Designing Political Order

Scott F Abramson[*] Emiel Awad[†] Brenton Kenkel[‡]

September 8, 2020

Abstract

Social scientists and political philosophers widely believe that the foundations of political order rest upon the existence of a sovereign agent endowed with a monopoly of coercive force. In this paper, we develop a formal model of anarchic competition and show that whenever it is possible to construct a peaceful political order based upon a monopoly of force, it is also possible to construct one where multiple agents maintain coercive abilities. What is more, we show that peaceful orders with multiple violence specialists generally require lower coercive investments than peaceful orders with a single violence specialist. This undercuts the notion that monopolistic domestic politics are inherently more efficient than the a competitive international system. Nevertheless, we identify why inefficient monopolies of violence might persist — any individual agent's payoff is maximized when she serves as a monopolist that invests more in coercion than is strictly necessary to maintain peace.

[*] Assistant Professor, Department of Political Science, University of Rochester.
[†] Fellow, Department of Government, The London School of Economics and Political Science.
[‡] Assistant Professor, Department of Political Science, Vanderbilt University.

...[T]here may be no more important service that scholars of international relations can provide than to subject the existence, causes and consequences of hierarchy to scrutiny.

— David Lake (2009)

1 Introduction

For many scholars the international system is definitionally anarchic (Waltz 1979; Keohane 1986; Mearsheimer 2001). And yet, various forms of hierarchy pervade (Lake 1996, 2009). Under what conditions can agents in anarchy construct rules, norms, or formal institutions that establish hierarchical relationships, and when might these hierarchies be welfare enhancing? In this paper we answer these questions by developing a formal model that characterizes the existence of two types of political order. First, we show when, in anarchy, agents can design institutions that preserve peaceful order by concentrating all coercive power in a monopolist of violence. Second, we describe the conditions under which order can be preserved by instead distributing coercive power across all actors. We then make welfare comparisons across these types of political order.

Our game begins in anarchy where there is no third party to enforce property rights and each player can use force to appropriate others' wealth. We examine the ability of these actors to construct institutions as "formal rules of the game" (North 1990, p. 3) with two characteristics. First, we want to know when agents in an institution-free society can develop rules that *prevent the use of violence*. Second, we seek to understand when these rules are *self-enforcing*. In other words, in an environment where agents can always resort to violence, we want to know when it is in the individual interest of each agent to participate in the institution and refrain from violence.

A key feature of the institutions we describe is that they produce peace by redistributing wealth across agents according to their relative investment in coercive abilities. That is, institutions must reward powerful agents (at the expense of the weak) to prevent the use of force. However, such an institution is only self-enforcing and peaceful if weaker agents willingly accept redistribution towards the powerful. In this sense, what we describe is the construction of political authority—hierarchy underpinned by the threat of force but where overt coercion is absent (Lake 2007, p. 50-53).

Our approach is agnostic with respect to the specific rules, protocols, or norms through which our agents interact. Rather, we are motivated by describing the existence of *any* institutional arrangement that makes peace sustainable. Of course, there may be second-order consequences of specific arrangements. That is, among the various institutions that limit the use of force, some may produce externalities that increase welfare and others that reduce it. We abstract away from these concerns to focus upon the first-order welfare effects that directly relate to the construction of peace.

In describing the political and economic conditions necessary to sustain various forms of peaceful political order, we obtain four main results. The first concerns the possibility of a peaceful political order wherein each agent refrains from appropriation even if they expect the other to invest nothing in coercive abilities. From an efficiency standpoint, this is the ideal form of political order—nobody invests in wasteful coercion. However, we find that this is sustainable as an equilibrium only if the expected costs of conflict are exceptionally large. As a consequence, we also look for self-enforcing political institutions that prevent all violence but whose participants nevertheless make costly investments in coercive abilities.

Second, if peace without coercion is unsustainable, then it is easier to support peace by having multiple agents maintain coercive abilities than by concentrating all ability to produce violence in a single actor. We contrast a *monopoly of violence*, in which a single actor maintains coercive capabilities, with *distributed coercion*, wherein all agents invest in coercion. We find that conflict must be sufficiently destructive in order for a monopoly of violence to preserve peace—otherwise, there will be too great a temptation for the monopolist to violently exploit the other, unarmed players. The greater the monopolist's inherent advantage in the use of coercive means, the less stringent this requirement becomes. By contrast, we find that peace is sustainable through a distribution of coercive powers even if violence is not very costly.

Our third key result is that peaceful distributed coercion requires strictly less investment in coercion than a peaceful monopoly of violence. Combined with the previous result, this means that not only can distributed coercion produce peace under a broader set of structural conditions, but that it does so at a lower cost. To be clear, we show that whenever a peaceful monopoly of violence exists, there also exists a more efficient alternative arrangement that distributes coercive power across agents. However, this does not imply that all peaceful institutions supported by distributed coercion are welfare enhancing. Indeed, we also show that the worst peaceful institution with distributed coercion is less efficient than the least efficient peaceful institution with a monopoly of violence.

Our model helps us identify an important tension between social and individual incentives. An individual player can extract more as a monopolist than that player would receive from any order based upon distributed coercive abilities. Thus, our final key result is that any individual agent, given the choice, would select a wasteful political order—one in which she is the monopolist and invests more in coercion than would be needed to sustain peace. In this way we highlight how social efficiency and individual incentives cut hard against each other in the construction of order. Monopolies of violence allow sovereign agents to extract rents proportional to their ability to coerce, resulting in a wasteful over-investment in force. In other words, monopolies of violence engender monopoly rents.

While our main results concern peaceful orders, we also examine institutional arrangements in which conflict sometimes occurs. Even when institutions that preserve peace would be sustainable, they may not be economically efficient relative to institutions that allow for a positive probability of violent conflict. The question of cost becomes particularly acute when the players have incomplete information about each other's coercive capabilities. When this is the case, peace can be assured only by investing enough to deter the strongest type of each player from conflict. As a consequence, the cost imposed by investing sufficiently in coercion to deter all violence may outweigh the costs of admitting occasional conflict.

Our results advance the set of critical responses to Realist theories (Ruggie 1982, 1983; Wendt 1987, 1992; Spruyt 1994). First, we blurry the bright line between domestic politics, where units "stand in relations of super-and subordination (Waltz 1979, p. 81)," from international politics, where units "are not formally differentiated (Ibid, p. 93)." Whenever agents can construct political order sustained by the sorting of actors into violence specialists and non-specialists, they can always support order without such specialization. This is to say, we show that the demarcation between unit and structure is neither obvious nor natural. Second, we show that the structural conditions of anarchy need not imply a "tragic" destructive rivalry (Mearsheimer 2001). Still, while costly and wasteful over-investment in power is not a necessary implication of anarchy, we also show that it may be sustainable as an equilibrium. Indeed, the worst peaceful outcome we identify (in terms of social welfare) involves all actors spending more than necessary on coercive capability.

In sum, starting from assumptions that comport with the realist view of the interstate system, we find that institutions provide no "false promise" of security (Mearsheimer 1994). That is, our model suggests that self-interested actors in anarchy can construct a wide range of self-enforcing institutions that preserve peace. However, this does not result from the introduction of market imperfections, transaction or information costs, repeated interaction, or uncertainty as is the neo-liberal institutionalist understanding (Kehoane 1984; Axelrod and Keohane 1985; Oye 1986). Rather, our result follows directly from the costs imposed by wasteful investment in—and the potential use of—coercive force. Put simply, our model suggests that the Realist perspective is flawed for viewing peace via self-enforcing institutions as unsustainable. However, it suggests that the neoliberal view of peace in the absence of coercion is similarly untenable. At a minimum, costly investment in coercive force and a latent threat to use said force are often requisites of peaceful order.

What is more, our analysis suggests that (monopolistic) domestic politics are not inherently less costly than the (distributed) international system. In this way, we take aim at apologists for the state's monopoly of coercive force like Hobbes and Locke who view a sovereign agent to be the necessary or efficient solution to the problem of order (Locke 1988; Hobbes 1994). This sentiment is echoed, near ubiquitously, in contemporary scholarship on the political economy of development, where statelessness is viewed as anathema to growth, security, and the protection of human rights (Olson 1993, 2000; Bates 2008; Besley and Persson 2011; Boix 2015). From this perspective, peace brought about by the construction of a monopoly of force is a precondition for subsequent development. Our model suggests the exact opposite. Whenever peaceful order can be sustained by a monopolist of coercion, it can similarly be supported at lower cost with diffuse centers of coercion.

Existing formal models typically treat the construction of political order in one of two ways. The first fixes a game form and sees the emergence of state-like institutions as an equilibrium to this predefined game (Skaperdas 1992; Calvert 1995, 1998; Hirshleifer 1995; Hafer 2006; Piccione and Rubinstein 2007; Mayshar, Moav and Neeman 2017). The second approach takes a set of games, often one describing a state and another characterized as anarchy, and makes welfare comparisons between them, allowing a planner or decisive actor to choose between them (Moselle and Polak 2001; Grossman 2002; Konrad and Skaperdas 2012). Our approach combines the self-enforcing features of the "institutions as an equilibrium" approach with the understanding of institutions as formal rules as in the "institutions as constraints" approach. That is, we want to know when it is possible for agents in a state of nature to construct a rule that solves the problem of order as long as the agents commit to following that rule. We then can make welfare comparisons between sets of feasible institutions.

Our model also closely relates to a set of formal studies in international relations that develop models that focus on the cost of maintaining peace through coercive threats. Powell (1993) examines an infinite-horizon model of conflict with sequential arming decisions, identifying a unique peaceful equilibrium. In his model, as in ours, the key tradeoff is that investments in coercion reduce the adversary's incentive to fight, but they also reduce one's own utility from peace (see also Coe 2011). Jackson and Morelli (2009) and Fearon (2018) consider similar strategic environments, but with simultaneous arming decisions and the possibility of transfers. Our model is closest to these, but we abstract away from dynamic considerations. This allows for greater coercive asymmetries and enables us to endogenize the shape of the transfer protocol.

A separate but related line of scholarship examines the relationship between arming, private information, and social welfare (Meirowitz and Sartori 2008; Meirowitz et al. 2019). Partially mirroring our results on the inefficiency of monopolies of violence, Gowa and Ramsay (2017) show that high arms spending under unipolarity may result from the incentive to deter new challengers. In contrast with these analyses, which focus on characterizing the equilibrium level of arms and the resulting level of conflict given a particular negotiating protocol, we inquire whether there is *any* arrangement of armaments and settlement that produces peace.[1] From there we characterize the welfare implications of the various peace-enabling institutions and investment decisions.

2 Model

The model consists of an interaction between two political actors representing individuals or self-organized political groups.[2] Each actor's motivation is to maximize her own consumption, either through peaceful division of the social surplus or through coercive appropriation. Our main goal is to characterize when peace is sustainable as an equilibrium and, if it is, to characterize which arrangements of political order can achieve peace at the lowest social cost.

There are two players, player 1 and player 2. At the outset, each actor's share of the

[1] In a recent working paper on optimal political institutions in the shadow of conflict, Canidio and Esteban (2020) consider a modeling technology similar to ours. Like us, they find that peace might require inefficient investments in coercion.

[2] In the Appendix, we extend the model to allow for $N \geq 2$ symmetric players and find that our main efficiency results hold up.

society's total wealth is $y_i > 0$, where $y_1 + y_2 = 1$. Before the two players interact with each other, each chooses an investment in coercion, which increases her chance of winning a violent conflict but reduces the amount of wealth available for eventual consumption. That is, the more resources a player invests for violent ends, the less she can devote to productive ends. Formally, in the investment stage, each player simultaneously chooses an amount $m_i \in M_i$ to invest in coercion.[3] These investments are non-negative and cannot exceed the player's initial wealth: $M_i = [0, y_i]$. Throughout the paper, we assume each player's initial wealth is large enough so that she can respond optimally to any anticipated investment by the other player.[4]

After making their coercive investments, the players decide whether to divide the remaining wealth through peaceful negotations or violent conflict. There is a *political institution* determining the division of spoils in case peace prevails. Formally, a political institution is a pair of functions, $V_1(m_1, m_2)$ and $V_2(m_1, m_2)$, denoting how much each player will receive from negotations, given how much each has invested in coercion. A political institution must meet two conditions. First, it cannot leave a player worse off than if she lost all her wealth: $V_1(m_1, m_2) \geq 0$ and $V_2(m_1, m_2) \geq 0$. Second, it cannot waste any wealth: $V_1(m_1, m_2) + V_2(m_1, m_2) = 1 - m_1 - m_2$.

If instead conflict occurs, then there is a violent contest over society's wealth, whose outcome is determined in part by the players' coercive investments. Each player's chance of winning the contest (or, equivalently, the proportion of wealth she attains) is given by the contest success function:[5]

$$p_i(m_i, m_j) = \frac{\theta_i m_i}{\theta_i m_i + \theta_j m_j}. \tag{1}$$

[3]The choices need not be literally simultaneous; what matters is that neither actor can condition her choice on the other's.

[4]Formally, the condition is that each $y_i > \frac{\sqrt{\theta_j}}{2(\sqrt{\theta_i}+\sqrt{\theta_j})}$, where θ_i (defined below) represents each player's coercive effectiveness. A sufficient condition is $y_1 = y_2 = 1/2$. In the Appendix, we discuss how extreme disparities in initial wealth would affect our results.

