Weapons, Wealth, and the Fates of Societies

Samuel Lee* Ilari Paasivirta† Alexander K. Zentefis*

March 26, 2026

Abstract

Why do weapons sustain durable peace in some societies but provoke perpetual violence in others? We develop a theory in which the value of human life and the frequency of violence are jointly determined by weapons technology and economic conditions. Lethal weapons deter conflict but raise mortality, taxing the future returns to investing in one's livelihood. When those returns are high, deterrence dominates: peace and investment reinforce each other. When those returns are low, the mortality tax dominates: agents divest from the future, the value of life falls, and violence deepens—a trap that deadlier weapons worsen. Whether weapons pacify or destabilize depends on the interaction between their offensive characteristics and the baseline prosperity of the society they enter. The theory explains four historical puzzles: why Tokugawa Japan (1603–1868), having mastered firearms, abandoned them and gained two centuries of peace and prosperity; why firearms traded into West Africa and among Native American nations (17th–18th century) produced escalating violence and persistent underdevelopment rather than deterrence; how Iceland (930–1262) sustained stateless order without a sovereign; and why Rome's late Republic descended into civil war as military reforms created a landless soldiery with no livelihood to protect and imperial plunder to fight over. The model also refines the logic of nuclear deterrence and generates testable predictions about urban gun violence in high-poverty neighborhoods.

JEL Classification: C73, D74, J17, K42, N40, O12

Keywords: deterrence, conflict traps, endogenous value of life, weapons technology, offense-defense balance, development traps

*Leavey School of Business, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053. Emails: slee@scu.edu, azentefis@scu.edu. Corresponding author: Zentefis.

†Columbia University, 116th St & Broadway, New York, NY 10027. Email: ilari.paasivirta@gmail.com

Disclosure. The authors received no external financial support for this research and declare no conflicts of interest.

1 Introduction

A foundational idea in economics is that raising the cost of an action reduces its frequency. Applied to violence, the logic predicts that weapons capable of killing should suppress conflict: an agent who faces lethal retaliation pays a higher expected cost for aggression and, at the margin, refrains. This reasoning underpins the canonical framework of Becker (1968), where increased punishment deters criminal entry. Lott (2000) extends it to firearms, arguing that higher gun prevalence reduces crime. In international relations, the same logic sustains nuclear deterrence strategy and the broader literature on offense-defense balance (Jervis, 1978). Chassang and Padró i Miquel (2010) formalize it in a model of conflict between two symmetric agents and show that under complete information, weapons unambiguously facilitate peace.

Yet the conflict-and-development literature documents the opposite. Whatley (2018) shows that European gun imports into West Africa and slave exports moved together in a self-reinforcing cycle across the eighteenth century. Nunn (2008) estimates that the slave trades account for a large share of African countries' income differences today; Nunn and Wantchekon (2011) show that descendants of affected populations exhibit lower interpersonal trust centuries later. Across the Atlantic and in the same era, the introduction of firearms into Native American societies triggered arms races and intensifying warfare rather than mutual deterrence (Silverman, 2016). The standard deterrence argument predicts none of this.

These two bodies of work—one predicting that weapons deter, the other documenting that weapons can destroy—coexist without a unifying theory. Existing models of deterrence treat the value of what an agent stands to lose as fixed. They do not allow that value to respond to the violent environment itself. If it does, the partial-equilibrium prediction that raising the cost of aggression curbs its frequency need not survive in general equilibrium. This paper provides such a framework, unifying both predictions to show that the same weapons that deter agents with a future worth protecting can destroy those without one.

We model interpersonal conflict as a sequential game in which agents choose between restraint and attack. Peace is sustained through the threat of retaliation: an agent who attacks and fails faces lethal consequences. We decompose weapons technology into three characteristics: *lethality* (q), the probability that a strike kills; *offensive advantage* (p), the probability that the first mover succeeds before the opponent can respond; and *multi-target capacity* (k), the number of opponents a weapon can engage per exchange.

Each parameter plays a distinct role. Lethality is the standard deterrent: a higher q raises the

cost of a failed attack. Offensive advantage works against deterrence: a high p insulates the attacker from retaliation, eroding the threat that sustains peace. When the offensive advantage is absolute ($p \to 1$), the retaliatory threat vanishes and lethality loses its deterrent power. Multi-target capacity functions as a substitute for numerical superiority in group conflict. It compresses the timeline of a fight, reduces the risk of return fire, and causes group-based deterrence to unravel. A firearm, for instance, combines high p, high q, and (in its modern variants) high k, which is a configuration that the model predicts will undermine deterrence rather than strengthen it.

We embed this micro-model in a dynamic population framework where agents face recurring conflict encounters. Before entering the environment, each agent makes a forward-looking investment in a livelihood (e.g., human capital, a farm, a trade, social ties) that generates a flow of utility over the agent's lifetime. The *conflict-adjusted hazard rate* Ω captures the rate at which agents die, combining background mortality with the lethal risk from realized violence. This hazard rate acts as a tax on the returns to investment: an agent who expects to die soon has less reason to build a future. And an agent who has not invested has less to lose from a fight.

This feedback is the paper's central mechanism. The value of human life is not a fixed parameter in the conflict problem; it is an equilibrium outcome, jointly determined by the technology of violence and the economic environment. When baseline prosperity is high and weapons allow credible retaliation (moderate p), a marginal increase in lethality deters a large share of conflicts. The hazard rate falls, agents invest more, the value of life rises, and the threshold for violence rises further. We call this a *virtuous cycle*. When baseline prosperity is low or offensive advantage is high, the same increase in lethality fails to deter enough violence to offset the higher death toll from the fights that still occur. The hazard rate rises, agents divest from the future, the value of life falls, and violence becomes easier to trigger. We call this a *vicious cycle*—a desperation trap that deadlier weapons deepen.

The model delivers a precise condition that determines which regime obtains. It depends on the interaction of weapons parameters (p, q, k) with the socioeconomic environment: baseline prosperity ($\bar{a}$), the cost of investment (c), the frequency of conflict encounters (λ), the distribution of stakes ($F(w)$), and background mortality (d). Neither weapons nor the economy alone determines the outcome. Indeed, weapons technology amplifies prevailing economic incentives. In societies where investment yields high returns, lethality reinforces order. Yet in societies at the margin of subsistence, it accelerates collapse.

The mechanism the model identifies is not specific to any single era, weapon, or continent. To show this, we apply the theory to four historical episodes, each mapping onto a distinct region of the model's parameter space. Taken together, they demonstrate that the same feedback between the technology of violence and the value of human life operated across centuries and civilizations that shared no institutions, no trade, and no cultural inheritance, suggesting that the mechanism

has broad applicability.

Tokugawa Japan (1603–1868). Firearms reached Japan in the 1540s and were adopted with striking speed. Within decades, Japanese armies deployed volley-fire tactics that raised multi-target capacity (k) alongside offensive advantage, and produced firearms in large quantities (Perrin, 1979). Civil war intensified. After political unification, the Tokugawa regime confiscated firearms and restricted their production, concentrating weapons in the hands of the samurai class. From the model's perspective, this intervention reduced both p and k. It sharply lowered the offensive advantage and multi-target capacity that firearms conferred and restored a weapon regime (swords and polearms) in which engagements were single-target and defense was privileged over offense. Two centuries of internal peace, rising urbanization, and economic development followed. The arrival of Commodore Perry's gunships in 1853, reintroducing high-p military technology, destabilized the equilibrium within two decades. Japan illustrates the virtuous cycle, and a deliberate policy intervention that achieved it by manipulating weapon parameters the model identifies as central.

The Atlantic slave trade (17th–18th century). Before the firearms trade, African societies sustained a stable decentralized order. Henn and Robinson (2025) document that in 1880, the continent was home to an estimated 45,000 independent polities, fewer than 2% of which could be classified as states, with inter-polity conflict rates well below those of contemporary Europe or Asia. These societies had deliberately organized to prevent political centralization, and, by their own objectives, succeeded. In terms of the model, this is the virtuous cycle operating at scale: decentralized communities sustaining order through local deterrence with low-p weapons.

The Atlantic slave trade disrupted this equilibrium. European traders exchanged firearms for captives along the West African coast. Whatley (2018) documents that gun imports and slave exports were cointegrated, rising together in a self-sustaining cycle. Firearms introduced a high-p weapon into these societies alongside large, expropriable prizes: captured people, traded for guns to capture more. The model predicts that this combination triggers the vicious cycle: violence raises the hazard rate, agents divest from productive activity, the value of life falls, and violence deepens. Nunn (2008) estimates that these effects persist: regions more exposed to the slave trades remain poorer today. Nunn and Wantchekon (2011) document that the affected populations still exhibit lower interpersonal trust. The model provides a candidate mechanism for these reduced-form findings. The technology of violence, interacting with baseline economic conditions, produced a self-reinforcing trap that operated through the endogenous value of life. The same mechanism functioned simultaneously among Native American nations, where firearms exchanged for furs triggered parallel cycles of inter-group violence (Silverman, 2016).

Medieval Iceland (930–1262). For over three centuries, armed households in Iceland sustained social order without a sovereign. Disputes were adjudicated at assemblies; enforcement relied on private threats of retaliation through the blood-feud system (Friedman, 1979). The

weapon technology—swords, axes, and shields—structurally privileged defense over offense (low p). Farmers had livestock, land, and trade networks worth protecting (moderate $\bar{a}$). The model predicts that these conditions sustain deterrence. Agents have both the means to retaliate and a future worth safeguarding. Iceland demonstrates that stateless order is possible under the conditions the model specifies.

The late Roman Republic (133–27 BC). Rome's descent from republic to civil war was not triggered by a change in weapon technology. Over the second and first centuries BC, conquest and the growth of slave-worked estates displaced the smallholding farmers who had formed the backbone of the citizen-militia (Brunt, 1971; Hopkins, 1978). A growing share of the armed population had no land, no trade, and no livelihood to protect ($\bar{a} \to 0$ for the military class). Soldiers' loyalty shifted from the state to the generals who promised them land and spoils upon discharge. Simultaneously, imperial expansion generated enormous expropriable prizes: provincial wealth, land distributions, and the plunder of conquest (fat-tailed $F(w)$). In the model's terms, this is the vicious cycle triggered through the economic channel rather than the technological one. A collapse in the baseline prosperity of the armed population combined with a surge in conflict stakes. Rome demonstrates the model's generality whereby the mechanism does not require a weapon shock.

The model also speaks to contemporary settings. Applied to nuclear strategy, it refines the Mutually Assured Destruction framework of Schelling (1960, 1966) by identifying a condition that framework leaves implicit: deterrence requires not only the physical capacity for retaliation but also baseline prosperity high enough that both parties have a future worth protecting. The model predicts that nuclear deterrence between high-$\bar{a}$ superpowers with credible second-strike capability (moderate p) is stable, but that proliferation to low-$\bar{a}$ states need not reproduce this stability.

Applied domestically, the model predicts that the effect of weapon availability on violence should be nonlinear in prosperity: deterrent above a threshold, destructive below it. Existing evidence is consistent with this prediction. Kennedy et al. (1998) show that income inequality is strongly correlated with firearm violent crime across US states, and that this relationship operates through the erosion of social capital and trust. Sampson, Raudenbush and Earls (1997) find that concentrated disadvantage predicts neighborhood violence in Chicago and that this association is mediated by collective efficacy (i.e, the willingness of residents to intervene on behalf of the common good). The model provides a micro-foundation for these patterns. In neighborhoods characterized by poverty (low $\bar{a}$), high encounter rates (high λ), and firearms with high offensive advantage (high p), the mortality tax on investment overwhelms the deterrent effect of lethality, producing a self-reinforcing trap.

Related Literature

The paper's contribution sits at the intersection of several literatures that have developed in parallel but not been formally connected. Existing models of deterrence, from the economics of crime to strategic interaction between armed agents, treat the stakes of peace as fixed. General equilibrium models of conflict endogenize the opportunity cost of violence through labor markets but not through the expected duration of life. The offense-defense literature in international relations distinguishes weapon types but does not connect them to investment or economic development. The conflict-and-development literature documents that violence causes persistent economic damage but lacks a formal theory of the technology of violence. The human capital literature recognizes that mortality risk affects investment incentives but has not endogenized that risk to the conflict environment. Our framework connects these literatures by making the value of life an equilibrium outcome, jointly determined by the physical characteristics of weapons and the economic conditions agents face. We discuss each body of work in turn and identify where our model departs from existing work.

Deterrence and the economics of conflict. Becker (1968) provides the canonical framework for criminal deterrence, in which the decision to offend depends on the expected penalty. Lee and McCrary (2017) test the deterrence effect of sentence severity. Lott (2000) applies the logic to firearms, arguing that gun prevalence reduces crime; Donohue (2004) contests the finding. Our model recovers the standard deterrence prediction as a special case (Proposition 1), then identifies the conditions under which it reverses (Propositions 7–8). The reversal occurs through a channel absent from the existing literature: the endogeneity of the value of life to the conflict environment.

Strategic interaction and weapons. Chassang and Padró i Miquel (2010) is the closest paper in the formal conflict-deterrence literature. They model conflict between two symmetric agents under strategic risk and show that weapons can have a nonmonotonic effect on peace. Weapons deter under complete information but can destabilize under strategic uncertainty, because weapons that protect against attack also shield an aggressor against retaliation. Our paper differs in three respects. First, we decompose weapon technology into three parameters (p, q, k) rather than treating weapons as a scalar, which generates predictions about *which* weapon attributes deter and which undermine deterrence. Second, we extend the analysis to group conflict and show that multi-target capacity (k) collapses the group-based retaliation that can restore deterrence even when p is high. Third, and most important, we embed the micro-model in a macro-framework with endogenous investment, so that the value of life responds to the conflict environment. This produces the virtuous and vicious cycles, which are absent from the Chassang and Padró i Miquel (2010) framework. Their model takes the stakes of peace as exogenous; ours makes them an equilibrium outcome.

