraw-Hill.
Johnston, B. and Y. Sun. 1997. "Some Evidence on Exchange Rate Determination." IMF Working Paper No. 97/98. Washington, DC: International Monetary Fund [accessed 21 April 2008].
Johnston, J. and J. DiNardo. 1997. *Econometric Methods*. New York: McGraw-Hill.
Krueger, A. 1983. *Exchange Rate Determination*. Cambridge, MA: Cambridge University Press.
Krugman, P. 1993. "Recent Thinking About Exchange Rate Determination and Policy." RBA Annual Conference Volume No. Acv1993-02 [Australia]. http://www.rba.gov.au/
Krugman, P. 1998. "What happened to Asia?" Mimeo. Cambridge, MA: Massachusetts Institute of Technology.
Mark, N.C. 1995. "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability." *American Economic Review*, No. 85, pp. 201-218.
Meese, R.A. and K.S. Rogoff. 1983. "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*. Vol. 14, pp. 3-24.
Mishkin, F. 1992. "Anatomy of a Financial Crisis." *Journal of Evolutionary Economics*, Vol. 2, pp. 115-130.
Mongelli, F.P. 2008. "European Economic and Monetary Integration and the Optimum Currency Area Theory." Economic Papers 302. Brussels, Belgium: European Commission.
Mundell, R.A. 1961. "A Theory of Optimum Currency Areas." *American Economic Review*, No. 51, pp. 509-517.
Mussa, M. 1981. "Sticky Individual Prices and the Dynamics of the General Price Level." Carnegie-Rochester Conference on Public Policy, Vol. 15, pp. 261-296.
Obstfeld, M. 1994. "The Logic of Currency Crises." *Cahiers Economiques et Monetaires*, Vol. 43, pp. 189-213.
Pindyck, R.S. and D.L. Rubinfeld. 1998. *Econometric Models and Econometric Forecasts*. New York: McGraw-Hill.
Ricci, L.A. 2008. "A Model of an Optimum Currency Area. *Economics*, Vol. 2, No. 2008-8, pp. 1-31.
Sachs, J., A. Tornell, and A.Velasco. 1996. "Financial Crises in Emerging Markets: the Lessons from 1995." *Brookings Papers on Economic Activity*, Vol. 1, pp. 147-215.
Salvatore, D. 2001. *International Economics*. New York: Wiley.
Salvatore, D. 1999. "Lessons from the Financial Crisis in East Asia." *Journal of Policy Modeling*, Vol. 21, No. 3, pp. 283-287.
Spiro, D.E. 1999. *The Hidden Hand of American Hegemony: Petrodollar Recycling and International Markets*. Ithaca: Cornell University Press.
Stiglitz, J.E. and L.J. Bilmes. 2008. *The Three Trillion Dollar War: The True Cost of The Iraq Conflict*. New York: Norton.

The democratic peace proposition: an agenda for critical analysis

Steve Chan

The proposition that democracies are more peaceful than nondemocracies in their international relations has gained a great deal of attention in academic as well as policy circles. Among scholars of international relations, this topic has arguably been the most popular subject of analysis in recent years. Its popularity reflects the considerable amount of empirical evidence pointing to an absence of war between democracies. Commenting on the democratic peace proposition, one scholar suggested that it came "as close as anything we have to an empirical law in international relations."[1] Echoing the same sentiment, another remarked that this phenomenon represented "one of the strongest nontrivial or nontautological generalization that can be made about international relations."[2] Among policymakers, U.S. President Bill Clinton propounded "democratic enlargement" to promote global peace. His successor, President George W. Bush, justified the invasion of Iraq in part on the grounds of introducing democracy to Iraq and the Middle East more generally, with the intent thereby of making this region more stable and peaceful.

The proposition that democracies are more peaceful than nondemocracies in their international relations has gained a great deal of attention in academic and policy circles. But is the proposition wholly warranted? This essay raises some conceptual, methodological, and causal difficulties of the proposition and points to resulting policy issues.

The discourse on democratic peace has reflected the tradition of international liberalism,[3] with its intellectual origin traced to U.S. President Woodrow Wilson and even the German philosopher Immanuel Kant.[4] It has encountered skeptical reaction from those adhering to the realist tradition. The realists argue that all states are compelled by the structure of international relations to pursue power and seek survival.[5] This systemic pressure, according to realists, implies that all states regardless of their regime characteristics (that is, regardless of whether they have a democratic or autocratic regime) are likely to behave similarly.

This essay pursues a different tack. In the spirit of constructive criticism, it takes on the literature on democratic peace on its own terms. This literature has been characterized primarily by statistical analyses using large datasets on states' regime characteristics and their conflict behavior. The intent is to suggest fruitful areas for further research in order to clarify, confirm, and extend the democratic peace proposition. It is not to undertake a comprehensive review of the relevant literature consisting of dozens of books and hundreds of articles,[6] but rather to point out areas of conceptual confusion and problematic inference that require attention. The remainder of this essay is organized into four sections, addressing, respectively, (1) conceptual, (2) methodological, (3) causal, and (4) policy issues pertaining to the democratic peace proposition.

Conceptual confusion

It is natural to begin by asking which ideas analysts have in mind when they speak of democracy and peace. Depending on one's conceptualization and operationalization of these terms, the democratic peace proposition can mean very different things. When one goes back to the supposed intellectual origins of this proposition, one is confronted with a discrepancy whereby popular sovereignty or majority rule was not necessarily the predominant characteristic hypothesized to cause a regime's peaceful orientation.

For instance, Immanuel Kant spoke of *Rechtstaat* and emphasized the ideas of republicanism and rule of law in his treatise on "perpetual peace."[7] A constitutional monarchy, such as that of the U.K., can offer an example *par excellence* for the rule of law — defined by the existence of an impartial and independent judiciary, legal transparency and equality, and respect for due process — but it is clearly not a republic. Conversely, Bismarkian Prussia as well as contemporary Singapore can exemplify the "administrative state," even though they may not quite meet some people's criterion of electoral competition and popular representation.

Some analysts have also correctly pointed out that Kant did not quite have popular rule in mind when he wrote about a possible relationship between a regime's characteristics and its foreign conduct.[8] Libertarianism rather than popular rule was a more important consideration for him and others. By libertarianism, one usually means personal freedom and individual liberty, such as the right to free expression and assembly. Yet public debate continues about whether abortion, divorce, or same-sex marriages are a matter of personal rights. Some countries restricted these rights until the recent past, and others continue to do so. There is clearly considerable historical variation even among the established democracies — and within these countries in the extent to which personal freedoms have been accorded to religious and racial minorities.

Naturally, libertarianism and popular rule do not always go hand in hand such that Victorian Britain and contemporary Hong Kong offer considerable personal liberties although not universal suffrage required for popular rule. A question arises, however, when some definition of libertarianism includes the right to private property and institutions of free enterprise.[9] On this dimension, Hong Kong can be rated more liberal than Britain and France (which have public enterprises), not to mention all the socialist states. With the adoption of a capitalist economy, countries such as China,

Russia, and Vietnam would presumably receive higher scores on this dimension even though democracy and capitalism are not synonymous in most people's mind.

Still, a different and presumably relevant definition of democracy, as alluded to already, can emphasize universal adult suffrage, electoral competition, alternation of power, and especially majority rule. These considerations would call into question much of the extant analyses that defined the United States, the United Kingdom, and other European countries as democracies before 1914 because except for four countries, female suffrage — not to mention actual voting rights for African Americans — was not instituted well afterwards. Moreover, one may well question the extent of electoral competition and political turnover for countries that have been (subjectively) rated as democracies, but ones that have had long stretches of single-party rule (Japan, Mexico, Italy, India, and even Sweden). At the same time, if majority rule is *the* defining character of democracy, Britain, Australia, New Zealand, and most recently, the United States (when Al Gore lost to George W. Bush) have all seen the election of a candidate or party which received fewer popular votes than its opposition. Conversely, current regimes in Venezuela, Iran, and the Palestinian territories (the Hamas) have been elected and arguably represent popular rule. Finally, peaceful regime changes or turnover of the office of the chief executive between the leaders of competing political parties has been suggested to offer another distinctive mark of a democracy.[10] The extent of such alteration or alternation, however, can be a matter of some controversy and naturally pertains to arguments about, for example, whether the August 2008 war between Russia and Georgia indicates any implications for the democratic peace proposition.

Much of the quantitative literature on democratic peace has relied on the Polity dataset, specifically on composite scores tallying disparate indicators such as constraints on executive power, competitiveness in selecting the chief executive, and openness of this process. The coders of this dataset rated (subjectively) each of the latter dimensions, and these individual ratings are usually summed by analysts to produce an aggregate measure of a country's "democraticness" or "autocraticness." This practice can often conceal more than reveal important differences,[11] because the same aggregate score can reflect very different underlying traits. It is as if a physician simply looks at a patient's body weight without considering this person's age, gender and physical stature (e.g., a short, fat person compared to a tall, thin person). As for the Freedom House's rating of freedom in different countries, another popular source of data, it is worthwhile to check whether more objective indicators of democracy, such as those offered by Vanhanen,[12] would produce convergent empirical results. Some scholars have sought to integrate different dimensions of democracy from several datasets in order to develop a more appropriate measure of this concept.[13]

Turning to the "dependent variable," one may ask whether one can study "peace" by counting the frequency of militarized interstate disputes. Recent work on the democratic peace proposition has focused overwhelmingly on these disputes rather than the occurrence of war per se. Significantly, other types of foreign conduct — such as covert subversion, economic sanction, and military intervention — are not necessarily reflected by the frequency count of militarized disputes.[14] Moreover, when a small country — say, Grenada, Panama, the Dominican Republic — is quickly overwhelmed in a military confrontation with a big country, battle fatalities never reach the 1,000 threshold of combat casualty customarily required for defining a war. Does this mean that a "war" did not happen, or just that due to the power disparity between the contestants, the weak was quickly subdued by the strong? Such "non-wars" reflect more the strength of the powerful than its peaceful disposition.

Significantly, the definition of militarized disputes and interstate wars used by most quantitative researchers of the democratic peace proposition systematically biases the results because it rules out from their analyses so-called extra-systemic conflicts. They are extra-systemic because, before the formation of the United Nations, in order to be recognized as a member of the interstate system, a regime must be diplomatically recognized by Britain and France in particular. Accordingly, all the imperial and colonial wars waged by these two "democratic" countries in the nineteenth century would not count against the monadic version of the democratic peace proposition because the other side of these conflicts did not technically qualify for statehood.[15]

Finally, should peace be defined simply as the absence of militarized disputes or of war? Should one consider a definition of "positive peace" that includes human well-being as well as a state's internal social conditions? Such a conception of peace offers a very different approach to conflict study. The Global Peace Index exemplifies this alternative vision and, if adopted by researchers of democratic peace, is likely to produce considerably different empirical patterns from those reported by the current literature.[16]

Methodological concerns

Wars are rare events — whether they are undertaken by democracies or autocracies. Indeed, there appears to be an "autocratic peace" in addition to "democratic peace."[17] On top of the analytic challenge introduced by this rarity, causal attribution of a state's peaceful foreign conduct to its democratic regime is complicated by the problem of multicollinearity. Democracy tends to be statistically associated with other socioeconomic and international variables, such as a country's wealth and industrial development, its alliance portfolio, its involvement in foreign trade and interstate organizations, and its status as a former colony. This multicollinearity invariably confounds attempts to discriminate among the influence of various candidate (independent) variables on a country's propensity to get into militarized disputes or wars. It is therefore not surprising that some scholars have questioned whether the democratic peace proposition is spurious and have suggested that shared capitalism, U.S. hegemony, or settled borders are actually responsible for the observed association.[18] Others have pointed out that the relationship between democracy and

peace is a two-way street. Whereas democracy may promote peace, countries located in a peaceful region (e.g., Scandinavia) are more likely to develop democratic institutions in the first place.[19]

Temporal and spatial interdependencies are another source of concern.[20] How helpful, or valid, is it to treat World War I, say, as dozens of bilateral conflicts (each taken to represent a separate independent observation), so that one records it as separate wars between Britain and Germany, France and Germany, Russia and Germany, and Belgium and Germany, and so on and so forth? This treatment seems to fly in the face of our historical knowledge. We know that Britain went to war against Germany because Germany had invaded Belgium, and that German invasion of Belgium (the Schlieffen plan) had something to do with Berlin's concern about fighting a two-front war against France and Russia simultaneously. And, of course, the initial impetus for the war in August 1914 came from a dispute between Serbia (a Russian ally) and Austria-Hungary (a German ally). There was therefore diffusion of conflict across borders, causing a chain effect that brought these belligerents to a military showdown.