[5]We let $p_i(0, 0) = \frac{1}{2}$ for the sake of symmetry, but all of our results would go through if there were an uneven distribution in case of no coercive investment.

The parameter $\theta_i > 0$ represents a player's *coercive effectiveness*: how much coercive force she can generate per unit of wealth she invests. The greater θ_i is, the cheaper it is for a player to build her forces to a given level. Without loss of generality, we assume $\theta_1 \geq \theta_2$.

We assume the outbreak of conflict is costly and inefficient—even beyond the inefficiency of the resources initially invested in preparation for conflict. Formally, we assume that conflict destroys a fraction $1-\gamma$ of the society's wealth, where $\gamma \in (0, 1)$. Letting $W_i(m_i, m_j)$ denote player i's expected utility from conflict given both players' coercive investments, we have

$$W_i(m_i, m_j) = \underbrace{\gamma}_{\text{fraction not destroyed}} \times \overbrace{p_i(m_i, m_j)}^{\text{prob. of winning}} \times \underbrace{[1 - m_i - m_j]}_{\text{wealth not used for coercion}}.$$

To close the model, we assume conflict prevails unless both players agree to divide wealth according to the political institution. Following the coercive investment decisions, each player simultaneously decides whether to opt for the political institution or for violent conflict.[6] We denote this decision with $w_i \in \{0, 1\}$, where $w_i = 0$ represents the institution and $w_i = 1$ represents conflict. Conflict occurs if either $w_i = 1$, meaning peace prevails if and only if $w_1 = w_2 = 0$.

2.1 Solution Concept

Taking the political institution (V_1 and V_2) as fixed for the moment, our solution concept is subgame perfect equilibrium. A strategy profile here consists of coercive investment choices m_1^* and m_2^*, as well as functions $w_1^*(m_1, m_2)$ and $w_2^*(m_1, m_2)$ indicating whether each player opts for conflict after all possible investment decisions. We are particularly interested in *peaceful equilibria*, in which the political institution prevails along the equilibrium path. Formally, an equilibrium is peaceful if $w_1^*(m_1^*, m_2^*) = w_2^*(m_1^*, m_2^*) = 0$.

However, in most of our analysis, we do not take the political institution as fixed. Instead,

[6]Our findings would not change if these decisions were sequential rather than simultaneous. A sequential model would eliminate equilibria in which each player opts for conflict simply because she expects the other to do the same, making her own choice immaterial. However, our analysis does not rely on such equilibria.

we ask whether it is possible to design some institutional framework to achieve a peaceful outcome. We say there exists an institution that *sustains peace* if there exist functions V_1 and V_2 such that the resulting game has a peaceful equilibrium. In a sense, the existence of a peace-sustaining institution is a weak result, as it holds as long as there is a single institutional specification with a peaceful equilibrium—even if many others wouldn't. By the same token, the non-existence of a peace-sustaining institution is a strong result: this means that out of the many conceivable bargaining protocols, none could lead to a peaceful result.

We highlight the distinction between political orders where coercive power is concentrated and those where it is dispersed. We say that a *monopoly of violence can sustain peace* if there exists an institution with a peaceful equilibrium in which $m_i^* > 0$ and $m_j = 0$. Similarly, a *distributed coercion can sustain peace* if there exists an institution with a peaceful equilibrium in which each $m_i^* > 0$. Again, in making these claims, we are searching across the set of all feasible institutional protocols, V_1 and V_2. If a monopoly of violence can sustain peace, that does not preclude the possibility that distributed coercion can sustain peace as well, as different institutional arrangements may lead to different choices of coercive investment.

2.2 Discussion of Model Framework

Before proceeding to the analysis, we briefly discuss the interpretation of the model and some of its notable features.

What Institutions Do. In the formal model, a political institution is simply a function that assigns a division of wealth to each player based on how much they invested in coercion. This might appear to be a quite restrictive conception of how institutions function. Indeed, our focus is on the redistributive aspects of the political interaction, much in line with the focus on bargaining in theories of both state formation (Levi 1988; North, Wallis and Weingast 2009) and international politics (Fearon 1995). That said, the minimal portrayal of bargaining in the model need not be taken literally. The V_i function may instead be interpreted as the reduced form of a much more complex political interaction between the players, representing what each player ultimately expects to gain from participation in the process. This process may be informal bargaining with few fixed rules, such as in classical international relations, or it could follow a rigid protocol, as in many contemporary domestic and international institutions. Our abstraction away from these details allows us to compare a diverse array of possible political orders.

Our analysis of institutions is similar in spirit to studies of mechanism design in economics and politics. These analyses reduce complex negotiating protocols to simple "direct mechanisms" to ask whether there is any way to achieve certain goals, such as the efficient provision of public goods (Groves and Ledyard 1977) or assured peace in crisis bargaining (Fey and Ramsay 2011). Formally, we follow these analyses in treating the rules of the game as a variable to solve for, rather than as a given feature of the strategic environment. Substantively, we mirror their concern for efficiency—which, in our case, amounts to identifying the institutions that can sustain peace at the lowest total cost.

We do not explicitly model the non-redistributive functions of institutions, such as the provision of public goods or the reduction of transaction costs. However, our model allows for the possibility of such collective benefits. In particular, the cost of non-participation—the fraction of consumption that is foregone if conflict occurs—may capture not just direct damage to output due to conflict, but also the loss of access to these collective benefits of political institutions.

How Conflict Happens. We assume that the peaceful institution prevails only if both players opt to participate in it. Conversely, conflict occurs if either player opts to fight. This reflects the common condition of interactions in anarchy—whether between states in the international system, or in the absence of established states altogether—that all actors reserve the ability to impose their will through force. Importantly, neither player can be forced to accept a political arrangement that leaves them with less than they could obtain through violence.

In our baseline model, there is complete information, and, accordingly, we find that it is always possible to identify a political institution that sustains peace. Howevever, this is a peace forged in the shadow of conflict and peace is not necessarily efficient in the sense of Fearon (1995). It may only be possible to maintain the peace through costly investments in coercion, which themselves sap the society of its wealth (Coe 2011; Fearon 2018). This inefficiency arises through a form of commitment problem, wherein players cannot commit themselves not to invest in armaments that enhance their future bargaining power.

In an extension following the main analysis, we relax the assumption of complete information. We show how uncertainty makes it harder to design institutions that sustain peace for sure, as these institutions now must guarantee enough surplus to satisfy very strong types even if the probability of observing such types is low. We also find that the goals of promoting peace and reducing inefficiency might be at odds when players do not fully understand each other's coercive capabilities.

Technology of Conflict. The contest success function in our baseline model, Equation 1, assumes constant returns to scale in the production of coercion. If we think of $\theta_i m_i$ as player i's coercive power, it costs no more or less for her to produce the second unit than it does to produce the first. Remember that a key goal of our analysis is to compare monopolies and distributions of violence in terms of their ability to produce peace at low cost. If we assumed increasing returns to scale in the production of coercion, this would naturally favor monopolies of violence; an assumption of decreasing returns would favor distributed coercion. Consequently, we focus on the case of constant returns to scale in order to provide a neutral baseline, where the deck is not stacked in favor of either type of political arrangement. We analyze an environment with increasing returns to scale following the main analysis.

3 Sustaining Peace

We begin the analysis by laying out the main conditions for peace to be sustainable as an equilibrium. The initial coercive investments determine the players' expected utilities from resorting to conflict, and thus play a large role in determining whether peace is sustainable.

The most obvious condition for a peaceful equilibrium is that the institution be *ex post* preferable to conflict. In terms of the model, the peaceful bargaining protocol must give each player at least as much as she would expect from resorting to force: $V_i(m_i, m_j) \geq W_i(m_i, m_j)$ for each player i. Because conflict is costly, it is always possible to design V_i such that this *ex post* condition holds. But the necessary conditions for peace are ultimately stronger than this, meaning there is often just a narrow range of coercive investments that can sustain peace.

In addition to this *ex post* condition, a peaceful equilibrium also requires the players' coercive investments to be optimal *ex ante*. At a minimum, neither player may have an incentive to deviate to investing a different amount and forcing a conflict, given how much she expects the other player to invest in coercion. This requirement makes it particularly difficult to sustain peace with low levels of coercive investment. To see why, imagine a player who expects her counterpart to invest nothing in coercion. By investing just a bit in coercion and opting for conflict, this player could capture all of the wealth that is left over after the costs of conflict are subtracted. Therefore, peace is sustainable only if this player expects to receive at least as much from the institution as she would if she were certain to win the conflict. Moreover, an equilibrium with no coercive investment at all is sustainable only if this condition is met for both players—which, as we will see, is a difficult condition to meet.

To formalize the *ex ante* necessary condition for peace, let a player's *reservation value* be the greatest payoff she could expect to receive by opting out and forcing conflict. Her expected value of doing so depends on how much she expects the other player to invest in coercion, so i's reservation value is a function of j's choice:

$$\mathrm{RV}_i(m_j) = \sup_{m_i \in M_i} W_i(m_i, m_j).$$

In order to have a peaceful equilibrium with coercive investments (m_i, m_j), each player must expect to receive at least her reservation value from the institution:

$$V_i(m_i, m_j) \geq \mathrm{RV}_i(m_j).$$

Our discussion so far has treated the institutional framework as fixed. However, we can combine the conditions developed so far into a necessary condition for there to exist *any* institution that sustains a peaceful equilibrium with coercive investments (m_1, m_2). Following the logic above, a peaceful equilibrium requires $V_1(m_1, m_2) \geq \mathrm{RV}_1(m_2)$ and $V_2(m_1, m_2) \geq \mathrm{RV}_2(m_1)$. Additionally, remember that the total payout from the institution is the wealth left over after the coercive investments: $V_1(m_1, m_2) + V_2(m_1, m_2) = 1 - m_1 - m_2$. Altogether, this means there exists an institution that sustains peace with these coercive investments only if

$$\mathrm{RV}_1(m_2) + \mathrm{RV}_2(m_1) \leq 1 - m_1 - m_2. \tag{2}$$

This condition is critical to our analysis of the efficiency of political orders. Our goal, in essence, is to find the lowest values of m_1 and m_2 that satisfy Equation 2.

Unless conflict is extraordinarily costly, it is impossible to sustain peace without at least one player making a coercive investment. To see why, remember that a player whose counterpart invests nothing can expect to receive the whole pie, less their own coercive investment and the overall cost of conflict: $p_i(m_i, 0) = 1$ and thus

$$W_i(m_i, 0) = \gamma(1 - m_i).$$

Evidently, then, the reservation value of such a player is $\mathrm{RV}_i(0) = \gamma$, as she can invest an infinitesimal amount in coercion and still win for certain. Substituting $m_1 = 0$ and $m_2 = 0$ into the necessary condition for a peace-sustaining institution, Equation 2, we see that peace is sustainable only if conflict would destroy a majority of society's wealth.[7]

Proposition 1. *Peace can be sustained without coercive investment only if* $\gamma \leq 1/2$.

In the remainder of the paper, we assume $\gamma > 1/2$, so as to focus on the most substantively interesting cases—those where at least a modicum of coercion is necessary to sustain peace. Because we are agnostic about the shape of institutions, Proposition 1 tells us that the absence of peace without coercive armament when $\gamma > 1/2$ is not a matter of the wrong bargaining protocol or poor institutional design. It is a fundamental property of a strategic interaction where players may invest in coercive means and resort to force. Our task in the remainder of the analysis is to discern how best to reduce the cost of the coercive investments necessary to sustain peace—and to explain why observed institutional frameworks may deviate from this ideal.

While we have just seen the difficulty of sustaining peace when coercive investments are too small, it is worth noting that high investments may also be unsustainable. The necessary condition given in Equation 2 expresses a fundamental strategic tension in the choice of coercive investments. If they are too low, then the players' reservation values will be too high for the condition to hold, as seen above. But if the investments are too large, then there might not be enough wealth left to satisfy both players. To take one extreme example, it would be impossible to sustain peace by investing all of society's wealth in coercion (i.e., with $m_1 + m_2 = 1$). Both players would then receive nothing from participating in the institution, meaning either would benefit by investing *less* in coercion and forcing a conflict.

So far we have found a necessary condition for peace to be sustainable—but is it sufficient? In fact, it is: as long as Equation 2 holds, it is possible to design an institution that sustains peace. We find that a simple and intuitive redistributive scheme is sufficient to preserve

[7] The proof is omitted, as the logic in the text establishes the result. Proofs of all subsequent propositions appear in the Appendix.

peace with coercive investments (m_1^*, m_2^*).[8] On the path of play, the institution must give each player at least her reservation value: $V_i(m_i^*, m_j^*) \geq \mathrm{RV}_i(m_j^*)$. If a player deviates to a lower investment in coercion, then the institution must lower her proportional share of the total surplus, so that her overall payoff from peace does not increase.[9] Similarly, if a player deviates to a greater investment, then her proportional share can increase at the other player's expense, but not so much that the deviating player's total payoff increases. Whether the players ultimately coordinate on this kind of peace-producing scheme is beyond the scope of our analysis. We are merely able to distinguish when peace is at least possible—and to characterize what it costs in terms of coercive investments.