In a related paper, Chassang and Padró i Miquel (2009) show that defensive weapons can be

destabilizing because they protect aggressors as well as victims. This insight is related to our result that offensive advantage (p) erodes deterrence by insulating the attacker from retaliation, but our mechanism operates through a different channel. The offense-defense balance shapes the threat of retaliation in the micro-game, which in turn shapes investment incentives in the macro-framework.

Conflict in general equilibrium. Dal Bó and Dal Bó (2011) incorporate appropriation activities into a general equilibrium trade model and show that shocks to capital-intensive industries increase conflict while shocks to labor-intensive industries reduce it, because the two types of shocks have opposite effects on the opportunity cost of violence. Our paper shares the insight that economic conditions shape the returns to conflict, but the mechanisms differ. In Dal Bó and Dal Bó (2011), the opportunity cost of violence is determined by the wage rate, and the model is static. In our framework, the relevant object is the forward-looking value of life, which depends on the expected stream of future utility discounted by the conflict-adjusted hazard rate. This produces a dynamic feedback. Violence reduces life expectancy, which reduces investment, which reduces the value of life, which increases violence. The self-reinforcing character of the trap is absent from opportunity-cost models.

Offense-defense balance. Jervis (1978) introduces the distinction between offensive and defensive technologies in international relations, arguing that a world favoring offense is more conflict-prone. Glaser and Kaufmann (1998) formalize offense-defense theory. Our micro-model provides an economic foundation for this distinction by decomposing p as a function of offensive and defensive weapon potentials, and derives the comparative statics that Jervis's framework describes verbally. We extend the analysis beyond the international-relations setting to interpersonal violence and connect it to investment behavior and economic development.

Conflict and development. Besley and Persson (2011) model how economic shocks and political institutions determine whether societies experience peace, repression, or civil war. Their framework focuses on state capacity and political incentives; ours focuses on the technology of violence and the endogenous value of life. The mechanisms are complementary. Rohner, Thoenig and Zilibotti (2013) study how conflict erodes trust and generates persistence through ethnic polarization. Our persistence mechanism operates through a different channel; namely, the feedback between the hazard rate and investment, which does not require ethnic heterogeneity.

Nunn (2008) and Nunn and Wantchekon (2011) document the persistent effects of the slave trades on African development and trust but do not provide a formal theory of the underlying dynamic. Whatley (2018) documents the gun-slave cycle as a proximate driver of slave exports. Henn and Robinson (2025) establish that pre-colonial African societies were deliberately organized to prevent centralization and sustained low levels of inter-polity conflict—the pre-shock equilibrium that our model characterizes as the virtuous cycle. Our model provides a candidate mechanism connecting these findings. The interaction of weapons technology (p, q) with baseline economic

conditions ($\bar{a}$) produces a self-reinforcing trap that works through the endogenous value of life and is consistent with the persistent underdevelopment these papers measure. Whatley and Gillezeau (2011) provide further evidence that the slave trades eroded institutional quality near points of extraction. Fenske and Kala (2017) show that the 1807 British abolition generated economic shocks and conflict, consistent with the model's prediction that removing an external demand shock need not immediately reverse the trap if the vicious cycle has become self-sustaining.

The value of life and mortality. The idea that mortality risk affects investment incentives has antecedents in the human capital literature. Ben-Porath (1967) shows that a longer time horizon increases the return to human capital investment. Our contribution is to endogenize the relevant mortality risk to the conflict environment and show that the resulting feedback can reverse the standard deterrence prediction. The conflict-adjusted hazard rate Ω in our model plays a role analogous to the discount rate in Ben-Porath's framework, but it is an equilibrium object that depends on weapons technology and investment decisions simultaneously.

2 A Model of Armed Conflict

We model an interpersonal conflict that may result in a violent altercation as an infinite-horizon sequential game of perfect information between two agents, $\mathscr{I} = \{1,2\}$. Agents alternate moves; player 1 is active in even periods ($t = 0,2,4,\dots$) and player 2 in odd periods. We view this game as capturing a transient situation providing the opportunity for a violent resolution, though the underlying conflict may be long-lived. Accordingly, we set the per-period discount factor within the altercation to $\delta = 1$.

In any period t, the active agent i's action space is $A_i = \{A,P\}$, where A and P denote "attack" and "pass," respectively. Agent i possesses a weapon characterized by offensive potential o_i and defensive potential d_i. An attack succeeds with probability $p_i = f(o_i,d_j) \in (0,1)$, which terminates the game. If an attack fails, the game transitions to $t+1$ and the opponent j becomes active. The game also terminates if both agents choose P in succession.

Payoffs are determined by the terminal state. Consecutive passes yield a baseline utility of peace, a, for both agents. A successful attack by i yields a winner's payoff $u_W = a+w$, where $w > 0$ represents the spoils of war. The loser dies with probability $q_i \in (0,1)$, yielding 0, or survives with probability $1-q_i$ to suffer an injury penalty $l > 0$, yielding an expected payoff $u_L = (1-q_i)(a-l)$.

Weapon Characteristics. To ground the weapon parameters in real-world technology, consider how different weapons map to the vector (o_i,d_i,q_i) and the implied offensive (i.e., first-mover) advantage, p_i. A firearm, for instance, represents extreme offensive potential (o_i) coupled with near-zero defensive utility (d_i). This implies $p_i \to 1$; an attacker is virtually guaranteed to succeed before the opponent can respond, and the resulting injury is highly lethal (large q_i). Con-

versely, traditional armored melee combat (e.g., sword and shield) structurally privileges defense over offense, yielding a low p_i but maintaining a high q_i in the event a strike connects. Modern incapacitants, such as pepper spray, feature a high p_i through rapid offense but near-zero lethality ($q_i \to 0$). Finally, unarmed combat features moderate bounds on both p_i and q_i, where deterrence is driven primarily by the non-lethal injury penalty, l. By varying these structural parameters, the model captures a broad spectrum of weapon classes.

Markov Equilibria: No Memoryless Peace. In Markov Perfect Equilibria (MPE), strategies depend only on the payoff-relevant state (in our game, the identity of the active player). If agents play memoryless strategies, the only punishment for a failed unilateral deviation from a "Mutual Restraint" profile (always play P) is the delay before the state resets and restraint resumes. Because we evaluate the limit $\delta = 1$, the cost of this delay vanishes.

Theorem 1. *No MPE of Mutual Restraint exists. The unique MPE is "Full Escalation," where both agents choose A at every information set.*

The economic intuition is stark: absent time discounting[1] or history-dependent punishment, a surprise attack constitutes a risk-free lottery ticket. There is no downside to failure, rendering memoryless peace structurally impossible.

Grim Trigger Equilibria: Peace by Retaliatory Threat. We allow for history-dependent play via Grim Trigger strategies to capture the logic of deterrence. Both players adopt a strategy to play P if the history of play consists entirely of P, and to play A otherwise. A deviation from mutual restraint plunges the game into a Full Escalation subgame.

Let V_i^* and W_i^* denote the expected continuation values of the active and passive player, respectively, in the Full Escalation subgame. The expected payoff of deviating from Mutual Restraint is the value of entering this subgame as the attacker, $U_i^{dev} = V_i^*$.

Theorem 2. *Under Grim Trigger strategies, the expected deviation payoff is given by*

$$U_i^{dev} = \frac{p_i u_W + (1-p_i) p_j u_L}{p_i + p_j - p_i p_j}$$

A Subgame Perfect Nash Equilibrium (SPNE) of Mutual Restraint exists if and only if $a \geq U_i^{dev}$ for all $i \in \{1,2\}$.

Unlike Markov strategies, history-dependent strategies can sustain peace without time discounting. Restraint holds when the threat of the Full Escalation subgame outweighs the immediate temptation of the spoils of war.

[1] For $\delta < 1$, there is the artificial pacifying effect that the cost of a failed attack delays consumption of a.

2.1 The Standard Deterrence Argument

The stability of the Grim Trigger equilibrium is sensitive to the technological parameters of the agents' weapons.

Lethality as Deterrent. Consider a symmetric setting in which $q_i = q_j = q$.

Proposition 1. *An increase in weapon lethality strictly decreases the expected deviation payoff* ($\frac{\partial U_i^{dev}}{\partial q} < 0$).

Restraint is driven by the fear of the Full Escalation following a failed attack. The retaliatory threat is a stronger deterrent when the consequence of a potential defeat is more severe. Thus, the Mutual Restraint equilibrium is more stable when the agents' weapons are more lethal. This is the argument that guns will deter altercations due to the higher potential harm to either side.

Fragility of Asymmetry. Compare an *asymmetric* environment where player i has a weak weapon (p_l) and player j has a strong weapon (p_h), with a *symmetric* environment where both players possess the strong weapon (p_h).

Proposition 2. *Violent conflict is strictly more likely in the asymmetric environment than in the symmetric environment, as* $U_j^{dev}(p_h, p_l) > U^{dev}(p_h, p_h)$.

The asymmetric environment incentivizes deviation; the strong player faces a low threat of effective retaliation, while the weak player may launch a preemptive strike—a Hobbesian trap. Symmetry supports mutual deterrence. Taken together, Propositions 1–2 suggest that Mutual Restraint is most likely to arise when all agents possess the most lethal weapons. This resonates with the argument that a proliferation of firearms would be pacifying.

2.2 The Limits of Deterrence

We identify two caveats to the standard argument in this baseline setting, and a third one when allowing for incomplete information.

Offensive Advantage and Limits of Lethality. The advantage of firearms over weaker weapons is twofold: firearms are significantly more lethal (q) and provide a stronger offensive advantage (p; see Jervis 1978), privileging offensive potential over defensive resilience.

Consider a symmetric setting where $p_i = p_j = p$ and $q_i = q_j = q$. We define a structural offensive advantage as a technological state where offensive potential heavily outweighs defensive potential, driving $p \to 1$.

A high offensive advantage means attacking first is decisive—the attacker is insulated from retaliation, creating pressure to preempt.

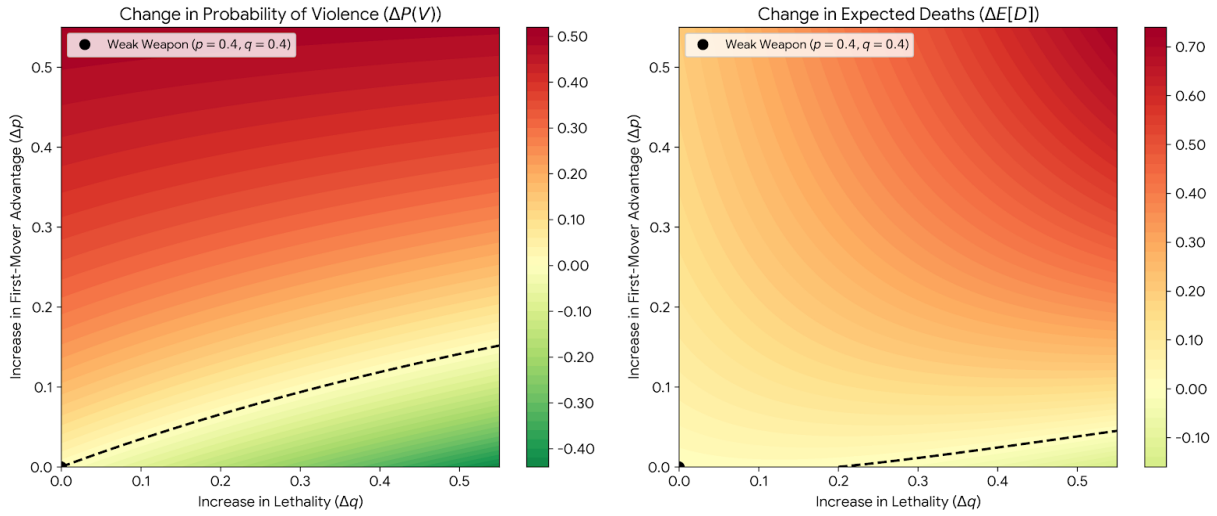


Figure 1: **The Dual Effects of Weapons Technology on Violence and Mortality.** The heatmaps illustrate the comparative statics of moving from a symmetric "weak" weapon baseline ($p = 0.4, q = 0.4$) to a "strong" weapon characterized by increases in the offensive advantage (Δp) and lethality (Δq). The left panel shows the change in the probability of a violent altercation, while the right panel shows the change in the expected number of deaths. The dashed black line indicates the isoquant of zero net change relative to the baseline. Parameters: $a = 10$, $l = 2$, with the temptation w distributed uniformly on $U[0, 15]$.

Proposition 3. *A greater offensive advantage increases the expected deviation payoff* $\left(\frac{\partial U_i^{dev}}{\partial p} > 0\right)$. *A growing offensive advantage weakens the pacifying effect of lethality* $\left(\frac{\partial^2 U^{dev}}{\partial p \partial q} > 0\right)$.

Deterrence fundamentally relies on the attacker facing retaliation following a failed strike. If weapons are overwhelmingly offensive, the attacker expects to win immediately. The shadow of retaliation vanishes and loses its capacity to sustain Mutual Restraint. The deterrent power of lethality is therefore a *second-order* effect: the threat of lethality deters attackers only to the extent that their strike might fail. As a weapon's offensive advantage grows ($p \to 1$), attacks are increasingly insulated from the consequences of retaliation, causing the deterrent power of lethality to collapse.

Consequently, even in a perfectly symmetric setting where Mutual Restraint is most viable, whether a societal transition to firearms decreases the overall probability of violence depends on which of these two technological shifts dominates. As the left panel of Figure 1 shows, the zero-change isoquant bends downward because as the offensive advantage increases, marginally higher lethality provides diminishing deterrent returns. Indeed, for deterrence to hold, the offensive advantage must remain moderate.

For weapons where the offensive advantage is significant, Propositions 1 and 2 are *moot*.

Lethality Paradox: Deterrence vs. Deaths. The standard argument does not distinguish between the frequency of conflicts and the frequency of casualties. Suppose the spoils of war, w, are drawn by Nature from a continuous distribution $F(w)$ with density $f(w)$, while all other parameters are symmetric and fixed. Let $w^*(q)$ denote the threshold above which Mutual Restraint collapses ($w > w^*$).

