

THE UNIVERSITY OF WARWICK

Original citation:

Harrison, Mark, 1949- (2013) The economics of coercion and conflict : an introduction. Working Paper. Coventry, UK: Department of Economics, University of Warwick. (CAGE Online Working Paper Series).

Permanent WRAP url:

<http://wrap.warwick.ac.uk/59303>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A note on versions:

The version presented here is a working paper or pre-print that may be later published elsewhere. If a published version is known of, the above WRAP url will contain details on finding it.

For more information, please contact the WRAP Team at: publicatons@warwick.ac.uk

warwick**publications**wrap

highlight your research

<http://wrap.warwick.ac.uk/>

Sep 2013

No.151

The Economics of Coercion and Conflict: an Introduction

Mark Harrison Department of Economics and CAGE,
University of Warwick Centre for Russian and East European Studies,
University of Birmingham

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

The Economics of Coercion and Conflict: an Introduction*

Mark Harrison**

Department of Economics and CAGE, University of Warwick
Centre for Russian and East European Studies, University of Birmingham
Hoover Institution on War, Revolution, and Peace, Stanford University

Abstract

This chapter introduces the author's selected papers on the economics of coercion and conflict. It defines coercion and conflict and relates them. In conflict, adversaries make costly investments in the means of coercion. The application of coercion does not remove choice but limits it to options that leave the victim worse off than before. Coercion and conflict are always political, but a number of key concepts from economics can help us understand them. These include rational choice, strategic interaction, increasing and diminishing returns, scale and state capacity, surplus extraction, and Type I errors. The chapter concludes that the economist's toolkit, although not complete, is useful.

Keywords: Coercion, Conflict, Games, Errors, Increasing Returns, Rational Choice, Scale, Surplus, Violence.

JEL Codes: D74, H56, N44, P26

* This paper introduces *The Economics of Coercion and Conflict*, by Mark Harrison, to be published by World Scientific in 2013/14. I thank David Hugh Jones and Peter Law for thoughtful comments on early drafts.

** Mail: Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. Email: mark.harrison@warwick.ac.uk.

The Economics of Coercion and Conflict: an Introduction

coercion, n. “1. a. Constraint, restraint, compulsion; the application of force to control the action of a voluntary agent ... b. Forcible restraint of (action) ... 2. Government by force, as opposed to that which rests upon the will of the community governed.”

conflict, n. “1. a. An encounter with arms; a fight, battle ... b. *esp.* A prolonged struggle ... c. (without *article* or *pl.*) Fighting, contending with arms, martial strife.”¹

The study of economics begins with trade. In an idealized market, sellers compete with each other for buyers. Under the rules of the market, competition is impersonal and non-violent. Each person makes their best choice, which can include staying out of the market; no one is forced to take part. As Adam Smith (1776/2005, p. 364) suggested, a result of the sellers’ pursuit of their own private profit is that resources are “led by an invisible hand” towards their best uses. In turn, the buyers’ pursuit of their own greatest satisfaction ensures that everyone gains who takes part willingly. In the outcome, the well-being of society is raised.

This model was first developed when England was still an agrarian society and is simplified in the extreme, yet it explains much of modern prosperity.

The world of coercion and conflict is, at first sight, utterly different. In this world “every man is enemy to every man ... and the life of man, solitary, poor, nasty, brutish, and short” (Hobbes 1651/1909, pp. 96-97). Each person’s behaviour is limited not by rules but only by their own conscience or, in its absence, by greed and fear. The immediate casualties are other people. In the longer term wealth is destroyed and society is impoverished. Even if a few lords or warlords retain a profit, there is never enough left over to compensate the losers.

Thus there are two worlds, one of free markets, industriousness, and enjoyment. The other is the world of tyrants, soldiers, slaves, and the lash. The gap between them is apparently unbridgeable. It turns out, however, that among the tools we have developed to analyze markets and

¹ OED Online. June 2013. Oxford University Press. <http://0-www.oed.com.pugwash.lib.warwick.ac.uk/view/Entry/35725> and <http://0-www.oed.com.pugwash.lib.warwick.ac.uk/view/Entry/38898> (accessed 27 August 2013).

corporations are some that can readily be turned to understand war and repression.

I am not the first to try to list these concepts and tools and explain their application. Economists, historians, and political scientists have recently contributed excellent applied and theoretical studies of the political economy of dictatorship and coercion (Wintrobe 2000; Gregory 2001; Lazarev and Gregory 2003; Dixit 2004; Gregory 2004, 2009). Others have done the same for interstate conflict (Offer 1989, pp. 7-20; Fearon 1995; Bueno de Mesquita 2006, Brauer and van Tuyl 2008, pp. 1-44, Smith 2009, pp. 19-53; Rockoff 2012, pp. 13-47; Garfinkel and Skaperdas 2012; Findlay and O'Rourke 2012).

If there is novelty in this chapter, it stems from examining coercion and conflict in the same framework. This is appropriate because the two are organically connected.

Adversaries, coercion, and conflict

To understand how coercion and conflict are connected, we need some definitions. What is conflict? There is a potential for conflict whenever two persons disagree, for example, about how to use or dispose of a resource. Resources can be of any kind – physical, financial, political, or emotional. Many such disagreements do not amount to conflict; they are fleeting, and are resolved quickly through compromise, so that they do not end in “an encounter with arms” as the Oxford English Dictionary puts it (cited above). Such readiness to compromise requires mutual recognition of the entitlements of the parties in disagreement, including the right not to agree. The fact that a particular dispute persists over a considerable period of time, however, is usually a signal that it is not of this transient and harmless nature; the parties cannot agree to each other’s right to disagree. They become adversaries in a “prolonged struggle” – one of the alternative meanings that the OED gives to conflict.

If unresolved disagreement over entitlements lies at the root of conflict, how is conflict resolved? Garfinkel and Skaperdas (2012) define conflict by the presence of “adversarial investments,” required for the “threat or use of coercive means.” In the words of the OED conflict is inherently violent: adversaries engage in “fighting, contending with arms, martial strife.” Here economists go further: violence is a possible correlate of conflict, but it is not necessary, because expectations matter: the threat of violence can be sufficient, provided it is backed by costly investments in “coercive means” (i.e. it is more than just cheap talk). In the presence of threats, conflict among adversaries can be worked out under duress but without violence, through negotiation or bargaining.

What is coercion? When I coerce you, I impose my choice on yours. As the OED puts it, it is “the application of force to control the action of a voluntary agent.” But the agent is still voluntary, and coercion does not deprive the agent of all choice. Rather, coercion means I force you to choose between the alternatives as I have defined them, not as you would see them. Moreover, each of the alternatives I allow you is inferior to the entitlement you have or seek. A highwayman points a gun at you and demands: “Your money or your life!” That is coercion. Within it, you have a choice: you’re free to choose whether to give me your wallet or die at my hand. The signal of coercion is that you cannot walk away and keep what you had. Whatever you choose, you will end up worse off than you were before. Thus, when I coerce you, it is intrinsic to the situation that you lose something to which you believe you were previously entitled.

(And, because entitlements rest ultimately on beliefs, conflicts over entitlement are always perceived asymmetrically. I stand up for my rights; you’re unreasonable.)

To summarize, coercion can be distinguished from free exchange, but the difference does not lie where many would naturally assume, in the victim’s absence of choice. The difference is that, when you and I trade freely, you do not suffer any loss of entitlement. Either the choices available to you through trade leave us both better off, or you can remain no worse off by choosing not to trade. Coercion differs from trade not in removal of the victim’s choice, but in the restriction of choice to a set of options that is strictly inferior to those available beforehand.

Coercion is often linked to repression, but the two are not the same. When I coerce you, I leave you worse off. When I repress you, I prevent you from signaling your protest – your dissatisfaction with the outcome. Repression can be political, but it can also be economic or financial. Applied to political markets, repression means the silencing of discontent. Economic and financial repression refers to the administrative controls that conceal market stress, which would otherwise be signaled by rising prices or interest rates. Such controls include price and interest caps and allocation by rationing and licensing.

Conflict, coercion, and repression have in common the exercise of force. Force requires “adversarial investments” and so is costly to produce and use. This is what makes them all negative-sum activities – as distinct from trade, which gives rise to a positive sum. After I have forced you to do what you would not have chosen to do willingly, our joint wealth is less than it was before. My power allows me to gain, only by imposing a loss on you. If you resist, and I impose my power on you by force or threats, our joint wealth is diminished, and the smallest loss I must impose on you in order to gain myself must be larger by that amount.

Just as conflict does not have to end in violence, coercion need not be violent either, but there is generally violence when coercion is resisted. It could be that you refuse the inferior choice set that I offer you (“Your money or your life”). This suggests that my power appeared inadequate at first, and required some demonstration or reinforcement. Or, conflict arises as a prelude to coercion, if I attack you (or you attack me) preemptively, to demonstrate superior coercive means. In other words, violence expresses actual or potential resistance to coercion and is therefore linked to the scope for coercion to fail.