4 Monopolies of Violence

We now consider monopolies of violence, in which coercive authority is concentrated in a single player. In order for a monopoly of violence to sustain peace, two conditions must be met. First, the monopolist must invest enough in coercion to deter the other player from resorting to force. Second, the political institution must promise enough surplus to the monopolist that she is not tempted to exploit her overwhelming advantage in the application of force. These conditions are in tension with each other: the more the monopolist spends on coercive power to deter conflict, the less wealth remains to satisfy the monopolist's demands.

Consider a monopoly of violence by player i, with coercive investments $m_i^* > 0$ and $m_j^* = 0$. We will call player i the monopolist and player j the subject. The monopolist, expecting the subject to invest nothing in coercion, could win the entire pie through violence even if she were to deviate to investing just a small amount herself. Consequently, the monopolist will prefer peace over violence only if the institution promises her the whole

[8] See Lemma 10 in the Appendix for formal details.

[9] In case of a monopoly of violence by player i, where $m_i^* > 0$ and $m_j = 0$, this scheme may violate the outside option principle (Binmore, Rubinstein and Wolinsky 1986), as i's deviation to $m_i < m_i^*$ increases j's payoff from bargaining even though j's outside option is unchanged ($W_j(m_j, 0) = W_j(m_j^*, 0) = 0$). However, this is an artifact of the simple contest success function we employ. For example, if we replaced Equation 1 with the almost-equivalent $p_i(m_i, m_j) = (\theta_i m_i + \epsilon)/(\theta_i m_i + \theta_j m_j + 2\epsilon)$, then a downward deviation by the monopolist would indeed raise the other player's outside option value.

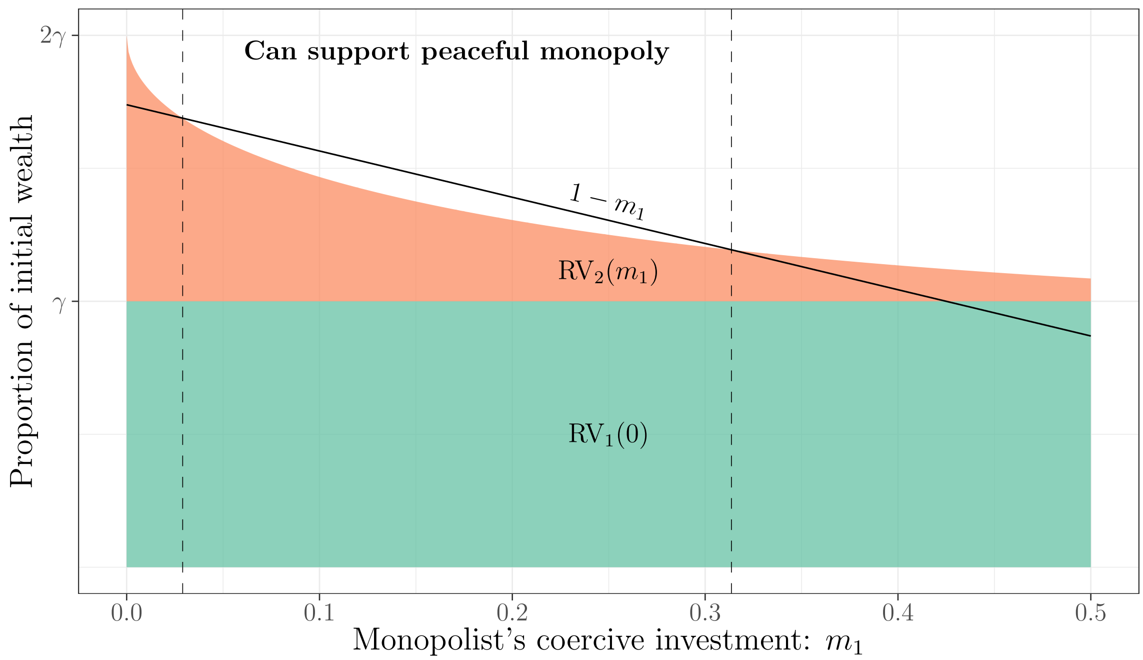


Figure 1. Range of coercive investments by player 1 that can support a peaceful monopoly of violence. (Parameters: $\theta_1 = \theta_2 = 1$, $\gamma = 0.575$.)

post-conflict pie: $\mathrm{RV}_i(0) = \gamma$. Meanwhile, the subject's reservation value, $\mathrm{RV}_j(m_i^*) < \gamma$, is a decreasing function of the monopolist's coercive investment. Using our necessary and sufficient condition for there to exist an institution that sustains peace (Equation 2), we know that peace is possible with this monopoly of violence if and only if

$$\gamma + \mathrm{RV}_j(m_i^*) \leq 1 - m_i^*. \tag{3}$$

If m_i^* is too low, this condition fails because the subject's reservation value is too great. On the other hand, if m_i^* is too high, it fails because there is not enough wealth left over to satisfy the monopolist. At most, then, a monopoly of violence is sustainable with an intermediate range of coercive investments.

The conditions for a monopoly of violence to sustain peace are illustrated in Figure 1. Remember that peace is sustainable when there is enough wealth (represented in the figure by the black line) to meet both players' reservation values (the colored shaded areas). An

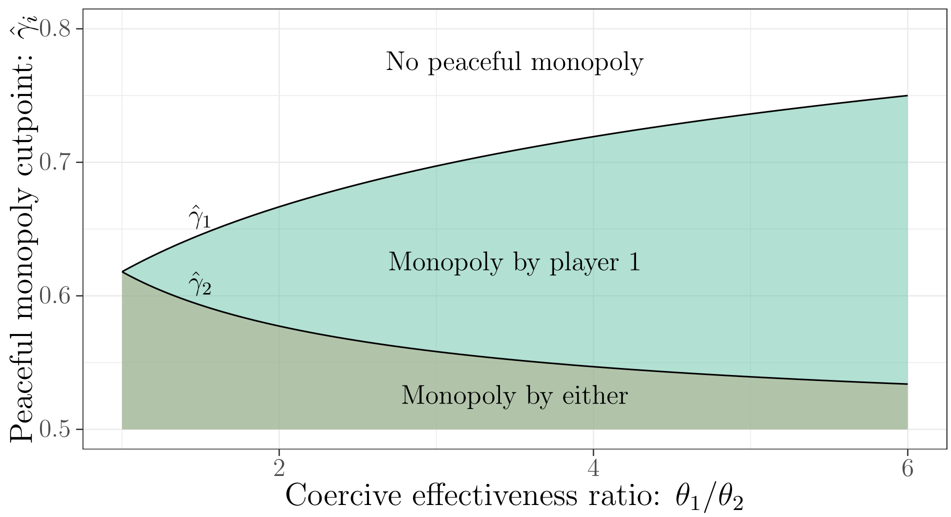


Figure 2. The sustainability of peace with a monopoly of violence, as a function of the imbalance in coercive effectiveness between players.

increase in the monopolist's coercive investment (moving right along the x-axis) decreases the subject's reservation value, but also tightens the wealth constraint. Critically, while the effect on the wealth constraint is constant, there are diminishing returns in the reduction of the subject's reservation value—at a certain point, additional coercive investment by the monopolist barely changes what the subject could expect to get from forcing a conflict. Therefore, at low levels of coercive investment by the monopolist, a small increase opens up the space for peace to be sustainable, while the opposite is true at higher levels.

The costs of conflict and the players' coercive effectiveness also determine whether a monopoly of violence can sustain peace. As violence becomes more costly—in terms of the model, as γ decreases—both players' reservation values decrease, making the conditions for peace easier to meet. Similarly, an increase in the monopolist's coercive effectiveness relative to the subject's, θ_i/θ_j, decreases the subject's reservation value, widening the space for peace.[10] Consequently, it is easier for a monopoly of violence by the stronger player (which we have labeled player 1) to yield a peaceful outcome, as illustrated in Figure 2. The following result summarizes these properties of monopolies of violence.

[10] The ratio is immaterial to the monopolist because she expects the subject to invest nothing, and thus can guarantee victory even with an arbitrarily small investment in coercion.

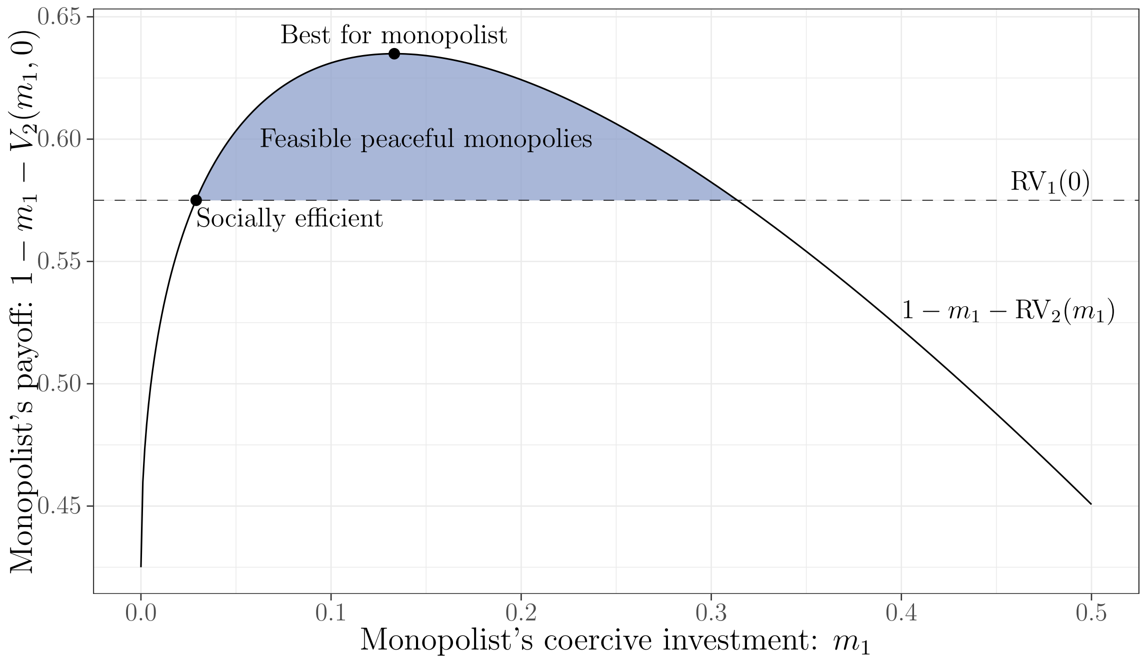


Figure 3. Monopolist's payoff from a peaceful equilibrium as a function of coercive investment. (Parameters: same as Figure 1.)

Proposition 2. *A monopoly of violence by player i can sustain peace if and only if $\gamma \leq \hat{\gamma}_i$, where $\frac{1}{2} < \hat{\gamma}_2 \leq \hat{\gamma}_1 < 1$.*

Our core concern is with the efficiency of different forms of political order. That is, we wish to identify the lowest total coercive investment at which peace can be sustained. It is evident from Figure 1 that there may be a range of coercive investments at which a monopoly of violence can sustain peace.[11] The low point of this range (marked in the figure by the dotted line on the left) is the efficiency frontier for a monopoly of violence. Mathematically, this is the lowest value of m_i^* at which Equation 3 holds.

In looking at the range of monopolies of violence, we see how social and individual incentives are at odds with one another in the construction of political order. If a monopoly of violence can sustain peace with a monopolistic investment of m_i^*, then the best institutional arrangement for the monopolist is one that gives the subject her reservation value and gives all remaining wealth to the monopolist: $V_i(m_i^*, 0) = 1 - m_i^* - \mathrm{RV}_j(m_i^*)$. Figure 3

[11] In the knife-edge case $\gamma = \hat{\gamma}_i$, there is exactly one level of coercive investment such that a monopoly of violence by player i can sustain peace. Otherwise, if $\gamma < \hat{\gamma}_i$, there is a range of such investments.

plots the monopolist's best-case consumption from a peaceful equilibrium as a function of her coercive investment. The arrangement that involves the lowest outlay on coercion also results in the lowest possible payoff for the monopolist, leaving her exactly indifferent between the institution and conflict. By investing slightly more and further reducing the subject's reservation value, the monopolist can come away with a surplus rather than breaking even. Therefore, given a choice among institutions that sustain a peaceful monopoly of violence, the monopolist would choose one that entails more investment in coercion than is strictly necessary to keep the peace.

Proposition 3. *If $\gamma < \hat{\gamma}_i$, then the peaceful monopoly of violence that maximizes player i's payoff entails more investment than the socially efficient level.*

This result may seem counterintuitive, as an over-investment in coercion shrinks the size of the pie that is redistributed in a peaceful equilibrium. But shrinking the pie, up to a certain point, is strategically advantageous for the monopolist. The less wealth there is left over after coercive investment, the less incentive the subject has to engage in costly conflict over the remaining wealth. Consequently, the subject's reservation value shrinks rapidly with the monopolist's investment, allowing the monopolist to extract more from redistribution while maintaining the peace. At the margin, the reduction in the subject's reservation value due to the monopolist's investment outweighs the reduction in the size of the pie, giving the monopolist an incentive to over-invest.