Proposition 4. *An increase in weapon lethality strictly lowers the likelihood of violent conflict. However, the expected number of deaths strictly rises if the hazard rate of the conflict stakes at the marginal threshold is sufficiently small:*

$$h(w^*) < \frac{1}{q(1-p)(a-l)}. \tag{1}$$

This condition is always satisfied as the offensive advantage becomes absolute ($p \to 1$).

If the hazard rate $h(w^*) = f(w^*)/[1-F(w^*)]$ is sufficiently low, increasing lethality deters only a negligible number of marginal altercations relative to the total volume of violence. Such a regime can arise for two distinct statistical reasons. First, the distribution of w may possess a fat right tail, where conflict is frequently driven by high stakes that remain profitable to contest despite heightened risk. Second, the violence threshold w^* may endogenously be far in the left tail, representing an environment where the baseline proclivity for violence $1-F(w^*)$ is high. In either case, a marginal increase in lethality q and thus the threshold w^* deters few but makes the vast number of "inframarginal" conflicts deadlier.

Proposition 4 also has a connection to Proposition 3. As $p \to 1$, the right-hand side of (1) grows to ∞ while the left-hand side goes to $h(0) = f(0)$ as $w^* \to 0$, ensuring that the inequality holds. When the offensive advantage p increases, the attacker becomes increasingly insulated from retaliatory risk. The (marginal) deterrent value of lethality vanishes. At the limit $p \to 1$, lethality no longer deters attacks, so any baseline probability of conflict ensures that introducing deadlier weapons mechanically translates into a higher expected death toll. That is, p influences whether the deterrent effect or the deadliness of lethality dominates.

Comparing the two panels of Figure 1 illustrate this potential divergence. Even in scenarios where the transition to firearms successfully enhances deterrence and reduces the frequency of violent altercations, the cost in human lives may nevertheless increase. Because the altercations that occur are more lethal, the increase in mortality per fight overwhelms the benefits of reduced conflict frequency.

The effects in Propositions 3–4 will play a crucial role in our analysis in Section 4.

Hidden Weapons and Failure of Deterrence. A third caveat is that deterrence is weaker when agents cannot reliably judge how dangerous the opponent is. If weapons are not visible until

they are used, strong agents have an incentive to probe others in search of weaker targets. In that environment, remaining passive is interpreted as a sign of weakness. Agents may initiate probing attacks, suspecting others may be weak, and encounters can escalate unexpectedly. Thus, hidden lethality undermines deterrence: when capabilities are visible, lethal weapons deter aggression from the outset, but when they are not, they do not prevent initial tests of strength and instead make mistaken confrontations—or warning shots that establish caution—far more deadly. A formal analysis of these dynamics, including conditions under which agents merely brandish their weapons rather than strike, is provided in Appendix B.[2]

3 Group Combat and the Collapse of Deterrence

The baseline analysis stresses a fragility in deterrence: a high offensive advantage (p) insulates an attacker from retaliation, thereby undermining the deterrent effect of lethality (q). We now extend the framework to group conflicts to examine how multi-agent combat qualifies this conclusion.

3.1 How Groups Restore Deterrence

Consider a sequential game of perfect information between two groups, $i \in \{1,2\}$, with sizes $N_1 \geq 1$ and $N_2 \geq 1$. Agents are armed with symmetric weapons (p,q).

In each active period for Group i, Nature selects one non-incapacitated member to choose between A and P. An attack targets one randomly chosen non-incapacitated member of Group j. The game ends after two consecutive P plays (Mutual Restraint) or if one group is entirely incapacitated (decisive victory).

Dead agents receive 0. If the game ends in Mutual Restraint, surviving agents receive utility a, or $a-l$ if incapacitated. In case of a decisive victory, survivors of the winning group receive $a+w$, or $a+w-l$ if incapacitated. Survivors of the losing group receive $a-w$, or $a-w-l$ if incapacitated. We assume parameters are restricted such that all realized utilities are positive.

To demonstrate that group fights can restore deterrence, we examine the limit of an absolute offensive advantage ($p \to 1$). In the bilateral model, this condition guarantees violence.

Proposition 5. *In a symmetric group altercation ($N_1 = N_2 = N \geq 2$), an absolute offensive advantage ($p \to 1$) does not eliminate the deterrent effect of lethality. Mutual restraint can be sustained as a subgame perfect equilibrium provided q is sufficiently high.*

[2] One might conjecture that hidden weapons have the benefit of keeping two weak agents from engaging in a fight for parameters where they would fight under complete information. This is not an equilibrium when neither agent can unilaterally walk away from the confrontation. A weak offensive attacker who plays P signals weakness and would be attacked by both a weak or a strong opponent.

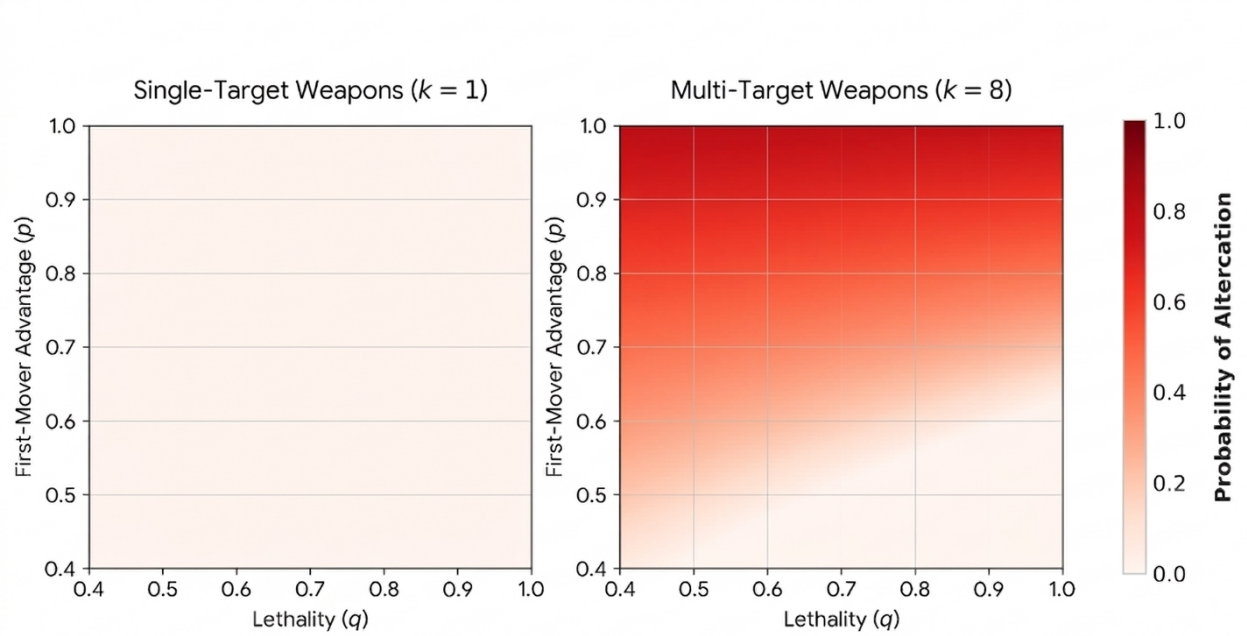


Figure 2: **Group Combat, Deterrence, and Multi-Target Technology.** This figure illustrates the impact of weapons technology on the probability of a violent altercation for a group of size $N = 10$. Both panels plot the probability of violence across the parameter space of lethality (q) and offensive advantage (p), restricted to the domain where $p, q \geq 0.4$. In the left panel ($k = 1$), single-target weapons preserve a large zone of perfect deterrence (the white region) because an attacker cannot eliminate the threat of group retaliation. In the right panel ($k = 8$), multi-target weapons allow the attacker to mitigate the attrition lottery, causing deterrence to collapse and the probability of violence to expand. In this case, increases in lethality (q) have no deterrent effect when the offensive advantage (p) is high, as in the single altercation model.

The intuition is that if an individual cannot instantly eradicate an entire group, surviving opponents are guaranteed to return fire. This introduces an attrition lottery, restoring the deterrent power of q. However, this stability is fragile: numerical superiority ($N_1 \gg N_2$) or coordinated strikes reduce the threat of retaliation, causing deterrence to collapse.

3.2 Multi-Target Capacity and the Collapse of Group Deterrence

We introduce a technological parameter $k \geq 1$, representing the maximum number of opponents a single weapon can target in one turn. Multi-target technology ($k > 1$) acts as a direct substitute for numerical superiority or coordinated strikes.

An agent with $k \geq N$ and $p \to 1$ can single-handedly eliminate the opposing group in one turn. Expected retaliation drops to zero, and deterrence collapses. The unraveling of deterrence does not require a weapon capable of wiping out the entire group instantly; it merely requires a weapon powerful enough to sufficiently compress the timeline of the conflict and thus the risk of absorbing

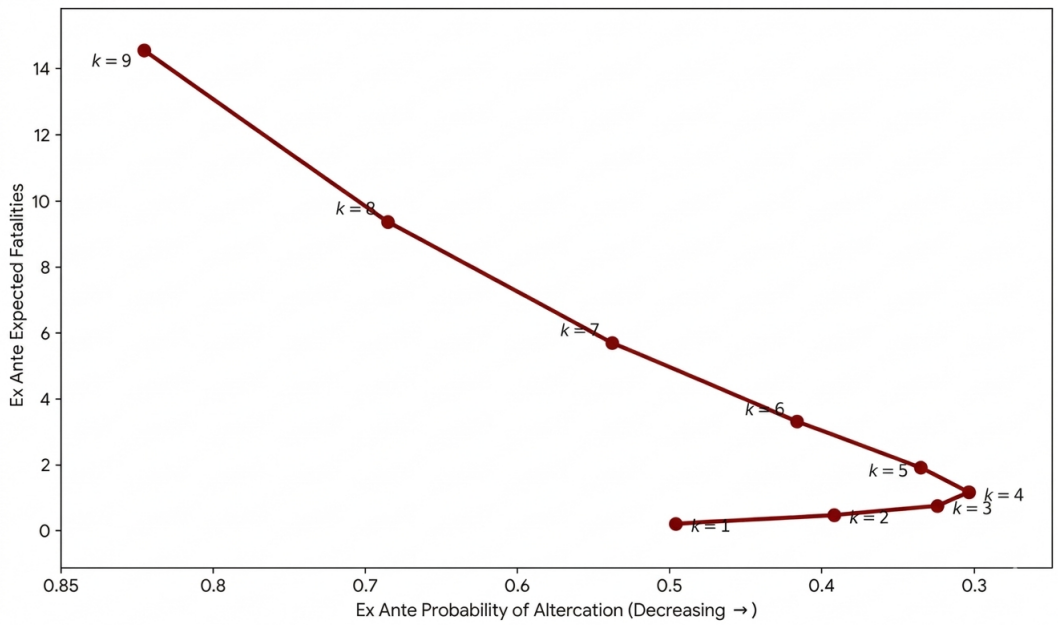


Figure 3: **The Ex Ante Risk Frontier.** This figure maps the trajectory of societal risk as multi-target capacity k and its associated lethality q increase. The horizontal axis shows the ex ante probability of an altercation (inverted, so safety increases to the right), while the vertical axis measures ex ante expected fatalities. The C-shaped curve illustrates that, in this example, initial technological advancements reduce the frequency of violence but strictly increase expected deaths, followed by an eventual collapse where both metrics worsen simultaneously. Parameters for this numerical illustration are set to $N = 10$, $p = 1.0$, $a = 15$, $l = 2$, lethality $q(k)$ scaling linearly from 0.2 to 0.95 for $k \in \{1, \ldots, 9\}$, and w distributed uniformly on $U[0, 10]$.

return fire.

Proposition 6. *In a symmetric group altercation ($N_1 = N_2 = N \geq 4$), for any state where Mutual Restraint holds under single-target weapons ($k = 1$), there exists a capacity $k < N$ that strictly increases the probability of Full Escalation. Furthermore, for sufficiently large w, there exists $k < N$ such that lethality unconditionally loses its deterrent effect as $p \to 1$.*

As the last part of the proposition emphasizes, high weapon capacity k renders the deterrent effect of lethality again second-order to the offensive advantage (see Figure 2).

Larger k amplifies the potential divergence between conflict frequency and human cost. To prevent a collapse of deterrence, lethality q must scale to compensate for reduced retaliation. If this equilibrium is achieved, conflict frequency remains suppressed, but any realized altercation is catastrophic. Technically, lethality q plausibly increases with weapon capacity k. In Figure 3, increasing k and q initially strengthens deterrence, but fatalities climb. Moreover, past a critical threshold for k, lethality loses its deterrent power (given a high p) and altercations become both

more frequent and more lethal.

4 Polite Societies and Desperation Traps

The previous analysis has shown that the offensive advantage p and multi-target capacity k of a weapon can neutralize the deterrent power of its lethality q. Still, the standard, more narrow argument remained valid: a *ceteris paribus* increase in lethality weakly raises the threshold for violence.

We now embed the altercation model into a dynamic population framework. We show that even the narrow version of the argument can break down in high-conflict environments due to a vicious cycle between lethal violence and the disincentive to invest in a valuable life.

To parsimoniously focus on the relevant dynamics, we restrict attention to one-on-one altercations as in the baseline model. Our results extend to multi-agent altercations with sufficiently high k.

Environment. Time is continuous. There is a continuum of ex-ante identical, risk-neutral agents. To maintain a stationary population, agents who exit (via non-violent death rate $d > 0$ or by being killed) are immediately replaced.

Upon entry, agents make a one-time sunk investment $e \geq 0$ at cost $c(e)$. This effort determines a permanent flow utility $a(e)$. Agents discount at rate $r > 0$ and face pairwise encounters at Poisson rate $\lambda > 0$, where a prize $w \sim F(w)$ is drawn. Nature assigns First-Mover (FM) and Second-Mover (SM) roles with equal probability. Peace requires Mutual Restraint. If an agent chooses to attack, Full Escalation ensues (as in the Grim Trigger equilibrium in Section 2).