In what I have written so far, coercion is never productive. That perspective is perhaps too narrow, or rather it accurately reflects a bias in my research. To explain further, in the example of the highwayman, coercion was used to steal your property, and this could be thought of as destructive not only of your personal entitlement but also of the general system of property rights. A system of property rights that can be enforced within the law is generally recognized as one of the hallmarks of a well-ordered society. In that case, it would seem a good idea for the public to apply some coercion to the highwayman. With the right combination of stick and carrot, a reformed highwayman might even gain in the long run. Then, coercion would be productive of a social benefit.²

In a well-ordered society, coercion enforces the law, but it is also restrained by the law. Here, law enforcement and legal restraint are two sides of the same coin. One importance of legal restraint is to underpin the community’s consent to the laws that are enforced. It is different in the other sort of society, the sort that is not “well-ordered.” There, coercion enforces the will of the government, not the rule of law. In that case government is, as the OED puts it, “by force, as opposed to that which rests upon the will of the community governed.”

Law enforcement and public coercion subject to the law are very important subjects for history and social science. On the whole, however, my research has addressed other aspects of public coercion, those that have lent support to arbitrary power including the unrestrained enforcement of arbitrary laws, when the government is relatively unconcerned about personal freedoms and private entitlements and does

² In a similar spirit Garfinkel and Skaperdas (2012) limit their definition of conflict to exclude those investments in adversarial means that bring external benefits to third parties. They intend this to exclude from the sphere of conflict the competitive tournaments and sporting events that raise productivity or provide enjoyment. Investments in policing under the rule of law also provide external benefits to third parties and on the same logic they should also be excluded from our definition of conflict.

not answer to the courts, the electorate, or anyone but itself. In the cases that I have studied, coercion was generally associated with repression, because the government did not regard the victims of coercion as entitled to any legal redress or legitimate expression of protest.

The chapters in this book are mainly about conflict and coercion that is organized among states and groups. The state is present in every chapter, sometimes as the body that seeks to monopolize violence over its subjects, sometimes as the main source of violence within global society.

In the long run of human history the monopolization of lawful violence by the state has been one factor in the great diminution of interpersonal violence in the world since early times (Gat 2006; Pinker 2011; North, Wallis, and Weingast 2011). But “lawful violence” must be understood carefully; it implies that alongside the state’s monopolization of violence come acceptance of the rule of law, including private property rights and other rights of the citizens vis à vis the state. A problem here is that the source of these laws is the sovereign state and the source of international law is agreement among sovereigns. All sovereigns are subject to the Weingast (1995) paradox: “A government strong enough to protect property rights and enforce contracts is also strong enough to confiscate the wealth of its citizens.” On the same reasoning, a state that is powerful enough to monopolize violence is also powerful enough to exercise it without restraint against its own citizens and against other states.

If public coercion and conflict among states are so closely related in theory, the connection should be reflected in the facts. And it is: Empirically, as discussed in Chapter 4, coercion and conflict often arise together because coercive political regimes are very often sources of conflict among states.

Before the modern era, virtually all states were authoritarian and warfare was the main function of the nation state. In European history, as Charles Tilly (1975, p. 42) observed, “War made the state and states made war.” This relationship between the nation state and warfare began to break down only as governments became democratic, and acquired additional functions, and at the same time the rule of law took the place of rule by men.

There are many reasons why states based on the rule of law (Kant 1795/1983) and democracy (de Tocqueville 1835/2000) might be expected to prefer peace. These range from moral considerations, such as reluctance to spill blood except in self-defence, to the self-interest of politicians and governments under democratic arrangements. Normatively, societies that subject themselves to the rule of law are likely to extend the same protection to foreigners as to their own citizens. The structure of democracy may also impose restraints on democratic leaders; empirically, leaders that lose wars are more likely to lose office in

democracies than in autocracies (Buono de Mesquita and Siverson 1995). Whatever the reason, it is “as close as anything we have to an empirical law in international relations” that “liberal or democratic states do not fight each other” (Levy 1988, p. 662).

Given the close links between coercion and conflict, it is not surprising that the toolkits that the economist applies to these two topics are largely the same.

Rational calculation

The single most important instrument that the economist brings to bear is the idea of rational choice. Without it, social science is nothing more than a psychological novel in which people are driven by inner forces they neither understand nor control. Are people sometimes driven by inner forces they neither understand nor control? Certainly. But there must be more to it than this, or the practice of social science becomes impossible.

When choices are rational, that is, calculated, each person is thought to weigh up the expected costs and benefits to themselves of the possible actions and to choose the one that offers the greatest surplus (or the smallest loss). As Clausewitz (1832/1982, p. 119) wrote:

War is a mere continuation of policy by other means ... War is not merely a political act, but also a real political instrument, a continuation of political commerce, a carrying out of the same by other means.

Consider the idea of optimization based on the computation of costs and benefits to oneself and the balancing of first order conditions, and call it “pure” rationality. Does that fully describe human nature as we find it expressed in empirical patterns of behaviour? Clearly not. It cannot account, for example, for people who destroy themselves or the things and people they love – at least, it cannot account for this in the form I have stated it. But is it a useful starting point? Very much so, because, as I wrote in another context (Harrison 2005):

If people do what they want, subject to the resource and information constraints that we can identify, and if we do not understand what they do, then we are missing something important and we should not be satisfied to throw up our hands.

Consider a country threatened with overwhelming force. For the sake of argument the country is Poland and the year is 1939. Hitler wishes to acquire the Polish homeland and this is something that most Poles would prefer to avoid. But avoidance is not on offer, so their real choice is

between surrender and fighting to defeat. Fighting is destructive and so must leave everyone worse off than in the case of surrender. Thinking on lines of pure rationality, a deal is available that would leave both sides better off than if war broke out, so both sides should accept it. The Poles should agree to give away the value of their homeland (which they will lose anyway), and accept from Germany in compensation a proportion of what will be saved by abstaining from resistance. For the Germans this should be an acceptable deal, since they will gain Poland and the Poles' compensation will cost them less than they would have spent on violence.

For some it is a problem that that's not what happened. The crazy Poles did not surrender, but fought until they were defeated. Our model of pure rationality is a failure, apparently.

Actually, no. The model has told us that we are missing something, and in that sense it is a success: it has extended our interface with the unknown. When we see human agents doing something that is costly, the principle of rationality tells us that they must expect to gain some future benefit (or avoid some future loss) that is equal or greater. If we missed it, we need to work out what it was.

What were we missing? I'll consider five different ways of understanding the rationality of Polish resistance to Germany in 1939. These are all factors in the motivation of conflict that are important for economists – and others – to understand, and at bottom they all point in the same direction. These factors are entitlement, reputation, identity, uncertainty, bounded rationality, and credibility. The first four would not automatically prevent coming to an agreement, although they would make coming to an agreement or less likely, more difficult, or more costly to Germany. The fifth, credibility, would seem to rule an agreement out, making war inevitable.³

Entitlement refers to the established fact that we value something more when it is already ours – when we consider that it belongs to us by right. This is one of the foundations of modern behavioural economics (e.g. Kahnemann, Knetsch, and Thaler 1991). This suggests a reason why the Poles were ready to fight: they considered Poland was theirs, and the value of resistance was the defence of their entitlement.

³ Fearon (1995) lays out further grounds that may prevent agreement and lead to war within a rational-choice framework:.. Even when agreement is possible, each side still wishes to secure the best agreement possible. To achieve this, the government may try to exploit private information and strategic misrepresentation in such a way that agreement is impeded. These do not seem applicable to Poland in 1939, but can still be relevant elsewhere.

Another factor can be *reputation*. According to the historian Anna Cienciala (2011), Polish leaders of the time such as foreign minister Józef Beck put honour above the avoidance of conflict. If they had given away what was theirs without a fight, accepting a vassal state for Poland, they would have lost their honour. In this view the value of resistance was the safeguarding of their honour, even if they lost everything else.

The concept of *identity* fills a gap in economics (Akerlof and Kranton 2000) Economists often describe economic behaviour as the rational pursuit of self-interest. But who are we, and how can we define our self-interest without first establishing who we are? "I could not live with myself if I did not ...": their own self-identification is often a person's most precious possession, and occasionally one for which life itself must be sacrificed. Thus, the paradox of behaviors that lead to self-destruction including suicide, suicidal terrorism, and suicidal heroism, cannot be understood except in terms of the self that such behaviours defend (Harrison 2006). If they had not fought for Poland, the Poles would not have been Poles. Thus the value of resistance was to safeguard a precious identity that would otherwise have been destroyed.

Our analysis would be grossly incomplete if we did not take into account that identity has many dimensions. The Poles were not just Poles. They were also men, women, and children; elites and poor; Jews and non-Jews; soldiers and civilians; religious and secular; nationalists and internationalists; and so on. They surely did not weigh up everything the same way. Poland was a young nation, and national entitlement, honour, and identity are the currency of nation-building leaders. Honour was quite possibly of more value to the elite than to the foot soldiers that were also going to have to die for it. Disputes are more likely to be resolved through violence when identities become polarized (Sen 2006).

This reminds us, finally, that it is not nations that optimize, but persons. Each person frames the choices of others; each person in the nation must decide their own self-interest, given the choices that they expect others to make. This interaction is something else that was missing from our first attempt at "pure" rationalism.

One factor in Polish resistance may have been *uncertainty* over future costs and benefits. Our model was framed by the assumption that if it came to a conflict the Polish defeat was certain. In reality war is a gamble: "No plan of operations can look with certainty beyond the first meeting with the major forces of the enemy," wrote the older Moltke (cited by Holborn 1986, p. 289). Every Pole remembered the "miracle on the Vistula" of 1920, when the Red Army's apparently unstoppable advance on Warsaw was halted by an unexpectedly successful Polish counter-attack. It is clear that Germany's Wehrmacht in 1939 was far more of an existential threat to Poland than Soviet Russia's Red Army in 1920. There

was no room for Polish optimism in any objective reading of the balance of forces. They were almost certainly going to lose.