There is also temporal dependency to consider. Some historians have viewed World War I and World War II as a continuous conflict with but a short interval of peace. Similarly, one can plausibly argue that the genesis of the U.S. and U.K. invasion of Iraq in 2003 can be traced back to the First Gulf War in 1989 when these countries and others confronted Baghdad over its invasion of Kuwait. As most dyads manage to avoid disputes year after year and because some conflicts' recurrences are interrelated, could the democratic peace phenomenon be shown to exist only because the same observations are compounded over and over?[21]

It is also pertinent to ask whether these observations have been compounded by the inclusion of marginal actors in the analysis and treating them as if they are equivalent to the major characters. For instance, does it make sense to give equal weight to the United States and Germany as a conflict dyad in World War II as to Brazil and Germany and to the USSR and Japan (the Soviet Union did not declare war on Japan until a few days before Japan's surrender)? Surely, these observations refer to very different substantive phenomena. Some (e.g., Brazil, Burma, Bolivia) are bit characters while others are major belligerents. Some states' participation in the war was minor and even only symbolic, whereas it involved a major undertaking for others. To take another example, it would be quite implausible for Poland, Georgia, or even Britain to join the "coalition of the willing" independent of a U.S. decision to invade Iraq. Similarly, Belgium, Greece, Colombia, and Turkey would hardly have fought in the Korean War had it not been for the U.S. intervention in this conflict. It seems important to distinguish these "apples" and "oranges" rather than simply lumping them as "fruits." The purchase of a larger number of observations may come at the expense of analytic precision and conceptual clarity.

Causal attributions

Space does not allow a full discussion of the various explanations that have been advanced for the democratic peace proposition. One especially cogent explanation, however, focuses on audience costs and popular sanction.[22] The basic idea is that democratic leaders will suffer negative political consequences if they fail to honor their declared policies (or if they encounter policy failures). Voters, the media, and the political opposition will punish them for this behavior. Therefore, democratic leaders face greater domestic audience costs — a tendency that in turn implies that when these leaders do make foreign policy threats, these threats are more credible. The logic of this implication, however, suggests that this greater credibility on the part of democratic leaders would incline authoritarian leaders to be more disposed to back down from a confrontation with them. As well, because authoritarian leaders do not face the same prospect of a domestic backlash as the democratic leaders, their threats are inherently less credible. They are more likely to be "pushed around" because they would not suffer the same domestic consequences if they renege on their policy or suffer foreign policy setback.[23] It points to authoritarian leaders' concessions as a reason why confrontations are sometimes avoided.[24] Note also that the attribution of domestic audience costs simply argues that democratic leaders are more prudent about taking on foreign conflicts due to their concerns about adverse domestic repercussions. It does not claim that these leaders are necessarily more peaceful in their disposition, only that they are more risk averse and require a higher threshold of expected success before they start a militarized dispute or go to war.

Another explanation of the democratic peace proposition focuses on structural or institutional features, specifically the fact that in democracies power is not concentrated in the hands of one person or institution. This sharing of power and the existence of many veto and oversight groups suggest that democracies must build a strong consensus before going to war. Some scholars have questioned how well institutional checks and balances have worked in the lead up to the U.S. invasion of Iraq. An assertive executive, an acquiescent legislature, uncritical media, and a generally passive public raise concerns about the institutional restraints on democracies' decision to start war (as opposed to fighting back after being attacked such as in the wake of Pearl Harbor).[25] Significantly, at least some of the hypothesized institutional mechanisms that are supposed to foster responsible government and checks against a rush to war are peculiar to specific democracies. For instance, in parliamentary systems the distinction between executive and legislative branches is irrelevant. Similarly, to the extent that opposition parties can serve as a watchdog, it is important to recognize differences among countries that have one dominant party, two competitive parties, and multiparty competition and coalition governments. The differentiation of these different types of democratic systems and the study of their different propensity to undertake war is a fruitful area for future research.[26] Note again that the institutional explanation of democratic peace simply argues that democratic

leaders face more domestic constraints in order to take their country to war; it does not claim that these leaders are inherently less bellicose in their nature.

The argument of institutional constraints naturally raises the question of whether authoritarian leaders may under certain circumstances face constraints not dissimilar from their democratic counterparts. For example, can authoritarian leaders also be punished by their domestic constituents and suffer audience costs if not from the media and voters, then from their political rivals? It is true that democratic leaders are accountable to a larger constituency and therefore more likely to represent the public good.[27] At the same time, just as not all democracies are alike (e.g., parliamentary versus presidential systems, two-party versus multiparty systems), not all autocracies are alike. Recent research shows that some authoritarian leaders can also suffer from domestic audience costs such as when they do not control their secret police or have exclusive power to make political appointments.[28] Therefore, more in-depth analysis of specific institutional mechanisms is necessary in order to understand better the propensity of different types of regimes to undertake war.

Yet another explanation of democratic peace has been offered with a focus on the norms and attitudes of democracies' mass public.[29] It is proposed that people living in democracies have been socialized to share a tolerant and pacific ethos that they extend to their relations with foreign countries, at least other countries that are democratic. This public opinion, however, has not restrained democracies from waging or considering preventive war against others, and even fighting against their own kind (e.g., the American Civil War, the Boer War, the War of American Independence) if one is willing to stretch the concept of democracy and go back in time. In fact, public opinion often rewards leaders for successful bellicosity (the so-called rally-behind-the-flag syndrome). Recent research has taken on the interesting question of whether public opinion turns against a war because of the perceived failure of an administration's policy (such as in Iraq) or the mounting costs (especially casualties) of this endeavor.[30] Research on public opinion suggests that the American people may be characterized as prudent but hardly pacific.[31]

Naturally, public opinion influences political accountability and official policies when people can cast ballots to express their preference. Electoral institutions and party systems, however, can intervene to impede or distort this expression.[32] When both major parties in a two-party system agree on a policy — as in Britain and the United States on going to war against Iraq in 2003 — voters may be denied a meaningful choice. Even when most voters are opposed to waging war as in Britain, public opinion may not be enough to restrain an incumbent government. Casting a vote for a minor party is discouraged because of the single-member district, plurality rule, which favors the two dominant parties. Conversely, popular sentiments are more likely to be reflected, and less likely to be distorted, when a country has a multi-member district, proportionate representation of election. Other institutional factors, such as term limitation of the length that an official can serve (the U.S. president is limited to two terms), surely also affect the prospect of audience costs and voter sanction. The U.S. presidential race is usually decided by elections in a few pivotal or battleground states. As a consequence, public opinion in some areas of the country is more salient for candidates than elsewhere. Due to the nature of the Electoral College, voters in some states are accorded more voice than in other states, and this electoral disproportionality tends to favor the more rural and conservative states with a smaller and less diverse population.

Policy implications

Until 2003, the democratic peace proposition has generally been taken to mean that democracies are unlikely to initiate war. This does not mean that democracies will not fight autocracies, but rather that they will not undertake war unless faced with imminent danger and given just cause.[33] This expectation suggests that democracies will not wage preventive wars when they are not faced with the danger of being attacked, and in violation of international norms recognizing other states' sovereignty and territorial integrity. The Iraq war therefore presents a critical challenge to the democratic peace proposition because it seemingly challenges both the institutional and normative explanations advanced for this proposition. Two leading democratic states, the United States and the United Kingdom, went to war in defiance of international norms and domestic opposition. The hypothesized institutional and normative constraints against initiating war did not appear to have worked or worked with great effect. It also appears that once in war, these democracies have a hard time extricating themselves from a protracted and unpopular conflict.

The Iraq war has increased interest in the debate about whether military intervention can foster democracy abroad. This burgeoning literature shows that the evidence bearing on this question is far from clear or decisive.[34] By its very nature, this research calls for counterfactual reasoning, posing the question of what would have been the political and economic conditions of a country had foreign intervention not happened. There appears to be some evidence that a hasty attempt to introduce capitalist democracy in a war-torn society can actually have a deleterious, or counterproductive, effect.[35] Moreover, the pre-existing socioeconomic conditions of the target of intervention would make a significant difference. That Germany and Japan, for example, had had experience with electoral competition at some point in their history prior to the U.S. occupation and that they had had an educated population and an industrial economy before 1945 augured well for their democratic prospects. Conversely, the chances for democracy taking roots in, say, Iraq and Afghanistan are much more problematic.

Recent scholarship has offered in-depth historical analysis of the preventive motivation. It has shown that democratic leaders are quite capable of contemplating the possibility of starting a preventive war.[36] Israel launched attacks against Iraqi and Syrian nuclear facilities, and the United States came very close to launching similar attacks against North Korea, China, and Cuba (against the Soviet missiles placed on

that island during the Cuban Missile Crisis). Similarly, Britain and France considered joint action against the looming threat from Nazi Germany in the 1930s. This and other actions did not come to pass but these "near misses" should not be counted as positive evidence for the democratic peace proposition.

More generally, the democratic peace proposition argues that democracy contributes to peace but it does not claim that democracy is the only approach to peace. Peace can be established by a variety of ways, including imperial rule and nuclear deterrence. It is, however, contrary to both the substance and spirit of the democratic peace proposition to wage war in the name of democracy. As just mentioned, whether one can impose democracy from outside is a topic that scholars have already started to study.[37] Whether one should is a different but equally important topic. It should be recalled that Immanuel Kant, the intellectual predecessor acknowledged by most liberal internationalists writing about the democratic peace proposition, envisioned like-minded political entities voluntarily joining a pacific league of autonomous members. Spreading democracy at the point of bayonets would be the farthest from his thinking.

Conclusion

The democratic peace proposition has been the focus of an impressive amount of research, much of which presenting evidence in support of at least the dyadic version of this proposition, namely, that democracies do not fight each other. This essay addresses mostly the monadic version of this proposition, raising a number of issues about whether democracies are more peaceful in general, that is, even when they are paired with autocracies in international interactions. The suggestions offered in this essay are intended to further advance the research agenda investigating the effects of regime characteristics on a country's bellicosity or pacifism in international relations. It has been deliberately provocative, but its purpose is to encourage further clarification and extension of this important area of research. It points out several areas pertaining to concept formation, methodological application, and causal attribution where greater clarity and care can be beneficial. Moreover, it warns against the potential misuse of this proposition for policy purposes or public justification in ways that are contrary to the spirit motivating it in the first place.

Notes

Steve Chan is Professor of Political Science at the University of Colorado, Boulder, U.S.A. He may be reached at steve.chan@colorado.edu. This is a revised version of a keynote address presented at Second Australasian Conference on Economics and Politics of War and Peace, 25-26 July 2008, Macquarie University, Sydney, Australia.

1. Levy (1989, p. 270).

2. Russett (1990, p. 123).

3. See, for example, Doyle (1983a, 1983b, 1986, 2005). For an outstanding empirical analysis, see Russett and Oneal (2001).

4. Kant (1795).

5. There are many strands of realism. Waltz (1979) offers the most prominent neorealist statement. For examples of realist work that challenges the democratic peace proposition, see Layne (1994) and Rosato (2005).

6. For literature reviews on this topic, see Rummel (1985); Morgan (1993); Maoz (1998); Ray (1998); Russett and Starr (2000); MacMillan (2000); Morrow (2002); Chan (1997, 2008).

7. Kant (1795).

8. Besides Doyle's articles, see Huntley (1996).

9. Rummel (1979, 1981, 1983).

10. Ray (1993).

11. Reiter and Tillman (2002).

12. Vanhanen (2000).

13. For instance, Gates, Hegre, Jones and Strand (2006) developed a Scalar Index of Politics (SIP) based on measures from both the Polity and Vanhanen datasets.

14. On democratic intervention, see Kegley and Hermann (1995, 1996) and Hermann and Kegley (1996).

15. The monadic version of the democratic peace proposition claims that democracies are more peaceful in general — that is, their peaceful disposition is not affected by the regime characteristic of the other party in a relationship, whether it is also a democracy or an autocracy. The dyadic version only claims that democracies are more peaceful toward each other.