We assume $c'(e) > 0$, $c''(e) > 0$, $a'(e) > 0$, and $a''(e) \leq 0$.

Win Probabilities. An encounter mirrors the baseline "duel" model analyzed in Section 1. In a Full Escalation subgame, each strike has an independent success probability $p \in (0,1]$. Because the FM holds the positional advantage of the initial shot, summing the infinite geometric series of alternating survival branches yields a FM win probability of:

$$\pi_{FM} = \sum_{k=0}^{\infty} p(1-p)^{2k} = \frac{1}{2-p} \tag{2}$$

The SM wins with the complementary probability:

$$\pi_{SM} = (1-p) \sum_{k=0}^{\infty} p(1-p)^{2k} = \frac{1-p}{2-p} \tag{3}$$

The winner claims the prize w. The loser faces immediate death with probability q or survives with an exogenous injury penalty $l > 0$ with probability $1-q$.

Bellman Equation. Following the logic of Section 2, the threshold for violence is:

$$w^* = (1-p)[qV(e) + (1-q)l] \tag{4}$$

(see equation (38) in the proof of Proposition 4).

The steady-state value of an agent's life includes the flow utility and expected gains or losses from encounters. Since an agent is FM or SM with equal probability, the Bellman equation is:

$$\begin{aligned}(r+d)V(e) = a(e) &+ \frac{\lambda}{2}\int_{w^*}^{\infty}(\pi_{FM}w - \pi_{SM}[qV(e)+(1-q)l])\,dF(w) \\ &+ \frac{\lambda}{2}\int_{w^*}^{\infty}(\pi_{SM}w - \pi_{FM}[qV(e)+(1-q)l])\,dF(w)\end{aligned} \tag{5}$$

As $\pi_{FM} + \pi_{SM} = 1$, the integral terms simplify, yielding

$$\Omega V(e) = a(e) + \frac{\lambda}{2}\int_{w^*}^{\infty}(w-(1-q)l)dF(w) \tag{6}$$

where the *conflict-adjusted* hazard rate $\Omega \equiv r + d + \frac{\lambda q}{2}(1-F(w^*))$ shows that each agent's life is discounted not just by the time discount rate (r) or background mortality (d), but also by the lethal hazard in a violent environment ($\frac{\lambda q}{2}(1-F(w^*))$).

Equilibrium Investment. Agents maximize $V(e) - c(e)$, yielding the first-order condition:

$$\frac{a'(e^*)}{\Omega} = c'(e^*) \tag{7}$$

The left-hand side $a'(e)/\Omega$ is the present value of the marginal flow utility from investing, given the conflict-adjusted hazard rate.

4.1 The Virtuous Cycle

We first characterize the conditions under which an increase in weapon lethality q generates a virtuous cycle of peace and investment. In this regime, the marginal deterrent effect of lethality outweighs its direct contribution to the hazard rate, leading to a safer environment that promotes investment in life-quality.

Define $h(w) = f(w)/(1-F(w))$ as the hazard rate of the distribution of the spoils of war.

Proposition 7. *An increase in weapon lethality q generates a Virtuous Cycle of Deterrence, characterized by higher equilibrium investment* $\left(\frac{de^*}{dq} > 0\right)$ *and a strictly lower conflict-adjusted hazard rate* $\left(\frac{d\Omega}{dq} < 0\right)$*, if and only if the marginal deterrence of the weapon satisfies the following distri-*

butional condition:

$$q \cdot h(w^*) \cdot \frac{dw^*}{dq} > 1 \tag{8}$$

where $\frac{dw^*}{dq}$ *is the equilibrium sensitivity of the violence threshold to lethality.*

The intuition for condition (8) rests on the tension between the deterrent power of deadlier weapons and their direct mortality cost. When the distribution of conflict stakes is concentrated around the marginal deterrence threshold (i.e., the hazard rate $h(w^*)$ is high), a slight increase in lethality q successfully deters a massive volume of potential altercations. If this reduction in conflict frequency more than compensates for the increased deadliness of the remaining fights, the society experiences a decline in the hazard rate Ω.

This initial safety improvement triggers a strategic complementarity between life expectancy and life quality. Agents respond to the longer expected life span by increasing their investment e^*. The magnitude of this response depends on the elasticity of the investment cost function, $c''(e)$. When investment is highly elastic (small $c''(e)$), the improved safety yields a substantial rise in investment. Consequently, agents enjoy a higher flow utility $a(e)$, which increases their endogenous value of life V. The higher continuation value acts as a heavier anchor for deterrence, pushing the violence threshold w^* even higher and deterring further conflict. Crucially, this virtuous cycle requires that the offensive advantage p remains sufficiently low so that the threat of lethal retaliation remains viable.[3] In environments with a high encounter rate (high λ), the safety and investment gains of this self-reinforcing cycle are amplified.

4.2 The Vicious Cycle

When condition (8) fails, the society may enter a "vicious cycle." In this regime, an increase in lethality q raises the hazard rate Ω, which triggers a divestment from the future. This divestment erodes the value of life, potentially lowering the threshold for violence further and leading to a collapse of deterrence.

The effect of lethality q reflects two opposing forces. The first is the standard deterrence effect: higher q raises the cost of defeat and discourages aggression. The second is a desperation effect. When the environment becomes more lethal, the shadow cost of conflict in Ω increases, triggering a rational divestment from life. If investment is sufficiently elastic, this desperation effect can offset the higher physical risk of defeat and lower the threshold for violence. In that case, more lethal weapons may paradoxically increase the frequency of conflict.

[3] This is the effect derived in Section 2 that a high offensive advantage erodes the deterrent power of lethality. As $p \to 1$, the violence threshold becomes insensitive to q: $dw^*/dq \to 0$.

Proposition 8. *An increase in weapon lethality q generates a Vicious Cycle of Desperation, characterized by a decrease in the violence threshold $\left(\frac{dw^*}{dq} < 0\right)$ and a collapse in equilibrium investment $\left(\frac{de^*}{dq} < 0\right)$, if condition (8) fails and the investment cost function is sufficiently elastic (c'' is small).*

This equilibrium outcome characterizes a "desperation trap" wherein investment responses overturn the standard deterrence effect. This regime takes hold when marginal deterrence is inherently weak—either because the distribution of the spoils is fat-tailed (low $h(w^*)$), or because a strong offensive advantage ($p \to 1$) largely insulates attackers from retaliation. Under these conditions, an increase in lethality q fails to deter enough marginal violence to offset its deadliness, thus raising the overall mortality hazard in Ω.

This triggers a rational divestment from life. The divestment is greater when effort is more elastic (small $c''(e)$ and high $a'(e)$), resulting in a sharp drop in the expected value of life V. Because deterrence is anchored by what agents stand to lose, a collapse in V lowers the violence threshold w^*, which further raises the mortality hazard and reinforces the divestment spiral.

The vulnerability of an environment to this trap further rises when quality of life is generally poor (low level of $a(e)$), which weakens the deterrence channel, and altercations are endemic (high λ), which magnifies the mortality "tax" on the future provided that deadliness dominates deterrence. In a high-a-low-λ environment, agents retain a valuable future to safeguard and are hence more deterred by lethal risk. In a low-a-high-λ environment, the persistent shadow of violence depresses the value of the future so much that it begets more violence—lethal weapons breed desperation rather than "politeness."

Figure 4 depicts an example with $a(e) = \bar{a} + \tilde{a}(e)$, where $\bar{a}$ allows us to manipulate the level of the flow utility independent of $a'(e)$, to demonstrate the importance of environmental factors. At the same time, Figure 5 shows us that the technological aspects of the weapons can dominate those factors.

The maxim that an "armed society is a polite society" is not an absolute law of deterrence, but a highly conditional equilibrium. As summarized in Table 1, the comparative statics of the model reveal that whether an increase in weapon lethality promotes a "polite society" or a "desperation trap" depends on (the interaction between) the technology of violence and the socio-economic environment. Technological parameters—such as the offensive advantage p and lethality q—determine the micro-dynamics of conflict. Environmental and socio-economic factors—such as quality of life $\bar{a}$, investment frictions $c(e)$, baseline mortality d, frequency of conflict encounters λ, and the expropriability of resources $F(w)$—determine to what extent agents find aggressions worthwhile or have a future worth protecting. Lethal weapons do not inherently impose peace, but may simply act as an amplifier of prevailing economic incentives. In an environment anchored by prosperity

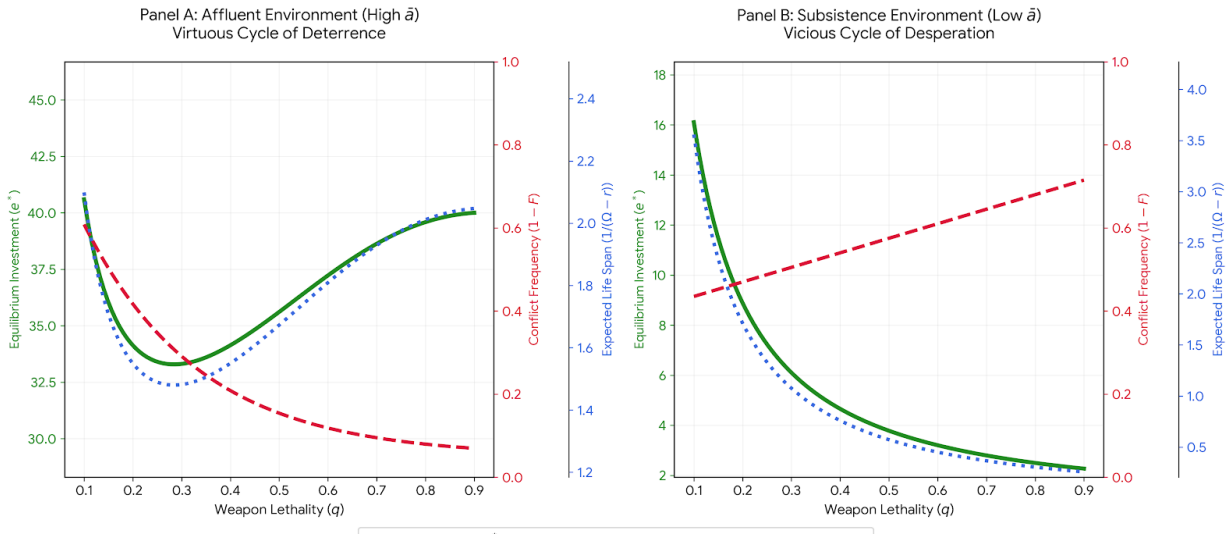


Figure 4: **The Effect of Lethality in Different Environments.** This figure illustrates how a baseline utility $\bar{a}$ dictates the equilibrium response to weapon lethality q. **Panel A (Affluence):** When baseline utility $\bar{a}$ is high, the deterrent effect of lethality is strong. Increasing q reduces conflict frequency, causing life expectancy to rebound and thus a virtuous cycle of investment. **Panel B (Subsistence):** When baseline utility $\bar{a}$ is low, the deterrent power of lethality is weak. As q increases, the environment becomes fatalistic: life expectancy falls and agents divest from the future. Because agents then have even less to lose, conflict frequency increases despite the deadliness of the weapons, resulting in a *vicious cycle* of desperation. *Parameters:* Offensive advantage $p = 0.5$, encounter rate $\lambda = 8.0$, loser penalty $l = 3.0$, baseline non-violent mortality $d = 0.02$, and discount rate $r = 0.05$. The spoils of conflict w are drawn from an exponential distribution with mean $\beta = 8.0$ (i.e., $F(w) = 1 - e^{-w/8}$). The flow utility is $a(e) = \bar{a} + 4\sqrt{e}$, and the cost of effort is $c(e) = 0.5e^2$. Crucially, $\bar{a} = 40$ in Panel A and $\bar{a} = 5$ in Panel B.

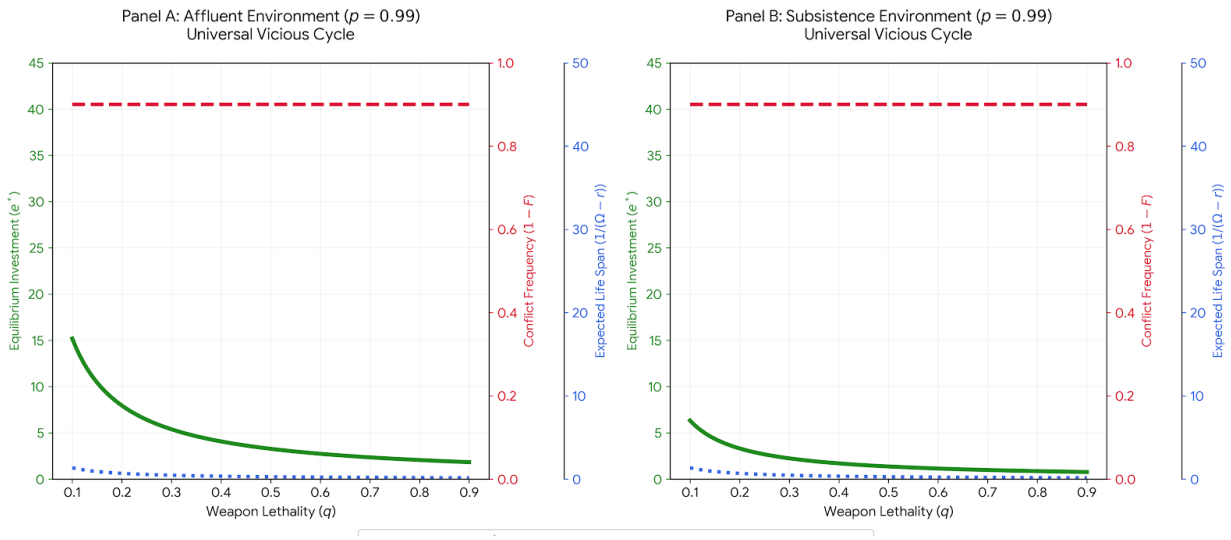


Figure 5: **The Insulated Attacker Limit** ($p = 0.99$)**: Universal Collapse of Deterrence.** This figure illustrates the equilibrium when the offensive advantage is near-absolute. Unlike in Figure 4, the stabilizing effect of a high $\bar{a}$ is completely neutralized. Because an attacker is virtually guaranteed a safe strike ($p = 0.99$), mutual deterrence fails. The threshold for violence collapses, regardless of the environment's baseline wealth. As weapon lethality (q) increases, it mechanically decrease the life expectancy, which in turn triggers a collapse in investment. Both affluent and subsistence environments fall into the vicious cycle of desperation. *Parameters:* Same as in Figure 4, except the offensive advantage is now set to $p = 0.99$.

and opportunity, they can reinforce mutual restraint; in communities at the margins of survival or fractured by systemic grievances, they accelerate a fatalistic spiral of violence and economic collapse. Moreover, at any rate, when the offensive advantage of weapons becomes absolute, the deterrence argument becomes moot.