Our *bounded rationality* restricts our capacity to make the right decision (for a recent summary see Kahneman 2012). The limits are cognitive as well as computational. Among other things, we overvalue small probabilities – for example, the chances of a second “miracle on the Vistula” within two decades of the first. Members of the Polish elite, with shared experience of service in the dragoons, were likely to see a cavalry charge as part of the solution to Poland’s problem, even if Poland’s problem had no solution at all.

Finally, any deal that Germany could have put on the table would not have been *credible*. In fact, the secrecy of Germany’s war preparations, including the secret pact with Stalin that opened Hitler’s way into Poland, shows that Hitler did not even try to offer the Poles a deal that could avoid war in 1939. Nor would the Poles have been wise to accept one if it was offered. Whatever agreement Hitler might have made before a Polish surrender, he would surely not have kept his word afterwards. This was the same Hitler that had recently swallowed Austria, the Czech Sudetenland, and Czechoslovakia itself. He had a clear record of making promises when it suited him to do so and breaking them afterwards when it suited him. Any deal that Hitler offered to the Poles would not have been credible. They would not have responded to the attempt, and he did not bother to try.

The model of rationality that we began with exemplifies the saying of the statistician George Box (1987, p. 74): “All models are wrong; the practical question is how wrong do they have to be to not be useful.” The idea of rational choice is wrong because it is oversimplified, but it is not so wrong that it is useless. Used properly, it can still guide us to a fuller understanding of the choices that people make, even in situations characterized by ignorance, anxiety, and existential threat.

Strategic interaction

Strategy has many meanings in the modern world, most of them only distantly related to its roots. Strategy begins with prediction of the future, and many public and private organizations claim to have “strategic plans” that are really just aspirational descriptions of the state in which they would like to be in a few years’ time. There the word “strategic” is empty, because any plan is about the future, and calling it strategic adds nothing.

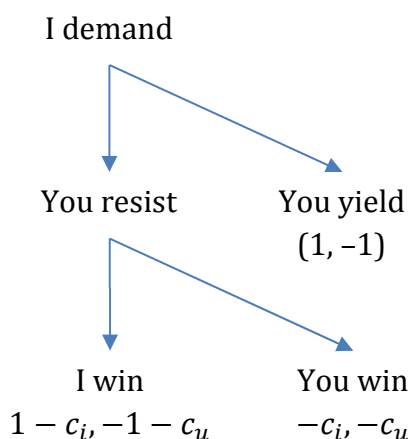
The origin of the word strategy lies with generalship. The main task of the general is to predict and counteract the action of some adversary or rival. He must decide his best action (for example, to advance along a given front), not knowing the enemy’s deployment but trying to predict it

based on the fact that the enemy is simultaneously trying to predict the general's line of advance in order to decide where to deploy his defences. "What does planning mean in war?" Answering his own question, Stalin's Marshal Ivan Konev (1970) wrote, "We make plans alone but carry them out, if one can put it so, in company with the enemy, that is, taking into account his counteraction." It is this simultaneity and mutual engagement that captures the spirit of strategic interaction.

When we act strategically, we may or may not reveal our true preferences, or we may even conceal them. We may line the streets to cheer a hated tyrant. We may prefer peace but nonetheless plan for war.

Modern social science uses game theory to think about strategic interaction. While many particular games are studied in the contexts of conflict and coercion, one is canonical (Fearon 1995). Suppose you have a resource that I want, that you would not give me freely, but only if I coerce you. In a dispute among states this resource might be your land; in a domestic setting it might be your labour. Set the value of this resource equal to 1. We can resolve our conflict with or without violence. Violence costs me c_i and you c_u (think of these costs as proportions of the value at stake, which is the unit of account).⁴ Then the structure of payoffs might look as in Figure 1.

Figure 1. Conflict over resources



The game played in the figure works like this. If I make my demand, and you yield, the resource changes hands from you to me. If you resist and I win, I gain the resource less the cost to me of violence; you lose the resource and the cost to you of violence. If you win, nothing changes hands and we both lose the costs of our own violence.

⁴ Other assumptions are that we are risk-neutral and we have common values and full information, so there is no scope for the private information and strategic misrepresentation issues discussed by Fearon (1995).

This game yields simple, intuitive results that also illustrate the principle of strategic interaction. Should I make my demand? On the face of it, if I expect you not to resist, or if I expect to gain from a fight. Both of these depend on the probability p of my winning if it comes to violence. Find our expected payoffs from violence by applying p to my winning payoff and $1 - p$ to yours. Then, you will be willing to fight me if you expect to gain from it, i.e. $1 - p - c_u > 0$; otherwise, you will yield without a fight. I will be willing to fight you if I expect to come out ahead from a conflict, i.e. $p - c_i > 0$. Unfortunately for the prospects of peaceful conflict resolution, both of these conditions can hold at the same time; they require only that the resource will not be completely consumed by the costs of violence, i.e. $c_i + c_u < 1$. Violence can be avoided only if it is so costly in prospect that a fight would destroy everything.

Missing from this setup is the option to bargain peacefully. That scope for negotiation ought to exist is shown by the fact that beforehand the value at stake is one unit whereas, after a value-destroying conflict, the most there can be to distribute between us is $1 - c_i - c_u$. By agreement we can save the value that otherwise would be destroyed ($c_i + c_u$) and share this among ourselves. My outcome can then be better than $p - c_i$ and yours can be better than $1 - p - c_u$. As we have already discussed, however, the likelihood of a bargain is dramatically reduced or eliminated altogether by the problem that rulers who are willing to use violence to get their way, whether domestically or internationally, may also be unable to commit to agreements that avoid it.

This game is sometimes called divide-the-dollar (two players bid for a dollar; if their bids sum to less than the dollar, they receive their bids, or otherwise nothing). While international relations provide its usual setting, the divide-the-dollar game can also be used to illustrate the outcomes of domestic conflict under a coercive regime. The failure of Soviet economic reforms provides an example. The “dollar” (or ruble) here is the additional resources that would have been supplied if Soviet managers had honestly reported their assets and capabilities to the government, if their workers had given their working time fully to meeting management goals, and if the government could have committed to reward the managers and workers in return, in other words, to divide the ruble with them.

In 1929, Stalin began to transform the Soviet economy into a centralized command system. By centralizing economic controls he was able to enforce a near-complete monopoly of capital and a monopsony of wage labour and farmers’ food surpluses. Between 1928 and 1940, increases in compulsory work norms and production quotas and reductions in real wages brought about a transfer of roughly 30 percent of Soviet GNP (measured at factor costs of 1937) from household

consumption to other uses, particularly investment and defence (Bergson 1961, p. 237).

In a market economy many workers would have responded to such a squeeze by looking elsewhere. They would have been free to find non-government buyers for their products or services or in the last resort to emigrate. Under communist rule these options were criminalized and heavily punished. In the Soviet Union most workers were limited to “inside” options which meant, for example, showing up to work but reducing effort and attention to quality. (Some important exceptions are the subject of Chapters 7 and 8.)

There was a vicious circle, summed up in the Soviet-era joke: “We pretend to work and they pretend to pay us.” At intervals over the next half century Soviet leaders made various efforts to unlock this vicious circle. The key, they believed, was to reward managers and workers for taking on more ambitious norms and quotas. Among these efforts, most prominent were the rate-busters’ movement inspired in 1935 by the coal miner Aleksei Stakhanov, and the postwar management and incentive reforms of 1965 sponsored by prime minister Aleksei Kosygin.

The problem was that government promises to divide the ruble by letting the workers keep the rewards for increased effort were not credible. Enterprises that played the government’s game tended to be rewarded only in the short term. Once they had revealed what they were truly capable of, they had given away their main bargaining chip in negotiating the next quota. In the longer term they were given higher norms and quotas and so lost their promised share of the ruble (Harrison 2002).

In this sense (if no other) the structure of the problem that Soviet workers faced with government promises in 1965 was identical to the one that the Poles faced with Hitler’s demands in 1939: after implementing any bargain, they could anticipate that the terms would be shifted against them, and they would have given away their power to resist, so it was better not to make any deal and slug it out. As a result the equilibrium outcome of seeking to divide the ruble in the Soviet version of the game was low effort, low productivity, and endemic, unresolved conflict over norms and quotas.

In game theory terms the models shown here are highly simplified.⁵ They serve mainly to illustrate the structure of strategic interaction. Although very simple, they have the virtue of suggesting how

⁵ In more advanced games we find that players will behave differently given opportunities to throw dice, anticipate repetition, learn about each other, form beliefs, build reputations, and purvey misinformation. For discussion of historical applications see Greif (2006).

governments and their adversaries had no choice but to put themselves in each other's shoes in order to identify their own best choices, whether the adversaries were foreign leaders or their own subjects. This is the essence of strategic interaction.

Increasing and diminishing returns

Strategic interaction has other features that make clear its affinity to the subject of coercion and conflict. One is the importance of activities in which there are increasing returns at the margin. Most economic activities are characterized by diminishing marginal returns, and most economic models assume it.