5 Distributed Coercion

We now examine political orders that sustain peace by distributing the means of coercion rather than concentrating them in a single actor. From a social welfare standpoint, we find that these arrangements are superior to monopolies of violence in two ways: the conditions for peace to be sustainable through a distribution of violence are less stringent and the total coercive investment required is lower. Nevertheless, in comparing monopolies of violence

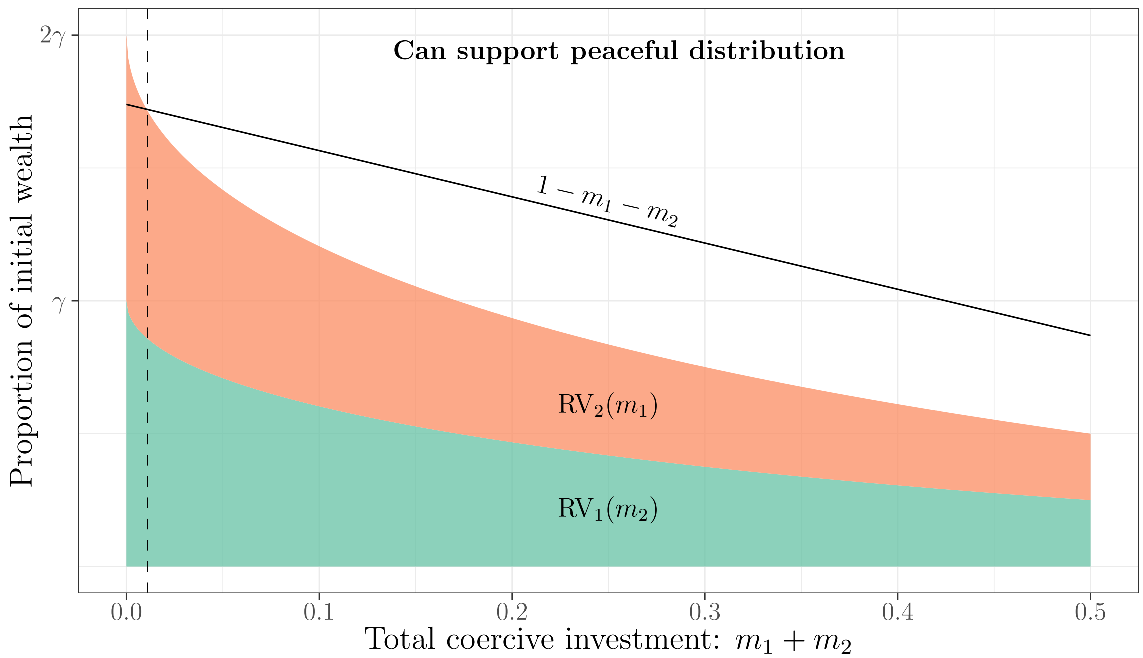


Figure 4. Range of coercive investments that can support peace through distributed coercion with equal investments, $m_1 = m_2$. (Parameters: same as Figure 1.)

and distributed coercion, we once again see a stark difference between social and individual incentives. Given a choice among institutions that sustain peace, any individual player would choose a monopoly of violence with an unnecessarily high level of coercive investment.

In terms of the model, distributed coercion is an equilibrium with coercive investments by both players: each $m_i^* > 0$. For such an arrangement to sustain peace, each player must invest enough in coercion to deter the other from resorting to violence. However, just in the case of a monopoly of violence, too large an investment in coercion may leave the players without enough wealth to bring each player's peaceful payoff up to her reservation value. To illustrate the sustainability of peace under distributed coercion, Figure 4 shows the wealth constraint and the individual players' reservation values as a function of the total investment, assuming both players invest equal amounts.[12] As before, distributed coercion can sustain peace with the given investments as long as the wealth constraint exceeds the sum of the players' reservation values.

[12] By our definition, distributed coercion need not entail equal investments by the two players. The assumption of equal investments is simply for convenience in displaying the figure.

Under distributed coercion, unlike in a monopoly of violence, neither player can expect to receive the entire pie by forcing a conflict. Remember that a monopoly of violence must reserve at least γ for the monopolist in order to sustain peace, lest she be tempted to employ violence to increase her consumption. Political orders with distributed coercion do not face this stringent requirement. This is evident in Figure 4: as the total investment (evenly split) goes up, each player's reservation value decreases. Consequently, it is easier to sustain peace through distributed coercion than through a monopoly of violence. Even when the costs of conflict are quite low ($\gamma \approx 1$), it is possible to design a political institution that sustains peace through distributed coercion. In sum, whenever conflict destroys even the smallest fraction of wealth, it is possible to construct an institution that promotes peace via distributed coercion.

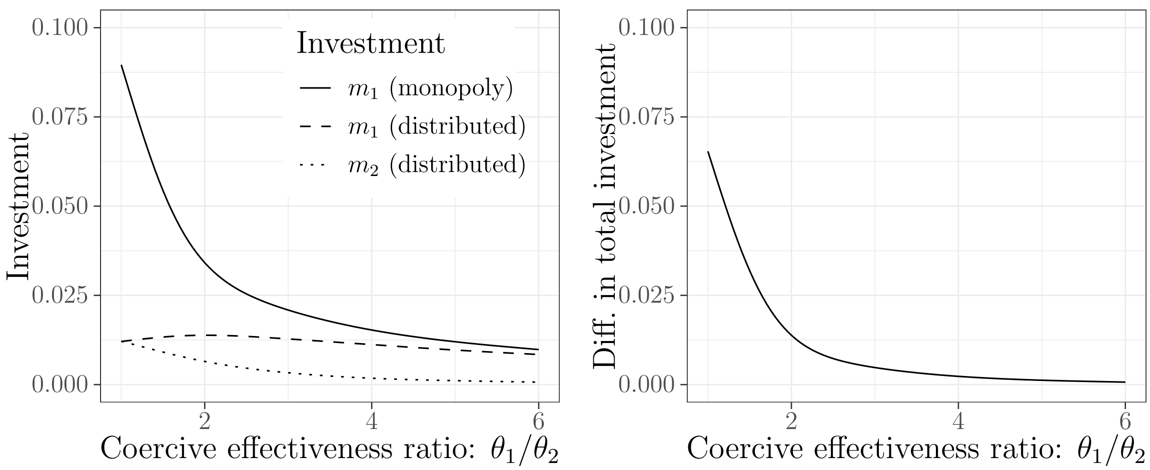
Proposition 4. *For all $\gamma < 1$, distributed coercion can sustain peace.*

Besides having less stringent conditions to sustain peace in the first place, distributed coercion can also do so at lower cost than a monopoly of violence. Comparing Figure 2 and Figure 4, we see that the lowest total coercive investment with distributed coercion is even lower than the most efficient monopoly of violence. Therefore, we obtain a clear answer to our question about which organization of political order is the most efficient at producing peace. From the standpoint of social welfare, any peaceful monopoly of violence could be improved by having the monopolist invest slightly less and the subject invest slightly more.

Proposition 5. *If a monopoly of violence can sustain peace, then distributed coercion can sustain peace with a strictly lower total investment in coercion.*

This result follows from the decreasing returns to coercive investment as an instrument of deterrence. The first unit of investment does much more to decrease a player's reservation value than does the second, which in turn does more than the third.[13] To produce peace

[13]This result emerges in part from our technological assumptions about the relationship between investment and coercive power (i.e., Equation 1). In an extension below we consider economies of scale in the production of coercion. In that setting, Proposition 5 continues to hold as long as the players are close enough in their coercive effectiveness: $\theta_1 \approx \theta_2$.



(a) Least costly equilibrium. **(b)** Inefficiency of a monopoly.

Figure 5. Efficient equilibria and the efficiency gap as a function of the stronger player's coercive advantage. (Parameters: $\gamma = 0.615$.)

at the lowest cost, per the condition identified in Equation 2, we must find the coercive investments that lower the players' reservation values just enough that the remaining wealth can compensate them for refraining from conflict. Because of the diminishing returns in the reduction of a player's reservation value, this can be accomplished more cheaply by having both players spend a bit than by having a single player spend a lot. If the goal is to maintain peace while wasting as few resources as possible on coercion, it is counterproductive to concentrate the costs of coercion in a single actor.

The magnitude of the inefficiency in a monopoly of violence depends on how imbalanced the players are in their coercive effectiveness. The closer they are to equality, the more inefficient a monopoly of violence will be, as illustrated in Figure 5. At parity, it requires a substantially larger investment to maintain a monopoly of violence than the most efficient arrangement with distributed coercion. However, as the coercive advantage of the stronger player grows, the equilibrium investment of the weaker player in the most efficient distribution of coercion shrinks. Consequently, the efficiency difference between this and the least expensive monopoly of violence becomes negligible.

We have shown that distributed coercion can sustain peace at lower cost than a monopoly.

But by the same token, peace is also sustainable at much greater costs when coercive force is distributed than when it is concentrated. For example, again comparing Figure 1 and Figure 4, we see that a monopoly of violence could not sustain peace with a total coercive investment of half of society's wealth ($m_1 = \frac{1}{2}$). However, distributed coercion with the same total investment ($m_1 = m_2 = \frac{1}{4}$) could sustain peace. In other words, the range of investments that can support peace is wider—at both ends—with a distribution of violence than with a monopoly.

Proposition 6. *If a monopoly of violence can sustain peace, then distributed coercion can sustain peace with a strictly greater total investment in coercion.*

This result follows from the monopolist's high consumption demand in a peaceful monopoly of violence. Remember that such an arrangement must promise the monopolist at least γ, as this is what she would expect to obtain from violently expropriating a defenseless subject. At a minimum, therefore, a monopoly of violence cannot sustain peace with $m_i^* > 1 - \gamma$. With distributed coercion, however, neither player can expect to receive much from forcing a conflict—particularly when each anticipates that the other will invest a high amount in coercion. Therefore, a greater degree of inefficiency is possible in peaceful political orders where coercive force is distributed among players.

In combination, Proposition 5 and Proposition 6 lead us to an ambiguous conclusion about the social welfare consequences of the international system. International politics is famously characterized by a diffusion of coercive authority across distinct actors, in contrast with domestic politics that more closely resemble monopolies of violence. Our model implies that concentrating international power in a single actor would require more costly investment in coercion than the most efficient distribution of force across individual states. However, to say that a greater concentration of power would necessarily cost more than the status quo would be to take the Panglossian view that the international system is currently in the best of all possible equilibria. Distributed coercion cannot attain peace at lower cost than a monopoly of violence unless the players interact in the exact right institutional framework and coordinate on an efficient allocation of coercive force.

We conclude the section by considering why political institutions might not achieve the efficiency bounds we have described here. If monopolies of violence have a greater coercive cost than necessary, and distributed coercion can sustain peace as an equilibrium, why should we ever observe the monopolization of force by a sovereign government? The problem is that the best equilibrium for the society as a whole is not necessarily the best for an individual player. If one player could dictate the choice of equilibrium, she would select a monopoly of violence. The following results extends Proposition 3 by showing that not only is the best monopoly of violence for a player one in which she invests more than necessary in coercion, but that the player prefers this monopoly over any other arrangement of political order.

Proposition 7. *If $\gamma < \hat{\gamma}_i$, then player i can obtain a greater payoff from a monopoly of violence with inefficiently high coercive investment than from any other institution that sustains peace.*

In summary, we have characterized the conditions for each type of peaceful equilibrium and uncovered some important implications for the efficiency of political orders. When some level of coercive investment is necessary to preserve peace, the cheapest way to do so involves investment by both players, with the majority coming from the player with the greater coercive effectiveness. Precisely because distributed coercion is cheaper to sustain, it is supportable under a wider set of conditions than is a monopoly of violence. However, social efficiency and individual incentives do not coincide. If a single player could dictate the nature of political order, she would choose a monopoly of violence where she is the monopolist and invests more than is necessary to deter the other player from conflict.

6 Economies of Scale

The baseline model analyzed above assumes constant returns to scale in the production of coercive force. We now relax this assumption and examine whether our main results hold up when there are economies of scale in coercion. In particular, we consider an environment in which it is cheaper, in terms of the total investment required, to have a single player produce a unit of coercive force than to split its cost across both. Intuitively, monopolies of violence ought to appear more attractive in this environment than in the baseline model. We confirm that this is the case: economies of scale mitigate the inefficiency of a monopoly of violence. Nevertheless, for players with symmetric initial wealth and coercive effectiveness, a monopoly of violence remains inefficient compared to the optimal institution with distributed coercion. Economies of scale alone do not make a monopoly of violence efficient—there must also be an initial imbalance in favor of the monopolist.

To model economies of scale, we simply replace the contest success function from the original model. Specifically, the model with economies of scale is identical to the original, replacing Equation 1 with

$$p_i(m_i, m_j) = \frac{\exp(\theta_i m_i)}{\exp(\theta_i m_i) + \exp(\theta_j m_j)}. \tag{4}$$

Unlike in the baseline model, even if $\theta_1 = \theta_2$, the cheapest way to produce $x > 0$ units of effective coercion is to have a single player take on the entire cost. Additionally, each player always recoups some fraction of the prize in case of conflict, even if she invests nothing. Consequently, unlike in the baseline model, it may be optimal for a player to choose $m_i = 0$ even if she anticipates conflict.

Even with economies of scale in the production of coercion, our finding on the efficiency of a monopoly of violence holds up as long as there is not too much inequality between the players at the outset of the interaction. The following proposition states this result formally.