16. See www.visionofhumanity.org [accessed 1 September 2008].

17. Peceny, Beer, and Sanchez-Terry (2002).

18. Farber and Gowa (1995, 1997); Gartzke (2007); Gibler (2007). A review of the hypothesis that it is not democracies but international trade that contributes to peace, see, e.g., Polachek (2007).

19. Thompson (1996); Gleditsch (2002).

20. Ward, Siversion, and Cao (2007).

21. Morrow (2002).

22. Fearon (1994a, 1994b, 1995) developed the idea of audience costs, and Gaubatz (1999) offered a persuasive analysis on electoral incentives and democracies' wars.

23. Bueno de Mesquita, Siverson, and Woller (1992) and Bueno de Mesquita and Siverson (1995) showed that democratic leaders' political tenure is likely to be shortened if they suffer defeat in war.

24. Schulz (1999, 2001a, 2001b) is especially pertinent in this regard.

25. For pertinent discussions, see Kaufman (2004, 2005) and Howell and Pevehouse (2007).

26. Reiter and Tillman (2002); Leblang and Chan (2003); Ireland and Palmer (2004); Palmer, London, and Regan (2004).

27. The term "selectorate" was used by Bueno de Mesquita, Morrow, Siversion, and Smith (1999), and Bueno de Mesquita, Smith, Sivreson, and Morrow (2003) to develop this line of argument.

28. See Weeks (2008).

29. Maoz and Russett (1993) presented normative and structural explanations of democratic peace.

30. See Mueller (1973) for a classic study on this topic, and Gelpi, Feaver, and Reifler (2008, 2005/06) and Berinsky (2007) for the more recent debate.

31. Jentleson (1992); Jentleson and Britton (1998); Eichenberg (2005).

32. See Chan and Safran (2006).

33. Bueno de Mesquita and Lalman (1992) and Maoz and Russett (1993), for example, recognized that democracies can wage war against autocracies. Schweller (1992) argued that democracies are unlikely to start preventive war unless under exceptional circumstances.

34. See, for example, Meernik (1996); Peceny (1999); Pickering and Kisangani (2006); Pickering and Peceny (2006).

35. Paris (2004).

36. Levy and Gochal (2001); Ripsman and Levy (2007); Silverstone (2007); Levy (2008); Trachtenberg (2008).

37. In addition to the studies already cited, see Bueno de Mesquita and Downs (2006) with particular relevance for the ongoing Iraq war.

References

Berinsky, A. 2007. "Assuming the Costs of War: Events, Elites, and American Public Support for Military Conflict." *Journal of Politics*, Vol. 69, No. 4, pp. 975-997.

Bueno de Mesquita, B. and G.W. Downs. 2006. "Intervention and Democracy." *International Organization*, Vol. 60, No. 3, pp. 627-649.

Bueno de Mesquita, B. and D. Lalman. 1992. *War and Reason: Domestic and International Imperatives*. New Haven, CT: Yale University Press.

Bueno de Mesquita, B., J.D. Morrow, R.M. Siverson, and A. Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review*, Vol. 93, No. 4, pp. 791-807.

Bueno de Mesquita, B. and R.M. Siverson. 1995. "War and the Survival of Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review*, Vol. 89, No. 4, pp. 841-855.

Bueno de Mesquita, B., R.M. Siverson, and G. Woller. 1992. "War and the Fate of Regimes: A Comparative Analysis." *American Political Science Review*, Vol. 86, No. 3, pp. 638-646.

Bueno de Mesquita, B., A. Smith, R.M. Siversion, and J.D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: The MIT Press.

Chan, S. 1997. "In Search of Democratic Peace: Problems and Promises." *Mershon International Studies Review*, Vol. 41, No. 1, pp. 59-91.

Chan, S. 2008. "Progress in the Democratic Peace Research Agenda." Mimeo. Boulder: University of Colorado.

Chan, S. and W. Safran. 2006. "Public Opinion as a Constraint against War: Democracies' Responses to Operation Iraqi Freedom." *Foreign Policy Analysis*, Vol. 2, pp. 137-156.

Doyle, M.W. 1983a. "Kant, Liberal Legacies, and Foreign Affairs: Part 1." *Philosophy and Public Affairs*, Vol. 12, No. 3, pp. 205-235.
Doyle, M.W. 1983b. "Kant, Liberal Legacies, and Foreign Affairs: Part 2." *Philosophy and Public Affairs*, Vol. 12, No. 4, pp. 323-353.
Doyle, M.W. 1986. "Liberalism and World Politics." *American Political Science Review*, Vol. 80, No. 4, pp. 1151-1169.
Doyle, M.W. 2005. "Three Pillars of the Liberal Peace." *American Political Science Review*, Vol. 99, No. 3, pp. 463-466.
Eichenberg, R.C. 2005. "Victory Has Many Friends: U.S. Public Opinion and the Use of Military Force." *International Security*, Vol. 30, No. 2, pp. 140-177.
Farber, H.S. and J. Gowa. 1995. "Polities and Peace." *International Security*, Vol. 20, No. 2, pp. 123-146.
Farber, H.S. and J. Gowa. 1997. "Common Interests or Common Politics?" *Journal of Politics*, Vol. 59, No. 2, pp. 393-417.
Fearon, J.D. 1994a. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review*, Vol. 88, No. 3, pp. 577-594.
Fearon, J.D. 1994b. "Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution*, Vol. 38, No. 2, pp. 236-269.
Fearon, J.D. 1995. "Rationalist Explanations of War." *International Organization*, Vol. 49, No. 3, pp. 379-414.
Gartzke, E. 2007. "The Capitalist Peace." *American Journal of Political Science*, Vol. 51, No. 1, pp. 166-191.
Gates, S., H. Hegre, M.P. Jones, and H. Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800-2000." *American Journal of Political Science*, Vol. 50, No. 4, pp. 893-908.
Gaubatz, K.T. 1999. *Elections and War: The Electoral Incentive in the Democratic Politics of War and Peace*. Stanford, CA: Stanford University Press.
Gelpi, C.F., P. Feaver, and J. Reifler. 2005/06. "Success Matters: Casualty Sensitivity and the War in Iraq." *International Security*, Vol. 30, No. 3, pp. 7-46.
Gelpi, C.F., P. Feaver, and J. Reifler. 2008. *Paying the Human Costs of War: American Public Opinion and Casualties in Military Conflicts*. Princeton, NJ: Princeton University Press.
Gibler, D.M. 2007. "Bordering on Peace: Democracy, Territorial Issues, and Conflict." *International Studies Quarterly*, Vol. 51, No. 3, pp. 509-532.
Gleditsch, K.S. 2002. *All Politics Is Local: The Diffusion of Conflict, Integration, and Democratization*. Ann Arbor: University of Michigan Press.
Global Peace Index: 2008 Methodology, Results & Findings. New South Wales, Australia. www.visionofhumanity.org.
Hermann, M.G. and C.W. Kegley, Jr. 1996. "Ballots, A Barrier against the Use of Bullets and Bombs: Democratization and Military Intervention." *Journal of Conflict Resolution*, Vol. 40, No. 3, pp. 436-459.
Howell, W.G. and J.C. Pevehouse. 2007. *While Dangers Gather: Congressional Check on Presidential War Powers*. Princeton, NJ: Princeton University Press.
Huntley, W. 1996 "Kant's Third Image: Systematic Sources of the Liberal Peace." *International Studies Quarterly*, Vol. 40, No. 1, pp. 45-76.
Ireland, M.J. and S.S. Gartner. 2001. "Time to Fight: Government Type and Conflict Initiation in Parliamentary Systems." *Journal of Conflict Resolution*, Vol. 21, No. 2, pp. 547-568.
Jentleson, Bruce W. 1992. "The Pretty Prudent Public: Post-Vietnam American Opinion on the Use of Military Force." *International Studies Quarterly*, Vol. 36, No. 1, pp. 47-74.
Jentleson, B.W. and R.L. Britton. 1998. "Still Pretty Prudent: Post-Cold War American Public Opinion on the Use of Military Force." *Journal of Conflict Resolution*, Vol. 42, No. 4, pp. 395-417.
Kant, I. 1795 (1957). *Perpetual Peace*. Translated by L.W. Beck. New York: Bobbs-Merrill.
Kaufman, C. 2004 "Threat Inflation and the Failure of the Marketplace of Ideas: The Selling of the Iraq War." *International Security*, Vol. 29, No. 1, pp. 5-48.
Kaufman, C. 2005. "Selling the Market Short: The Marketplace of Ideas and the Iraq War." *International Security*, Vol. 29, No. 4, pp. 196-207.
Kegley Jr., C.W. and M. Hermann. 1995. "Military Intervention and the Democratic Peace." *International Interactions*, Vol. 21, No. 1, pp. 1-21.
Kegley Jr., C.W. and M. Hermann. 1996. "How Democracies Use Intervention: A Neglected Dimension in Studies of the Democratic Peace." *Journal of Peace Research*, Vol. 33, No. 3, pp. 309-322.
Layne, C. 1994. "Kant or Cant: The Myth of Democratic Peace." *International Security*, Vol. 19, No. 2, pp. 5-49.
Leblang, D. and S. Chan. 2003. "Explaining Wars Fought by Established Democracies: Do Institutional Constraints Matter?" *Political Research Quarterly*, Vol. 56, No. 4, pp. 385-400.
Levy, J.S. 1989. "The Causes of War: A Review of Theories and Evidence," pp. 209-333 in P.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Stern, and C. Tilly, eds. *Behavior, Society, and Nuclear War*. New York: Oxford University Press.
Levy, J.S. 2008. "Preventive War and Democratic Politics." *International Studies Quarterly*, Vol. 52, No. 1, pp. 1-24.
Levy, J.S. and J.R. Gochal. 2001. "Democracy and Preventive War: Israel and the 1956 Sinai Campaign." *Security Studies*, Vol. 11, No. 2, pp. 1-49.
MacMillan, J. 2003. "Beyond the Separate Democratic Peace." *Journal of Peace Research*, Vol. 40, No. 2, pp. 233-243.
Maoz, Z. 1998. "Realist and Cultural Critiques of the Democratic Peace." *International Interactions*, Vol. 24, No. 1, pp. 1-89.
Maoz, Z. and B.M. Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946-1986." *American Political Science Review*, Vol. 87, No. 3, pp.

