

The Laws of War as an International Institution

by James D. Morrow

University of Michigan

Very preliminary and rough draft, but it should give you an idea of how I am thinking

Prepared for presentation at Cooperation and Conflict conference at Northwestern University,
November 7-8, 2008.

War is often thought of as lawless violence. On the contrary, conduct during wartime has often been regulated by agreements to restrict violence. Such agreements can be either explicit or tacit; they may be the product of prewar negotiations or just commonly accepted military practices. These agreements have actually succeeded in limiting the range of weapons employed, how weaponry is used, and the legitimate targets of force during wartime. Still, many of these agreements fail during wartime. The record of the laws of war is mixed at best (Morrow 2007). How can we understand which conventions are observed and which violated and when such violations are likely to occur?

International politics in modern times generally recognizes no authority above the nation-state. Agreements among states are enforceable only by the agreeing states themselves. This assumption of anarchy poses a paradox for agreements to limit violence during wartime. War is ultimate form of enforcement available to states. What can states do to enforce their agreement if they are already at war?

Reciprocity serves as the main tool to enforce agreements in international politics. Enforcement of an agreement is devolved to the parties themselves. Damaged parties have the option to respond with retaliatory sanctions to a violation of an agreement. The threat of reciprocal sanctions may be sufficient to deter violations, and so agreements can be enforced in international politics.

Agreements to limit violence during wartime can be enforced through a form of reciprocity as well. If a warring state violates an agreement, the damaged side can respond in kind. Sometimes, as in the case of chemical weapons during the Second World War, that threat is sufficient to sustain a convention against violence. In other cases, such as treatment of

prisoners of war between Nazi Germany and the Soviet Union during the Second World War, the threat is not sufficient to deter proscribed violence. Such reciprocal threats can succeed only when the reciprocal threat is clear and of sufficient magnitude to deter violations.

However, this reciprocity differs from the reciprocal sanctions familiar in trade negotiations and repeated game models. In the latter, the sanction imposes a punishment that is meant to exceed the gain from the original violation, and so deter the violation. In the former, retaliation abandons the same restraint as the violation. If a state uses gas, then its opponent is no longer held to its promise to not use gas. Rather than being a punishment, reciprocity in the laws of war means that the war will now be fought without the restraint that was violated.

In practice, reciprocity on the laws of war also exists on the battlefield between soldiers fighting one another. When the soldiers of one side kill those attempting to surrender or commit perfidy, that is, use the protections of the laws of war to gain an unfair advantage on the battlefield, soldiers of the other side often respond with violations of their own. If one side uses ambulances to transport ammunition, an act of perfidy abusing the protection of medical services displaying the Red Cross, the other side may simply fire on all ambulance of the first side. The scope for individual violations varies with the issue-area. Every armed soldier on the battlefield can commit violations against civilians, but they cannot use chemical weapons if their military does not provide them with those weapons. The laws of war must restrain the acts of individual soldiers as well as those of state leaders to be effective institutions for the limitation of violence.

The laws of war create rights and obligations for soldiers on the battlefield in addition to the restraints they place on state policy during war. Reciprocity on the battlefield could control some behavior on the battlefield, particularly perfidy. But this form of reciprocity is weak in

other areas because the soldiers of one side lack the information to target violators on the other side individually (Fearon and Laitin 1996). Alternatively, each side's military could attempt to police and control their own soldiers. The combination of training in the laws of war and effective military justice for violators holds some hope for controlling the acts of individual soldiers on the battlefield. Command authority has access to violators and witnesses, making prosecution of violators possible. This is not to say that such control is easy, only that it is possible.

Violations by individual soldiers pose another problem for compliance in state policy. Governments at war rarely state that they are violating agreements. Each side must infer the compliance of the other side from what happens on the battlefield. Widespread violations by individual soldiers of one side could lead the government of the other side to conclude that the government of the first chose to break the agreement, and so retaliation in kind is in order. Individual violations then produce a problem of reciprocal enforcement under noise (Green and Porter 1984).

The laws of war pose three related strategic problems:

- State to state deterrence of deliberate violations,
- Individual to individual violations on the battlefield, and
- State control of individual soldiers.