Proposition 8. *Assume $y_1 = y_2 = \frac{1}{2}$ and $\theta_1 = \theta_2$ in the model with economies of scale. For any institution that sustains peace with $m_1^* \neq m_2^*$, there is an institution that sustains peace with strictly less total investment in which $m_1^* = m_2^*$.*

To see why this finding holds, it is important to understand that there is not a fixed price

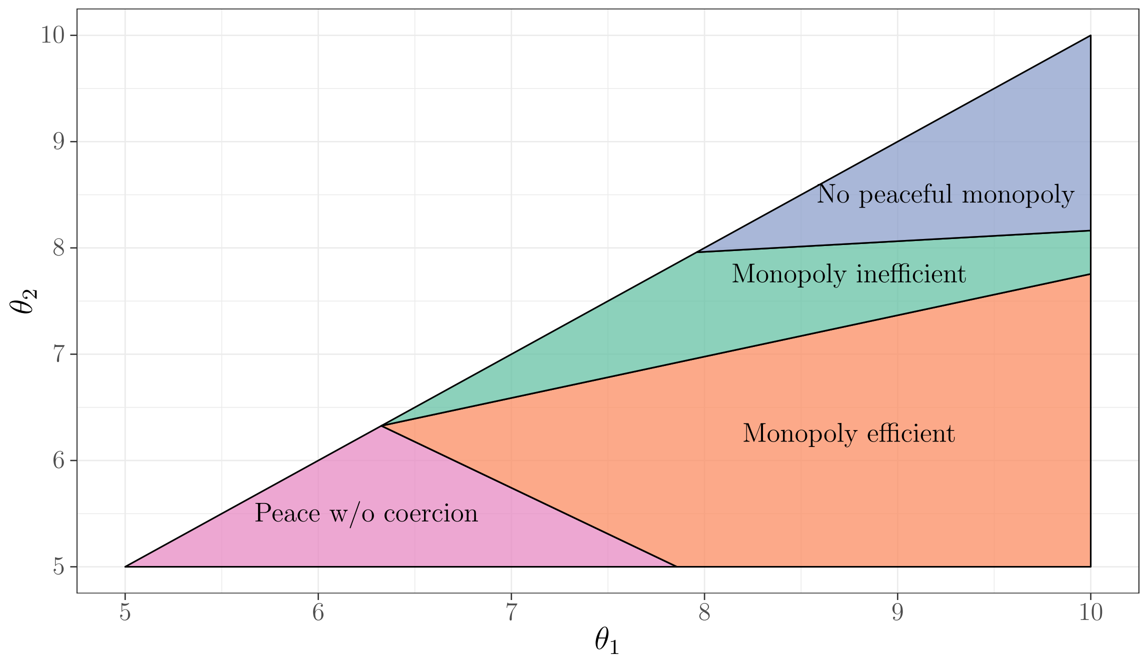


Figure 6. Existence and efficiency of equilibria in the game with economies of scale in the production of coercive force. (Parameters: $\gamma = 0.8$.)

for the preservation of peace. If that were the case, then it would be optimal to concentrate coercion in a single authority even if the players were *ex ante* identical. Instead, the price of preserving the peace depends on how much each player could expect to receive from conflict, which in turn is a function of the equilibrium choices of coercive investment. Whether or not there are economies of scale in the production of coercion, there is a steep price to deter a monopolist from exercising her coercive advantage through conflict. Therefore, it remains cheaper to have both players spend just enough to deter each other from defecting to conflict.

On the other hand, the introduction of economies of scale substantively changes our results when the players differ in their coercive effectiveness. In particular, if one player has a significant advantage in the application of coercion ($\theta_1 > \theta_2$), then the lowest-cost peaceful political order might entail having that player as the monopolist. Figure 6 illustrates the existence and efficiency of institutions that sustain peace as a function of the coercive imbalance between players. When player 1's coercive effectiveness substantially outweighs player 2's (the bottom-right of the figure), a monopoly of violence can achieve peace at lower cost than distributed coercion. However, when the players are roughly equal (the 45-degree line), a monopoly of violence is either inefficient or fails to sustain peace altogether.

To sum up, the results of the extension confirm the intuition that the technology of coercion affects the relative efficiency of various political orders. However, while increasing returns to scale in the production of coercion appear to be a necessary condition for a monopoly of violence to be efficient, they are not sufficient on their own. Unless there is also an initial imbalance in the technology of coercion, distributed coercion remains the most economical arrangement of political order.

7 Equilibria with Conflict

Our focus up to now has been on how actors in an initial state of anarchy may design institutions to sustain peace. Even in the absence of conflict, the cost of the coercive investments necessary to sustain peace is a drag on the society's wealth. We now consider an environment with private information and show that the most efficient political order sometimes involves a positive probability of violence.

When the players have private information about their own coercive effectiveness, the cost of sustaining peace is even greater than in the baseline case. Peace is guaranteed only if the strongest type of each player receives as much from the institution as they could expect to win from fighting. For example, imagine a situation in which player 1 is very likely, but not certain, to be much stronger than player 2. In order to ensure peace, player 1 must make a considerable coercive investment to deter even the stronger type of player 2 from fighting. *Ex post*, however, such an investment will prove to be wasteful when player 2 turns out to be weak—as is likely the case. In expectation, it may be much less costly for player 1 to invest just enough to deter the weaker type of player 2, accepting that conflict will occur in case player 2 turns out to be strong. An efficient political order may therefore entail a chance of open conflict on the equilibrium path.

To formalize the logic of this example, we consider a simple extension of the baseline model to include private information.[14] In the *private information model*, it is common knowledge that player 1's coercive effectiveness $\theta_1 = 1$. However, player 2's coercive effectiveness θ_2 is private information. We assume a simple type space, $\theta_2 \in \{0, 1\}$, so that either the players have equal coercive effectiveness ($\theta_2 = 1$) or player 1 is guaranteed to win any conflict ($\theta_2 = 0$).[15] Let $\pi \in (0, 1)$ denote the prior probability that $\theta_2 = 1$, and assume this distribution is common knowledge. Finally, to avoid problems due to boundary constraints, assume each $y_i = \frac{1}{2}$.

In the private information model, player 1 could deter the weaker type of player 2 from conflict by investing an infinitesimal $m_1 > 0$. An institutional arrangement that assigned $V_1 \geq \gamma$ and $V_2 \geq 0$ would result in peace at an arbitrarily low coercive cost. However, if $\gamma > \frac{1}{2}$, such an arrangement would not be satisfactory to the strong type of player 2. In particular, if $m_1 \approx 0$, then the strong type could expect to receive close to γ by forcing a conflict. Therefore, peace cannot be assured without player 1 investing as much as it would take to sustain a peaceful equilibrium if it were *certain* that $\theta_2 = 1$. If the prior probability of a strong type of player 2 is high, this may be indeed be the most efficient solution to the problem of order. But if $\pi \approx 0$, then the social welfare optimum entails player 1 investing very little, thereby accepting that conflict will occur in the rare case that player 2 is strong.

Proposition 9. *In the private information model, if $\gamma > \frac{1}{2}$ and $\Pr(\theta_2 = 1)$ is sufficiently low, then any peaceful equilibrium is* ex ante *inefficient.*

Figure 7 illustrates this result. If the prior probability of a strong type of player 2 is great enough, then assured peace is optimal from the standpoint of social welfare. However, when π becomes low enough, so that strong types are rare, it is more efficient to allow for conflict. In fact, as π approaches zero, there is close to zero social welfare loss in equilibrium,

[14] We focus on one-sided private information here to make the logic as clear as possible. In the Appendix, we show that similar results may apply with two-sided private information.

[15] We assume this for ease of exposition in the extension. Similar results would hold with a richer type space.

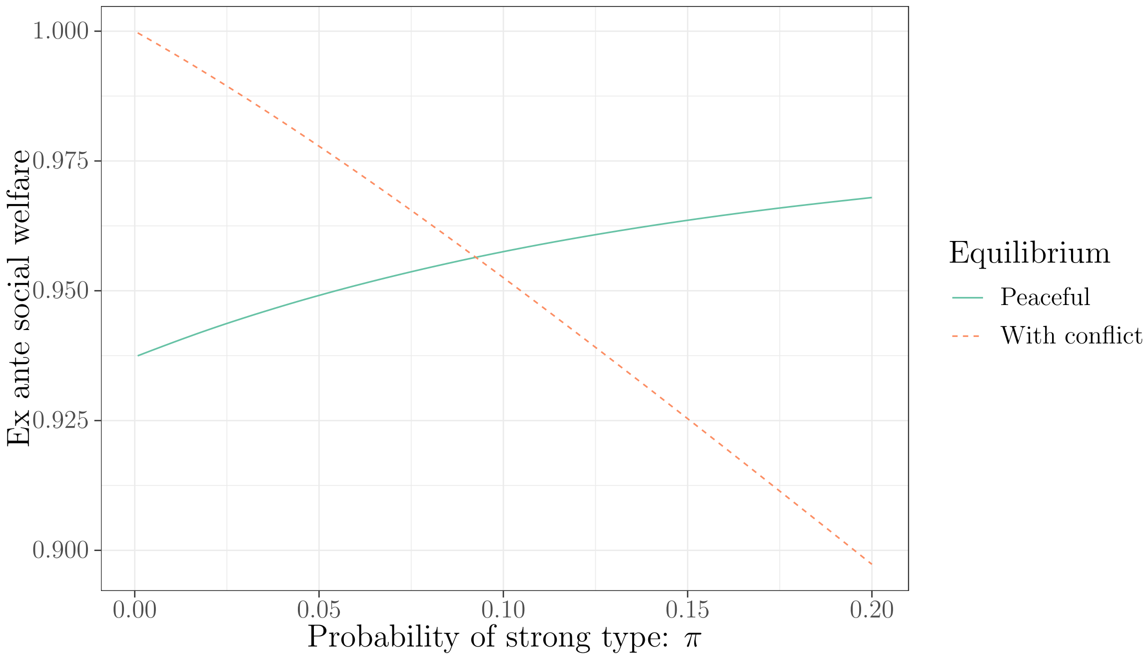


Figure 7. *Ex ante* social welfare with peaceful and conflictual equilibria in the game with private information. (Parameters: $\gamma = 0.6$.)

as conflict occurs with probability π and player 1's coercive investment is close to nil.

Conflict among nations is widely—and correctly—seen as a source of waste and inefficiency. Our result here shows that the alternative might be even worse for social welfare. In order to conclude that a world with assured peace would be an improvement, we need to compare the expected losses of conflict with the additional coercive investment that would be necessary to guarantee peace. Indeed, prior scholarship has identified the high cost of arming as a potential cause of interstate war (Coe 2011). This result is also closely related to the well-known finding in economics that there are generally not *ex post* efficient trading mechanisms that are compatible with individual incentives (Myerson and Satterthwaite 1983).

8 Conclusion

When is peaceful political order self-enforcing? Under what conditions can monopolies of violence be sustained? If these conditions are met, what are the welfare implications of order? The model we developed in this paper answers these questions. In contrast with both conventional wisdom and extant scholarship we have shown that orders characterized by a monopoly of force are generally inefficient relative to political orders where multiple agents maintain coercive abilities. Furthermore, even within the set of peace-preserving institutions backed by a monopoly of violence, the institution most preferred by the monopolist requires an inefficiently high investment in coercion. In other words, monopolies of violence beget monopoly rents.

Our results suggest that organizing principles defined by sovereign constituent units that monopolize violence are not “natural” in the way many contemporary observers of international politics might assert. Indeed, political order based upon diffuse coercive abilities is, as we have shown, sustainable whenever there is peace based upon a monopolist of force. Moreover, we find that this diffuse coercive power is welfare-improving.

Why then does the state persist? Our model suggests two plausible answers. First, our model admits multiple equilibria. Both monopolized and distributed coercion are supported under a wide range of structural conditions and, moreover, our approach is agnostic about selection between them. As such, sociological factors – ideas, beliefs, and identities – may constrain actors from coordinating on other, potentially more efficient, equilibria. Nevertheless, our model suggests a second also likely answer: monopolies of violence, though inefficient, persist because they endow the most powerful actors with the greatest payoff. If the powerful are capable of establishing the rules of the game, we expect them to select socially inefficient yet individually optimal institutional arrangements.

References

Axelrod, Robert and Robert O Keohane. 1985. "Achieving cooperation under anarchy: Strategies and institutions." *World politics* 38(1):226–254.

Bates, Robert H. 2008. "State Failure." *Annual Review of Political Science* 11.

Besley, Timothy and Torsten Persson. 2011. *Pillars of prosperity: The political economics of development clusters*. Princeton University Press.

Binmore, Ken, Ariel Rubinstein and Asher Wolinsky. 1986. "The Nash Bargaining Solution in Economic Modelling." *The RAND Journal of Economics* pp. 176–188.

Boix, Carles. 2015. *Political Order and Inequality: Their Foundations and their Consequences for Human Welfare*. Cambridge University Press.

Calvert, Randall. 1995. "The rational choice theory of social institutions: Cooperation, coordination, and communication." *Modern political economy: Old topics, new directions* pp. 216–268.

Calvert, Randall L. 1998. "Explaining Social Order: Internalization, External Enforcement, or Equilibrium?" *Institutions and social order. The University of Michigan Press, Ann Arbor* pp. 131–161.

Canidio, Andrea and Joan Esteban. 2020. "Optimal Political Institutions in the Shadow of Conflict.". Working paper.
URL: `http://andreacanidio.com/research`

Coe, Andrew J. 2011. "Costly peace: A new rationalist explanation for war." *Working Paper, University of Southern California* `https://www.andrewjcoe.com/research`.