624-638.
Meernik, J. 1996. "U.S. Military Interventions and the Promotion of Democracy." *Journal of Peace Research*, Vol. 33, No. 4, pp. 391-402.
Morgan, C.T. 1993. "Democracies and War: Reflections on the Literature." *International Interactions*, Vol. 18, No. 3, pp. 197-203.
Morrow, J.D. 2002. "Understanding International Conflict: Microfoundations, the Democratic Peace, and Offense-Defense Theory," pp. 172-196 in I. Katznelson and H.V. Milner, eds. *Political Science: The State of the Discipline*. New York: Norton.
Mueller, J.E. 1973. *War, Presidents, and Public Opinion*. New York: Wiley.
Palmer, G.G., R.R. London, and P.M. Regan. 2004. "What's Stopping You? The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies." *International Interactions*, Vol. 30, No. 1, pp. 1-24.
Paris, R. 2004. *At War's End: Building Peace After Civil Conflict*. Cambridge: Cambridge University Press.
Peceny, M. 1999. *Democracy at the Point of Bayonets*. University Park, PA: Pennsylvania State University Press.
Peceny, M., C. Beer, and S. Sanchez-Terry. 2002. "Dictatorial Peace?" *American Political Science Review*, Vol. 96, No. 1, pp. 15-26.
Pickering, J. and E. Kisangani. 2006. "Political, Economic, and Social Consequences of Foreign Intervention." *Political Research Quarterly*, Vol. 59, No. 3, pp. 363-37.
Pickering, J. and M. Peceny. 2006. "Forging Democracy at Gunpoint." *International Studies Quarterly*, Vol. 50, No. 3, pp. 539-560.
Polachek, S.W. "How Trade Affects International Interactions." *Economics of Peace and Security Journal*, Vol. 2, No. 2 (2007), pp. 60-68.
Ray, J.L. 1993. "War Between Democracies: Rare, or Nonexistent?" *International Interactions*, Vol. 18, No. 3, pp. 251-276
Ray, J.L. 1998. "Does Democracy Cause Peace?" *Annual Review of Political Science*, Vol. 1, pp. 27-46.
Reiter, D. and E.R. Tillman. 2002. "Public, Legislative, and Executive Constraints on the Democratic Initiation of Conflict." *Journal of Politics*, Vol. 64, No. 3, pp. 810-826.
Ripsman, N.M. and J.S. Levy. 2007. "The Preventive War that Never Happened: Britain, France, and the Rise of Germany in the 1930s." *Security Studies*, Vol. 16, No. 1, pp. 32-67.
Rosato, S. 2003. "The Flawed Logic of Democratic Peace Theory." *American Political Science Review*, Vol. 97, No. 4, pp. 585-602.
Rummel, R.J. 1979. *War, Power, Peace*. Beverly Hills, CA: Sage.
Rummel, R.J. 1981. *The Just Peace*. Beverly Hills, CA: Sage.
Rummel, R.J. 1983. "Libertarianism and International Violence." *Journal of Conflict Resolution*, Vol. 27, No. 1, pp. 27-71.
Rummel, R.J. 1985. "Libertarian Propositions on Violence Within and Between Nations: A Test Against Published Research Results." *Journal of Conflict Resolution*, Vol. 29, No. 3, pp. 419-455.
Russett, B.M. 1990. *Controlling the Sword: The Democratic Governance of National Security*. Cambridge, MA: Harvard University Press.
Russett, B.M. and J.R. Oneal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York: Norton.
Russett, B.M. and H. Starr. 2000. "From Democratic Peace to Kantian Peace: Democracy and Conflict in the International System," pp. 93-128 in M.I. Midlarsky, ed. *Handbook of War Studies*, Vol. II. Ann Arbor: University of Michigan Press.
Schultz, K.A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization*, Vol. 53, No. 2, pp. 233-266.
Schultz, K.A. 2001a. "Looking for Audience Costs." *Journal of Conflict Resolution*, Vol. 45, No. 1, pp. 32-60.
Schultz, K.A. 2001b. *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.
Schweller, R.L. 1992. "Domestic Structure and Preventive War: Are Democracies More Pacific?" *World Politics*, Vol. 44, No. 2, pp. 235-269.
Silverstone, S.A. 2007. *Preventive War and American Democracy*. London: Routledge.
Thompson, W. 1996. "Democracy and Peace: Putting the Cart Before the Horse?" *International Organization*, Vol. 50, No. 1, pp. 141-174.
Trachtenberg, M. 2007. "Preventive War and U.S. Foreign Policy." *Security Studies*, Vol. 16, No. 1, pp. 1-31.
Vanhanen, T. 2000. "A New Dataset for Measuring Democracy, 1810-1998." *Journal of Peace Research*, Vol. 37, No. 2, pp. 251-265
Waltz, K.N. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.
Ward, M.D., R.M. Siverson, and X. Cao. 2007. "Disputes, Democracies, and Dependencies: A Reexamination of the Kantian Peace." *American Journal of Political Science*, Vol. 51, No. 3, pp. 583-601.
Weeks, J. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization*, Vol. 62, No. 1, pp. 35-64.

Friedman's First Law fails: oil prices do not predict freedom

Steve Townsend

One issue that has received much attention as a factor in conflict is the presence of natural resources in less-developed countries, oil in particular, and what is called the resource curse. Research is emerging that casts doubt on the resource curse and the way in which oil wealth is supposed to impede democratization.[1] To be useful, the term oil wealth needs to be separated into its components such as oil reserves, production levels, prices, revenues, profits, royalties, and ownership. This articles focuses on one of these components, price, and how it may affect the level of freedom in countries where oil revenues make up a large share of gross domestic product (GDP). In particular, it will examine the First Law of Petropolitics, a rather extreme extension of the resource curse, put forward by the influential *New York Times* columnist, Thomas L. Friedman. This first appeared in *Foreign Policy* in 2006.[2] The "law" claims that high oil prices cause a decline in the levels of freedom. It contends that freedom flourishes in less-developed countries as soon as oil prices fall.

This article investigates Friedman's claims for the four countries for which he provides data as well as other countries where he claims the law applies. In some cases his data is incomplete and when more complete data is used his claim is refuted. In some cases his data is inappropriate such as when he uses oil prices in current dollars rather than have them converted to constant dollar values. In one important case, Iran, he misinterprets the data to support his thesis but it is clear that, for this example, the correlation is positive, not negative as Friedman's law would require.

It is important to bring these shortcomings to light. Friedman's idea is gaining increasing currency and acceptance among political commentators and even academics. Friedman continues to promote the idea in his *New York Times* articles, and in television and radio interviews. It has been discussed on several weblogs and has received credence among bloggers on the Private Sector Development Blog maintained by the World Bank. *The Economist* also concurs with the idea and even accepts that it is a law.[3]

So far, there is not a lot of mention of Friedman's First Law in peer-reviewed journals but it is beginning to appear and has already been used in an article in *International Political Science Review* to support a new analysis of governance issues. It has also been cited twice in journal articles proposing changes in energy policy and has been reprinted as a reading in a respected text on international relations.[4] Given the influence Friedman's claims are having, it is important that they be examined and appraised.

What Friedman claims

According to Thomas Friedman's First Law of Petropolitics, "the price of oil and the pace of freedom always move in opposite directions in oil-rich petrolist states." This law is false.

According to Friedman's First Law of Petropolitics, "the price of oil and the pace of freedom always move in opposite directions in oil-rich petrolist states." He defines petrolist states as those that depend on oil production for most of their exports or gross domestic product and that also have weak state institutions or outright authoritarian governments. He claims that as the average price of oil increases, freedom of speech, free elections, and the rule of law will decline in these states as will the independence of political parties and the judiciary. When the price of oil decreases, these states will be forced to become more responsive to internal and international criticism and reverse the deterioration.[5]

Much of Friedman's *Foreign Policy* article consists of anecdotes; these will not be discussed here. The main quantitative thrust of the article is provided in data for four countries: Venezuela, Nigeria, Russia, and Iran. Friedman does no more than present graphs to establish his correlations. In this article, pairwise correlations and probability values are calculated for the strength, direction, and statistical significance of each correlation (employing Pearson's *r*).

How good are Friedman's correlations?

Venezuela

For Venezuela, Friedman graphed Freedom House scores and nominal oil prices for the period 1986 to 2005.[6] Using nominal oil prices rather than inflation-adjusted ones, especially when using data stretching over 20 years, is inappropriate. Friedman offers no explanation for doing so. (Emails to Friedman and to the editor of *Foreign Policy* on this and other issues have not had a reply.) As to the measure of freedom, Friedman uses the simple average of two Freedom House scores, for political rights and for civil liberties.[7]

Examining Friedman's 1986-2005 data in Figure 1, a negative correlation between oil prices and freedom levels does appear to exist. The correlation coefficient, *r*, equals -0.60, a reasonably strong negative correlation, and it shows a strong statistical significance as well ($p = 0.004$). Of course, a negative value of *r* is what would be expected when variables are moving in opposite directions.

But Freedom House data are available for Venezuela as from 1972 and inflation-adjusted oil prices go even further back in time. The oil prices used here are British Petroleum's mean world crude oil prices in constant 2006 U.S. dollars per barrel.[8] When these numbers are plotted in Figure 1 as well, it is readily apparent that Friedman's claim falls apart. First, using the inflation-adjusted oil prices, the

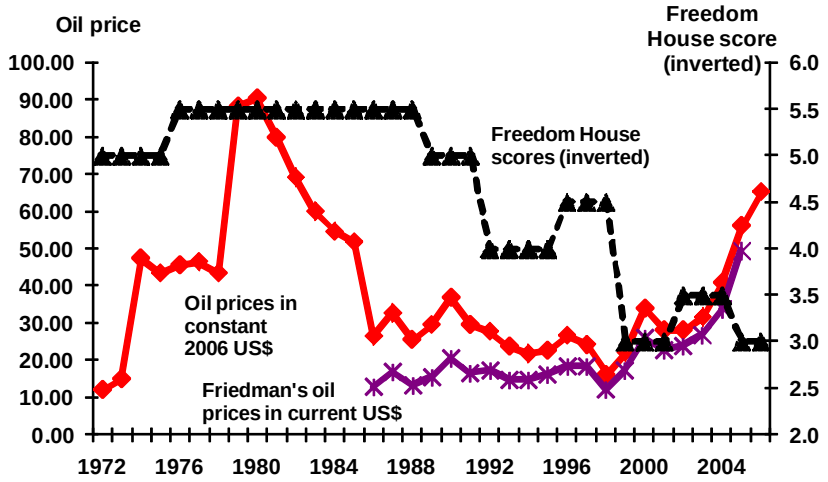

Figure 1: Venezuela, 1972-2006: oil prices and Freedom House scores (inverted).

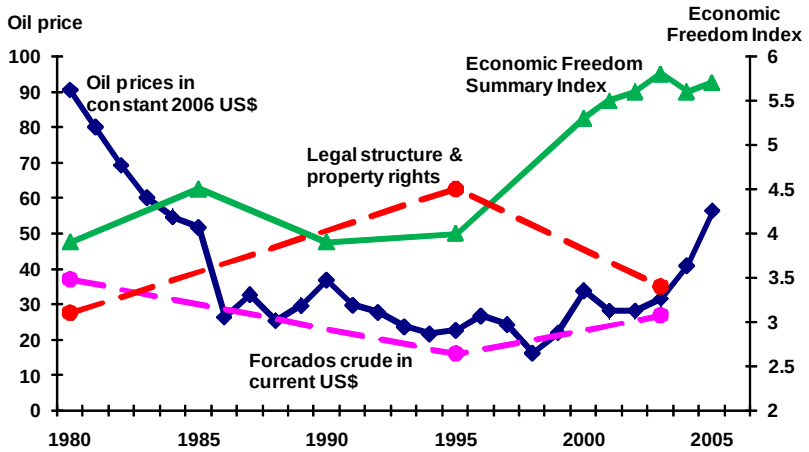

Figure 2: Nigeria, 1980-2005, oil prices and economic freedom.

correlation for Friedman's 1986 to 2005 time period is no longer statistically significant (r = -0.26, p = 0.27) and, second, when the time period is extended to 1972 to 2006, the correlation becomes a statistically significant *positive* one (r = +0.32; p = 0.06). Friedman's law is turned on its head.

Table 1: Venezuela, 1972-2006: correlations for oil price in 2006 U.S. dollars and Freedom House scores

Lags (years)	*Pearson's r*	*p-value*
1	+0.45	0.007
2	+0.55	0.001
3	+0.57	0.001
4	+0.57	0.001

A change in oil prices is unlikely to affect a country's level of freedom instantly. Friedman does not discuss the likelihood that the effects of any price changes will be delayed, but obviously this should be considered. Table 1 presents the correlation (and probability) values when the dependent variable (freedom) is lagged by between one to four years. All show reasonably strong positive correlations and with high statistical significance, again turning Friedman's law on its head.

Nigeria

The graph Friedman presents for Nigeria — the two dashed lines in Figure 2 — also appear to strongly supports his case. The measure of legal structure and property rights and the measure of oil prices certainly seems to move in opposite directions. But while r = -0.98, statistical significance is low at p = 0.12, so a correlation is not certain. This is to be expected when using only three data pairs, as Friedman does. Also, once more there are problems with the data he employed. First, it is puzzling why Friedman provides data only for 1980, 1995, and 2003 when the sources he used had data for the intervening years available. Second, as for Venezuela, he should have converted nominal prices into real prices. And third, we should have examined the effect of inflation-adjusted oil price on lagged measures of freedom.

The oil prices Friedman used are attributed to BP's *Statistical Review of World Energy 2005*. Friedman seems to have chosen the average yearly spot price per barrel for Nigerian Forcados blend in current U.S. dollars. The figures for economic freedom are more difficult to specify but seem to be those for Legal Structure and Security of Property Rights, taken from the Fraser Institute's *Economic Freedom of the World* report. For the period 1980 to 2003, five more data pairs are available for the economic freedom measure used by Friedman as well as for the current-dollar oil prices he used. When these are added to the data set, bringing the total number of observations to eight per variable, the correlation is r = -0.51; p = 0.201, far weaker than the result using only three data pairs. When another two data points are added to each variable (for 2004 and 2005; not available to Friedman at the time), the relationship deteriorates even further: r = -0.25; p = 0.492.