Table 1: **Exogenous Parameters, Equilibrium Effects, and Economic Interpretations**

Parameter	Equilibrium Effect	Economic Interpretation
Panel A: Endowments and Production		
$\bar{a}$	**Shifts toward Virtuous Cycle:** Increases the baseline value of life (V), shifting the violence threshold (w^*) upward.	Baseline environmental wealth, unconditional institutional support, subsistence quality, or geographically endowed public goods.
$\tilde{a}(e)$	**Amplifies Prevailing Cycle:** A higher marginal return to effort raises V, but only if the expected life span is long enough to justify investing e.	Productivity of capital/labor, quality of property rights over produced goods, or technological efficiency.
$c(e)$	**Shifts toward Vicious Cycle:** Higher marginal costs reduce investment (e^*) and lower V, shifting the violence threshold (w^*) downward.	Frictions in investment, cost of capital/credit, or barriers to entry in formal markets.
Panel B: The Technology and Structure of Conflict		
$F(w)$	**Determines Cycle Elasticity:** A fat right tail (high variance/mean) ensures a persistent supply of violent shocks that override deterrence.	Expropriability of wealth (e.g., point-source resources vs. human capital), baseline inequality, or the intensity of non-material grievances.
p	**Shifts toward Vicious Cycle:** As $p \to 1$, attacker is insulated, driving $w^* \to 0$ and guaranteeing a Vicious Cycle regardless of other parameters.	The offense-defense balance of technology, weapon range, advantage of surprise.
λ	**Amplifies Prevailing Cycle:** In a Vicious Cycle, high λ accelerates mortality and divestment. In a Virtuous Cycle, multiplies deterrence.	Population density, frequency of social interactions, urbanization without segregation, or state capacity to police encounters.
Panel C: Discounting and Mortality		
d	**Shifts toward Vicious Cycle:** Higher baseline mortality reduces the expected horizon, directly lowering V and making agents more fatalistic.	Endemic disease burden, malnutrition rates, or exposure to natural disasters.
r	**Shifts toward Vicious Cycle:** Heavier discounting shrinks the present value of future, lowering V and reducing the incentive to avoid lethal risks.	Time preference, institutional instability, inflation, generalized economic uncertainty, myopia.

5 Historical Evidence

The model identifies a feedback between the technology of violence and the endogenous value of human life that determines whether weapons sustain peace or produce conflict. This section shows that the mechanism is consistent with four historical episodes spanning different centuries, continents, and institutional settings. We organize the evidence around the model's two regimes. Sections 5.1 and 5.2 illustrate the virtuous cycle. Sections 5.3 and 5.4 illustrate the vicious cycle. Each episode is driven by a distinct combination of the model's parameters. Because these societies shared no institutions, no trade networks, and no common cultural inheritance, the consistency of the pattern across episodes suggests that it reflects the mechanism itself rather than an unobserved common factor. Each episode, moreover, documents an outcome that existing formal models of deterrence and conflict do not predict.

We emphasize at the outset that these applications are illustrative, not causal. The model provides a candidate mechanism consistent with documented outcomes; it does not claim to identify the mechanism to the exclusion of all alternatives. We note where the model's predictions align with the historical record, where the fit is imperfect, and where competing explanations exist.

These four episodes do not exhaust the model's empirical scope. The American frontier, for instance, combined high-p weapons with moderate-to-high $\bar{a}$ and rapid institutional development—a setting where the model predicts the virtuous cycle could survive despite firearms, broadly consistent with the outcome but confounded by simultaneous state-building that the model does not capture. We selected episodes where the relevant parameter shifts are historically documented and the long-run outcomes have been measured independently in the existing literature.

5.1 Tokugawa Japan: Disarmament and the Virtuous Cycle

Japan provides the cleanest historical illustration of the model's predictions. The same society experienced both regimes in sequence. Firearms triggered escalating violence, a deliberate policy intervention reduced the offensive advantage of available weapons, and two centuries of peace and economic development followed.

The firearms shock. Portuguese traders introduced matchlock firearms (*tanegashima*) to the island of Tanegashima in 1543. Japanese gunsmiths reverse-engineered the weapons within a decade. By the 1570s, Japan was producing firearms in quantities that likely exceeded contemporary European output (Perrin, 1979; Chase, 2003). Warlords integrated firearms into large-scale battles during the Sengoku (Warring States) period (1467–1615), deploying organized volley-fire tactics that predated their European adoption (Perrin, 1979).

In terms of the model, firearms raised both p (the offensive advantage of a first strike) and q (lethality) relative to the sword-and-polearm baseline. Volley fire also raised k: coordinated ranks

of gunners could engage multiple opponents per exchange, compressing the timeline of battle and reducing the threat of return fire (Proposition 6). Proposition 3 predicts that the increase in p erodes the deterrent value of q. The historical record is consistent—the introduction of firearms did not produce mutual deterrence among the competing domains. Civil war intensified. The Sengoku period produced over a century of near-continuous conflict among hundreds of rival warlords (Totman, 1993; Jansen, 2000).

The policy intervention. Political unification under Toyotomi Hideyoshi (1590) and the Tokugawa shogunate (1603) was followed by measures that restricted firearms. Hideyoshi's "sword hunt" (*katanagari*) of 1588 disarmed the peasantry. The Tokugawa regime subsequently centralized firearms production, confined gun ownership to the samurai class, and imposed regulations that progressively reduced the number of gunsmiths (Perrin, 1979). The dominant weapon technology reverted to swords and polearms, which structurally privilege defense over offense in close combat. A swordsman must close to arm's length, giving the defender time to react and retaliate, and a shield or armor provides meaningful protection.

Per the model, this intervention reduced p and returned k to 1. Proposition 7 predicts that reducing p while maintaining moderate q restores the conditions for the virtuous cycle, because the retaliatory threat becomes credible and the deterrent value of lethality is restored.

The virtuous cycle. The Tokugawa period (1603–1868) produced over two centuries of internal peace. Population grew from roughly 12 million in 1600 to over 30 million by 1720. Agricultural productivity rose through labor-intensive techniques, market networks expanded across domains, and urban centers developed into commercial hubs (Totman, 1993; Jansen, 2000). Hayami (2015) documents what he terms an "Industrious Revolution": a sustained increase in labor productivity, market participation, and living standards during the Tokugawa period that predated and prepared the ground for Japan's rapid industrialization after the Meiji Restoration. This pattern is consistent with the model's investment response: a lower conflict-adjusted hazard rate Ω raises the return to investment in human capital and productive activity (equation 7), which raises the value of life V and further stabilizes the peace equilibrium.

The out-of-sample test. Commodore Perry's arrival in 1853 with steam-powered warships reintroduced high-p military technology from outside the system. The Tokugawa equilibrium unraveled within fifteen years. Domains rearmed, civil conflict erupted (the Boshin War, 1868–1869), and the shogunate fell. The model predicts this: an exogenous increase in p can destabilize the virtuous cycle by eroding the credibility of retaliatory deterrence. When the offensive advantage becomes sufficiently large, Figure 5 shows that the virtuous cycle collapses regardless of baseline prosperity.

Limitations. The model does not capture several features of the Tokugawa case. The disarmament was imposed by a centralized state, and the model has no state actor. The Tokugawa system

created a sharp asymmetry between the armed samurai class and the disarmed peasantry, which was itself a source of social control beyond anything the model's symmetric-agent framework captures. The Industrious Revolution had many causes beyond the reduction in violence, including domain-level governance, infrastructure investment, and Japan's geographic insularity (Totman, 1993). Our claim is not that the weapon-technology channel was the sole cause of Tokugawa prosperity, but that the deliberate reduction of p and k is consistent with the model's prediction and that no existing formal model in the conflict-deterrence literature predicts why a society would voluntarily abandon a superior weapon technology.

5.2 Medieval Iceland: Stateless Order

The Icelandic Commonwealth (930–1262) sustained social order for over three centuries without a centralized state. Disputes were adjudicated at local and national assemblies (*thing*); enforcement relied on private threats of retaliation through the blood-feud system (Friedman, 1979; Miller, 1990). Friedman (1979) observes that the institutional arrangement functioned as if designed to test how far decentralized enforcement could substitute for government.

The weapon technology of the period consisted of swords, axes, spears, and shields. This configuration privileges defense: a shield wall is effective against edged weapons, and combat requires closing to arm's length, giving the defender time to react. In terms of the model, p was low, q was moderate, and $k = 1$ (melee weapons engage a single opponent per exchange). Farmers held livestock, arable land, and access to trade networks with Norway and the British Isles, providing moderate baseline prosperity $\bar{a}$. The blood-feud system meant that an attack on any individual triggered potential retaliation from the victim's entire kinship group, a form of group deterrence consistent with Proposition 5. Even if an individual attacker has a moderate offensive advantage, surviving group members are guaranteed to respond, restoring the deterrent power of lethality.

Propositions 1 and 7 predict that this parameter combination sustains deterrence. Agents have both the means to retaliate credibly (low p ensures the retaliatory threat is viable) and a future worth protecting (moderate $\bar{a}$ keeps the value of life V high enough to anchor the violence threshold w^* above most conflict stakes). The historical record is consistent—while feuds occurred, they were regulated through compensation payments and arbitration at assemblies, and large-scale warfare was rare during the Commonwealth's first two centuries (Byock, 2001; Miller, 1990).

The system's eventual collapse is also informative. By the early thirteenth century, a small number of chieftain families had consolidated control over the majority of chieftaincies (*goðorð*), concentrating armed power asymmetrically (Friedman, 1979; Byock, 2001). Proposition 2 predicts that this asymmetry destabilizes deterrence: the strong player faces a lower retaliatory threat and has a greater incentive to deviate. Civil conflict (the *Sturlungaöld*, 1220–1264) ensued, ending

with Iceland's submission to the Norwegian crown in 1262.

Iceland demonstrates that stateless order can be sustained under the conditions the model specifies and that it is fragile to the asymmetries the model identifies. The example is limited by its small population (roughly 50,000–70,000 during the Commonwealth) and geographic isolation, which restricted both the encounter rate λ and the range of conflict stakes $F(w)$. Whether the mechanism scales to larger, denser populations is an open question, though the pre-colonial African evidence discussed below suggests that it can.

5.3 The Atlantic Slave Trade and the Gun Frontier: Firearms and the Vicious Cycle

The introduction of European firearms into West Africa and North America during the seventeenth and eighteenth centuries produced escalating violence and persistent underdevelopment rather than mutual deterrence. The model identifies the conditions under which this outcome occurs, and it provides a candidate mechanism linking the technology of violence to the long-run consequences that the empirical literature has documented.

The pre-shock baseline. Henn and Robinson (2025) document that in 1880, the African continent was home to an estimated 45,000 independent polities, fewer than 2% of which could be classified as states. These societies had organized deliberately to prevent political centralization and sustained inter-polity conflict rates below those of contemporary Europe or Asia. Through the lens of the model, this is the virtuous cycle operating at scale: decentralized communities sustaining order through local deterrence with low-p weapons (spears, shields, and edged weapons where defense could meaningfully counter offense). The parallel to Iceland is direct, though the African case operated across a continent rather than a single island.

The firearms-for-captives cycle. The Atlantic slave trade disrupted this equilibrium through two simultaneous shocks. First, European traders exchanged firearms for captives along the West African coast, raising p and q for groups that obtained guns. Inikori (1977) estimates that between 283,000 and 394,000 guns per year were imported into West Africa during the second half of the eighteenth century. Whatley (2018) documents that gun imports and slave exports were cointegrated across the eighteenth century, rising together in a self-sustaining cycle. Thornton (1999) provides a military history of how firearms transformed warfare along the West African coast from the sixteenth century onward. Second, the slave trade created large expropriable prizes: captured people could be exchanged for more weapons. Within the model's framework, this fattened the right tail of $F(w)$, ensuring a persistent supply of conflict stakes above any deterrence threshold.

Proposition 8 predicts that this combination triggers the vicious cycle. The increase in p erodes the deterrent value of lethality (Proposition 3). The fat-tailed distribution of stakes ensures that

violence remains frequent despite higher q (Proposition 4). The resulting increase in the conflict-adjusted hazard rate Ω reduces the return to investment in productive activity, lowers the value of life V, and further depresses the violence threshold w^*.

The empirical literature is consistent with these predictions. Nunn (2008) estimates that the slave trades account for a substantial share of cross-country income differences within Africa today, and that regions more exposed to the slave trades remain poorer. Nunn and Wantchekon (2011) document that descendants of populations more heavily affected by the slave trades exhibit lower interpersonal trust, consistent with the erosion of the cooperative equilibrium the model describes. The model provides a candidate mechanism for these empirical findings. The technology of violence, interacting with baseline economic conditions, produced a self-reinforcing trap that operated through the endogenous value of life.

The parallel in North America. The same mechanism operated simultaneously among Native American nations. Silverman (2016) documents that firearms, traded by Dutch, English, and French colonists in exchange for furs, triggered arms races across eastern North America beginning in the 1630s. The Haudenosaunee (Iroquois), who gained early access to Dutch firearms at Fort Orange, launched campaigns against neighboring nations to control trade routes and replenish populations lost to war and disease. The Huron confederacy was destroyed; the Erie and Neutral nations were dispersed. As the "gun frontier" advanced westward over two centuries, each region experienced the same pattern: early adopters gained temporary advantage, rivals armed in response, and escalating violence followed (Silverman, 2016).