Fearon, James D. 1995. "Rationalist explanations for war." *International organization* 49(03):379–414.

Fearon, James D. 2018. “Cooperation, conflict, and the costs of Anarchy.” *International Organization* 72(3):523–559.

Fey, Mark and Kristopher W Ramsay. 2011. “Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict.” *American Journal of Political Science* 55(1):149–169.

Gowa, Joanne and Kristopher W Ramsay. 2017. “Gulliver Untied: Entry Deterrence Under Unipolarity.” *International Organization* 71(3):459–490.

Grossman, Herschel I. 2002. “Make us a king: anarchy, predation, and the state.” *European Journal of Political Economy* 18(1):31–46.

Groves, Theodore and John Ledyard. 1977. “Optimal Allocation of Public Goods: A Solution to the ‘Free Rider’ Problem.” *Econometrica* pp. 783–809.

Hafer, Catherine. 2006. “On the origins of property rights: Conflict and productionin the state of nature.” *The Review of Economic Studies* 73(1):119–143.

Hirshleifer, Jack. 1995. “Anarchy and its breakdown.” *Journal of Political Economy* pp. 26–52.

Hobbes, Thomas. 1994. *Leviathan: with selected variants from the Latin edition of 1668*. Vol. 8348 Hackett Publishing.

Jackson, Matthew O. and Massimo Morelli. 2009. “Strategic Militarization, Deterrence and Wars.” *Quarterly Journal of Political Science* 4(4):279–313.

Kehoane, R. 1984. “After Hegemony: cooperation and discord in the world political economy.” *After Hegemony: Cooperation and Discord in the World Political Economy* .

Keohane, Robert O. 1986. *Theory of world politics: structural realism and beyond*. Columbia University Press New York pp. 190–97.

Konrad, Kai A and Stergios Skaperdas. 2012. "The Market for Protection and the Origin of the State." *Economic Theory* 50(2):417–443.

Lake, David A. 1996. "Anarchy, hierarchy, and the variety of international relations." *International organization* 50(1):1–33.

Lake, David A. 2007. "Escape from the state of nature: Authority and hierarchy in world politics." *International Security* 32(1):47–79.

Lake, David A. 2009. *Hierarchy in international relations*. Cornell University Press.

Levi, M. 1988. "Of Rule and Revenue (Berkeley." *Los Angeles, London (University of California Press), Kap* 5:95–121.

Locke, John. 1988. *Locke: Two Treatises of Government Student Edition*. Cambridge University Press.

Mayshar, Joram, Omer Moav and Zvika Neeman. 2017. "Geography, Transparency, and Institutions." *American Political Science Review* 111(3):622–636.

Mearsheimer, John J. 1994. "The false promise of international institutions." *International security* 19(3):5–49.

Mearsheimer, John J. 2001. *The tragedy of great power politics*. WW Norton & Company.

Meirowitz, Adam and Anne E. Sartori. 2008. "Strategic Uncertainty As a Cause of War." *Quarterly Journal of Political Science* 3(4):327–352.

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsay and Francesco Squintani. 2019. "Dispute Resolution Institutions and Strategic Militarization." *Journal of Political Economy* 127(1):378–418.

Moselle, Boaz and Benjamin Polak. 2001. "A model of a predatory state." *Journal of Law, Economics, and Organization* 17(1):1–33.

Myerson, Roger B and Mark A Satterthwaite. 1983. "Efficient mechanisms for bilateral trading." *Journal of economic theory* 29(2):265–281.

North, D.C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

North, Douglass C, John Joseph Wallis and Barry R Weingast. 2009. *Violence and social orders: a conceptual framework for interpreting recorded human history*. Cambridge University Press.

Olson, Mancur. 1993. "Dictatorship, democracy, and development." *American Political Science Review* pp. 567–576.

Olson, Mancur. 2000. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. Basic books.

Oye, Kenneth A. 1986. *Cooperation under anarchy*. Princeton University Press.

Piccione, Michele and Ariel Rubinstein. 2007. "Equilibrium in the Jungle*." *The Economic Journal* 117(522):883–896.

Powell, Robert. 1993. "Guns, butter, and anarchy." *American Political Science Review* 87(1):115–132.

Ruggie, John Gerard. 1982. "International regimes, transactions, and change: embedded liberalism in the postwar economic order." *International organization* 36(2):379–415.

Ruggie, John Gerard. 1983. "Continuity and transformation in the world polity: Toward a neorealist synthesis." *World Politics* 35(2):261–285.

Skaperdas, Stergios. 1992. "Cooperation, conflict, and power in the absence of property rights." *The American Economic Review* pp. 720–739.

Skaperdas, Stergios. 1998. "On the Formation of Alliances in Conflict and Contests." *Public Choice* 96(1-2):25–42.

Spruyt, Hendrik. 1994. "Institutional selection in international relations: state anarchy as order." *International Organization* 48(4):527–557.

Waltz, Kenneth N. 1979. *Theory of international politics*. Waveland Press.

Wendt, A. 1992. "Anarchy is what states make of it: The social construction of power politics." *International Organization* 46(2):391–425.

Wendt, Alexander E. 1987. "The agent-structure problem in international relations theory." *International organization* 41(3):335–370.

A Proofs

Contents

A.1 Preliminaries

Before proving individual results listed in the text, this section establishes the basic mathematical structure of the problem. In this set of results, we relax the condition on the wealth constraints imposed in the text, allowing for any $y_i, y_j \in (0, 1)$ such that $y_i + y_j = 1$. When a result only holds under a particular restriction on the budget constraints, we state this explicitly.

Throughout this section, we will work with functions $\psi_i : [0, y_j] \to \mathbb{R}_+$, each defined as

$$\psi_i(m_j) = \mathrm{RV}_i(m_j) + m_j.$$

We define $\Psi : [0, y_1] \times [0, y_2] \to \mathbb{R}_+$ as $\Psi(m_1, m_2) = \psi_1(m_2) + \psi_2(m_1)$, so that our necessary condition for a peaceful equilibrium with investments (m_1, m_2), Equation 2, is equivalent to $\Psi(m_1, m_2) \leq 1$. Our main goal in this preliminary section is to prove that Ψ is convex (and hence so are its lower contour sets) and to show that it attains a minimum no greater than 1.

A.1.1 Contest Success Function Derivatives

We begin with two basic results on the partial derivatives of the contest success function defined in Equation 1.

Lemma 1. *For all $m_i > 0$,*

$$\frac{\partial p_i(m_i, m_j)}{\partial m_i} = \frac{p_i(m_i, m_j)(1 - p_i(m_i, m_j))}{m_i}.$$

Proof. Immediate from computing

$$\frac{\partial p_i(m_i, m_j)}{\partial m_i} = \frac{\theta_i(\theta_i m_i + \theta_j m_j) - \theta_i(\theta_i m_i)}{(\theta_i m_i + \theta_j m_j)^2} = \frac{\theta_i \theta_j m_j}{(\theta_i m_i + \theta_j m_j)^2}.$$

□

Corollary 1. *For all $m_i > 0$,*

$$\frac{\partial p_i(m_i, m_j)}{\partial m_i} = -\frac{m_j}{m_i} \frac{\partial p_i(m_i, m_j)}{\partial m_j}.$$

Proof. Immediate from Lemma 1 if $m_j = 0$. Otherwise, by Lemma 1 and the fact that $p_j = 1 - p_i$, we have

$$\frac{\partial p_i(m_i, m_j)}{\partial m_j} = -\frac{\partial p_j(m_i, m_j)}{\partial m_j} = -\frac{m_i}{m_j} \frac{\partial p_i(m_i, m_j)}{\partial m_i}.$$

□

A.1.2 Reservation Value and Best Responses

In this section, we further characterize the reservation value function,

$$\mathrm{RV}_i(m_j) = \sup_{m_i \in [0, y_i]} W_i(m_i, m_j).$$

The following lemma demonstrates that a unique best response exists for all $m_j > 0$ and derives the expression for this maximizer.

Lemma 2. *For all $m_j > 0$, the maximization problem*

$$\max_{m_i \in [0, y_i]} W_i(m_i, m_j)$$

has a unique solution $\mu_i(m_j)$, where

$$\mu_i(m_j) = \min \left\{ y_i, \frac{\sqrt{(\theta_j m_j)^2 + \theta_i \theta_j m_j (1 - m_j)} - \theta_j m_j}{\theta_i} \right\}.$$

Proof. For any $m_j > 0$, W_i is differentiable in m_i, with

$$\begin{aligned}\frac{\partial W_i(m_i, m_j)}{\partial m_i} &= \gamma \left[\frac{\partial p_i(m_i, m_j)}{\partial m_i} [1 - m_i - m_j] - p_i(m_i, m_j) \right] \\ &= \gamma \left[\frac{\theta_i \theta_j m_j}{(\theta_i m_i + \theta_j m_j)^2} [1 - m_i - m_j] - \frac{\theta_i m_i}{\theta_i m_i + \theta_j m_j} \right].\end{aligned}$$

This expression is strictly decreasing in m_i, so W_i is strictly concave and thus has a unique maximizer. The necessary and sufficient condition for unconstrained maximization of W_i with respect to m_i is

$$\theta_i m_i^2 + 2\theta_j m_j m_i + \theta_j m_j(m_j - 1) = 0.$$

The quadratic formula then implies

$$m_i = \frac{\sqrt{(\theta_j m_j)^2 + \theta_i\theta_j m_j(1 - m_j)} - \theta_j m_j}{\theta_i}.$$

where we select the positive root to satisfy the constraint $m_i \geq 0$. The claimed definition of the constrained maximizer $\mu_i(m_j)$ then follows from the concavity of W_i in m_i. $\square$

It will prove helpful to observe that a player's reservation value can be written as an affine function of their best response and the other player's investment whenever the wealth constraint does not bind.

Lemma 3. *If $\mu_i(m_j) < y_i$, then*

$$\mathrm{RV}_i(m_j) = \gamma[1 - 2\mu_i(m_j) - m_j].$$

Proof. Take any $m_j > 0$ such that $\mu_i(m_j) < y_i$. The first-order conditions for maximization of W_i imply

$$\frac{\partial p_i(\mu_i(m_j), m_j)}{\partial m_i}[1 - \mu_i(m_j) - m_j] = p_i(\mu_i(m_j), m_j).$$

Rearranging terms and applying Lemma 1, this implies

$$1 - \mu_i(m_j) - m_j = \frac{p_i(\mu_i(m_j), m_j)}{\partial p_i(\mu_i(m_j), m_j)/\partial m_i} = \frac{\mu_i(m_j)}{1 - p_i(\mu_i(m_j), m_j)}.$$

Consequently, we have

$$\begin{aligned}
\mathrm{RV}_i(m_j) &= \gamma p_i(\mu_i(m_j), m_j)[1 - \mu_i(m_j) - m_j] \\
&= \frac{\gamma p_i(\mu_i(m_j), m_j)\mu_i(m_j)}{1 - p_i(\mu_i(m_j), m_j)} \\
&= \frac{\gamma\theta_i\mu_i(m_j)^2}{\theta_j m_j} \\
&= \frac{\gamma\theta_i\left[\sqrt{\theta_j m_j}(\sqrt{\theta_j m_j + \theta_i(1 - m_j)} - \sqrt{\theta_j m_j})\right]^2}{\theta_i^2\theta_j m_j} \\
&= \gamma\left[1 + 2\left(\frac{\theta_j m_j - \sqrt{(\theta_j m_j)^2 + \theta_i\theta_j m_j(1 - m_j)}}{\theta_i}\right) - m_j\right] \\
&= \gamma[1 - 2\mu_i(m_j) - m_j].
\end{aligned}$$

$\square$

It will also prove helpful to observe that the best response function is strictly concave in m_j whenever the wealth constraint does not bind.

Lemma 4. *If $\mu_i(m_j) < y_i$, then μ_i is strictly concave in a neighborhood of m_j.*

Proof. At any m_j such that $\mu_i(m_j) < y_j$, we have

$$\mu_i'(m_j) = \frac{1}{\theta_i}\left[\frac{2(\theta_j^2 - \theta_i\theta_j)m_j + \theta_i\theta_j}{2\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j}} - \theta_j\right]. \tag{5}$$

The sign of $\mu_i''(m_j)$ is therefore equal to that of

$$\begin{aligned} 2(\theta_j^2 - \theta_i\theta_j)&\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j} - \frac{(2(\theta_j^2 - \theta_i\theta_j)m_j + \theta_i\theta_j)^2}{2\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j}} \\ &= \frac{-\theta_i^2\theta_j^2}{2\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j}} \\ &< 0. \end{aligned}$$

□

A straightforward corollary of the two prior results is that RV_i is strictly convex whenever the wealth constraint does not bind. This will prove important to proving the convexity of the set of investment profiles that meet the necessary condition for equilibrium.

Corollary 2. *If $\mu_i(m_j) < y_i$, then RV_i is strictly convex in a neighborhood of m_j.*

Proof. Immediate from Lemma 3 and Lemma 4. □

Finally, we show that the marginal decrease in the reservation is arbitrarily large for small values of m_j.

Lemma 5. $\lim_{m_j \to 0^+} \mathrm{RV}_i'(m_j) = -\infty$.