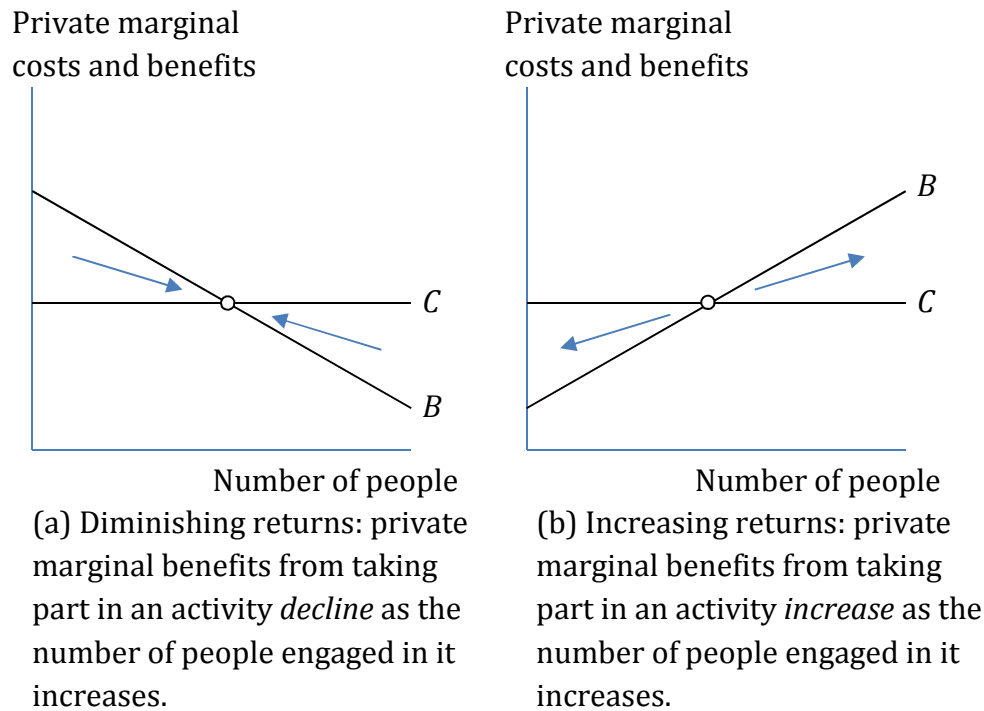
"Diminishing returns" sounds like a gloomy sort of thing, but in practice it is what keeps society on the rails most of the time.⁶ I like a glass of wine, but I enjoy the second glass less than the first. As a result I limit my consumption. Beyond a point, I stop. In other words, processes characterized by diminishing returns tend to be self-limiting and arrive sooner or later at a point of stability (which economists call equilibrium). The same processes would become explosive in the presence of increasing returns. If I enjoyed every glass more than the one before, I would quickly drink myself into unconsciousness. At the same time, if I did not enjoy the *first* glass very much, I might never start. In other words, my life would go to extremes: I would either be completely sober or there would be no stopping me.

In conflict and coercion there are often increasing returns, although in a sense that differs a little from the example of my drinking habit. In this case it is not my effort that makes returns diminish or increase, but the efforts of others. Figure 2 has two panels. Panel (A) illustrates diminishing returns. Think of an industry that has a given market for its product. Working in that industry has certain costs, which are also taken as given. When few workers are engaged in the industry they command a high wage, so it is profitable for others to join them. Workers flow into the industry, but this also reduces the wage. Conversely, with too many workers engaged in the industry the wage is driven below the cost of working in the industry, and workers are driven out. Whichever side of the equilibrium point you begin from, the process stops in the middle, where the number of people engaged is such that the benefit to the

⁶ David Hugh Jones highlighted this point for me in a blog post, "Very simple thoughts about the politics of crisis," (7 April 2013), available at <http://davidhughjones.blogspot.co.uk/2013/04/very-simple-thoughts-about-politics-of.html> (accessed 7 April 2013).

joining worker just matches the cost of doing so. The result is a stable equilibrium. This is the normal world of everyday production and trade.

Figure 2. Increasing and diminishing returns



Panel (B) illustrates increasing returns. This covers situations where the private benefit to a person of doing something increases with the number of other people that are doing the same thing at the same time. Imagine a war of attrition, in which two armies grind each other down. The battle is won by the last man standing. In that battle, the more soldiers join in on my side, the more likely I am to be that last man standing. But if you, the soldier next to me, turn and run, that exposes me to the enemy and reduces my chance of survival. As each additional soldier falls out of the ranks, they reduce the benefit to me of continuing to fight. As I watch you, I ask myself continually: Should I run too? If that calculation ripples along the line, the army unravels suddenly into a fleeing mob, and the battle is lost.

Under these circumstances war becomes a game of all or nothing: we all join in together, or we all run for it. There is no balance between the two.

A similar example works for political coercion. My block leader orders me to fall into line with my neighbours and sing with them: "All hail our glorious leader!" How should I respond? My first response may not be to question whether the leader is really glorious but to find out what others will do. If my neighbours are evenly divided, I might feel free to go with my inner beliefs. If I hear they will all march and sing, my private feelings about our leader may be overwhelmed by the likely cost to me if I make a

stand on my own and register dissent. If other people show resistance, however, even if I am the biggest coward on the street, I should bear in mind that it is also dangerous to be a hated tyrant's last supporter.

In other words, for anyone who is not an instinctive rebel, conforming has returns that increase with the number of other people that conform. The trick is to predict the standard to which one should conform. Hence the common phenomenon of "weathervanes" – people whose professed opinions follow the prevailing orthodoxy.

In the explosive world of increasing returns there is no longer just one equilibrium. Rather than getting slightly inebriated, I will either be completely sober or unconscious. The army will either hold or collapse. The regime will be stable until it falls suddenly; the crowd will be fickle, and everyone will be strongly "for" until they turn "against." To make my own best choices I must continually try to work out how everyone else will behave, knowing incidentally that they are also trying to predict my own behaviour, which is more or less the definition of strategic interaction. This is, finally, a world of agonizing dilemmas: every day I must weigh integrity against survival.

The role of increasing returns should not be overstated. In wartime, many returns continue to diminish, and this accounts for important regularities. The airplane was an invention of the twentieth century, and in the 1930s most major powers gave a lot of resources to this new weapon. But beyond a point they stopped, and also modernized their armies and navies. This recognized diminishing returns to air power: sooner or later, they got a point where they figured the military value of one more plane would fall below the same value assigned to one more of the traditional guns or ships. They also limited rearmament somewhere around the point where the expected value of one more dollar given to the army fell below the value of the same given to agriculture or public works. Even in total war, this principle continued to apply: even when the threat was existential, no country gave literally everything to the war.

The returns to coercion also diminish beyond a point. An extreme case of coercion is provided by the residents of Stalin's labour camps. Behind the barbed wire of the Gulag, men and women were stripped of nearly all possessions, nearly all family contact, and nearly all control over time and movement. All their choices left them worse off. Yet, what did experience show? You can put a man in these conditions, and point a gun at his head, and it can still be hard to make him work. As Chapter 10 shows, the expansion of the Gulag saw the evolution of ever more varied and complex incentive schemes that were designed for no other purpose than to persuade the slaves to work. The stick and the carrot were not alternatives. Even when the main emphasis fell on the stick, Stalin still had to give out a few carrots. In other words, there were limits to

coercion. Not even a totalitarian dictator could overcome diminishing returns.

In short, both diminishing and increasing returns are essential instruments in the toolkit that everyone needs to understand conflict and coercion, and the problem is largely to understand which one to apply in each context.

Scale and state capacity

In conflict, scale might seem unambiguously good. Since Clausewitz, modern strategists have thought of the ideal military campaign as the application of overwhelming force to the enemy's weak points. Sometimes it worked quickly, sometimes not. When it dragged out, modern conflict often became a war of attrition in which victory would be claimed by the last man standing on the field of battle. Whether warfare took the form of a lightning blow or a bloody slog, why would less ever be worth more than more?

In fact, the wars of the twentieth century have been historically exceptional. If we go back through history it is relatively easy to find battles in which smaller armies were able to inflict overwhelming defeats on much larger forces. At Marathon (490BC) a Greek army numbered in thousands routed a Persian force of five to ten times the size. The Carthaginians achieved a similar victory over Rome at Cannae (216BC). The English did it to the French twice in a century at Crécy (1346) and Agincourt (1415).

In such cases the smaller force defeated the larger one not so much because of an advantage of equipment or terrain but by disrupting the opposing force to a point where it suddenly changed state, from a proper army to an uncoordinated rabble that could easily be slaughtered. At times, it seems as though the smaller force was easier to hold together by ties of discipline and comradeship than a larger force where each had a smaller chance of a part in the fighting and felt less commitment and obligation to the cause of the whole.

Something that defines modern war, dating it from the time of Napoleon, is the emergence of mass armies that could be held together in defeat. Napoleon's retreat from Moscow, the Confederate retreat from Gettysburg, the Japanese retreat to the home islands from South East Asia and the South Pacific, and the German retreat from Stalingrad and the Caucasus are all examples of vast bodies of men that somehow failed to disintegrate despite pulverizing defeats that were piled one on top of another for months and years.