It is odd that Friedman confined his measure of economic freedom to just one of the Fraser Institute's 41 freedom measures that the report gathers and compresses into a single index number. It seems reasonable to assume the summary number would give a better indication of the level of economic freedom than any one of its many components. And as already discussed, for an analysis spanning 23 years, oil prices

Table 2: Nigeria, 1980-2005: correlations for oil price in 2006 U.S. dollars and Economic Freedom index

Lags (years)	*Pearson's r*	*p-value*
1	+0.01	0.98
2	+0.21	0.59
3	+0.51	0.20
4	+0.66	0.10

ought to be adjusted for inflation. When these variables are graphed in Figure 2 — the two solid lines — and the correlation is computed, then there is insufficient evidence for Friedman's thesis (r = -0.34; p = 0.332). Moreover, when the economic freedom variable was lagged by one or more years, the correlations turn out to be positive (although not statistically significant; see Table 2).

Friedman's claim for a strong correlation between oil prices and economic freedom in Nigeria is not supported by a wider view of the facts. Friedman cannot claim support for his First Law of Petropolitics in Nigeria.

Russia

Friedman's graph for Russia (not shown here) provides good support for his claim. It clearly shows the measure of a variable called Electoral Process declining as oil prices increase, and it is spread over the years 1998 to 2005 (with 1999 and 2000 having to be reduced to a single point due to difficulties in gaining data). But picking the Freedom House score for Electoral Process as his measure of freedom in Russia is, again, an idiosyncratic choice because it is only one of the seven components used in the Freedom House study *Nations in Transit* to develop a Democracy Score (the mean of electoral process, civil society, independent media, national democratic governance, local democratic governance, judicial framework and independence, and corruption.)[9] Friedman's choice of oil price seems to be the spot price for Dubai crude in current dollars. For these variables, the correlation for 1998 to 2005 is excellent (r = -0.96; p < 0.001).

Table 3: Treating the USSR/Russia as a single country, 1976-2006: examining the effect of lagging on the correlations for oil price in 2006 U.S. dollars and Freedom House scores (inverted)

Lags (years)	*Pearson's r*	*p-value*
0	-0.55	0.0002
1	-0.61	0.0002
2	-0.61	0.0002
3	-0.58	0.0005
4	-0.59	0.0005

The correlation with the Electoral Process scores holds up very well even when using inflation-adjusted oil prices for the same period. Moreover, when the composite Freedom House score is used, along with inflation-adjusted oil prices, the correlation still looks very good r = -0.88; p = 0.004). And when the computations are redone for a longer time period, 1972 to 2006, the correlation is still moderately good (r = -0.55; p < 0.001). Table 3 shows that the strength of the correlation stays at much the same level and statistically remains highly significant when the Freedom House scores are lagged between one and four years (with the no lag correlation shown for comparison).

Table 4: USSR and Russia considered separately: correlations are for average oil prices in constant 2006 U.S. dollars and Freedom House scores (inverted)

	USSR 1972-1990		*Russia 1991-2006*	
Lags (years)	*Pearson's r*	*p-value*	*Pearson's r*	*p-value*
0	-0.27	0.28	-0.71	0.002
1	+0.08	0.77	-0.78	0.003
2	+0.36	0.19	-0.70	0.002
3	+0.36	0.21	-0.53	0.034
4	+0.03	0.38	-0.36	0.174

However, because two very different regimes existed during this time period, it is reasonable to split it into two parts, the USSR from 1972 to 1990, and Russia from 1991 to 2006. The results are shown in Table 4. The USSR clearly did not conform to Friedman's Law but Russia, over the period 1991 to 2006, showed a reasonably strong negative correlation with high significance for lags of 0 to 3 years. So, it would seem that Russia during this period could be an instance where the First Law of Petropolitics is satisfied.

Iran

In the case of Iran (Figure 3), Friedman graphs oil prices and a measure he calls Freedom to Trade, taken from the Fraser Institute's *Economic Freedom of the World* report. Taking the best estimates of the data from his chart, it seems this figure represents what the report calls Freedom to Exchange with Foreigners.[10] His prices for crude oil seem to be, as with Venezuela, the average Dubai spot crude price in current U.S. dollars.[11] But this example of evidence in support of the First Law of Petropolitics is the most perplexing of all. Even using exactly the same data as provided in Friedman's graph, it clearly contradicts the Law and shows the opposite of what it is meant to show. For most of the given time span, 1995 to 2003, the freedom to trade follows oil price, and the two measures do not begin to diverge until 2002 (r = +0.79; p = 0.11). It is difficult to explain this situation. Friedman risked being accused of cherry-picking by choosing a narrow time-frame, selecting an unusual measure for economic freedom, and using current dollars for the price of oil, yet he still arrives at a strong argument against his own thesis. It is possible that he has been diligently even-handed by presenting some counter-evidence. But if that

were the case he should state it clearly, otherwise the reader can conclude only that it was an example of careless research or editing.

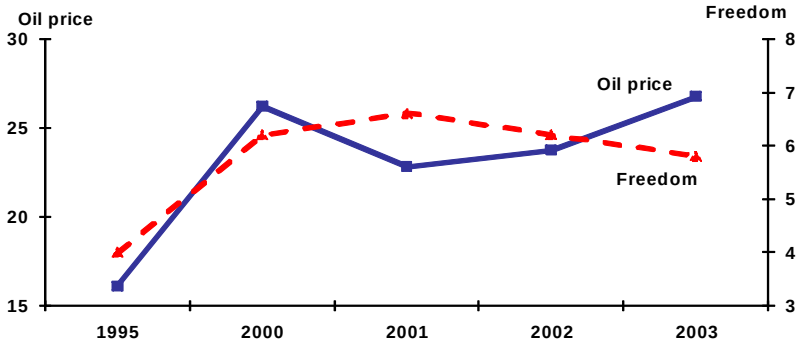

Figure 3: Iran, 1995-2003.

It is strange that Friedman presented figures only for the period 1995 to 2003 which tend to disprove his theory. Extending the time period to include numbers for 1970, 1975, 1980, 1985, 1990, 2004, and 2005 suggests a weak negative, but statistically insignificant, correlation (r = -0.2669; p = 0.4018).

Using average crude oil prices in 2006 dollars and the Fraser Institute's more general measure, the Summary Index,[12] over the same period shows a stronger negative correlation (r = -0.4995, p = 0.0982). However, with only 12 data points, this relationship is heavily influenced by the numbers for 1980. The economic freedom index fell sharply to reach its lowest point, 4.0, in 1980 and the oil price that year reached its highest value, $90.46. This would suggest that here was a case when rising oil prices did cause a decline in freedom. But the correlation is spurious. Both changes were caused by the same event: the 1979 Iranian Revolution.

Thus, the Freedom House score for Iran also seems to negate the First Law of Petropolitics. There is a weak correlation between oil price and the Freedom House score but it is in the wrong direction: the higher the oil price, the higher the level of freedom (r = +0.37; p = 0.0294). These various ways of looking at the relationship between oil price and freedom in Iran show there is no support for Friedman's Law.

Petrolist states in general

Thus far the analysis deals with only four of Friedman's "petrolist" states. Nine other countries are included in Friedman's list of petrolist states, albeit without data. They are: Azerbaijan, Angola, Chad, Egypt, Equatorial Guinea, Kazakhstan, Saudi Arabia, Sudan, and Uzbekistan. As a final exercise, then, data for all 13 countries on the list were collected. The freedom measures were the inverted Freedom House scores and the Fraser Institute's summary index of economic freedom, as discussed previously. In addition, a measure derived from the Polity IV data set was used as well.[13]

Table 5: 13 petrolist countries, and how the price of oil correlates with three measures of freedom, 1972-2006

	FH (inverted)				*EFI*				*PS*			
Lag (years)	*0*	*1*	*2*	*4*	*0*	*1*	*2*	*4*	*0*	*1*	*2*	*4*
Angola	Y								Y	Y	Y	Y
Azerbaijan												Y
Chad	Y				Y				N	N		
Egypt	N	N			Y				N			
Equat. Guinea				N					Y	Y	Y	
Iran	N	N	N									
Kazakhstan									Y	Y	Y	Y
Nigeria	N	N	N						N	N	N	
USSR-Russia	Y	Y	Y	Y	N	N	N	N	Y	Y	Y	Y
Saudi Arabia	N	N	N									
Sudan	N	N	N									
Uzbekistan												
Venezuela	N	N	N		N							

Note: FH = Freedom House; EFI = Economic Freedom Index; PS = Polity score; Y = instances which lend some support to Friedman's claim ($r < 0$; $p < 0.1$); N = instances which tend to contradict Friedman's claim ($r > 0$; $p < 0.1$); empty cells = instances where the data cannot be used to support or contradict the claim or data is not available.

A summary of the results is given in Table 5 (and full details are available from the author). With 13 petrolist states, three measures of freedom, and lags of 0, 1, 2, and 4 years, there are 156 opportunities for Friedman's claim to be substantiated. However, only 24 instances support his claim, that is, when the statistical significance is better than 90 percent and the correlation is negative. In 66 cases, the results neither support nor contradict the claim, and in 36 cases the data is not available. In 30 instances, the claim is contradicted: that is, the statistical significance is better than 90 percent but the correlation is positive. Only one country, Angola, showed support for the claim across more than one measure of freedom without having any contradicting cases. Russia showed strong support for the claim using all lag periods of Freedom House and Polity scores but was contradicted by all lag periods for economic freedom scores.

Conclusion

As has been shown, there is almost no support for Friedman's First Law of Petropolitics in the four countries for which he provided data. The claims he draws from his data fail to hold when the time period is widened or if oil prices are presented in inflation-adjusted terms. A lag in the effect of changes in oil price is highly likely yet when this is introduced into the analysis, this tends to contradict the Law. In the case of Iran, Friedman's own figures contradict his claim.

It should not be surprising that a single factor such as price cannot explain a complex process like political reform. Many other influences need to be considered. As in the case of Iran, an apparent correlation can be shown to be spurious because the price rise and the decline in freedom were both due to another factor, the Iranian Revolution. It should be pointed out that there could still be a causal link between oil price and political reform but it is being disguised by other factors. Using multiple regression analyses that include factors such as GDP growth, GDP per capita, military expenditure, and levels of taxation, no sign of clear relationships emerge. Also, for a proper analysis of the effects of oil on political reform other measures of oil wealth need to be considered such as the extent of reserves, production levels, production costs, revenues, profits, and ownership. But at this point it can be confidently stated that price alone is not been able to explain the level of democracy in the countries studied here.

Notes

Steve Townsend is a Ph.D. candidate in the Department of Government and International Relations at the University of Sydney, Australia. He may be reached at st@recognition.com.au.

1. Horiuchi and Waglé (2008); Ramsay (2006); John (2007).

2. Friedman (2006).

3. Anonymous (2007); Paul (2006); *The Economist* (2007).

4. Cole (2007); Mignone (2007); Stevens and Dietschea (2008); Mingst and Snyder (2007).

5. Friedman (2006).

6. Actually, it is difficult even to determine just what measure Friedman used for nominal oil prices. He offers the data not as numbers but only as positions on his graph. He attributes the data to "BP Statistical Review of World Energy 2005 and IEA" (Friedman, 2006). The International Energy Agency (IEA) reports do provide information on current energy balances and conversion factors for all types of energy production and consumption but they do not provide historical data on crude oil prices that are as complete as the BP data. BP's *Review* provides several series of oil prices. After comparing the graphed data with the sources cited, it appears that Friedman used the series for yearly average spot price for crude oil sold in Dubai, and that the units are U.S. dollars per barrel in current prices.