The model maps this pattern through the same channels. Firearms raised p relative to the bow-and-arrow baseline. The fur trade created expropriable stakes (access to European goods markets) that fattened $F(w)$. Proposition 2 is also relevant—nations that acquired firearms first held a temporary asymmetric advantage over unarmed neighbors, precisely the configuration the model predicts is most destabilizing. There is a structural parallel to West Africa. Guns traded for furs in North America, guns traded for captives in West Africa, and in both cases the introduction of high-p weapons into societies with moderate baseline prosperity produced escalating conflict rather than deterrence.

Limitations. The model provides a proposed channel for the empirical findings in Nunn (2008) and Nunn and Wantchekon (2011). Moreover, Whatley and Gillezeau (2011) document that proximity to slave-trade ports is associated with weaker institutional quality and greater ethnic fragmentation, consistent with the erosion of cooperative equilibria the model predicts. But the model does not claim to be the only mechanism. European colonial institutions, the direct demographic destruction of the slave trades, the disruption of pre-existing political structures, and the ecological consequences of depopulation all contributed to long-run underdevelopment. The model isolates one channel—the feedback between weapon technology, investment incentives, and the endoge-

nous value of life—and shows that this channel alone is sufficient to generate a self-reinforcing trap. Whether this channel was quantitatively dominant relative to alternatives is an empirical question the model does not answer. The gun-slave cycle thesis has itself been debated. Some historians argue that firearms played an ancillary rather than primary role in African warfare during this period, and that the relationship between gun imports and slave exports was more complex than a simple causal loop.[4]

5.4 The Late Roman Republic: When Prosperity Collapses

The preceding examples involve shocks to weapon technology through changes in p, q, or k. Rome illustrates the model's generality by showing that the vicious cycle can be triggered through the economic channel alone. The trigger was not a new weapon but a collapse in the baseline prosperity of the armed population and a simultaneous surge in the stakes available through violence.

The baseline. For roughly four centuries (509–133 BC), the Roman Republic sustained internal stability. The citizen-militia system (*dilectus*) tied military service to property ownership. Soldiers were farmers who fought during campaign season and returned to their land afterward (Brunt, 1971). In the eyes of the model, these citizen-soldiers had moderate $\bar{a}$ (land, family, civic standing) and faced conflict stakes $F(w)$ bounded by the scale of Italian and Mediterranean warfare. The combination of a livelihood worth protecting and credible capacity for collective retaliation sustained cooperation, consistent with the virtuous cycle.

The $\bar{a}$ collapse. Over the second and first centuries BC, two forces eroded this equilibrium. First, prolonged overseas campaigns displaced smallholders. Extended military service left farms untended and vulnerable to acquisition by wealthy elites operating slave-worked estates (*latifundia*). Brunt (1971) documents the progressive dispossession of the Italian peasantry. Hopkins (1978) traces the economic mechanisms through which conquest, slave imports, and elite land accumulation hollowed out the smallholding class that had formed the backbone of the citizen-militia. The Marian reforms of 107 BC formalized the change: military recruitment was opened to the landless poor (*capite censi*), creating a professional soldiery with no property, no trade, and no independent livelihood (Gruen, 1974). Per the model's parameters, $\bar{a} \to 0$ for the military class.

Second, imperial expansion generated enormous expropriable prizes. Provincial wealth, land grants, triumphal plunder, and the spoils of civil war offered stakes that dwarfed anything available in the Italian countryside. In the model's terms, $F(w)$ acquired a fat right tail. Moreover, the shift from a citizen-militia to professional legions under individual commanders created a sharp asymmetry of armed force. A Sulla or Caesar commanding loyal professional legions held con-

[4] See the discussion in Thornton (1999) and the quantitative analysis in Inikori (1977). Our model does not require that firearms were the sole driver of the slave trade, only that the introduction of high-p weapons into an environment with large expropriable prizes is sufficient to trigger the vicious cycle.

centrated military power, while the Senate had no equivalent capacity for retaliation—precisely the asymmetric configuration that Proposition 2 identifies as most destabilizing.

Proposition 8 predicts that the combination of low $\bar{a}$ and fat-tailed $F(w)$ triggers the vicious cycle even without a change in weapons technology. Soldiers with nothing to protect have a low value of life V, which depresses the violence threshold w^*. The presence of large expropriable stakes ensures that conflict remains frequent. Soldiers' loyalty shifted from the Republic to the individual generals (Marius, Sulla, Pompey, Caesar) who promised them land and wealth upon discharge. The result was a century of civil wars (133–27 BC) that destroyed republican institutions.[5]

Rome demonstrates that the model's mechanism is not restricted to firearms or to any particular weapons technology. The legionary *gladius* and *pilum* did not change between the middle and late Republic. What changed was the economic condition of the armed population and the distribution of prizes available through violence. Without Rome, the preceding examples might suggest a theory specific to firearms. Rome shows that either channel—technological or economic—can trigger the same dynamic, and that the economic channel can operate in isolation.

6 Conclusion

Why do weapons sustain peace in some societies but trap others in perpetual violence? This paper provides an answer grounded in the interaction between the technology of violence and the economic environment. We decompose weapons technology into lethality (q), offensive advantage (p), and multi-target capacity (k), and embed a micro-model of armed conflict in a dynamic framework where agents make forward-looking investments in their livelihoods. The key object is the conflict-adjusted hazard rate Ω, which taxes the returns to investment. An agent who expects to die soon has less reason to build a future, and an agent who has not invested has less to lose from a fight.

This feedback produces two regimes. When baseline prosperity is high and weapons allow credible retaliation (moderate p), lethality deters violence, agents invest, the value of life rises, and peace deepens: a virtuous cycle. When baseline prosperity is low or offensive advantage is high, lethality fails to deter enough violence to offset its direct mortality cost, agents divest from the future, the value of life falls, and violence becomes easier to trigger: a vicious cycle. The model delivers a precise condition (equation 8) that determines which regime obtains, depending on the interaction of weapon parameters with the socioeconomic environment.

[5] Historians debate whether the Marian reforms *caused* the Republic's collapse or merely reflected deeper structural changes already underway. See Brunt (1971) for the structural view and Gruen (1974) for the argument that republican institutions retained vitality longer than the standard narrative suggests. The model does not resolve this debate, but it identifies a precise mechanism: conditional on $\bar{a}$ falling below the threshold implied by Proposition 8, the vicious cycle becomes individually rational for the military class regardless of individual commanders' ambitions.

The four historical episodes examined in Section 5 illustrate the mechanism across radically different contexts. Tokugawa Japan shows the virtuous cycle achieved through deliberate reduction of offensive advantage and multi-target capacity. A society that, having mastered firearms, abandoned them and gained two centuries of peace and documented economic development. Medieval Iceland shows the virtuous cycle sustained through low-p weapon technology and decentralized enforcement, and its fragility to the asymmetric concentration of armed power. The Atlantic slave trade and the gun frontier among Native American nations show the vicious cycle triggered by the introduction of high-p weapons into societies where the resulting arms races and expropriable prizes overwhelmed deterrence. The late Roman Republic shows the vicious cycle triggered through the economic channel alone, without a change in weapons technology, when the baseline prosperity of the armed population collapsed and imperial expansion generated enormous stakes.

These episodes are separated by centuries, continents, and entirely unrelated social orders. The fact that a single mechanism illuminates all four is the case for the model's generality. But the mechanism is not confined to historical settings. It generates predictions about two domains of contemporary policy.

Nuclear weapons represent the limiting case of the model's parameter space: $q \to 1$ and, without second-strike capability, $p \to 1$. The standard Mutually Assured Destruction (MAD) framework holds that the certainty of mutual annihilation deters first use (Schelling, 1960, 1966). Our model refines this logic by identifying a condition the standard framework leaves implicit. Proposition 7 predicts that deterrence holds when baseline prosperity $\bar{a}$ is high enough that agents invest heavily in the future, anchoring the violence threshold above available conflict stakes. During the Cold War, both superpowers were advanced industrial economies with high $\bar{a}$. The development of submarine-launched ballistic missiles and hardened silos established credible second-strike capabilities, reducing p below 1. The combination of extreme q, moderate p, and high $\bar{a}$ places Cold War deterrence in the region of the parameter space where the virtuous cycle is strongest. The model's contribution relative to MAD is to identify what MAD takes as given: deterrence depends not only on the physical capacity for retaliation but also on the value of what each party stands to lose. There is a direct policy implication. Nuclear proliferation to states with low baseline prosperity and without secure second-strike capability satisfies the conditions for the vicious cycle, regardless of the deterrence logic that stabilized superpower competition. The model predicts that MAD is not portable—the same weapons technology that sustained decades of peace between high-$\bar{a}$ superpowers need not produce stability between adversaries where one or both parties have low returns to investment in the future.[6]

[6]We do not claim the model captures the full complexity of nuclear strategy, which involves alliance structures, extended deterrence, command-and-control reliability, and domestic political constraints. The model isolates one channel—the interaction between weapon parameters and baseline prosperity—and shows that it generates predictions consistent with observed patterns of nuclear stability and instability.

The model also generates a testable prediction about domestic settings. The effect of firearm availability on violence should be nonlinear in neighborhood-level prosperity, deterrent above a threshold of $\bar{a}$ and destructive below it. In neighborhoods characterized by concentrated poverty (low $\bar{a}$), limited returns to legal employment, and frequent interpersonal encounters in dense urban environments (high λ), handguns at close range confer a high offensive advantage (p near 1). The model predicts that the vicious cycle condition is satisfied. The mortality tax on investment overwhelms the deterrent effect of lethality, producing a self-reinforcing trap. Existing evidence is consistent with this prediction: Kennedy et al. (1998) document that income inequality is correlated with firearm homicide rates across US states and that this relationship operates through the erosion of social capital. Sampson, Raudenbush and Earls (1997) find that concentrated disadvantage predicts neighborhood violence in Chicago, mediated by collective efficacy. The model provides a micro-foundation for these patterns and generates a falsifiable prediction: the marginal effect of firearm prevalence on violence should change sign as neighborhood prosperity crosses the threshold defined by condition (8). To our knowledge, this interaction effect has not been directly tested.

These contemporary applications underscore both the reach of the mechanism and its boundaries. The model is a theory of how the technology of violence and economic conditions jointly determine whether societies achieve peace or fall into conflict traps. It is not a theory of institutions, and it has no state. It does not explain how societies build the institutional capacity to intervene on weapon parameters (as Tokugawa Japan did) or why some societies develop such capacity and others do not. It does not model the role of information asymmetries in sustained conflicts, nor does it address ethnic or religious dimensions of violence that feature in the work of, for example, Rohner, Thoenig and Zilibotti (2013). The historical applications are illustrative, not causal, and each episode admits alternative explanations that the model does not rule out. These are directions for future work.

The paper's core claim is that weapons technology is neither inherently pacifying nor inherently destructive. It is an amplifier. In societies where investment yields high returns, lethality reinforces order. In societies at the margin of subsistence, it accelerates collapse. The value of human life is the channel through which amplification operates, and it is an equilibrium outcome, not a parameter. Whether an armed society is a polite society depends on whether its members have a future worth protecting.

References

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75(4): 352–365.

Besley, Timothy, and Torsten Persson. 2011. "The Logic of Political Violence." *Quarterly Journal of Economics*, 126(3): 1411–1445.

Brunt, P. A. 1971. *Italian Manpower 225 B.C.–A.D. 14.* Oxford:Oxford University Press.

Byock, Jesse L. 2001. *Viking Age Iceland.* London:Penguin.

Chase, Kenneth. 2003. *Firearms: A Global History to 1700.* Cambridge:Cambridge University Press.

Chassang, Sylvain, and Gerard Padró i Miquel. 2009. "Defensive Weapons and Defensive Alliances." *American Economic Review*, 99(2): 282–286.

Chassang, Sylvain, and Gerard Padró i Miquel. 2010. "Conflict and Deterrence under Strategic Risk." *Quarterly Journal of Economics*, 125(4): 1821–1858.

Dal Bó, Ernesto, and Pedro Dal Bó. 2011. "Workers, Warriors, and Criminals: Social Conflict in General Equilibrium." *Journal of the European Economic Association*, 9(4): 646–677.

Donohue, John J. 2004. "Guns, Crime, and the Impact of State Right-to-Carry Laws." *Fordham Law Review*, 73(2): 623–652.

Fenske, James, and Namrata Kala. 2017. "1807: Economic Shocks, Conflict and the Slave Trade." *Journal of Development Economics*, 126: 66–76.

Friedman, David D. 1979. "Private Creation and Enforcement of Law: A Historical Case." *Journal of Legal Studies*, 8(2): 399–415.

Glaser, Charles L., and Chaim Kaufmann. 1998. "What Is the Offense-Defense Balance and Can We Measure It?" *International Security*, 22(4): 44–82.

Gruen, Erich S. 1974. *The Last Generation of the Roman Republic.* Berkeley:University of California Press.

Hayami, Akira. 2015. *Japan's Industrious Revolution: Economic and Social Transformations in the Early Modern Period. Studies in Economic History*, Tokyo:Springer.

Henn, Soeren J., and James A. Robinson. 2025. "Africa as a Success Story: Political Organization in Pre-Colonial Africa." NBER Working Paper No. 34546.

Hopkins, Keith. 1978. *Conquerors and Slaves: Sociological Studies in Roman History.* Cambridge:Cambridge University Press.

Inikori, Joseph E. 1977. "The Import of Firearms into West Africa 1750–1807: A Quantitative Analysis." *The Journal of African History*, 18(3): 339–368.

Jansen, Marius B. 2000. *The Making of Modern Japan.* Cambridge, MA:Harvard University Press.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics*, 30(2): 167–214.

Kennedy, Bruce P., Ichiro Kawachi, Deborah Prothrow-Stith, Kimberly Lochner, and Vanita Gupta. 1998. "Social Capital, Income Inequality, and Firearm Violent Crime." *Social Science & Medicine*, 47(1): 7–17.