An effective system for limiting violence during wartime must address all three problems at once. A breakdown in any one could lead to a breakdown in the others. The laws of war primarily focus on the first of the three problems above. They create detailed state obligations, describe some individual rights, and lay out general norms and specific rules of conduct on the

battlefield. Implementation of the laws of war requires limiting individual violations through effective military training and justice. Finally, soldiers must observe those limits on the battlefield, albeit under the shadows of military justice and the reactions of their opponents. This paper presents three simple models of each of these strategic problems in the laws of war and shows how they depend upon one another.

Political Institutions as Constellations of Equilibria

Political institutions must persist if they are to shape what political actors do. If actors are free to change institutions at will, how could those institutions induce them to take actions they would not in their absence? This observation leads to the central tenet of rational choice approaches to political institutions, that an institution is an equilibrium of some game (Schotter 1981, Calvert 1994a,b, Grief 2006). An equilibrium requires two conditions; one, that all actions are in the best interest of the parties taking those actions, including actions that may not occur--best replies; and two, that all parties understand what actions one another will take during the competition, particularly in response to violations of an agreement--common conjecture. Successful conventions then are self-enforcing combinations of behavior provided that all anticipate that one another will follow the convention, including reciprocal threats.

This argument leads directly to the role of the laws of war. Multiple equilibria typically exist in games. In the case of conduct during war, there are multiple conventions under which states could fight wars. They could use chemical weapons or not, for instance. The war that would result from a convention not to use chemical weapons would be different than a war where

such weapons were used, assuming that the convention was self-enforcing. All parties might choose such a convention before war broke out, believing that fighting under the convention to be better than fighting without it. The treaties of the laws of war codify the common conjecture about what conventions to limit violence during wartime. If the actions delineated by that convention are self-enforcing, then the convention will be honored in wartime.

The equilibrium view of political institutions leaves open several questions. First is the selection of institutions (cf. Shepsle 1984 on equilibrium institutions). Some political institutions appear to be inefficient in the sense that they are Pareto dominated by other institutional arrangement. Then it should be in the interest of all actors to agree to switch to a Pareto superior institution. If the institution is the common conjecture alone, it should be simple to change that common conjecture through negotiation. Second is the completeness of the shared understanding. International treaties and law generally are typically incomplete in the sense that they cannot anticipate all possibilities. Instead, they rely on some general principles which can be used to judge novel cases and situations. But the common conjecture in an equilibrium provides full detail of optimal strategies by all players in all situations, even those off-the-equilibrium-path. Third is what is institutional about this arrangement? Players might arrive at playing an equilibrium through shared experience of playing the game. They may correctly anticipate one another's strategies without having an understanding why other than that common experience. While this evolution of common understanding may describe some informal, social institutions, the laws of war have been negotiated, as is the case for most political institutions. Now we could claim preplay negotiation as the source of the common conjecture, we are again left with

Institutions are properly thought of as constellations of equilibria in a group of related games. Within each separate game, the institution defines which equilibrium the players will play in that game. Different players play in each game with each player active in only one or two of the related games. The players assume that all the other games will be played according to the equilibrium specified in the institution. Play in these other games is fixed then in the context of each game individually. Because the outcomes of one game depend in part on the outcomes of other games, optimal play in each of them depends on how the others will be played. The collection of interdependent equilibria across the games is institutional in that it is difficult for any set of players playing one game to change the equilibrium they are playing.

The case of the laws of war developed in this paper clarifies this concept of equilibrium. The actors are the two governments at war and the soldiers in their armies. There are three interrelated games: one between the two government about whether they will comply with their obligations under the laws of war, a second on the battlefield where the soldiers of the two armies fight and may engage in atrocities, and a third pair of parallel games where each government monitors and disciplines its soldiers' conduct on the battlefield. Figure 1 shows the four actors with an arrow representing the social relationship modeled in each game. The equilibrium played in each game depends on the others. There is little reason for soldiers on the battlefield to discipline themselves against an enemy practicing atrocity and perfidy because their government chooses not to discipline those violations. What happens inside the military of the other country is outside their control. Similarly, even governments that wish to comply with their legal obligations may find it difficult to control their soldiers on some issues. Each player, or set of players in the case of the soldiers of the two armies, plays in only two of the four games,

On the Battlefield

The soldiers of the two armies meet on the battlefield. The combat environment is represented by randomly matched all soldiers into pairs with one soldier from each side. They then play the simultaneous game in Figure 2 below. Each soldier chooses between H —honoring the existing treaty standard, and V —violating that standard by committing an atrocity. The game has two parameters, T with $0 < T < I$ for the temptation to commit violations and $V < 0$ for the vulnerability to violations. These parameters vary with the issue-area the game covers. T would be higher for issues where an individual soldier could benefit from a violations, such as looting on the issues of treatment of civilians. I restrict T to be less than I so that some a_i, b_j prefer not committing violations. V would be greater in magnitude for issues where the soldier suffers greater consequences as a victim of a violation, such as on treatment of prisoners of war or enemy wounded.