7. In this article, the Freedom House scores are presented in inverse order. This is because the Freedom House scores range from 1 (democratic) to 7 (authoritarian). That is, the more freedom, the lower the score, exactly opposite to the way other measurements of freedom are handled. The tag "(inverted)" will appear with all Freedom House scores used here to indicate that this transformation has been made. That Freedom House's original scores run in the opposite direction to other measures of freedom makes it particularly confusing when looking at negative correlations. So, throughout this article a Freedom House score of x will be presented as its additive inverse with respect to eight. That is, it will be converted to $8-x$. Thus we will be dealing with scores ranging from 7 (democratic) to 1 (authoritarian). Note also that during the period 1982 to 1989, Freedom House made changes in the timing of its collection years. It used several different collection periods before settling on calendar years in 1990. I have overcome this inconsistency by regarding all years as calendar years and interpolating values for 1989 as the mean of the values for 1988 and 1990.

8. Freedom House (2007); BP (2008).

9. Freedom House (2006). As with other Freedom House scores, the composite score ranges from 1 (democratic) to 7 (authoritarian), and in the reported calculations the inverse of the score has been used.

10. Gwartney, Lawson, and Gartzke (2005).

11. BP (2007).

12. The figures for Freedom to Exchange with Foreigners for 1970 and 1975 come from the Fraser Institute's Economic Freedom of the World Report 2003 while the figures for other years are from the 2007 report. BP reports Dubai spot crude prices only as far back as 1972 when the price was $1.90 in current dollars. I have put the 1970 price at $1.80, which is actually the world average spot crude oil price in current dollars. This would seem to be a reasonable approximation.

13. The Polity database uses a variable called POLITY which results when the AUTOC score is subtracted from the DEMOC score. It gives values ranging from -10 to +10. The AUTOC and DEMOC variables contain "standardized authority scores," -66, -77, and -88, to indicate periods of transition, war, etc. These scores are carried over into the POLITY variable, so that the Polity IV project provides another variable, POLITY2, which modifies these scores to make the data more usable in time series analyses. Following Ross's example, my Polity score variable is POLITY2 adjusted to give positive values ranging from 0 to 10. That is, 10 is added to the POLITY2 score to make all the numbers positive, then the result is halved to bring it into the range 0 to 10.

References

Anonymous. 2007. "First Law of Petropolitics." Private Sector Development Blog, Vol. 6. http://psdblog.worldbank.org/psdblog/2006/04/first_law_of_pe.html [accessed 13 March 2008].

BP. 2008. *Statistical Review of World Energy 2007*. http://www.bp.com/productlanding.do?categoryId=6848&contentId=7033471 [accessed 30 October 2008].

Cole, N.S. 2007. "Hugo Chavez and President Bush's Credibility Gap: The Struggle Against U.S. Democracy Promotion." *International Political Science Review*, Vol. 28, No. 4, pp. 493-507.

[The] Economist. 2007a. "All puffed up and stalling on reform." *The Economist*, Vol. 382, No. 8518, p. 53.

Freedom House. 2006. *Nations in Transit*. Lanham, MD: Rowman & Littlefield and Freedom House.

Freedom House. 2007. *Freedom in the World: Country Ratings, 1972-2007*. New York: Freedom House.

Friedman, T.L. 2006. "The First Law of Petropolitics." *Foreign Policy* (May/June), pp. 28-36.

Gwartney, J., R. Lawson, and E. Gartzke. 2005. *Economic Freedom of the World: 2005. Annual Report*. Vancouver, B.C.: Fraser Institute.

Gwartney, J., *et al*. 2007. *Economic Freedom of the World: 2007. Annual Report*. Vancouver, B.C.: Fraser Institute.

Horiuchi, Y. and S. Waglé. 2008. "Is Oil Bad for Democracy? Replication and New Analysis." Crawford School of Economics and Government, Australian National University, Canberra, Australia.

John, J.D. 2007. "Oil Abundance and Violent Political Conflict: A Critical Assessment." *Journal of Development Studies*, Vol. 43, No. 6, p. 26, pp. 961-986.

Marshall, M.G., K. Jaggers, and T.R. Gurr. 206. *Polity IVd Project Dataset*. Center for International Development and Conflict Management, University of Maryland.

Mignone, B. 2007. "The National Security Dividend of Global Carbon Mitigation." *Energy Policy*, Vol. 35, No. 11, p. 5403-5409.

Mingst, K.A. and J.L. Snyder, eds. 2007. *Essential Readings in World Politics*. New York: W.W. Norton: New York.

Paul. 2006. "Mr. Metaphor and the First Law of Petropolitics." Truck and Barter Blog. http://truckandbarter.com/mt/archives/2006/05/mr_metaphor_and.html [accessed 13 March 2008].

Ramsay, K.W. 2006. *Natural Disasters: Price of Oil and Democracy*. Princeton, NJ: Princeton University Press.

Stevens, P. and E. Dietschea. 2008. "Resource Curse: An Analysis of Causes, Experiences, and Possible Ways Forward." *Energy Policy*, Vol. 36, No. 1, pp. 56-65.

Short-term versus long-term military planning

Ronen Bar-El, *Kobi Kagan*, and *Asher Tishler*

In this article, we analyze two rival countries that are involved in an arms race. We compare the consequences of myopic (period-by-period) planning versus rational (long-term) planning and show that although myopic planning is always favorable for both countries, they are likely to become locked in a prisoners' dilemma equilibrium in which they plan rationally but which results in overinvestment in arms procurement and underspending on civilian services. Some options to break out of the dilemma are explored.

Countries involved in an ongoing (military) conflict are usually caught up in an arms race and spend considerable resources to ensure an acceptable level of security vis-à-vis their rivals. For example, the base defense budget of the United States for 2009 is $515.4 billion,[1] and Israel's defense budget in 2008 was $11 billion, about 7.4 percent of its GDP. Both the U.S. and Israel allocate considerable shares of their defense budgets to procurement (mostly of sophisticated weapon systems) and to military R&D. For example, in 2009 the U.S. plans to spend $79.6 billion on military R&D activities and $104.2 billion on procurement.[2] The considerable resources that are committed to military build-up around the world and the substantial efforts of rival countries to achieve a military edge over their adversaries indicate that in allocating their resources between civilian (education, welfare, health, etc.) and military expenditure governments account for long-term considerations, at least when they plan their military order of battle (arrays).[3]

In this article, we analyze two rival countries that are involved in an arms race. We compare the consequences of myopic (period-by-period) planning versus rational (long-term) planning and show that although myopic planning is always favorable for both countries, they are likely to become locked in a prisoners' dilemma equilibrium in which they plan rationally but which results in overinvestment in arms procurement and underspending on civilian services. In general, they would be well-advised to consider other strategies to improve the welfare of their citizens without compromising their required security levels. A dynamic version of Kagan, Levkowitz, Tishler, and Weiss (2008) is employed, with real-world data, to show the likely existence of a prisoners' dilemma in the current Israeli-Syrian arms race.

The nature of arms race planning strategies

Since Richardson's seminal contribution,[4] economists have been analyzing arms races as a noncooperative game between two or more rival countries, each intent on accumulating weapon systems to build up their respective military power. The discussion has focused, among other issues, on the nature of the dynamic strategies adopted by the decisionmakers of the rival countries when allocating the government budget between arms procurement and civilian expenditure and on the characteristics of the resulting (Nash) equilibrium. Some researchers have considered open-loop Nash equilibrium strategies while others have considered closed-loop Nash equilibrium strategies.[5] The choice between them may not be critical as they exhibit similar properties, although a closed-loop equilibrium results in lower arms stocks and higher welfare than does an open-loop equilibrium.[6] Furthermore, the static Nash equilibrium exhibits the same properties as the open-loop Nash equilibrium.[7] There is also the conclusion of Brito and Intriligator (1995) that "since the actual mechanism involved in the allocation of resources in the countries involved is a complex combination of political and bureaucratic behavior, there is some virtue in simplicity."

Previous literature considers the strategies undertaken by the participants in the arms race as given (open-loop or closed-loop). In this article, as is the case in reality, we let the decisionmakers of the rival countries decide on the planning strategy, in addition to the allocation of the government budget between arms procurement and civilian expenditure. The decisionmakers can be either myopic, planning only one period at a time (time-step planning), or rational, determining, at the beginning of the first period, the allocation of the government budget, taking the rival's decisions into account, for the whole planning horizon (e.g., open-loop strategy).

Generally, this article suggests that solving military/political conflicts that evolve into an arms race by relying only on military might is an expensive and suboptimal solution. The better approach in an arms race setup is to consider political and economic strategies, in addition to the military option.

The growing importance of planning ahead

As weapon systems are becoming ever more sophisticated, the planning horizon plays an important role in their accumulation and in the overall military power build-up.[8] The development of a major weapon system involves several consecutive stages, each dependent on the previous one: technology feasibility study, pre-development, full-scale development, testing, integration, prototype, serial production, field deployment, and achieving full operational capabilities. This process may take 20 years or more to complete.[9] Even weapon systems that do not require full-scale development (weapon systems developed in the past) may require a considerable amount of time to upgrade and modernize. For example, the time from the procurement stage of a submarine to its full integration into the navy may be up to 10 years. Thus, the time it takes for a new weapon system to be fully deployed forces the

decisionmaker to take into account past procurement and the availability of an existing stock of weapons, as well as to envision the characteristics of the future battle field, including possible reactions of the adversary to future deployments of the weapon system.

The arms race and the prisoners' dilemma

The strategies of the rival countries may give rise to a prisoners' dilemma equilibrium in which both countries over-invest in arms procurement to counter each other's stocks of weapon systems, thus lowering civilian expenditure and, hence, the welfare (utility) of both countries. Smith, Sola, and Spagnolo (2000) present empirical evidence of a prisoners' dilemma equilibrium in the Greek-Turkish conflict. Snyder (1971) implements a prisoners' dilemma setup in several international conflicts, and Plous (1993) analyzes the nuclear arms race as a perceptual dilemma.

Here we follow Bar-El, Kagan, and Tishler (2008) and identify a different kind of prisoners' dilemma, one that emerges from the planning strategies of the rival countries. We show that although myopic (short-term) planning is always favorable for both rivals, they are more likely to be locked into a prisoners' dilemma in which they plan rationally (long-term), which results in higher stocks of weapon systems and lower welfare for both countries. Moreover, we find some evidence of the existence of a prisoners' dilemma equilibrium in the current Israeli-Syrian arms race.

The analytic framework

For expositional simplicity consider an arms race between two identical countries.[10] The level of security of each country is equal to its military capability (defined as its own stock of weapon systems) divided by the military capability of its rival.[11] Again, for brevity we assume that each country employs only one (aggregate) type of weapon system which depreciates at a constant rate over time. The objective of both countries is to maximize the discounted stream of utilities defined over its security level and civilian expenditure. The decisionmakers in the rival countries decide, in each period, on the allocation of their government budget between civilian services (which yield utility for only one period) and weapon systems. In addition, they decide on their planning strategy. That is, they can be either myopic, planning each period at a time (time-step planning) or rational, determining, in the first period, the allocation of the government budget and taking the rival's decisions for the whole planning horizon into account. The myopic decisionmaker plans one period (each planning period consists of five years, for example) at a time and in her decision in the first period, say, does not take into account the benefits (utility) of the weapon systems in future periods. The rational planner procures more weapon systems than the myopic planner since, at the beginning of the planning horizon, she takes into account future benefits (utility), which are positive, to be received from the procurement of weapon systems in earlier periods (weapon systems yield security over their lifetime, which may last many periods). In our simplified example the two rival countries are identical. Thus, both procure identical quantities of weapon systems regardless of whether they are myopic or rational, implying that the ratio of security levels equals one for both countries under either strategy. However, for a given government budget, larger quantities of arms procurement under rational planning imply a lower level of civilian services. Therefore, when the decisionmakers are rational they allocate fewer resources to civilian services than myopic decisionmakers, but supply the same security level as do myopic ones. Clearly, the utility levels of the citizens of both countries are lower (the same security level and fewer civilian services) under rational decisionmaking. Bar-El, *et al*. (2008) show a greater likelihood of being locked into a rational planning Nash equilibrium when the decisionmakers' discount factor is higher, the perceived benefit from security is higher, and the depreciation rate of weapon systems is lower.

The Israeli-Syrian arms race

To test our model we employ here a dynamic version of Kagan, Levkowitz, Tishler, and Weiss (2008) (henceforth KLTW) to determine the equilibrium strategies of the Israeli-Syrian arms race.