Lee, David S., and Justin McCrary. 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence." *Journal of Law and Economics*, 60(2): 301–324.

Lott, John R. 2000. *More Guns, Less Crime: Understanding Crime and Gun-Control Laws.* . 2 ed., Chicago:University of Chicago Press.

Miller, William Ian. 1990. *Bloodtaking and Peacemaking: Feud, Law, and Society in Saga Iceland.* Chicago:University of Chicago Press.

Nunn, Nathan. 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics*, 123(1): 139–176.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review*, 101(7): 3221–3252.

Perrin, Noel. 1979. *Giving Up the Gun: Japan's Reversion to the Sword, 1543–1879.* Boston:David R. Godine.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti. 2013. "Seeds of Distrust: Conflict in Uganda." *Journal of the European Economic Association*, 11(s1): 1–40.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science*, 277(5328): 918–924.

Schelling, Thomas C. 1960. *The Strategy of Conflict.* Cambridge, MA:Harvard University Press.

Schelling, Thomas C. 1966. *Arms and Influence.* New Haven:Yale University Press.

Silverman, David J. 2016. *Thundersticks: Firearms and the Violent Transformation of Native America.* Cambridge, MA:Harvard University Press.

Thornton, John K. 1999. *Warfare in Atlantic Africa, 1500–1800.* London:UCL Press.

Totman, Conrad. 1993. *Early Modern Japan.* Berkeley:University of California Press.

Whatley, Warren, and Rob Gillezeau. 2011. "The Impact of the Transatlantic Slave Trade on Ethnic Stratification in Africa." *American Economic Review: Papers & Proceedings*, 101(3): 571–576.

Whatley, Warren C. 2018. "The Gun-Slave Hypothesis and the 18th Century British Slave Trade." *Explorations in Economic History*, 67: 80–104.

A Proofs

Proof of Theorem 1

Let the Markov state space be $S = \{1,2\}$, where $s \in S$ denotes the identity of the currently active player. A pure-strategy Markov strategy $\sigma_i(s)$ for player i maps the state to an action $A_i \in \{A,P\}$. Since i only moves when $s = i$, we denote the strategy profile as (σ_1, σ_2).

First, expected payoffs are well-defined and finite at $\delta = 1$. The game ends in a terminal state either through consecutive passes (instantaneous) or through a successful attack. Because $p_i \in (0,1)$ for all i, any strategy profile containing at least one A terminates the game almost surely. Thus, expected payoffs are finite and well-defined.

Non-existence of Mutual Restraint. Assume there exists an MPE $(\sigma_1, \sigma_2) = (P,P)$. Under this profile, any state s leads to two consecutive passes, yielding the baseline payoff $V_i^{restraint} = a$ for both players. Consider a unilateral deviation by player i at state $s = i$ to action A. The expected deviation payoff is:

$$U_i^{dev} = p_i u_W + (1 - p_i) W_i(j) \tag{9}$$

where $W_i(j)$ is player i's value when player j is active. In the hypothesized MPE, $\sigma_j(j) = P$ and the subsequent move is $\sigma_i(i) = P$, terminating the game. Thus, $W_i(j) = a$. Substituting $u_W = a + w$:

$$U_i^{dev} = p_i(a + w) + (1 - p_i)a = a + p_i w \tag{10}$$

Since $p_i > 0$ and $w > 0$, it follows that $U_i^{dev} > a$. Therefore, (P,P) is not an MPE.

Uniqueness of Full Escalation. We now show that $(\sigma_1, \sigma_2) = (A,A)$ is the unique MPE. Let V_i be the value to the active player and W_i the value to the passive player. Under (A,A), these values satisfy:

$$V_i = p_i u_W + (1 - p_i) W_i \tag{11}$$

$$W_i = p_j u_L + (1 - p_j) V_i \tag{12}$$

Solving this system yields the unique payoffs:

$$V_i^* = \frac{p_i u_W + (1 - p_i) p_j u_L}{p_i + p_j - p_i p_j}, \quad W_i^* = \frac{p_j u_L + (1 - p_j) p_i u_W}{p_i + p_j - p_i p_j} \tag{13}$$

Since $u_W > a > u_L$, it can be verified that $u_W > V_i^* > a > W_i^* > u_L$. Now consider the best response of player i given $\sigma_j = A$:

- If i plays A, the payoff is V_i^*.

- If i plays P, the game transitions to state $s = j$. Since $\sigma_j = A$, the payoff is W_i^*.

Because $V_i^* > W_i^*$, A is the unique best response to A. Finally, we check if player i would ever play P if $\sigma_j = P$. As shown in the first part of the proof, i strictly prefers A (payoff $a + p_i w$) over P (payoff a) when the opponent plays P. Thus, A is a strictly dominant Markov strategy for both players, making Full Escalation the unique MPE. ■

Proof of Theorem 2

We consider a strategy profile $\sigma = (\sigma_1, \sigma_2)$ of Grim Trigger. Let h_t be the history of play up to period t. The strategy for player i is defined as:

$$\sigma_i(h_t) = \begin{cases} P & \text{if } a_\tau = P \text{ for all } \tau < t \\ A & \text{otherwise} \end{cases} \tag{14}$$

This profile is a Subgame Perfect Nash Equilibrium (SPNE) if the strategies constitute a Nash Equilibrium in every subgame.

The Punishment Subgame. If a deviation A occurs at any t, the history h_{t+1} triggers the "Full Escalation" phase. By Theorem 1, the unique Markov Perfect Equilibrium of this subgame is for both agents to play A at every subsequent opportunity. Since an MPE is a refinement of SPNE, the threat of reversion to Full Escalation is credible.

Expected Payoff of the Full Escalation Subgame. Let V_i^* denote the expected continuation value for the active player i in a subgame where both players choose A indefinitely. Let W_i^* be the value for the passive player i. These values satisfy the recursive system:

$$V_i^* = p_i u_W + (1 - p_i) W_i^* \tag{15}$$

$$W_i^* = p_j u_L + (1 - p_j) V_i^* \tag{16}$$

Substituting (16) into (15):

$$V_i^* = p_i u_W + (1 - p_i) \left[p_j u_L + (1 - p_j) V_i^* \right]$$

$$V_i^* = p_i u_W + (1 - p_i) p_j u_L + (1 - p_i)(1 - p_j) V_i^*$$

Collecting V_i^* terms:

$$\begin{aligned} V_i^* \left[1-(1-p_i)(1-p_j)\right] &= p_i u_W + (1-p_i)p_j u_L \\ V_i^* \left[1-(1-p_i-p_j+p_i p_j)\right] &= p_i u_W + (1-p_i)p_j u_L \\ V_i^*(p_i+p_j-p_i p_j) &= p_i u_W + (1-p_i)p_j u_L \end{aligned}$$

Solving for V_i^*, we obtain the expected deviation payoff U_i^{dev}:

$$U_i^{dev} = V_i^* = \frac{p_i u_W + (1-p_i)p_j u_L}{p_i + p_j - p_i p_j} \tag{17}$$

Incentive Compatibility on the Equilibrium Path. On the equilibrium path (h_t consists only of P), the value of the game is a, as $\delta = 1$ and peace is maintained indefinitely. A player chooses not to deviate if the value of peace meets or exceeds the value of launching a surprise attack and entering the Full Escalation subgame as the first mover:

$$a \geq U_i^{dev} \quad \forall i \in \{1,2\} \tag{18}$$

Substituting the expression for U_i^{dev}, the condition for a Mutual Restraint SPNE is:

$$a \geq \frac{p_i u_W + (1-p_i)p_j u_L}{p_i + p_j - p_i p_j} \tag{19}$$

for both $i = 1$ and $i = 2$. If this condition fails for either player, the temptation of the spoils of war w (embedded in u_W) or the fear of a low p_j (weak retaliation) will trigger a preemptive strike. ■

Proof of Proposition 1

Consider the symmetric case where $q_i = q_j = q$ (but any fixed p_i, p_j). From Theorem 2, the expected deviation payoff is:

$$U_i^{dev} = \frac{p_i u_W + (1-p_i)p_j u_L}{p_i + p_j - p_i p_j} \tag{20}$$

Recall the terminal payoffs: $u_W = a + w$ and $u_L = (1-q)(a-l)$, where $q \in (0,1)$ is the probability of death and $l > 0$ is the injury penalty. Substituting $u_L(q)$ into the expression for U_i^{dev}:

$$U_i^{dev}(q) = \frac{p_i(a+w) + (1-p_i)p_j(1-q)(a-l)}{p_i + p_j - p_i p_j} \tag{21}$$

To determine the effect of an increase in lethality, we differentiate U_i^{dev} with respect to q:

$$\frac{\partial U_i^{dev}}{\partial q} = \frac{\partial}{\partial q}\left[\frac{p_i(a+w)+(1-p_i)p_j(a-l)-(1-p_i)p_jq(a-l)}{p_i+p_j-p_ip_j}\right] \tag{22}$$

Since p_i, p_j, a, w, and l are independent of q, the derivative simplifies to:

$$\frac{\partial U_i^{dev}}{\partial q} = \frac{-(1-p_i)p_j(a-l)}{p_i+p_j-p_ip_j} \tag{23}$$

By the model's assumptions, $p_i \in (0,1)$, $p_j \in (0,1)$, and $a > l$ (the utility of injured peace is positive). Therefore:

- The denominator $(p_i + p_j - p_ip_j)$ is strictly positive (as it represents the probability that at least one of the first two strikes succeeds).
- The numerator $-(1-p_i)p_j(a-l)$ is strictly negative.

Thus, $\frac{\partial U_i^{dev}}{\partial q} < 0$. An increase in weapon lethality strictly decreases the expected payoff of deviating from mutual restraint, thereby strengthening the stability of the peace equilibrium. ■

Proof of Proposition 2

Consider an asymmetric environment where player j has a strong weapon (p_h) and player i has a weak weapon (p_l), with $p_h > p_l$. We compare the deviation payoff of the strong player, $U_j^{dev}(p_h, p_l)$, to the symmetric case, $U^{dev}(p_h, p_h)$.

From Theorem 2, the deviation payoff for player j when facing i is:

$$U_j^{dev}(p_h, p_l) = \frac{p_hu_W + (1-p_h)p_lu_L}{p_h+p_l-p_hp_l} \tag{24}$$

To determine how this changes with the opponent's strength, we differentiate U_j^{dev} with respect to p_l:

$$\frac{\partial U_j^{dev}}{\partial p_l} = \frac{(1-p_h)u_L(p_h+p_l-p_hp_l)-(1-p_h)(p_hu_W+(1-p_h)p_lu_L)}{(p_h+p_l-p_hp_l)^2} \tag{25}$$

Factoring out $(1-p_h)$:

$$\frac{\partial U_j^{dev}}{\partial p_l} = \frac{(1-p_h)\left[u_Lp_h+u_Lp_l-u_Lp_hp_l-p_hu_W-u_Lp_l+u_Lp_hp_l\right]}{(p_h+p_l-p_hp_l)^2} \tag{26}$$

The terms involving $u_L p_l$ cancel out, leaving:

$$\frac{\partial U_j^{dev}}{\partial p_l} = \frac{(1-p_h)p_h(u_L - u_W)}{(p_h + p_l - p_h p_l)^2} \tag{27}$$

Since $u_W > u_L$ (spoils of war exceed injury/death payoffs) and $p_h \in (0,1)$, it follows that:

$$\frac{\partial U_j^{dev}}{\partial p_l} < 0 \tag{28}$$

Because the derivative is strictly negative, the deviation payoff is strictly higher when the opponent is weak (p_l) than when the opponent is strong (p_h):

$$U_j^{dev}(p_h, p_l) > U^{dev}(p_h, p_h) \tag{29}$$

The strong player j faces a lower "deterrence price" when attacking a weak opponent because the weak opponent's retaliatory strike is likely to fail. This lowers the IC threshold $a \geq U_j^{dev}$ for the strong player. Furthermore, knowing that player j is more likely to attack, the weak player i may also abandon restraint to avoid being the passive target of a high-p_h strike (the preemptive "Hobbesian Trap"). Thus, asymmetry strictly expands the parameter space for violent conflict compared to symmetric armament. ■

Proof of Proposition 3

Consider a symmetric setting where $p_i = p_j = p$ and $q_i = q_j = q$. From the derivation of U_i^{dev} in Theorem 2, we have:

$$U_i^{dev}(p,q) = \frac{pu_W + (1-p)pu_L}{p + p - p^2} = \frac{pu_W + p(1-p)u_L}{p(2-p)} = \frac{u_W + (1-p)u_L}{2-p} \tag{30}$$

where $u_W = a + w$ and $u_L = (1-q)(a-l)$.

Effect of Offensive Advantage (p): Differentiating U_i^{dev} with respect to p:

$$\frac{\partial U_i^{dev}}{\partial p} = \frac{-u_L(2-p) - (-1)(u_W + (1-p)u_L)}{(2-p)^2} = \frac{-2u_L + pu_L + u_W + u_L - pu_L}{(2-p)^2} = \frac{u_W - u_L}{(2-p)^2} \tag{31}$$

Since $u_W > u_L$ by the definition of the spoils of war and the injury penalty, $\frac{\partial U_i^{dev}}{\partial p} > 0$. A higher strike efficacy increases the temptation to deviate.

Interaction Effect (q and p). We now examine how the deterrent effect of lethality, $\frac{\partial U_i^{dev}}{\partial q}$,

changes as the offensive advantage p increases. From Proposition 1, we know:

$$\frac{\partial U_i^{dev}}{\partial q} = \frac{\partial U_i^{dev}}{\partial u_L} \cdot \frac{\partial u_L}{\partial q} = \left[\frac{1-p}{2-p}\right] \cdot [-(a-l)] \tag{32}$$

Taking the cross-partial derivative with respect to p:

$$\frac{\partial^2 U_i^{dev}}{\partial p \partial q} = -(a-l)\frac{\partial}{\partial p}\left[\frac{1-p}{2-p}\right] \tag{33}$$

Using the quotient rule on the bracketed term:

$$\frac{\partial}{\partial p}\left[\frac{1-p}{2-p}\right] = \frac{-1(2-p)-(-1)(1-p)}{(2-p)^2} = \frac{-2+p+1-p}{(2-p)^2} = \frac{-1}{(2-p)^2} \tag{34}$$

Substituting this back into the cross-partial expression:

$$\frac{\partial^2 U_i^{dev}}{\partial p \partial q} = -(a-l)\left[\frac{-1}{(2-p)^2}\right] = \frac{a-l}{(2-p)^2} \tag{35}$$

Since $a > l$, we have $\frac{\partial^2 U_i^{dev}}{\partial p \partial q} > 0$.