If every great army of the twentieth century had disintegrated on suffering a great defeat, as the French army did in 1940, both world wars

would have been over within months. The fact that they often did not, but were held together and retained their military capabilities, ensured that modern war became a war of resources, in which those armies won that could throw the greater number of men and munitions into battle and still had men standing on the battlefield when the battle was over.

In short, to have a war of resources, two great armies must meet, fight, and neither of them fall apart – something that seems to have happened quite rarely before the twentieth century. The question that follows is: When do armies fall apart, and when not? This is not an economic question, but it is one that economists have considered. Brennan and Tullock (1982) cast the soldier's motivation in the framework of a prisoner's dilemma: unless some other factor enters the soldier's calculation, whether his comrades stand and fight or turn and run, in the face of battle his best choice is always to desert. Officers build comradeship and discipline to prevent this from happening.

In history defeat has come most often, not at the point when all the soldiers are killed, but at the point when discipline and mutual obligation fail so that the army is suddenly pulled apart and is turned into a leaderless rabble. Modern armies, however, have evolved into resilient networks based on morale and commitment that can overcome the prisoner's dilemma and hold the ranks together even in quite extreme circumstances. When this factor was equally present on both sides, improved leadership and morale could no longer make the difference between victory and defeat. It was under these conditions, and only then, that scale and resources could become the decisive factor in two world wars.

Scale in modern warfare has implications for the state. To assemble an army of modest size for a season, fight a battle, and then disperse is one thing. To keep a mass army supplied and moving from battle to battle in one season after another is another. As shown in Chapters 1 and 2, the demands of twentieth century warfare have sometimes required the mobilization of half or more of a country's resources for years on end. To achieve this required immense state capacity.

State capacity has many dimensions, financial and administrative as well as coercive (Besley and Persson 2009). The raising of war loans and taxes and the coordination and direction of human and physical resources into the supply of modern wars produced a vast derived demand for professional, non-corrupt administrative services. Autocracies of the pre-modern era had very limited capacity of this type. It was the seventeenth century revolutions, guaranteeing private property and constitutional rule in Northwestern Europe, that also gave unprecedented power to government to tax and borrow in the event of war. In military and naval power the liberal democracies pulled ahead of their rivals, keeping the

advantage through World War I. As shown in Chapters 1 to 4, however, by World War II some non-democracies had caught up and closed the gap, exploiting modern nationalism and repression to mobilize resources with still greater intensity, at least for short periods of time.

If pure scale has had the edge in modern warfare, does the same carry over to the exercise of coercion? That is, is it better to be the tyrant of a big country compared with a small one? To rule a large country is certainly more complex, and complexity is costly. As ruler of the world's largest territory, Stalin had to commit major resources to the control of the vast Soviet space and to border regions inhabited by ethnic minorities from the Baltic to Mongolia and China. To ensure the implementation of his orders, he had to rely on ever lengthening chains of command and multiple, overlapping agencies of control (Markevich 2012). Compare Stalin's problem with that of Alyaksandr Lukashenka. Once a political instructor of the KGB border troops, Lukashenka rose to become the autocratic ruler of the former Soviet republic of Belarus. In Belarus, he governs no more than three percent of the population and one percent of the territory of the old Soviet Union. That is surely a simpler task than Stalin's.

Perhaps it is better to be a pike in a small pond than a minnow in a lake. But it is better still to rule the lake. The reasons are things we have already seen. First, dictators appear to have a propensity for war. Second, in modern war, scale counts.

Surplus extraction

The economic analysis of coercion and conflict encourages, in some ways, a return to Marxian ideas about the economy. One of Marx's (1867/1974, pp. 173-182) central obsessions was his idea of surplus value: the value that a worker creates for the employer, above and beyond his own value (or maintenance cost). This idea was so important to Marx because he regarded all economies as, in essence, mechanisms whereby an elite of one kind or another extracted a surplus from the working poor.

Modern economics has abandoned this view as oversimplified to an extreme. The model is not only wrong; it is not useful in most contexts. In the Marxian perspective there is only one kind of surplus, called profit, one source of surplus, called labour, and one recipient, the capitalist class. In the competitive market economy of today's textbooks every transaction gives rise to a surplus, and these accrue to the producers and consumers that are party to every transaction. In other words, there are surpluses everywhere and they accrue to everyone; they are not the monopoly of one class.

At the same time, there are some economies that are undeniably extractive. In these economies there is a restricted class of proprietor, or perhaps a dictator, that controls production and distribution in a coercive way, for its own exclusive benefit. In extractive economies there is often no clear distinction between politics and economics, because the security of the proprietors rests on forcible restriction of the choices available to members of the underclass, not on universal property rights and the rule of law. In other words, the same people that want to monopolize the economy must also monopolize the state.

Manorial economies, plantation and slave economies, and communist command economies seem to fit this template better than others. Inspired by the example of Stalin's Soviet Union, Mancur Olson (1993) developed the idea of a "proprietary dictator" that defeats rivals, monopolizes a territory, and extracts a surplus from it. Where competing predators would simply ruin the territory, a monopolistic ruler would rationally seek to expand the territory and even be willing to pay for public goods and infrastructure to make it more productive, because this will raise the output of the territory and so increase his revenues in the long run.

In this model economic development is not the dictator's purpose, but it is an incidental by-product of the dictator's desire to increase the surplus over time. As the example of the Soviet Union suggests, extractive regimes have sometimes been associated with prolonged growth spurts, but they have never given rise to the modern economic growth that persists unbroken for a century. One important reason may be that the monopolization of resources by an elite is antithetical to the "creative destruction" of market competition (Açemoglu and Robinson 2012, p. 94). The analysis of the downfall of the Soviet model in Chapters 11 and 12 strongly suggests a link with its failure to respond to the opportunities presented by new technologies and new industries that were spurring market-economy development and globalization in the 1970s and 1980s.

Marxists used to lay great emphasis on international movements of economic surpluses. Economic historians somewhat lost interest in this subject after it became clear that international trade before the nineteenth century was simply not large enough to be a channel for significant transfers from one country to another (O'Brien 1982).

Quantitative history suggests that this topic became much more important in twentieth century conquest (Lieberman 1996). Germany's occupation of much of Europe in World War II arose out of long-term plans for territorial expansion but the design of the occupation regime was motivated more narrowly by plans to extract a surplus of food and resources from the occupied countries by coercive means and transfer it to Germany or the German Army. A new study by Klemann and Kudryashov (2012, p. 104) estimates the actual contribution at around

one third of Germany's war costs. (Further calculations would be required to estimate the additional surplus extracted within Germany from the slave labourers imported from occupied Europe.)

The importance of the economic surplus for our subject goes beyond its role in communist and fascist states. In total war, all economies became extractive. Even in democracies, the elected government had to find resources for the army and its equipment. The resources had to come from somewhere. Beyond a point, the working poor had to be squeezed because they were the only class of society numerous enough to supply the necessary resources.

In short, to understand what overlords and dictators maximize, and how they maximize it, and to understand what all governments must maximize in total war, Marx's intuition remains useful. Although neither of the authors is a Marxist, the surplus model that we apply to Soviet labour coercion in Chapter 10 is consistent with a Marxian approach. To repeat George Box, all models are wrong – and that certainly includes the model that Marx developed. But wrong models can still yield insight, and the Marxian model is one of these. It too has a place in the economist's toolkit.

Type I Errors

One way that social science advances is by formulating and testing hypotheses against data. Evolution has programmed us all to formulate hypothesis all the time, that is, to look for patterns around us and find significance in chains of events. Our ability to do this has given us huge advantages as a species. Without it there would be no science, technology, religion, or culture: we could not understand the universe, manipulate the laws of physics, build social relationships, or peer into the future. Imposing significance upon chains of events so as to look into the future is something from which an extraordinarily wide range of people can make a living both for themselves and for others, for example politicians, priests, tipsters, entrepreneurs, novelists, journalists, economists, and historians.⁷

Our propensity to discover patterns in what we observe is not only creative. It is also a source of risk. The risk is realized when we make

⁷ Here I draw on material that I wrote for a blog post asking "How Can We Get to See What's Coming Round the Corner?" (August 6, 2009), available at http://blogs.warwick.ac.uk/markharrison/entry/can_we_see/, and also (from the time before blogs) a short unpublished paper, "The War Against Terrorism: Type I Versus Type II Errors" (1 November 2001), available at <http://warwick.ac.uk/markharrison/comment/>.

mistakes. The mistakes we make can be of two kinds, which statisticians call Type I and Type II. These concepts arise from statistics rather than economics, but they are of fundamental importance to the economics of coercion and conflict.

- An error of Type I is to see a pattern in the data where in reality there is none. On a dark night you hear the footsteps of an attacker behind you, but actually no one is there.
- An error of Type II is to miss a pattern in the data that is actually present. Another night, an attacker comes behind you. You're distracted and unaware until it's too late.

Which is worse? Both carry dangers. People that persistently make errors of Type II can be described as complacent. They fail continually to see the risks of terrorism and war and the environmental and health hazards around the corner. As a result, bad things are allowed to happen before we take action. Complacent people wake up in the end, only after they and others have suffered significant damage which, given better foresight, could have been avoided. They lock the stable door after the horse has bolted. Thus Type II errors can be corrected but correction is costly.