Proof. We begin by observing that

$$\lim_{m_j \to 0^+} \frac{\sqrt{(\theta_j m_j)^2 + \theta_i\theta_j m_j(1 - m_j)} - \theta_j m_j}{\theta_i} = 0,$$

so $\mu_j(m_j) < y_j$ for sufficiently small m_j. Therefore, by Lemma 2, Lemma 3, and Equation 5,

$$\begin{aligned} \lim_{m_j \to 0^+} \mathrm{RV}_i'(m_j) &= \lim_{m_j \to 0^+} \left[-2\gamma\mu_i'(m_j) - \gamma\right] \\ &= \frac{\theta_j}{\theta_i} - \gamma - \frac{2\gamma}{\theta_i} \underbrace{\lim_{m_j \to 0^+} \frac{2(\theta_j^2 - \theta_i\theta_j)m_j + \theta_i\theta_j}{2\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j}}}_{\to +\infty} \\ &= -\infty. \end{aligned}$$

□

A.1.3 Upper Bound on Best Response

We now characterize an upper bound on each player's best response function. The following result shows that the boundary conditions on players' wealth never bind best responses when the condition of footnote 4 is met.

Lemma 6. *For all $m_j > 0$, $\mu_i(m_j) \leq \bar{\mu}_i < 1/2$, where*

$$\bar{\mu}_i = \frac{\sqrt{\theta_j}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})}.$$

Proof. Let $\tilde{\mu}_i : [0, 1] \to \mathbb{R}_+$ denote the unconstrained value of μ_i,

$$\tilde{\mu}_i(m_j) = \frac{\sqrt{(\theta_j m_j)^2 + \theta_i \theta_j m_j (1 - m_j)} - \theta_j m_j}{\theta_i}.$$

By Lemma 4, a necessary and sufficient condition for maximization of $\tilde{\mu}_i$ is $\tilde{\mu}_i'(m_j) = 0$. Applying Equation 5, this is equivalent to

$$2(\theta_j - \theta_i)m_j + \theta_i = 2\sqrt{(\theta_j m_j)^2 + \theta_i \theta_j m_j (1 - m_j)}.$$

Squaring both sides, rearranging terms, and applying the quadratic theorem yields

$$m_j = \frac{\sqrt{\theta_i}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})},$$

so this is the unique maximizer of $\tilde{\mu}_i$. As $\mu_i \leq \tilde{\mu}_i$ by construction, for all $m_j \in [0, y_j]$ we have

$$\begin{aligned}
\mu_i(m_j) &\leq \tilde{\mu}_i(m_j) \\
&\leq \tilde{\mu}_i \left(\frac{\sqrt{\theta_i}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})} \right) \\
&= \frac{1}{\theta_i} \left[\sqrt{\frac{(\theta_j^2 - \theta_i \theta_j)\theta_i}{4(\sqrt{\theta_i} + \sqrt{\theta_j})^2} + \frac{\theta_i \theta_j \sqrt{\theta_i}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})}} - \frac{\theta_j \sqrt{\theta_i}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})} \right] \\
&= \frac{\sqrt{\theta_j}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})} \\
&= \bar{\mu}_i.
\end{aligned}$$

□

A.1.4 Convexity and Minimization of Ψ

We conclude the preliminaries by characterizing the minimizer of each ψ_i when each $y_i > \bar{\mu}_i$. This will allow us to show that there exist m_1, m_2 such that $\Psi(m_1, m_2) \leq 1$. We begin with an observation on the convexity of ψ_i when boundary constraints do not bind.

Corollary 3. *If $\mu_i(m_j) < y_i$, then ψ_i is strictly convex in a neighborhood of m_j.*

Proof. Immediate from Corollary 2. $\square$

This allows us to use first-order conditions to characterize the minimizer of ψ_i, as in the following result.

Lemma 7. *If $y_i > \bar{\mu}_i$, the minimization problem*

$$\min_{m_j \in [0, y_j]} \psi_i(m_j)$$

has a unique solution ρ_j, where

$$\rho_j = \min\left\{ y_j, \frac{2\gamma^2\theta_i\theta_j}{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j + [(1-\gamma)\theta_i + 2\gamma\theta_j]\sqrt{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j}} \right\}.$$

Proof. By Corollary 3, the necessary and sufficient condition for unconstrained minimization of ψ_i is $\psi_i'(m_j) = 0$. Consequently, to prove the claim, it suffices to show that $\psi_i'(\rho_j) = 0$ when $\rho_j < y_j$. Letting $g(m_j) = (\theta_j - \theta_i) + \theta_i/m_j$, it follows from Lemma 3 and Equation 5 that

$$\begin{aligned}
\psi_i'(m_j) &= \mathrm{RV}_i'(m_j) + 1 \\
&= 1 - \gamma - 2\gamma\mu_i'(m_j) \\
&= 1 - \gamma - \frac{2\gamma}{\theta_i}\left[\frac{2(\theta_j^2 - \theta_i\theta_j)m_j + \theta_i\theta_j}{2\sqrt{(\theta_j^2 - \theta_i\theta_j)m_j^2 + \theta_i\theta_j m_j}} - \theta_j\right] \\
&= 1 - \gamma + \frac{2\gamma\theta_j}{\theta_i} - \frac{\gamma\sqrt{\theta_j}}{\theta_i}\left[\sqrt{g(m_j)} + \frac{\theta_j - \theta_i}{\sqrt{g(m_j)}}\right]
\end{aligned}$$

for all m_j such that $\mu_i(m_j) < y_j$. If $\rho_j < y_j$, we have

$$\begin{aligned}
g(\rho_j) &= \frac{2\gamma^2\theta_j(\theta_j - \theta_i) + (1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j + [(1-\gamma)\theta_i + 2\gamma\theta_j]\sqrt{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j}}{2\gamma^2\theta_j} \\
&= \left(\frac{(1-\gamma)\theta_i + 2\gamma\theta_j + \sqrt{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j}}{2\gamma\sqrt{\theta_j}}\right)^2,
\end{aligned}$$

and therefore

$$\begin{aligned}
\sqrt{g(\rho_j)} + \frac{\theta_j - \theta_i}{\sqrt{g(\rho_j)}} &= \frac{\left[(1-\gamma)\theta_i + 2\gamma\theta_j + \sqrt{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j}\right]^2 + 4\gamma^2(\theta_j - \theta_i)\theta_j}{2\gamma\sqrt{\theta_j}\left[(1-\gamma)\theta_i + 2\gamma\theta_j + \sqrt{(1-\gamma)^2\theta_i^2 + 4\gamma\theta_i\theta_j}\right]} \\
&= \frac{(1-\gamma)\theta_i + 2\gamma\theta_j}{\gamma\sqrt{\theta_j}}.
\end{aligned}$$

Consequently, if $\rho_j < y_j$,

$$\begin{aligned}
\psi_i'(\rho_j) &= 1 - \gamma + \frac{2\gamma\theta_j}{\theta_i} - \frac{\gamma\sqrt{\theta_j}}{\theta_i}\left[\sqrt{g(\rho_j)} + \frac{\theta_j - \theta_i}{\sqrt{g(\rho_j)}}\right] \\
&= 1 - \gamma + \frac{2\gamma\theta_j}{\theta_i} - \frac{(1-\gamma)\theta_i + 2\gamma\theta_j}{\theta_i} \\
&= 0.
\end{aligned}$$

□

We now provide a sufficient condition for the boundary condition on ρ_j never to bind. Specifically, we show that unconstrained minimization of ψ_i is feasible as long as each $y_i > \bar{\mu}_i$.

Lemma 8. *If $y_i > \bar{\mu}_i$, then $\rho_j < \bar{\mu}_j$.*

Proof. If $y_i > \bar{\mu}_i$, then it follows from the proof of Lemma 7 that ρ_j is defined by $\mathrm{RV}_i'(\rho_j) + 1 = 0$, which by Lemma 3 is equivalent to

$$\mu_i'(\rho_j) = \frac{1-\gamma}{2\gamma} > 0.$$

Recall from the proof of Lemma 6 that $\mu_i'(\bar{\mu}_j) = 0$. Therefore, as μ_i is strictly concave (per Lemma 4), $\rho_j < \bar{\mu}_j$. □

This brings us to the result that will be instrumental in establishing the existence of a peaceful equilibrium with distributed coercion. As long as each player's initial wealth is great enough that their budget constraint never binds, we can find a pair of feasible initial investments such that $\Psi(m_1, m_2) \leq 1$.

Lemma 9. *If each $y_i > \bar{\mu}_i$, then $\Psi(\rho_1, \rho_2) \leq 1$.*

Proof. It is obvious that each RV_i (and thereby each ψ_i) is increasing in γ, so it will suffice to prove the claim for the limiting case $\gamma = 1$. Notice that each

$$\rho_i = \frac{2\theta_i\theta_j}{4\theta_i\theta_j + 2\theta_i\sqrt{4\theta_i\theta_j}} = \frac{\sqrt{\theta_j}}{2(\sqrt{\theta_i} + \sqrt{\theta_j})} = \bar{\mu}_i,$$

and recall from the proof of Lemma 6 that each $\mu_j(\bar{\mu}_i) = \bar{\mu}_j$. Therefore, by Lemma 3, we have

$$\begin{aligned}
\Psi(\rho_1, \rho_2) &= \mathrm{RV}_1(\rho_2) + \mathrm{RV}_2(\rho_1) + \rho_1 + \rho_2 \\
&= [1 - 2\mu_1(\rho_2) - \rho_2] + [1 - 2\mu_2(\rho_1) - \rho_1] + \rho_1 + \rho_2 \\
&= 2 - 2[\mu_1(\rho_2) + \mu_2(\rho_1)] \\
&= 2 - 2(\bar{\mu}_1 + \bar{\mu}_2) \\
&= 1.
\end{aligned}$$

□

The final result of the section succinctly states the existence and convexity of a set of feasible investments meeting the necessary condition for an equilibrium, so long as each player's initial wealth is great enough.

Corollary 4. *If each $y_i > \bar{\mu}_i$, then $\{(m_1, m_2) \in [0, y_1] \times [0, y_2] : \Psi(m_1, m_2) \leq 1\}$ is nonempty and convex.*

Proof. Immediate from Corollary 3 and Lemma 9. □

A.2 Proof of Proposition 2

We begin by proving that for any pair of investments meeting the necessary condition, Equation 2, we can construct an institutional mechanism that supports a peaceful equilibrium. This will also be useful in the proof of Proposition 4 below.

Lemma 10. *If (m_1', m_2') satisfies Equation 2 and each $m_i' \leq y_i$, then there exists an institution that supports a peaceful equilibrium with these investments.*

Proof. Assume (m_1', m_2') satisfies Equation 2, with each $m_i' \leq y_i$, and define V_1 as follows:[16]

$$V_1(m_1, m_2) = \begin{cases} \mathrm{RV}_1(m_2') & m_1 \leq m_1', m_2 = m_2', \\ \min\{\mathrm{RV}_1(m_2'), 1 - m_1 - m_2\} & m_1 > m_1', m_2 = m_2', \\ \max\{\mathrm{RV}_1(m_2') + (m_2' - m_2), 0\} & m_1 = m_1', m_2 \neq m_2', \\ (1 - m_1 - m_2)/2 & m_1 \neq m_1', m_2 \neq m_2'. \end{cases} \tag{6}$$

Given this institution, we claim that the following strategy profile constitutes a peaceful equilibrium:

- In the investment stage, each player invests $m_i = m_i'$.
- In the institutional stage, following investments (m_1, m_2), each player opts for the institution ($w_i = 0$) if and only if $V_i(m_i, m_j) \geq W_i(m_i, m_j)$.

The strategies in the institutional stage are mutual best responses by construction, so we need only check that the initial coercive investments are optimal. There is no incentive for either player to force a conflict, as $V_1(m_1', m_2') = \mathrm{RV}_1(m_2')$ and, by Equation 2,

$$\begin{aligned} V_2(m_1', m_2') &= 1 - m_1' - m_2' - V_1(m_1', m_2') \\ &= 1 - m_1' - m_2' - \mathrm{RV}_1(m_2') \\ &\geq \mathrm{RV}_2(m_1'). \end{aligned}$$

Therefore, we must only check that neither player can receive more from the institution by deviating. It is obvious from Equation 6 that player 1 cannot benefit by choosing $m_1 \neq m_1'$. For player 2, notice that choosing $m_2 < \mathrm{RV}_1(m_2') + m_2'$ results in

$$\begin{aligned} V_2(m_1', m_2) &= 1 - m_1' - m_2 - V_1(m_1', m_2) \\ &= 1 - m_1' - m_2' - \mathrm{RV}_1(m_2') \\ &= V_2(m_1', m_2'), \end{aligned}$$

[16]The specification of $V_1(m_1, m_2)$ when $m_1 \neq m_1'$ and $m_2 \neq m_2'$ is immaterial to the result, as we only need to check for unilateral deviations.

so no such deviation is profitable. Similarly, choosing $m_2 > \mathrm{RV}_1(m_2') + m_2'$ yields

$$\begin{aligned} V_2(m_1', m_2) &= 1 - m_1' - m_2 \\ &< 1 - m_1' - m_2' - \mathrm{RV}_1(m_2') \\ &= V_2(m_1', m_2'). \end{aligned}$$

Therefore, the proposed strategy profile is an equilibrium. Finally, observe that peace prevails on the equilibrium path, as each $V_i(m_i', m_j') \geq \mathrm{RV}_i(m_j') \geq W_i(m_i', m_j')$. □

Throughout the remainder of this section, we write $W_i(m_i, m_j; \gamma)$, $\mathrm{RV}_i(m_j; \gamma)$, $\psi_i(m_j; \gamma)$, and $\rho_j(\gamma)$ to make explicit the dependence of these quantities on γ. As a precursor to proving Proposition 2, we have a helpful result showing that the minimized value of ψ_i increases with γ.