Figure 2
Battlefield Violation Game

		b_j	
		Honor	Violate
a_i	Honor	(0,0)	(V, T-t _j)
	Violate	(T-t _i , V)	(-1,-1)

This game has multiple equilibria, with the character of the equilibria depending on whether $T + V \geq -1$. In each case below, b_j plays the same strategy based on its type that a_i , so I omit their strategy to save duplication.

Equilibria: Let $t_{-crit} = 1 + \frac{1}{2}(T+V) - \frac{1}{2}((T+V)^2+4V+4)^{\frac{1}{2}}$ and $t_{+crit} = 1 + \frac{1}{2}(T+V) + \frac{1}{2}((T+V)^2+4V+4)^{\frac{1}{2}}$.

Case 1: If $T + V \geq -1$ and $V \leq -1$, then there are three equilibria (ranked from the one with the lowest frequency of violations to the highest):

- a) a_i plays Honor when $t_i > t_{-crit}$, Violate otherwise;
- b) a_i plays Honor when $t_i > t_{+crit}$, Violate otherwise; and
- c) all a_i play Violate.

Case 2: If $T + V \geq -1$ and $V > -1$, then there is a unique equilibrium: a_i plays Honor when

$t_i > t_{crit}$, Violate otherwise.

Case 3: If $T + V < -I$, then all a_i, b_j play Violate.

The strategic logic of the cases is simple. The first and third cases have prisoners' dilemma payoffs for low types, those who are willing to commit violations. They always prefer Violate. The risk of facing such a type leads some types to play Violate to avoid the worse payoff of being a victim (you play Honor, opponent plays Violate, payoff T) even though they prefer the outcome where both play Honor (payoff θ) to their own unilateral violation (payoff $T - t_i < \theta$). Case 3 is the extreme case where the risk that the other side will violate is so high that all types play Violate to protect themselves. The second case has dynamics closer to Chicken in the sense that being a victim is preferable to both playing Violate. The use of chemical weapons could be such a case where from the individual soldiers' view, it is preferable to run a small risk of suffering a chemical attack to the certainty of a battlefield where both sides use chemical weapons regularly.

Military Discipline

Each government and its soldiers have a principal-agent relationship, where soldiers are agents of the state. For ease, we will assume the state is A . The government A would like its soldiers $\{a_i\}$ to comply by playing Honor rather than Violate. Many but not all a_i would play Honor left to their own devices; they have the same payoff structure as in the Battlefield game if the opponent plays Honor. A can monitor the soldiers, choosing a probability of detection of those who Violate, denoted by p with $0 \leq p \leq 1$. Monitoring is costly for A in two ways. One, creating the capability for monitoring costs the probability of successful detection, pD , where D

represents the cost of detecting violations on that given issue. D is assumed to average I across issues. Second, A suffers losses for every soldier a_i caught committing violations, with the total loss equal to the fraction of soldiers caught committing violations. You don't have an army if you arrest them all for war crimes. For convenience, I assume the punishment of any soldier a_i detected committing violations is $-I$.

The time line of the monitoring and discipline game is as follows:

1. A chooses p and announces it publicly.
2. All $\{a_i\}$ choose Honor or Violate.
3. Nature determines whether each a_i who played Violate is detected with probability p of detection. Payoffs are received as follows:

a_i played Honor: 0

a_i played Violate and not detected: $T - t_i$

a_i played Violate and detected: $T - t_i - I$

A : $\#(a_i \text{ played Honor})/\#(a_i) - \#(a_i \text{ played Violate and detected}) - pD$

The equilibrium of the monitoring game is straightforward. It depends on the cost of monitoring as follows:

A : Choose $p = T$ if $D < 1$ and $p = 0$ otherwise

a_i : Violate if $t_i < T - p = 0$ if $p = T$ or T if $p = 0$

Simply, A engages in no monitoring when monitoring is more costly than average because the dual cost of setting up the monitoring and of catching some of its soldiers in violations exceeds

the benefit of deterring violations. When monitoring is less costly than average, this calculation reverses, and A engages in enough monitoring to deter all violations.