Type I errors are bad too. The people that persistently make errors of Type I can be called paranoid. They see enemies and conspiracies everywhere. The world is complicated; they think the world is a complicated gun that someone, somewhere is pointing at them. They don't recognise that many things are random; they won't accept that bad stuff often happens by accident or by mistake. Nor are they able to accept that a lot of what happens is not really about them at all. Trying to forestall the bad things they anticipate, people who are gripped by Type I errors are often themselves responsible for doing things that are as bad or worse, such as mobilizing society to solve problems that do not exist or attacking people who are not enemies and would prefer to keep themselves to themselves.

Finally, there is a trade-off between the errors of Type I and Type II. If we try harder to avoid one, other things being equal, we will inevitably end up making more of the other. The only way to reduce both types of error at the same time is to have more and better data. As long as data collection is costly, we will never have complete information and we will always have to balance one kind of error against the other.

Attitudes to the consequences of Type I and Type II errors have proved to be important markers of social attitudes. Understanding the balance that is struck between Type I and Type II errors can tell you a lot about the nature of the society under study. In a court of law, the research hypothesis is that the defendant is guilty. Let's say the crime is a public order offence. Here the court can make two mistakes. It can convict the

truly innocent (a Type I error) or acquit the truly guilty (a Type II error). In most liberal democracies, the courts are stacked against the prosecution, which must prove guilt beyond reasonable doubt that the defendant incited a riot. In other words, prioritizing the rights of the accused implies that we tolerate Type II errors in order to avoid errors of Type I, which we see as particularly costly.

In contrast the historical record of dictators is that, since they are not particularly concerned about the rights of citizens, but very concerned to protect their own skins, they tend to operate criminal justice in a way that scoops up and penalizes many fairly innocent people in order to ensure that every guilty person is taken out of circulation. This has important implications for the study of coercion. As discussed in Chapter 10, the coercive regimes of the dictators that preside over extractive economies typically accept large numbers of Type I errors as the price of avoiding the Type II errors that might undermine their own security. In contrast, more inclusive economies that respect the rights of individual citizens tend to have arrangements, such as presuming innocence in the absence of proof of guilt, that avoid Type I errors at the price of sometimes making Type II errors.

The asymmetry in the consequences of Type I and Type II errors has strongly affected the research strategies of natural and social scientists. Most scientists would rather not be personally responsible for precipitate action that wastes resources and destroys lives. Therefore, science is cautious, and standard criteria for statistical significance put much more weight on avoiding Type I errors than avoiding Type II errors. So the data support your hypothesis only weakly? Go back and get a bigger sample, or let somebody else try; don't prematurely announce that you've identified a cure for cancer or the root of some other evil.

Oddly enough, many historical commentators (including politicians, pundits, and some historians) do not seem to feel this self-restraint. In fact, the quality of historical writing is often judged by the ability of the historian to weave a few random threads into an interesting pattern. Most readers will be impressed by a story that is logical and is supported by some data, even though the data points are selected and not at all representative.

In relation to our subject matter, wars are especially suited to biased treatment because they comprise many great stories in which we can identify directly with the actors and we are emotionally compelled by the drama. The same is also true of the history of societies where there has been victimization on a large scale. Every victim has a story to tell and a need for acknowledgement.

Story-telling is powerful. Its intuitive appeal is much greater than models, charts, and numbers. It speaks the language of nations and

politics: shared experiences, common destinies, collective rights and wrongs. It is easily voiced by leaders and heard by followers untrained in statistical thinking about trends and standard errors. As a result, while politicians may turn to economists for technical advice, they get historians to help write their speeches -- Arthur Schlesinger Jr (John F. Kennedy), Richard Pipes (Ronald Reagan), and Norman Stone (Margaret Thatcher).

What makes the power of story-telling? First, story-telling is deliberately selective. It does not try to be representative. When we scan history for stories, we look by definition for sequences of events that have a beginning, a middle, and an end. In the middle, something happens that is out of the ordinary, dramatic, and unexpected. Invariably, we rule out all those much more representative past circumstances out of which something might have come, but after which there were no surprises and nothing much happened.

Second, stories give us a way of thinking about how to handle rare and unpredictable events. Houghton (1996), for example, has shown how U.S. leaders used historical parallels to make decisions in the “novel emergency” of the 1979 Iranian hostage crisis. In this sense a well-chosen story (in other words, one data point) can meet the needs of decision makers in a much more satisfying way than a statistician who just says: “Well, we need to collect more data.”

Third, story-telling typically sounds an alarm. In history, dramatic events are rarely good news. The good news in history has generally been made up of the slow, steady progress of emancipation, literacy, and prosperity. Such good news is easily illustrated by statistics and trends, but does not make good stories. It is the bad news of tragedies and crises that makes good stories.

In fact, an entirely legitimate purpose of a good story may sometimes be to sound the alarm about the risks we face and so avert their realization. George Orwell’s novel *1984*, for example, warned western societies of the dangers of totalitarian rule much more effectively than a thousand learned treatises on the subject.

There is a downside to this. Some stories can be self-fulfilling. There is a particular kind of collective story, for example, that communal identity politicians like to tell (Glaeser 2005). These are stories of past hate crimes allegedly committed by some other ethnic or religious group against their own group: Black against White, Reds against Whites, Germans against Jews, Jews against Palestinians, Protestants against Catholics, Sunni against Shia – and, in all cases, vice versa. Such powerful stories can be extrapolated into predictions of future hate crimes yet to be committed, and then into justifications for hateful and violent action to preempt the future crimes.

When we buy a story that is untrue or unrepresentative, we make a Type I error. We accept as meaningful a coincidence between the pattern in the story and a pattern inside our brains, despite the fact that the pattern is not matched in the external world. For example, we may prefer some stories to others, not because they are true, but because they are more comfortable or more affirming of our identity. One result is that the propensity to Type I errors is strongly present in the history of conflict, where is where national identities are made and affirmed.

Consider the following popular hypotheses, which all relate to the outcome of World War II:

- Britain won the war because the British way of life inspires loyalty.
- Germany (or Japan) almost won the war because Germans (or Japanese) are disciplined and follow orders.
- France lost the war because French society was rotten.
- Italy lost the war because Italians are cowards.
- America won the war because Americans are free and equal.
- The Soviet Union won the war because Russians love their country.

These ideas all carry a high risk of Type I error, as a moment's reflection will indicate. To explain the scope for error in each case: It is doubtful that Britain could have won the war alone, in other words, without the help of the Soviet Union, the United States, and the British Empire where millions lived in ways that were quite different from those of the British at home. If Germany and Japan nearly won the war, it is likely that this was because their leaders planned it that way and it took time for others to rise to the challenge; strategy and timing were more important than the innate characteristics of the citizens. If France lost the war, its society was hardly more rotten than that of others. If Italy lost the war it was not from cowardice; quite apart from anything else the courage of Italians who resisted the dictator Mussolini must also have been a factor. If America won the war, it was with the participation of large numbers of slave descendants who were less free and less equal, and it was also with the help of the Soviet Union whose citizens were ruled by a totalitarian dictator. Finally, if the Soviet Union won the war it was despite the best efforts of millions of Russians who collaborated with the occupier or fought actively against the Red Army.

In fact, Chapter 2 will show that in the twentieth century "our nation" won wars most often by belonging to an alliance that was richer and larger than the adversary's alliance. This finding does not particularly help to affirm anyone's national identity, not does it contribute an exciting story. Nonetheless it fits most of the patterns in the data much better than any hypothesis based on national exceptions or peculiarities.

A lesson is that the student of coercion and conflict needs to approach all good stories equipped with a strong sense of scepticism. A deep

understanding and continuous awareness of the risks of the Type I error is an essential element of the economist's toolkit.

A chapter-by-chapter summary

Table 1 sets out the basic tools used in this book, chapter by chapter. Not every tool is used in every chapter. Notably, no chapter could have been written without reference to the economist's idea of rationality.

Table 1. The economist's toolkit, chapter by chapter

						Type I errors		
						Surplus extraction		
						Scale and state capacity		
						Diminishing and increasing returns		
						Strategic interaction; multiple equilibria		
						Rational actors; expected costs and benefits		
1.	War and Disintegration, 1913 to 1945	x				x		
2.	Why the Wealthy Won	x				x	x	
3.	Why Didn't the Soviet Economy Collapse	x	x	x		x	x	
4.	The Frequency of Wars	x	x			x	x	
5.	Soviet Industry and the Red Army Under Stalin	x	x				x	
6.	Contracting for Quality under a Dictator	x	x				x	
7.	A Soviet Quasi-Market for Inventions	x	x				x	
8.	The Political Economy of a Soviet Military R&D Failure	x	x				x	
9.	The Fundamental Problem of Command	x	x	x			x	
10.	Accumulation and Labor Coercion Under Late Stalinism	x		x			x	x
11.	Economic Information in the Life and Death	x	x	x			x	
12.	Coercion, Compliance, and the Collapse	x	x	x			x	

The economics of global war

The book is divided into three parts. The first part is devoted to the economics of global war. A common theme of the chapters in this part of the book is the importance of scale, resources, and state capacity in the outcomes of wars.

In Chapter 1, "War and Disintegration, 1914-1945," Jari Eloranta and I examine Europe in the first half of the twentieth century. Globalization came to an abrupt halt, replaced by protectionism, nationalism, war, and killing and destruction on an immense scale. In mid-century, globalization was resumed, and the European economies began to converge on much higher and more uniform income levels. After two world wars, three things had changed. First, European economic growth, integration, and prosperity had lost its association with empire. No doubt to the surprise of Europe's nineteenth century leaders, had they lived to see it, it proved possible to acquire wealth and wield influence without claiming imperial sway over vast stretches of faraway peoples and their lands and oceans.