The KLTW model describes an asymmetric arms race between a developed (wealthy) Western country (Israel in KLTW's application) and a (relatively poor) less developed country (Syria in KLTW's application). Due to insufficient financial resources, human capital, and technological infrastructure, the less developed country cannot purchase sufficient quantities of expensive (and effective) modern weapon systems to achieve what it considers a proper security level. Therefore, this country may arm itself with cheaper weapons of mass destruction (WMD), in addition to conventional weapon systems. KLTW assume that the less developed country intends to use its WMD in future wars against its (stronger) rival and, possibly, other potential rivals. The objective of the government of each country is to maximize its discounted stream of utilities, which depends on its expenditure on civilian services and on its security level. The (different) attitudes of the Syrian and Israeli governments to security are embedded in the parameters of their welfare functions. KLTW describe each country's budget allocation between civilian services (education, municipal authorities, legal system, health, etc.) and security, where the latter is a function of the quantities and types of weapon systems in each country's arsenal and those of its adversary. More specifically, the less developed country purchases some conventional weapon systems and some (relatively cheap) WMD. The wealthy developed country purchases conventional weapon systems and, in addition, modern (and expensive) weapon systems which can effectively counter the WMD of its rival. As before, we assume that the central planners in both countries can be either myopic or rational. By using current data on government budgets, growth rates, and real prices of weapon

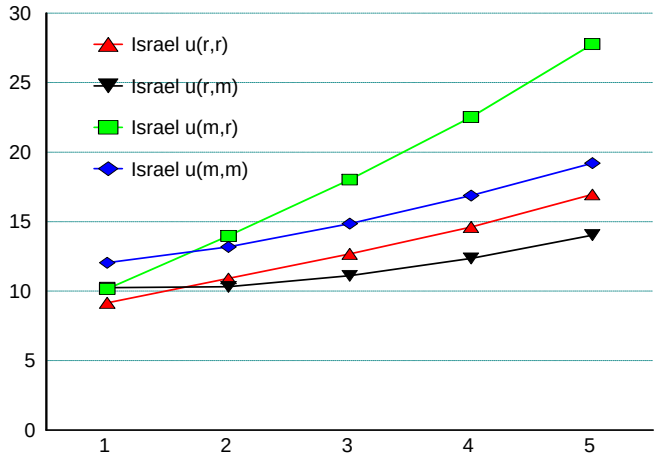

Figure 1: Israeli utility levels over 5 periods for 4 types of equilibria (m,m; r,r; m,r; r,m). Israel's strategy is listed second.

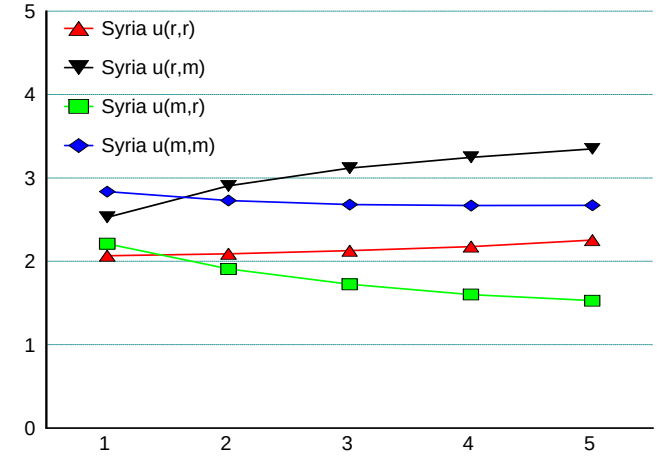

Figure 2: Syria's utility levels over 5 periods for 4 types of equilibria (m,m; r,r; m,r; r,m). Syria's strategy is listed first.

systems we compute the equilibrium solution of the Israeli-Syrian arms race for eight periods (each period consisting of five years) when both countries plan myopically (*m*,*m*), both plan rationally (*r*,*r*), when Syria plans rationally while Israel plans myopically (*r*,*m*), and vice versa (*m*,*r*). Figures 1 and 2 show the utility levels of both countries for the first five periods (25 years) of the planning horizon.[12]

During all five planning periods, Israel's utility level is higher under myopic planning than under rational planning (see Figure 1, where Israel's strategies are listed second). The same is true for Syria (Figure 2, where Syria's strategies are listed first).[13] Israel's utility is, generally, maximal when it plans rationally and Syria plans myopically (Figure 1). The same is true for Syria: its utility is, generally, maximal when it plans rationally and Israel plans myopically. This phenomenon explains why both countries will tend to plan rationally.

Table 1 presents the discounted values of the equilibrium utilities at the equilibrium solutions during the planning horizon (five periods). Clearly, although planning myopically is Pareto-preferred to planning rationally (compare the two numbers in the top-left corner of Table 1 to those in the bottom-right corner), the rational-rational (*r*,*r*) equilibrium is a dominant strategies Nash equilibrium. That is, because 11.2 > 10.3 and 7.9 > 6.9, Syria always prefers to choose rational planning if Israel first chooses either myopic or rational planning. Similarly, because 43.1 > 37.5 and 31.1 > 28.7, Israel always prefers to choose rational planning if Syria first chooses either myopic or rationally planning. The phenomenon of a prisoners' dilemma in the Syrian-Israeli arms race is robust. That is, it is present under a wide range of values that are centered around the parameters reported by KLTW.

Table 1: Discounted value of utility over 5 periods for 4 equilibria types (*m*,*m*; *r*,*r*; *m*,*r*; *r*,*m*)

Israel \ Syria	*myopic (m)*	*rational (r)*
myopic (m)	37.5 \ 10.3	28.7 \ 11.2
rational (r)	43.1 \ 6.9	31.1 \ 7.9

Note: The first-listed number is for Israel, the second for Syria.

Policy implications

Our results suggest that countries engaged in a noncooperative dyadic arms race will likely find themselves in an inferior (rational-rational) equilibrium in which each holds too high a stock of weapon systems without gaining the sought after military advantage on their respective rival.

This somewhat surprising result is due to the fact that (a) a dollar spent on weapon system procurement yields positive returns over several periods (as long as the weapon system lasts) while civilian government services benefit the public only once (most of government civilian expenditures are payments for salaries, social security, etc.) and (b) security is dependent on the rival's actions, that is, it equals one's military capability divided by the rival's capability. Hence, each dollar invested in arms procurement by one country will cause an increase in the procurement of weapon systems by the rival country (the "countering effect"). This phenomenon is at the heart of the arms race dilemma. A similar interpretation is given by Mendez (1997) in the context of a regional security organization. He argues that an increase in military power and deterrence by any member of the security organization is a "public good" for all member states of the organization, while it appears to be a "private bad" for the organization's opponents since it lowers their security.

Each (social welfare maximizing) country should identify the arms race dilemma and attempt to take it into account in its decisionmaking by using appropriate policy alternatives. There are two major policy options in this case. The first option is to reach some kind of agreement with the rival (directly or through a third party). The

second option is to institute force multipliers techniques.

Arms limitation agreements

In this article we demonstrate that the countering effect, where both countries engage in an arms race, results in a "rat race."[14] Acknowledging this outcome, a social welfare-maximizing policymaker should seek an arms limitation agreement to be monitored by a third party (a superpower such as the United States, or the U.N., for example) or a bilateral settlement with the rival.

KLTW provide an extensive discussion of two types of arms limitation agreements. The first type involves a third party that will compensate the less developed country for halting the procurement of weapon systems (or even reducing the existing stocks); the second type is an agreement in which the developed country compensates its rival for utility loss due to halting the procurement of weapon systems. In the latter case an agreement will exist only if both countries enjoy higher utility levels at the arm-limitation solution.[15]

Clearly, an arms limitation agreement in an asymmetric arms race is not a simple matter, and in reality the less developed country tends to cheat in these situations (particularly when the agreement was enforced by political and economic pressures and was not designed to its advantage). For example, in the agreement between North Korea and the United States (U.S.-North Korea Agreed Framework of October 1994) that set guidelines for the disarmament of the North Korean nuclear weapons, North Korea committed to freeze its nuclear proliferation policy. In return, the four parties to the agreement (North Korea, South Korea, China, and the U.S.) agreed to construct two light-water reactors to compensate North Korea for power supply lost and provide it with a yearly supply of 500,000 metric tons of crude oil. This agreement encountered major criticism in the United States and was labeled "surrendering to blackmail" on the grounds of high costs and lack of trust in the North Korean government. In November 2003, the construction of the two light-water nuclear reactors in North Korea was suspended in response to Pyongyang's failure to meet "the conditions necessary for continuing the project." Examination of the outcome of this agreement shows that although North Korea continued its proliferation efforts, it had to do so covertly and, hence, very slowly. That is, the United States achieved a substantial delay in the development of North Korean nuclear capabilities for a very low price (the annual cost of supplying the crude oil to North Korea was about $150 million in current prices). The total amount that was spent on this arms limitation agreement during the 11 years from its beginning to its termination was about $2.5 billion, donated by the 31 participating countries.[16] We believe that today, almost 15 years after the inauguration of the agreement, it can be considered a success despite its formal failure.[17]

Generally, the analysis here suggests that solving military/political conflicts by relying only on the military is an expensive and suboptimal solution.[18] The better approach in an arms race setup is to increase the array of options by adding new dimensions — political, economic, and other — to the menu of all possible solutions. We therefore conclude that policymakers should consider the option of arms limitation agreements. But arms races may evolve in unexpected ways. If a peaceful solution cannot be reached, policymakers should consider military strategies that ensure their country's military advantage.

Generally, the analysis here suggests that solving military/political conflicts by relying only on the military is an expensive and suboptimal solution. The better approach in an arms race setup is to increase the array of options by adding new dimensions — political, economic, and other — to the menu of all possible solutions.

Force multiplier methods

Several force multiplier methods are available. First, *security and deception*. A military authority that acknowledges the arms race dilemma may try to gain advantage by classifying its relevant military information, mainly its order of battle and its tactics. Deception is also a method of concealing information from an adversary. The first method (security) is a passive one and the second (deception) is an active one. In our context, a country that succeeds in misleading its opponent may gain an advantage on the battlefield without overinvesting in an open long-term arms race.

Second, *training, force motivation, readiness, and effective operational concepts*. The right size of the order of battle is crucial for the country's military effectiveness. However, a shortfall of training, personnel motivation, and readiness, or weak operational concepts, are also likely to yield less than optimal military might. Clearly, a strategy of investing in R&D and spending on procurement has effects that are different to one of investing in training, forces motivation, readiness, and proper operational concepts. While the first (R&D and procurement) is long-term in nature, the second can be considered as short-term. During routine periods the military is better off spending sufficient resources on procurement and R&D. Once it is forced to prepare for a war or an active conflict on short notice, it is better off spending resources on short-term investments. This distinction between short-term and long-term planning is particularly relevant for countries with compulsory service, where most of the military personnel are retained only for short periods of time.

Third, *intelligence*. High quality intelligence is essential for cost effectiveness of the military. Countering the opponent's deception and investment in weapon systems can be achieved, among other ways, by increasing investment in improved intelligence (the effective use of a wide variety of techniques to obtain and assess information).

Fourth, *asymmetric response*. Military authorities should seek low-cost responses

to both the low-cost routine operations of less developed countries and terror organizations and their high-cost weapon systems. Using the appropriate conceptual response to the rival's military investments and actions will likely coerce the rival into "too high" military expenditure and, thus, a reduction in its allocation of resources to its military apparatus.

Fifth, *technology modularization*. Investing in all-purpose military technology seems to be a force multiplier. If a country is forced into long-term planning, the military authorities, recognizing the arms race dilemma, should invest in developing technology modules that can be assembled in the future into as yet unknown coherent weapon systems which will be able to respond to future (not yet known) threats.[19] There are plenty of examples of very expensive weapon systems that were abandoned due to huge cost overruns (the Israeli Lavi fighter plane, the U.S. DDG-1000 guided missile destroyer, the U.S. Comanche helicopter, etc.) and should not have been developed in the first place. There are counter examples in which investments in modular systems proved to be winners (various UAVs, SAR capability for improved radars, GPS, and more).