State-to-State Deterrence

The war between A and B engage is represented by a model in the spirit of Smith (1998). The sides fight a series of battles over an ordered set of fixed points in discrete time. The endpoints of set represent total victory for one side or the other. The initial state of the war is one of these points. Each battle moves the state one position closer to the victory outcome for the winning side. The winner of each battle is determined by a given and known probability. There are then $N+1$ states to the game, the set $\{0, 1, \dots, N\}$, where state N represents complete victory for A and 0 complete victory for B and initial state k . Refer to any state as S_i . The transition probability from state to state depends on how many of the country's soldiers Violate rather than Honor the treaty appropriate standard: $p(S_{i+1}) = q + r(\#a_i \text{ violate}/\#a_i) - s(\#b_j \text{ violate}/\#b_j)$ with $p(S_{i-1}) = 1 - p(S_{i+1})$ and $r, s > 0$. In each round, A and B simultaneously decide how many of their soldiers will Violate. Both players also pay costs of $c_A(\#a_i \text{ violate}/\#a_i)$, $c_B(\#b_j \text{ violate}/\#b_j)$ with $c_A, c_B > 0$ per round. Each player's strategy maps the state into how many soldiers it will order to Violate in the current round.

Let $V_A(s, \#a_i \text{ violate}, \#b_j \text{ violate})$ be A 's continuation value for the game from state s given current strategies as specified. In equilibrium, neither player wishes to deviate from its strategy for how many soldiers Violate. For each player then (A is shown for convenience; parallel conditions holds for B),

$$V_A(s, \#a_i \text{ violate}, \#b_j \text{ violate}) \geq V_A(s, x, \#b_j \text{ violate}) - c_A(x - \#a_i \text{ violate})$$

$$c_A(x - \#a_i \text{ violate}) \geq r(x - \#a_i \text{ violate})[V_A(s+1) - V_A(s-1)]$$

If this condition holds for some $\#a_i \text{ violate}$, I conjecture that it is likely to hold for more than one.

There would then be multiple equilibria of levels of violations driven by state policy.

Interactions of the Equilibria

Each equilibrium of each separate games relies on what happens in the other games (even though I have not solved for those interactions yet). The monitoring equilibria depend in part on the battlefield equilibrium. If the soldiers of the two armies are playing the equilibrium where violations are likely, it will be more difficult and costly to monitor and discipline those violations by a side's own troops. Similarly, the monitoring of each state changes the calculations of the soldiers on their side on the battlefield. It could prove very difficult to sustain the equilibrium with low levels of violations if the other side's government does not monitor and discipline its soldiers. Whether the government choose to break out of an equilibrium of their strategic competition depends in part on what levels of violations occur on the battlefield outside their control.

References

- Calvert, Randall L. 1994a. The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication. In Jeffrey S. Banks and Eric A. Hanushek, eds., *Modern Political Economy: Old Topics, New Directions*. New York: Cambridge University Press.
- Calvert, Randall L. 1994b. Rational Actors, Equilibrium, and Social Institutions. In Jack Knight and Itai Sened, eds., *Explaining Social Institutions*. Ann Arbor, MI: University of Michigan Press.
- Fearon, James D., and David D. Laitin. 1996. Explaining Interethnic Cooperation. *American Political Science Review* 90 (4):715-735.
- Green, Edward J., and Robert H. Porter. 1984. Noncooperative Collusion under Imperfect Price Information. *Econometrica* 52 (1):87-100.
- Grief, Avner. 2006. *Institutions and the Path to the Modern Economy*. New York: Cambridge University Press.
- Morrow, James D. 2007. When Do States Follow the Laws of War? *American Political Science Review* 101(3):559-572.
- Schotter, Andrew. 1981. *The Economic Theory of Social Institutions*. New York: Cambridge University Press.
- Shepsle, Kenneth A. 1983. Institutional Equilibrium and Equilibrium Institutions. In Ada Finifter, ed., *Political Science: State of the Discipline*. Washington, DC: American Political Science Association.