Second, Europe's leaders had a new sense of the importance of cooperation. They now cooperated with the United States in economic recovery, exchange rate coordination, and tariff reduction, with each other in laying new foundations for European integration, and with developing countries in decolonization and development assistance. Third, Europe's leaders had learned to use the power of the state to regulate economic life. Learning the appropriate limits of government control over the market economy in turn became a major challenge of the postwar period.

In Chapter 2, "Why the Wealthy Won," I look more closely at the roots of victory and defeat in two world wars. The chapter starts from the variation that we observe in the degrees of economic mobilization of different countries for total war in the twentieth century. Most of this variation is explained by differences in the level of economic development of each country, though not all of it and there some exceptions. There are several good reasons that help to explain why mobilization capacity should depend significantly on economic development. The empirical record is to some extent a puzzle since it seems to leave little room for other factors that would feature prominently in narrative accounts, such as national differences in war preparations, war leadership, or military organization and morale. The chapter looks at ways of solving this puzzle.

Chapter 3 asks "Why Didn't the Soviet Economy Collapse in 1942?" It appraises the economic dimensions of World War II both generally and with specific reference to the eastern front. When the Soviet war effort is examined more closely, it becomes surprising that the Soviet economy did not collapse in 1942. A rational-choice model is developed to illustrate the economic conditions under which a wartime collapse of the economy is rendered more and less likely. The possible effects of policy interventions by Stalin, Hitler, and Roosevelt on the stability of the Soviet war effort are defined.

In Chapter 4, "The Frequency of Wars," Nikolaus Wolf and I step back from particular countries and conflicts to the puzzles that arise from a quantitative overview of world history. Conflicts among states are, it seems, increasingly frequent, and the trend has been steadily upward since 1870. The main tradition of Western political and philosophical thought suggests that extensive economic globalization and democratization over this period should have reduced appetites for war far below their current level. This view is clearly incomplete: at best, confounding factors are at work. Here, we explore the capacity to wage war. Most fundamentally, the growing number of sovereign states has been closely associated with the spread of democracy and increasing commercial openness, as well as the number of bilateral conflicts. Trade and democracy are traditionally thought of as goods, both in themselves,

and because they reduce the willingness to go to war, conditional on the national capacity to do so. But the same factors may also have been increasing the capacity for war, and so its frequency. The chapter concludes that we need better understanding of how to promote these goods without incurring adverse side-effects on world peace.⁸

Communism and defence

The middle part of the book focuses on the relationship between communism and defence. In all countries where communists ruled, the mechanisms of power were designed on a similar template of coercion and repression. The two best accounts of this template are the sociological findings of Inkeles and Bauer (1959) on the Soviet Union and Anne Applebaum's (2012) new history of the imposition of communist rule on Eastern Europe.

The more we know about communist rule from the inside, the more important it seems to be that we understand how preparations for external conflict were linked to coercion at home. A common theme of the chapters in this part of the book is strategic interaction, conflict, and bargaining over the surpluses that the communist rulers were willing to share with the military and industrial interests that had the job of supplying defence needs.

Chapter 5, "Soviet Industry and the Red Army Under Stalin," considers some views of the Soviet "military-industrial complex" that are current in the literature. The economic weight of the defence sector in the economic system is summarized in various aspects. The lessons of archival research are used as a basis for analysing the army-industry relationship under Stalin as a prisoners' dilemma in which, despite the potential gains from mutual cooperation, each party faced a strong incentive to cheat on the other. The chapter concludes that the idea of a Soviet military-industrial complex is not strictly applicable to the Stalin period, but there may be greater justification for the Soviet Union after Stalin.

In Chapter 6, "Contracting for Quality under a Dictator," Andrei Markevich and I examine military procurement in the Soviet economy under Stalin. This provides a novel historical context for a standard problem of market organization, that of contracting for quality. The Soviet ministry of defense was engaged in the procurement of military goods from Soviet industry. An internal market was formed and contracts were made. In the market, the contractor had power over the buyer and

⁸ This chapter came under critical fire from leading political scientists. For their criticisms and our replies see Gleditsch and Pickering (2013) and Harrison and Wolf (2013).

typically used this power to default on quality. In the background loomed a dictator, imperfectly informed, but with the power to tear up any contract and impose unlimited penalties on the contractor. The buyer's counter-action took the form of deploying agents through industry with the authority to verify quality and reject substandard goods. The final compromise restored quality at the expense of quantity. Being illicit, it had to be hidden from the dictator. Our case provides an historical illustration of the limits of dictatorship.

Chapter 7, "A Soviet Quasi-Market for Inventions," is also about the allocation of resources in the Soviet economy under Stalin's dictatorship. In this chapter the resources were designated for military research and development. The context was formed by the rapidly approaching limits of an existing aviation propulsion technology, the need to replace it with another, and profound uncertainty as to how to do so. In the process we observe the formation of an internal "market" in which rival designers proposed projects and competed for funding to carry them out. We observe rivalry and rent seeking by agents, imperfectly regulated by principals. As rent seeking spread and uncertainty was reduced, the chapter shows, the market was closed down and replaced by strict hierarchical allocation and monitoring.

Chapter 8 examines "The Political Economy of a Soviet R&D Failure," the prewar attempt to create a new aeroengine technology based on the steam turbine. From this example we find out more about the motivations, strategies, and payoffs of principals and agents in the Soviet command economy. Alternative approaches to the evaluation of R&D failure are outlined. New archival documentation shows the scale and scope of the Soviet R&D effort in this field. The allocation of R&D resources resulted from agents' horizontal interactions within a vertical command hierarchy. Project funding was determined in a context of biased information, adverse selection, and agents' rent seeking. Funding was rationed across projects and through time. Budget constraints on individual projects were softened in the presence of sunk costs, but were hardened periodically. There is no evidence that rents were intentionally distributed through the Soviet military R&D system to win trust or reward loyalty; the termination of aviation steam power R&D in 1939 despite the sunk costs they represented was timely.

Communism and coercion

Communist rule rested ultimately on command and coercion, and the third part of the book focuses directly on this. Again, there is a lot about strategic interaction between the state and the citizens, this time conceptualized more broadly as producers and even consumers.

Chapter 9, “The Fundamental Problem of Command,” studies the problems that arose under communist rule when the dictator gave an order to an agent and advanced resources for its implementation. The temptation was for the agent to shirk or steal from the principal rather than comply; this constitutes the fundamental problem of command. Historically, partially centralized command economies enforced compliance in various ways, assisted by nesting the fundamental problem of exchange within that of command. The Soviet economy provides some relevant data. The Soviet command system combined several enforcement mechanisms in an equilibrium that shifted as agents learned and each mechanism’s comparative costs and benefits changed. When the conditions for an equilibrium disappeared, the system collapsed. An extractive state had the task of aligning the incentives of its citizens to work together for the surplus that the state required to finance its own goals. The citizens faced continual temptations to shirk and steal. Keeping everything together was at least a feasible problem in the age of mass production of things. With the information revolution and the rise of the services economy this task became more and more complex and ultimately, perhaps, infeasible.

In Chapter 10, “Accumulation and Labour Coercion,” Paul Gregory and I examine how the level of coercion affected Soviet wages and the surplus available for investment under Stalin. From 1940 until the death of Stalin the conditions of Soviet labour were highly regimented under draconian legislation which prescribed imprisonment and forced labour for even minor violations. How successful was this experience? A successful use of coercion implied that fewer resources would be needed for consumption, without compromising the amount of effort supplied by the workers. In the desired outcome, accumulation would be enabled to grow. A major issue was the problem of mistakes and unintended consequences of various kinds. In the face of harsh punitive measures Soviet workers shifted effort from production to mutual insurance through informal horizontal and vertical networks. The Stalinist leaders often interpreted attempts to protect subordinates as proof of conscious or unconscious betrayal of the party and Stalin, which then became an occasion for new repression. In this atmosphere, the likelihood of errors in the allocation of both labour and capital was only increased. In such conditions it is possible coercion did not contribute to an increase in the level of accumulation. In fact, it is quite likely that the Soviet economy’s accumulation capacity improved after Stalin died and his successors reduced the level of violence in the economy.

Chapter 11, “Economic Information in the Life and Death of the Soviet Command System,” shifts the focus to information problems under communism. In market economies information adds value to transactions

in three ways: it supports reputations, permits customisation, and provides yardsticks. In the Soviet economy such information was frequently not produced; if produced, it was often concealed; whether concealed or not, it was often of poor quality; regardless of quality, it often suffered from low credibility outside the ruling circle. In short, the Soviet command system forced economic growth on the basis of a relatively low-value information stock. This might help explain aspects of Soviet postwar economic growth and slowdown, the collapse of the command system, and the persistence of low output after the collapse.

Chapter 12, “Coercion, Compliance, and the Collapse of the Soviet Command Economy,” concludes the book. It addresses directly the problem of how and why the Soviet economy collapsed in 1991. It begins by asking whether command systems that rest on coercion are inherently unstable, and whether the Soviet economy collapsed for this reason. The problem is that, until it collapsed, the Soviet economy did not look unstable. Why did it then collapse? A game between a dictator and a producer shows that a high level of coercion may yield a stable high-output equilibrium, that stability may rest in part on the dictator’s reputation, and that a collapse may be brought about by adverse trends in the dictator’s regime costs and a loss of reputation. The facts of the Soviet case are consistent with a collapse that was triggered by the Russian workers’ strike movement of 1989.