Limit. The positive cross-partial indicates that an increase in p makes the (negative) derivative $\frac{\partial U_i^{dev}}{\partial q}$ less negative. As $p \to 1$, the coefficient on u_L in the deviation payoff vanishes:

$$\lim_{p\to 1} \frac{\partial U_i^{dev}}{\partial q} = \lim_{p\to 1} \frac{-(1-p)(a-l)}{2-p} = 0 \tag{36}$$

Thus, a high offensive advantage effectively "insulates" the attacker from the consequences of their own weapon's lethality should they fail, rendering the deterrent effect of q moot. ■

Proof of Proposition 4

Let $w \sim F(w)$ with density $f(w)$. Mutual restraint holds if $w \leq w^*(q)$, where $w^*(q)$ is the threshold derived from the IC constraint $a = U_i^{dev}(w,q)$. From Theorem 2, rearranging for w:

$$w^*(q) = \frac{(p_i + p_j - p_i p_j)a - (1-p_i)p_j u_L}{p_i} - a \tag{37}$$

Substituting $u_L = (1-q)(a-l)$ and simplifying for the symmetric case ($p_i = p_j = p$):

$$w^*(q) = \frac{(2-p)a - (1-p)(1-q)(a-l)}{1} - a = (1-p)[qa + (1-q)l] \tag{38}$$

Note that $\frac{\partial w^*}{\partial q} = (1-p)(a-l) > 0$. As lethality q increases, the threshold for violence w^* rises, meaning the probability of conflict $P(w > w^*) = 1 - F(w^*)$ strictly decreases.

Expected Deaths. The expected number of deaths $E[D]$ is the product of the probability of a conflict and the expected deaths per conflict. In a symmetric Full Escalation subgame, the probability that at least one person dies is q (for the offensive's target) plus the probability the first strike fails and the second-mover's strike kills the offensive, and so on. Summing the geometric series for total expected deaths D_{fight} in a fight:

$$D_{fight} = \sum_{t=0}^{\infty} p(1-p)^t \cdot q = q \cdot p \sum_{t=0}^{\infty} (1-p)^t = q \tag{39}$$

Thus, the ex-ante expected deaths are $E[D] = q[1 - F(w^*)]$.

Condition for Increasing Mortality. We examine the sign of $\frac{\partial E[D]}{\partial q}$:

$$\frac{\partial E[D]}{\partial q} = \left[\frac{\partial (1 - F(w^*))}{\partial q} q + (1 - F(w^*)) \right] \tag{40}$$

Using the chain rule $\frac{\partial (1-F(w^*))}{\partial q} = -f(w^*) \frac{\partial w^*}{\partial q}$:

$$\frac{\partial E[D]}{\partial q} = -qf(w^*)(1-p)(a-l) + (1 - F(w^*)) \tag{41}$$

For total mortality to rise with lethality ($\frac{\partial E[D]}{\partial q} > 0$), we require:

$$1 - F(w^*) > qf(w^*)(1-p)(a-l) \tag{42}$$

Conclusion. The left-hand side represents the "survival" of violence (conflicts that still occur despite high q), while the right-hand side represents the "deterrence" at the margin. If the distribution of temptations w is sufficiently "fat-tailed" (small $f(w^*)$ relative to $1 - F(w^*)$), the increase in deaths per realized conflict outweighs the reduction in conflict frequency. ■

Proof of Proposition 5

Let $p = 1$ and $k = 1$. In the Full Escalation subgame starting from (N,N), the game follows a deterministic sequence of state transitions $(n,m) \to (n,m-1) \to (n-1,m-1)$. An agent survives to the next round of group-activity only if they are not the target of the opponent's strike during the intervening passive phase.

Survival Probability Coefficients. Let s_j be the probability that the agent survives up to the j-th time their group is active. For the agent to be active when the opponent group has $N - j$

members, they must have survived j previous rounds of being passive. The probability of survival at each passive step j (where the group size is $N-j+1$) is $\frac{N-j}{N-j+1}$. Thus:

$$s_j = \prod_{i=1}^{j} \frac{N-i}{N-i+1} = \frac{N-1}{N} \cdot \frac{N-2}{N-1} \cdots \frac{N-j}{N-j+1} = \frac{N-j}{N} \tag{43}$$

for $j \in \{1,\ldots,N-1\}$. The probability of being killed exactly at the j-th passive round is the probability of surviving to that round and then being targeted:

$$d_j = s_{j-1} \cdot \frac{1}{N-j+1} = \frac{N-(j-1)}{N} \cdot \frac{1}{N-j+1} = \frac{1}{N} \tag{44}$$

Explicit Deviation Payoff. The expected value $V(N,N)$ is the weighted sum of the payoff in every possible terminal state. The agent is either killed in one of the $N-1$ passive rounds or survives to the final victory:

$$V(N,N) = \underbrace{s_{N-1}(a+w)}_{\text{Victory State}} + \sum_{j=1}^{N-1} \underbrace{d_j u_L}_{\text{Death States}} \tag{45}$$

Substituting $s_{N-1} = \frac{N-(N-1)}{N} = \frac{1}{N}$ and $d_j = \frac{1}{N}$:

$$V(N,N) = \frac{1}{N}(a+w) + \sum_{j=1}^{N-1} \frac{1}{N}(1-q)(a-l) \tag{46}$$

Simplifying the summation of the $N-1$ identical terms:

$$V(N,N) = \frac{a+w+(N-1)(1-q)(a-l)}{N} \tag{47}$$

Deterrence Threshold. Mutual restraint is an SPNE if $a \geq V(N,N)$. Substituting the explicit expression:

$$a \geq \frac{a+w+(N-1)(1-q)(a-l)}{N} \implies Na \geq a+w+(N-1)(1-q)(a-l) \tag{48}$$

Rearranging to isolate the effect of lethality q:

$$(N-1)a - (N-1)(1-q)(a-l) \geq w \tag{49}$$

$$(N-1)[a-(1-q)(a-l)] \geq w \tag{50}$$

As q increases, the term in the brackets increases. In the limit $q = 1$, this becomes $(N-1)a \geq w$. Because the left-hand side is strictly increasing in q, there exist values of a and w such that deterrence holds if q is sufficiently high—even though $p = 1$. ■

Proof of Proposition 6

Let $k \geq 1$ be the number of opponents eliminated in each active period. We analyze the deviation payoff $V(N,N;k)$ under absolute offensive advantage $p = 1$. The conflict ends in $T = \lceil N/k \rceil$ active periods for the aggressor.

Survival Probability with Capacity k. The agent survives to the final victory only if they are not targeted during any of the intervening passive rounds. Let j index the passive rounds ($j = 1, 2, \ldots, T-1$). At each passive round j, the opponent group has been reduced to $m_j = N - jk$ survivors, but the agent's group has $n_j = N - (j-1)$ survivors (having absorbed $j-1$ previous retaliatory strikes).

The probability of surviving passive round j is $\frac{n_j - 1}{n_j}$. The cumulative probability of surviving to victory, $P(S;k)$, is:

$$P(S;k) = \prod_{j=1}^{T-1} \frac{N-j}{N-(j-1)} = \frac{N-1}{N} \cdot \frac{N-2}{N-1} \cdots \frac{N-(T-1)}{N-(T-2)} \tag{51}$$

This telescopes to:

$$P(S;k) = \frac{N-(T-1)}{N} = \frac{N - \lceil N/k \rceil + 1}{N} \tag{52}$$

Note that $P(S;k)$ is strictly increasing in k. In the limit $k \geq N$, $T = 1$, the product is empty, and $P(S) = 1$.

Explicit Deviation Payoff. As in Proposition 5, the agent is either killed in one of the $T-1$ passive rounds or survives to win. Since the probability of being killed in any passive round j is $d_j = 1/N$, the payoff is:

$$V(N,N;k) = \underbrace{\frac{N-T+1}{N}(a+w)}_{\text{Victory State}} + \underbrace{\frac{T-1}{N}u_L}_{\text{Death States}} \tag{53}$$

Substituting $u_L = (1-q)(a-l)$, we see that the coefficient on the "Victory State" grows with k, while the coefficient on the "Death State" (the deterrent) shrinks.

Collapse of Deterrence. The IC constraint $a \geq V(N,N;k)$ simplifies to:

$$Na \geq (N-T+1)(a+w) + (T-1)(1-q)(a-l) \tag{54}$$

Rearranging to isolate the required lethality for peace:

$$(T-1)[a-(1-q)(a-l)] \geq (N-T+1)w \tag{55}$$

As k increases, T decreases. This has a dual effect:

1. The "total weight of deterrence" on the left $(T-1)$ decreases.
2. The "temptation multiplier" on the right $(N-T+1)$ increases.

For a sufficiently large k, $T-1$ becomes small relative to $N-T+1$. For any finite q, there exists a k such that the inequality fails, guaranteeing violence. In the extreme case $k=N$, the left side becomes 0 and the right side becomes $Nw>0$, rendering lethality q entirely incapable of sustaining peace. ■

Proof of Proposition 7

The equilibrium (e^*, w^*) is determined by the system of equations $H(e, w^*; q) = 0$:

$$H_1 = a'(e) - c'(e)\Omega(w^*, q) = 0 \tag{56}$$

$$H_2 = w^* - (1-p)[qV(e, w^*, q) + (1-q)l] = 0 \tag{57}$$

where $\Omega = r + d + \frac{\lambda q}{2}(1 - F(w^*))$ and V is defined by the Bellman equation (6). The Jacobian of the system is:

$$J = \begin{pmatrix} a''(e) - c''(e)\Omega & -c'(e)\frac{\partial \Omega}{\partial w^*} \\ -(1-p)q\frac{\partial V}{\partial e} & 1-(1-p)q\frac{\partial V}{\partial w^*} \end{pmatrix} \tag{58}$$

By the Envelope Theorem (and the agent's optimization), $\frac{\partial V}{\partial e} = a'(e)/\Omega = c'(e) > 0$. Differentiating the Bellman equation with respect to w^* yields $\frac{\partial V}{\partial w^*} = \frac{\lambda p w^* f(w^*)}{2\Omega(1-p)} \geq 0$. Note that $\frac{\partial \Omega}{\partial w^*} = -\frac{\lambda q}{2} f(w^*) < 0$.

Stability requires $\det(J) > 0$. Let $\Delta = \det(J)$. We use Cramer's Rule to find $\frac{de^*}{dq}$:

$$\frac{de^*}{dq} = \frac{1}{\Delta} \det \begin{pmatrix} -\frac{\partial H_1}{\partial q} & \frac{\partial H_1}{\partial w^*} \\ -\frac{\partial H_2}{\partial q} & \frac{\partial H_2}{\partial w^*} \end{pmatrix} \tag{59}$$

where $\frac{\partial H_1}{\partial q} = -c'(e)\frac{\partial \Omega}{\partial q}$ and $\frac{\partial \Omega}{\partial q} = \frac{\lambda}{2}(1 - F(w^*)) > 0$. Note that $\frac{\partial H_2}{\partial q} = -(1-p)\left[(V-l) + q\frac{\partial V}{\partial q}\right]$. Working through the Bellman equation, one finds $\frac{\partial V}{\partial q}\Big|_{e,w^*} = \frac{\lambda(1-F(w^*))(l-V)}{2\Omega}$. Substituting and

simplifying:

$$\frac{\partial H_2}{\partial q} = -(1-p)(V-l)\left[1 - \frac{q\lambda(1-F(w^*))}{2\Omega}\right] = -\frac{(1-p)(V-l)(r+d)}{\Omega} < 0 \tag{60}$$

where the last equality uses $\Omega = r + d + \frac{\lambda q}{2}(1-F(w^*))$. This is the direct deterrence effect, scaled by the non-conflict share of the hazard rate.

Investment rises $\left(\frac{de^*}{dq} > 0\right)$ if the numerator is positive. Expanding the determinant:

$$c'(e)\left[\frac{\partial \Omega}{\partial q}\left(1-(1-p)q\frac{\partial V}{\partial w^*}\right) + \frac{\partial \Omega}{\partial w^*}(1-p)(V-l)\right] > 0 \tag{61}$$

Substituting $\frac{\partial \Omega}{\partial w^*}$ and $\frac{dw^*}{dq}|_e$, one sees that this collapses to $qh(w^*)\frac{dw^*}{dq} > 1$. This condition implies $\frac{d\Omega}{dq} < 0$. When the hazard rate falls, investment e^* must rise to satisfy the FOC (7). ■

Proof of Proposition 8

Consider the Jacobian J in equation (58). As $c''(e) \to 0$ (highly elastic investment) and $a''(e) \to 0$, the $(1,1)$ entry of J vanishes. The stability determinant Δ is then dominated by the off-diagonal feedback terms. The total derivative $\frac{dw^*}{dq}$ is:

$$\frac{dw^*}{dq} = \frac{1}{\Delta}\left[-(a'' - c''\Omega)\frac{\partial H_2}{\partial q} - (1-p)qc'(e)\left(-c'(e)\frac{\partial \Omega}{\partial q}\right)\right] \tag{62}$$

When c'' is small, the first term (direct deterrence) is overwhelmed by the second term (the desperation effect). Since $\frac{\partial \Omega}{\partial q} > 0$, the second term is negative. Thus, $\frac{dw^*}{dq} < 0$. By the FOC, $\frac{de^*}{dq}$ must also be negative, completing the Vicious Cycle. ■