Sixth, *late response effect*. Armies tend to respond very late to signals from a dyadic opponent. Early acknowledgment of such signals may facilitate reduction in the investment in all types of defense expenditure. Two examples supporting the late response effect are, first, the Strategic Defense Initiative (SDI) initiated by President Reagan in March 1983 that was set to use ground and space-based systems to deter the USSR from nuclear attack on the U.S. and achieve a decisive advantage in the dyadic arms race between the two superpowers. The USSR collapsed in 1991, partly because it could not compete with the SDI initiative. And yet the United States continued its vast investment in developing the SDI and still does so today. Second, Kagan, Tishler, and Weiss (2005) argue (and show) that the Israeli response to Syria's declining conventional weapon systems, as well as to the Syrian build-up of WMD, came very late. Nevertheless, Israel continues to spend vast resources to counter the diminishing threat of Syria's conventional army.

Seventh, *spiral development*.[20] This project management methodology allows the military to obtain technological and other capabilities faster and at a lower cost, its major advantage being the reduction of the lead-time from the laboratory to deployment. It consists of producing and deploying systems based on mature technologies. When deployed, the first modules of capability will meet some, but not all, of the weapon systems required specifications. Future modules (and improvements) will incorporate new technologies that have, in the meantime, matured and can be fielded at a later stage. The series of modules represents a spiral of increasing capability of the final weapon system.

Summary and crisis management

In this article we argue that an arms race is likely to result in a prisoners' dilemma equilibrium in which the rival countries plan rationally and, thus, overinvest in arms procurement. The likelihood of a war breaking out may be higher, as may the damages due to war, when the rivals in an arms race overinvest in arms procurement.[21] Thus, our discussion indicates the importance of conflict management that can be enhanced by the involvement of a third party (like the U.S., the EU, or the U.N.). The role of the third party may involve negotiating the arms limitation agreement and assisting in inspection, as well as guaranteeing the agreement, and offering economic incentives to the two adversaries in return for concessions in their arms development and procurement. As is the case in the arms race described by KLTW, the main purpose of the inspection, guarantees, and economic incentives is to unlock the prisoners' dilemma.[22]

Notes

Ronen Bar-El is at The Open University of Israel. **Kobi Kagan** is at the Department of Economics, Bar-Ilan University, Israel. **Asher Tishler**, the corresponding author, is Dean, Faculty of Management, Tel Aviv University, Tel Aviv, Israel. He may be reached at atishler@post.tau.ac.il.

1. Actual U.S. military expenditure in 2009 will be much higher; see Cordesman and Kaeser (2008b).

2. DoD (2008).

3. See, for example, the 30-year plan of the U.S. Navy (Kaeser, 2008), and the long-term U.S. Air Force plan (Cordesman and Kaeser, 2008a).

4. Richardson (1960).

5. Open-loop: See, for example, Brito (1972) and Deger amd Sen (1984). Closed-loop: See, for example, Simaan and Cruz (1975) and van der Ploeg and Zeeuw (1990).

6. See van der Ploeg and Zeeuw (1990).

7. See Panagariya and Shibata (2000).

8. See, for example, DoD (2003); Cordesman and Kaeser (2008b); Kaeser (2008).

9. Setter and Tishler (2006).

10. It is straightforward, although tedious, to extend the analysis to two different countries (see Bar-El, Kagan, and Tishler, 2008).

11. National security is usually measured as some function of the country's military capabilities relative to those of its rivals. Most studies define security as either the difference or the ratio between the country's stock of weapon systems and its adversaries' stock of weapon systems (see, e.g., Bolks and Stoll, 2000; Levine and Smith, 1997; Mantin and Tishler, 2004; Garcia-Alonso and Levine, 2007).

12. We solved the rational equilibrium solution for eight periods to eliminate the "last period effect" in the comparison of the two strategies.

13.Syria's strategies are always listed first and Israel's strategies are always listed second.

14. Akerlof (1976).

15. KLTW show that both options apply to the Israeli-Syrian conflict. Only the first option (involvement of a third party) applies to the North vs. South Korean conflict.

16. See KLTW.

17. The agreements that were reached in the Six Party Talks in September 2005 and February 2007 (Niksch, 2007) can be considered as a direct continuation of the earlier agreement and included a North Korean obligation for complete denuclearization in return for political and economic compensation by all the parties to the 2007 agreement.

18. See, for example, Brauer (2004).

19. Integrative technologies are good examples of these capabilities (see Setter and Tishler, 2006).

20. See, for example, DoD (2003) and Farkas and Thurston (2003).

21. See Brito and Intriligator (1984) and Wallace (1982).

22. See Levy (1985) for an analysis of the terms for mediating prisoners' dilemma conflicts and its application to the Namibian-South African conflict. See Sandler and Hartley (1995) for the role of monitoring and inspection in resolving prisoners' dilemma type arms races.

References

Akerlof, G. 1976. "The Economics of the Caste and of the Rat Race and Other Woeful Tales." *The Quarterly Journal of Economics*, Vol. 90, No. 4, pp. 599-617.

Bar-El, R., Kagan, K. and A. Tishler. 2008. "Myopic or Rational? Defense Budget Allocation and the Prisoners' Dilemma." Working Paper. The Institute of Technology and Society, Faculty of Management, Tel Aviv University.

Bolks, S. and R.J. Stoll. 2000. "The Arms Acquisition Process." *Journal of Conflict Resolution*, Vol. 44, No. 5, pp. 580-603.

Brauer, J. 2004. "Developing Peace Making Institutions: An Economist's Approach," pp. 137-153 in G. Harris, ed. *Achieving Security in Sub Saharan Africa: Cost-effective Alternatives to the Military*. Pretoria, South Africa: Institute for Security Studies.

Brito, D. L. 1972. "A Dynamic Model of an Arms Race." *International Economic Review*, Vol. 13, pp. 359-375.

Brito, D.L. and M.D. Intriligator. 1984. "Can Arms Races Lead to the Outbreak of Wars?" *The Journal of Conflict Resolution*, Vol. 28, No. 1, pp. 63-84.

Brito, D.L. and M.D. Intriligator. 1995. "Arms Races and Proliferation," pp. 109-163 in K. Hartley and T. Sandler, eds. *Handbook of Defense Economics*, Vol. 1. Amsterdam: Elsevier.

Cordesman, A.H. and H.U. Kaeser. 2008a. "America's Self-Destroying Air Power: Becoming Your Own Threat." Report. Center for Strategic and International Studies, Washington, D.C. http://www.csis.org/component/option,com_csis_pubs/task,view/id,4990/ [accessed 29 November 2008].

Cordesman, A. and H.U. Kaeser. 2008b. "Defense Procurement by Paralysis." Report. Center for Strategic and International Studies, Washington, D.C. http://www.csis.org/component/option,com_csis_pubs/task,view/id,5125/ [accessed 29 November 2008].

Deger, S. and S. Sen. 1984. "Optimal Control and Differential Game Models of Military Expenditure in Less Developed Countries." *Journal of Economic Dynamics and Control*, Vol. 7, pp. 153-169.

[DoD] United States, Department of Defense. 2003. "Instruction 5000.2: Operation of the Defense Acquisition System, Department of Defense." Washington, D.C.: Office of the Undersecretary of Defense.

[DoD] United States, Department of Defense. 2008. "RDT&E Programs (R-1): Department of Defense Budget Fiscal year 2009." Washington, D.C.: Office of the Undersecretary of Defense (comptroller).

Farkas, K. and M.P. Thurston. 2003. "Evolutionary Acquisition Strategies and Spiral Development Processes." *Program Manager Magazine*, July-August, 10.

Garcia-Alonso, M.D.C. and P. Levine. 2007. "Arms Trade and Arms Races: A Strategic Analysis," pp. 941-971 in T. Sandler and K. Hartley, eds. *Handbook of Defense Economics*, Vol. 2. Amsterdam: Elsevier.

Kaeser, H.U. 2008. "Abandon Ships." Report. Center for Strategic and International Studies, Washington, D.C. http://www.csis.org/component/option,com_csis_pubs/task,view/id,4816/type,1/ [accessed 29 November 2008].

Kagan, K., A. Tishler, and A. Weiss. 2005. "On the Use of Terror Weapons vs. Modern Weapon Systems in an Arms Race between Developed and Less Developed Countries." *Defense and Peace Economics*, Vol.16, No.5, pp. 331-346.

Kagan, K., A. Levkowitz, A. Tishler, and A. Weiss. 2008. "Evaluating Strategic Arms Limitation Agreements, with Applications to the Israeli-Syrian and the North vs. South Korean Conflicts." Forthcoming in *Defense and Peace Economics*.

Levine, P. and R. Smith. 1997. "The Arms Trade." *Economic Policy*, Vol. 12, pp. 336-370.

Levy, M.A. 1985. "Mediation of Prisoners' Dilemma Conflicts and the Importance of the Cooperation Threshold." *The Journal of Conflict Resolution*, Vol. 29, No. 4, pp. 581-603.

Mantin, B. and A. Tishler. 2004. "The Structure of the Defense Industry and the Security Needs of the Country: A Differentiated Products Model." *Defense and Peace Economics*, Vol.15, No. 5, pp. 397-419.

Mendez, R. 1997. "War and Peace from a Perspective of International Public Goods," pp. 307-336 in J. Brauer and W.G. Gissy, eds. *Economics of Conflict and Peace*. Aldershot, U.K. and Brookfield, VT: Avebury Press.

Niksch, L.A. 2007. "North Korea's Nuclear Weapons Development and Diplomacy." Report. RL33490. Washington, D.C.: Congressional Research Service. http://assets.opencrs.com/rpts/RL33590_20080121.pdf [accessed 29 November 2008].

Panagariya, A. and H. Shibata. 2000. "Defense and Welfare under Rivalry." *International Economic Review*, Vol. 41, No. 4, pp. 951-969.

Ploeg, van der F., and A.J de Zeeuw. 1990. "Perfect Equilibrium in a Model of Competitive Arms Accumulation." *International Economic Review*, Vol. 31, No. 1, pp. 131-146.

Plous, S. 1993. "The Nuclear Arms Race: Prisoner's Dilemma or Perceptual Dilemma?" *Journal of Peace Research*, Vol. 30, No. 2, pp. 163-179.

Richardson, L.F. 1960. *Arms and Insecurity*. Pittsburgh: Boxwood.

Sandler, T. and K. Hartley. 1995. The Economics of Defense. Cambridge, U.K.: Cambridge University Press.

Setter, O. and A. Tishler. 2006. "Introduction: Defense Research and Development in the 21st Century." *Defense and Peace Economics*, Vol. 17, No. 3, pp, 65-167.

Simaan, M. and J.B Cruz, Jr. 1975. "Formulation of Richardson's Model of Arms Race From a Differential Game Viewpoint." *Review of Economic Studies*, Vol. 42, pp. 67-77.

Smith, R., M. Sola, and F. Spagnolo. 2000. "The Prisoner's Dilemma and Regime-Switching in the Greek-Turkish Arms Race." *Journal of Peace Research*, Vol. 71, No. 6, pp. 737-750.

Snyder, H.G. 1971. "'Prisoner's Dilemma' and 'Chicken' Models in International Politics." *International Studies Quarterly*, Vol. 15, No. 1, pp. 66-103.

Wallace, M.D. 1982. "Armaments and Escalation: Two Competing Hypotheses." *International Studies Quarterly*, Vol. 26, pp. 37-56.

The Economics of Peace and Security Journal, ISSN 1749-852X
© www.epsjournal.org.uk – Vol. 4, No. 1 (2009)

Economists for Peace and Security

Economists for Peace and Security is a network of affiliated organizations. Affiliates exist in Australia, Canada, Chile, France, Germany, India, Israel, Japan, the Netherlands, Russia, South Africa, Spain, and the United Kingdom. Each is legally independent and determines its own membership criteria and activities. (This journal is an activity of the affiliate in the United Kingdom.)

A group of prominent economists have agreed to serve as trustees for EPS. They are: George Akerlof*, Oscar Arias*, Kenneth J. Arrow*, William J. Baumol, Barbara Bergmann, Andrew Brimmer, Sir Clive Granger*, Walter Isard, Sir Richard Jolly, Lawrence R. Klein*, Eric Maskin*, Daniel L. McFadden*, Robert S. McNamara, Douglass C. North*, Robert Reich, Thomas Schelling*, Amartya Sen*, William Sharpe*, Robert M. Solow*, Joseph E. Stiglitz*, and Janet Yellen (*Nobel Laureate).

For more information about EPS and its affiliates, please visit the U.S. affiliates' web site at www.epsusa.org.