Conclusion

In this book I apply the concepts and methods of analysis that I have found most useful for study of the economics of coercion and conflict. The economist’s toolkit is versatile, and can be applied to many problems that are not economic at first sight. Each of the tools is useful provided it is applied to the right context. The problem is to recognize the context and select the right tool. If you want to fix together two pieces of wood, there are times when nails will do the job and sometimes it requires screws. Hammering a screw and trying to turn a nail will get you nowhere.

Sometimes the economist’s toolkit must fall short. Is economics adequate by itself for a rounded understanding of everything that is wrapped up in the subject of coercion and conflict? Clearly not. Power is about politics before it is about economics. War is a political act. Force and violence are the most multidisciplinary of human activities. They engage our emotions as well as our reason. In my own work, whether I have acknowledged it or not, I am sure that I have also borrowed many ideas drawn from psychology, political and strategic studies, international relations, and historical narrative. Nonetheless, any activity that involves

the expenditure of effort for an expected return has an economic dimension, and that is the justification of the present book.

References

- Açemoglu, Daron, and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown.
- Akerlof, George A., and Rachel E. Kranton. 2000. Economics And Identity. *Quarterly Journal of Economics* 115(3), pp. 715-753.
- Applebaum, Anne. 2012. *Iron Curtain: The Crushing of Eastern Europe*. London: Allen Lane.
- Besley, Timothy, and Torsten Persson. 2009. The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review* 99(4): 1218-44.
- Box, George E. P., and Draper, Norman R. 1987. *Empirical Model Building and Response Surfaces*. New York: Wiley.
- Brauer, Jurgen, and Hubert van Tuyl. 2008. *Castles, Battles, and Bombs: How Economics Explains Military History*. Chicago: University of Chicago Press.
- Brennan, Geoffrey, and Gordon Tullock. 1982. An Economic Theory of Military Tactics: Methodological Individualism at War. *Journal of Economic Behavior and Organization* 3:2-3, pp. 225-42.
- Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability. *American Political Science Review* 89:4, pp. 841-855.
- Bueno de Mesquita, Bruce. 2006. Game Theory, Political Economy, and the Evolving Study of War and Peace. *American Political Science Review* 100:4, pp. 637-642.
- Cienciala, Anna M. 2011. The Foreign Policy of Józef Piłsudski and Józef Beck, 1926-1939: Misconceptions and Interpretations. *The Polish Review* 56(1-2), pp. 111-152.
- Clausewitz, Carl von. 1832/1982. *On War*. London: Penguin Books.
- de Tocqueville, Alexis. 1835/2000. *Democracy in America*. Chicago: University of Chicago Press.
- Dixit, Avinash K. 2004. *Lawlessness and Economics: Alternative Modes of Governance*. Princeton: Princeton University Press.
- Edelstein, Michael. 2000. War and the American Economy in the Twentieth Century. In *The Cambridge Economic History of the United States*. Vol. 3, *The Twentieth Century*, pp. 329-405. Edited by Stanley L. Engerman and Robert E. Gallman. Cambridge: Cambridge University Press.
- Fearon, James D. 1995. Rationalist Explanations for War. *International Organization* 49(3), pp. 379-414.
- Findlay, Ronald, and Kevin H. O'Rourke. 2012. War, Trade, and Natural Resources: A Historical Perspective. In *The Oxford Handbook of the Economics of Peace and Conflict*. Edited by Garfinkel, Michelle R., and Stergios Skaperdas. Oxford: Oxford University Press.

- Garfinkel, Michelle R., and Stergios Skaperdas. 2012. Economic Perspectives on Peace and conflict. In *The Oxford Handbook of the Economics of Peace and Conflict*. Edited by Garfinkel, Michelle R., and Stergios Skaperdas. Oxford: Oxford University Press.
- Gat, Azar. 2006. *War in Human Civilization*. Oxford: Oxford University Press.
- Glaeser, Ed. 2005. The Political Economy of Hatred. *Quarterly Journal of Economics* 120:1, 45-86.
- Gleditsch, Kristian Skrede, and Steve Pickering. 2013. Wars are Becoming Less Frequent: a Response to Harrison and Wolf. *Economic History Review* (forthcoming).
- Gregory, Paul R., ed. 2001. *Behind the Façade of Stalin's Command Economy: Evidence from the Soviet State and Party Archives*. Stanford: Hoover Institution Press.
- Gregory, Paul R. 2004. *The Political Economy of Stalinism: Evidence from the Soviet Secret Archives*. Cambridge: Cambridge University Press.
- Gregory, Paul R. 2009. *Terror by Quota: State Security from Lenin to Stalin (an Archival Study)*. New Haven: Yale University Press.
- Gregory, Paul R., and Valery Lazarev, eds. 2003. *The Economics of Forced Labor: The Soviet Gulag*. Stanford: Hoover Institution Press.
- Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge: Cambridge University Press
- Harrison, Mark. 2002. Economic Growth and Slowdown. In *Brezhnev Reconsidered*, pp. 38-67. Edited by Edwin Bacon and Mark Sandle. London and Basingstoke: Palgrave.
- Harrison, Mark. 2005. Stalin and Our Times. In *Stalin: His Times and Ours*, pp. 67-84. Edited by Geoffrey Roberts. Dublin: Irish Association for Russian and East European Studies.
- Harrison, Mark. 2006. An Economist Looks at Suicide Terrorism. *World Economics* 7:3, pp. 1-15.
- Harrison, Mark, and Nikolaus Wolf. 2012. The Frequency of Wars. *Economic History Review* 65:3, pp. 1055-1076.
- Harrison, Mark, and Nikolaus Wolf. 2013. The Frequency of Wars: Reply to Gleditsch and Pickering. *Economic History Review* (forthcoming).
- Hobbes, Thomas. 1651/1909. *Hobbes's Leviathan*. Reprinted from the Edition of 1651 with an Essay by the Late W. G. Pogson Smith. Oxford: Clarendon Press.
- Holborn, Hajo. 1986. The Prusso-German School: Moltke and the Rise of the General Staff. In *Makers of Modern Strategy From Macchiavelli to the Nuclear Age*. Oxford: Oxford University Press.
- Houghton, David Patrick. 1996. The Role of Analogical Reasoning in Novel Foreign-Policy Situations. *British Journal of Political Science* 26(4), pp. 523-552.
- Inkeles, Alex, and Raymond A. Bauer. 1959. *The Soviet Citizen: Daily Life in a Totalitarian Society*. Cambridge, Mass.: Harvard University Press.
- Kahneman, Daniel. 2011. *Thinking Fast and Slow*. London: Allen Lane.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives* 5(1), pp. 193-206.

- Kant, Immanuel. 1795/1983. *Perpetual Peace and Other Essays*. Indianapolis: Hackett Publishing.
- Klemann, Hein, and Sergei Kudriashov. 2012. *Occupied Economies: An Economic History of Nazi-Occupied Europe, 1939-1945*. London and New York: Berg.
- Konev, I. S. 1970. *Sorok piatyi*. 2/e. Moscow: Voenizdat.
- Levy, Jack S. 1988. Domestic Politics and War. *Journal of Interdisciplinary History* 18:4, pp. 653-673.
- Lieberman, Peter. 1996. *Does Conquest Pay? The Exploitation of Occupied Industrial Societies*. Princeton: Princeton University Press.
- Markevich, Andrei. 2011. How Much Control is Enough? Monitoring and Enforcement under Stalin. *Europe-Asia Studies* 63(8), pp. 1449-1468.
- Marx, Karl. 1867/1974. *Capital: A Critical Analysis of Capitalist Production*, vol. 1. London: Lawrence & Wishart.
- North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2011. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.
- O'Brien, Patrick K. 1982. European Economic Development: The Contribution of the Periphery. *Economic History Review* 35(2), pp. 1-18.
- Offer, Avner. 1989. *The First World War: An Agrarian Interpretation*. London: Clarendon.
- Olson, Mançur. 1993. Dictatorship, Democracy, and Development. *American Political Science Review*, 87:3, pp. 567-76
- Pinker, Steven. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Viking.
- Rockoff, Hugh. 2012. *America's Economic Way of War: War and the US Economy from the Spanish-American war to the Persian Gulf War*. Cambridge: Cambridge University Press.
- Sen, Amartya. 2006. *Identity and Violence: The Illusion of Destiny*. London: Allen Lane.
- Smith, Adam. 1776/2005. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Jim Manis. An Electronic Classic Series publication, available at <http://www2.hn.psu.edu/faculty/jmanis/adam-smith/wealth-nations.pdf> (accessed 28 August 2013).
- Smith, Ron P. 2009. *Military Economics: the Interaction of Power and Money*. Basingstoke: Palgrave Macmillan.
- Tilly, Charles. 1975. Reflections on the History of European State-Making. In *The Formation of National States in Western Europe*, pp. 3-83. Edited by Charles Tilly. Princeton: Princeton University Press.
- Weingast, Barry R. 1995. The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development. *Journal of Law, Economics, and Organization* 11:1, pp. 1-31.