Lemma 11. *Assume $y_i > \bar{\mu}_i$. If $\gamma'' > \gamma'$, then $\psi_i(\rho_j(\gamma''); \gamma'') > \psi_i(\rho_j(\gamma'); \gamma')$.*

Proof. First note that for any fixed $m_j > 0$,

$$\frac{\partial \psi_i(m_j; \gamma)}{\partial \gamma} = \frac{\partial \,\mathrm{RV}_i(m_j; \gamma)}{\partial \gamma} = \frac{\partial W_i(\mu_i(m_j), m_j; \gamma)}{\partial \gamma} > 0,$$

where the final equality follows from the envelope theorem. Recall from Lemma 7 that $\psi_i(\rho_j(\gamma); \gamma) = \min_{m_j \in [0, y_j]} \psi_i(m_j; \gamma)$, so another application of the envelope theorem gives us

$$\frac{d\psi_i(\rho_j(\gamma); \gamma)}{d\gamma} = \frac{\partial \psi_i(\rho_j(\gamma); \gamma)}{\partial \gamma} > 0. \qquad \square$$

With this result in hand, we can prove the proposition.

Proposition 2. *A monopoly of violence by player i can sustain peace if and only if $\gamma \leq \hat{\gamma}_i$, where $\frac{1}{2} < \hat{\gamma}_2 \leq \hat{\gamma}_1 < 1$.*

Proof. We begin with the necessity claim. By Equation 2 and Lemma 7, a necessary condition for there to exist a peaceful monopoly of violence by player i is

$$\psi_i(0; \gamma) + \psi_j(\rho_i(\gamma); \gamma) \leq 1. \tag{7}$$

Because $\psi_i(0; \gamma) = \mathrm{RV}_i(0; \gamma) = \gamma$, this condition is equivalent to $\psi_j(\rho_i(\gamma); \gamma) \leq 1 - \gamma$. The LHS of this condition is strictly increasing in γ per Lemma 11 and the RHS is strictly decreasing in γ, so there is a unique point $\hat{\gamma}_i$ at which it holds with equality. To prove that $\hat{\gamma}_i < 1$, observe that

$$\psi_i(0; 1) + \psi_j(\rho_i(1); 1) > \psi_i(\rho_j(1); 1) + \psi_j(\rho_i(1); 1) = 1$$

by Lemma 7 and the proof of Lemma 9. To prove that $\hat{\gamma}_i > 1/2$, observe that

$$\psi_i(0; 1/2) + \psi_j(\rho_i(1/2); 1/2) < \psi_i(0; 1/2) + \psi_j(0; 1/2) = 1,$$

again by Lemma 7. To prove that $\hat{\gamma}_2 \leq \hat{\gamma}_1$, observe that $\theta_1 \geq \theta_2$ implies $\mathrm{RV}_1 \geq \mathrm{RV}_2$ and thus $\psi_1 \geq \psi_2$. Combining Lemma 7 with the definition of $\hat{\gamma}_2$, we have

$$\psi_2(\rho_1(\hat{\gamma}_2); \hat{\gamma}_2) \leq \psi_2(\rho_2(\hat{\gamma}_2); \hat{\gamma}_2) \leq \psi_1(\rho_2(\hat{\gamma}_2); \hat{\gamma}_2) = 1 - \hat{\gamma}_2,$$

so $\hat{\gamma}_1 \geq \hat{\gamma}_2$. Finally, the sufficiency claim follows from Lemma 10. □

A.3 Proof of Proposition 3

Proposition 3. *If $\gamma < \hat{\gamma}_i$, then the peaceful monopoly of violence that maximizes player i's payoff entails more investment than the socially efficient level.*

Proof. Assume $\gamma < \hat{\gamma}_i$, so there exists a range of coercive investments that can support a peaceful monopoly of violence by player i (Proposition 2). The monopolist's greatest possible payoff from such an equilibrium is

$$V_i(m_i, 0) = 1 - m_i - \mathrm{RV}_j(m_i) = 1 - \psi_j(m_i).$$

By Lemma 10, the monopoly of violence that is best for the monopolist can be identified as the solution to the constrained maximization problem

$$\begin{aligned} \max_{m_i} \quad & 1 - \psi_j(m_i) \\ \text{s.t.} \quad & \psi_j(m_i) \leq 1 - \gamma, \\ & 0 \leq m_i \leq y_i. \end{aligned}$$

Clearly, as this is equivalent to minimization of ψ_j, the solution to this problem is $m_i = \rho_i$ (Lemma 7, Lemma 8). However, $\gamma < \hat{\gamma}_i$ implies $\rho_i > \min\{m_i \in [0, y_i] : \psi_j(m_i) \leq 1 - \gamma\}$ (see the proof of Proposition 2). □

A.4 Proof of Proposition 4

Proposition 4. *For all $\gamma < 1$, distributed coercion can sustain peace.*

Proof. If each player invests $m_i = \rho_i$, then Equation 2 is satisfied, per Lemma 9. Each $\rho_i < y_i$ per Lemma 8, so the result follows from Lemma 10. □

A.5 Proof of Proposition 5

Proposition 5. *If a monopoly of violence can sustain peace, then distributed coercion can sustain peace with a strictly lower total investment in coercion.*

Proof. Lemma 10 allows us to identify an efficient peaceful equilibrium with a solution to the constrained maximization problem

$$\begin{aligned}\max_{m_1,m_2} \quad & 1 - m_1 - m_2 \\ \text{s.t.} \quad & \Psi(m_1, m_2) \leq 1, \\ & 0 \leq m_1 \leq y_1, \\ & 0 \leq m_2 \leq y_2.\end{aligned}$$

This is a concave maximization problem with a nonempty and convex constraint set (Corollary 4), so first-order conditions are necessary and sufficient for maximization. Because each $\lim_{m_i \to 0^+} \partial\Psi(m_i, m_j)/\partial m_i = -\infty$ (Lemma 5), any solution to this problem must entail $m_1 > 0$ and $m_2 > 0$. □

A.6 Proof of Proposition 6

Proposition 6. *If a monopoly of violence can sustain peace, then distributed coercion can sustain peace with a strictly greater total investment in coercion.*

Proof. The proof mirrors that of Proposition 5. Lemma 10 allows us to identify the *least* efficient peaceful equilibrium with a solution to the constrained maximization problem

$$\begin{aligned}\max_{m_1,m_2} \quad & m_1 + m_2 \\ \text{s.t.} \quad & \Psi(m_1, m_2) \leq 1, \\ & 0 \leq m_1 \leq y_1, \\ & 0 \leq m_2 \leq y_2.\end{aligned}$$

Once again, this is a concave problem with a convex constraint set, so the first-order conditions are necessary and sufficient for a solution. Because each $\lim_{m_i \to 0^+} \partial\Psi(m_i, m_j)/\partial m_i = -\infty$ (Lemma 5), any solution must entail $m_1 > 0$ and $m_2 > 0$. □

A.7 Proof of Proposition 7

Proposition 7. *If $\gamma < \hat{\gamma}_i$, then player i can obtain a greater payoff from a monopoly of violence with inefficiently high coercive investment than from any other institution that sustains peace.*

Proof. The proof largely follows that of Proposition 5. Assume $\gamma \leq \hat{\gamma}_i$, so there exists a coercive investment that can support a peaceful monopoly of violence by player i (Proposition 2). The monopolist's greatest possible payoff from any peaceful equilibrium is

$$V_i(m_i, m_j) = 1 - m_i - m_j - \mathrm{RV}_j(m_i) = 1 - \psi_j(m_i) - m_j.$$

By Lemma 10, the coercive investments supporting the peaceful equilibrium that is best for player i can be identified as the solution to the constrained maximization problem

$$\begin{aligned}\max_{m_i,m_j} \quad & 1-\psi_j(m_i)-m_j \\ \text{s.t.} \quad & \Psi(m_i,m_j)\le 1, \\ & 0\le m_i\le y_i, \\ & 0\le m_j\le y_j.\end{aligned}$$

As $m_i=\rho_i$ minimizes ψ_j (Lemma 7) and $\Psi(\rho_i,0)\le 1$ (see the proof of Proposition 2), the solution to this problem is $(m_i,m_j)=(\rho_i,0)$. That this entails more coercive investment than is socially efficient follows from Proposition 5. □

A.8 Proof of Proposition 8

Here we consider the model with economies of scale in the production of coercion, in which the contest success function is defined by Equation 4. Let the reservation value functions RV_i, best response functions μ_i, and equilibrium criterion functions ψ_i be defined as in the baseline model. We begin with a result on the strict quasiconcavity of the conflict payoffs.

Lemma 12. *In the model with economies of scale, W_i is strictly quasiconcave in m_i.*

Proof. It suffices to prove that p_i is strictly log-concave in m_i, as this implies W_i is strictly log-concave (and thus strictly quasiconcave) in m_i as well. Observe that

$$\frac{\partial p_i(m_i,m_j)}{\partial m_i}=\frac{\theta_i\exp(\theta_i m_i)\exp(\theta_j m_j)}{[\exp(\theta_i m_i)+\exp(\theta_j m_j)]^2}=\theta_i p_i(m_i,m_j)p_j(m_j,m_i).$$

Consequently, we have

$$\frac{\partial \log p_i(m_i,m_j)}{\partial m_i}=\frac{\partial p_i(m_i,m_j)/\partial m_i}{p_i(m_i,m_j)}=\theta_i p_j(m_j,m_i),$$

which is strictly decreasing in m_i, proving strict log-concavity of p_i in m_i. □

This allows us to prove that the budget constraint will not bind in the symmetric game with economies of scale. Under symmetry, each player's best response function is identical: $\mu_1=\mu_2=\mu$.

Lemma 13. *In the symmetric model with economies of scale, $\mu<1/2$.*

Proof. Notice that

$$\begin{aligned}\frac{\partial W_i(m_i,m_j)}{\partial m_i}&=\gamma\left[\frac{\partial p_i(m_i,m_j)}{\partial m_i}[1-m_i-m_j]-p_i(m_i,m_j)\right]\\&=\gamma p_i(m_i,m_j)\left[\theta p_j(m_j,m_i)[1-m_i-m_j]-1\right].\end{aligned}$$

Consequently, by Lemma 12, $\mu(m_j)\ge 1/2$ if and only if

$$p_j(m_j,1/2)\left[\frac{1}{2}-m_j\right]\ge\frac{1}{\theta}.$$

As $\mu(1/2)$ is the maximizer of the above expression with respect to m_j, a necessary condition for the above inequality to hold is

$$p_j(\mu(1/2), 1/2)\left[\frac{1}{2} - \mu(1/2)\right] \geq \frac{1}{\theta}. \tag{8}$$

Equation 8 obviously fails if $\mu(1/2) \geq 1/2$, as then the LHS is negative. But if $\mu(1/2) < 1/2$, then Lemma 12 implies

$$p_j(1/2, \mu(1/2))\left[\frac{1}{2} - \mu(1/2)\right] = \frac{1}{\theta},$$

which in turn implies

$$p_j(\mu_j(1/2), 1/2)\left[\frac{1}{2} - \mu_j(1/2)\right] < p_i(1/2, \mu_j(1/2))\left[\frac{1}{2} - \mu_j(1/2)\right] = \frac{1}{\theta}.$$

Therefore, $\mu(m_j) < 1/2$ for all m_j. □

We next prove a helpful lemma showing that a player's probability of victory if she best-responds is strictly decreasing in the other player's coercive investment. In what follows, let $p(m, n) = \exp(\theta m)/(\exp(\theta m) + \exp(\theta n))$.

Lemma 14. *In the symmetric model with economies of scale, if $\mu(m) > 0$, then $p(\mu(m), m)$ is strictly decreasing in a neighborhood of m.*

Proof. The implicit function theorem guarantees the existence of $\mu'(m)$. The claim obviously holds if $\mu'(m) < 0$. So to complete the proof, assume $\mu'(m) \geq 0$. Since $\mu(m) > 0$, the first-order condition

$$p(m, \mu(m))\left[1 - m - \mu(m)\right] = \frac{1}{\theta} \tag{9}$$

holds in a neighborhood of m, per Lemma 13. Because $1 - m - \mu(m)$ is strictly decreasing in a neighborhood of m, $p(m, \mu(m))$ must be strictly increasing, which implies $p(\mu(m), m)$ is strictly decreasing. □

As with our baseline model, the convexity of the ψ_i functions is important to prove Proposition 8. In the symmetric case, $\mathrm{RV}_i = \mathrm{RV}_j$, and we therefore have $\psi_i = \psi_j = \psi$.

Lemma 15. *In the symmetric model with economies of scale, ψ is strictly convex.*

Proof. It will suffice to prove that RV is strictly convex. By the envelope theorem,

$$\begin{aligned}\mathrm{RV}'(m) &= \frac{\partial W_i(\mu(m), m)}{\partial m_j} \\ &= -\gamma p(\mu(m), m)\left[\theta p(m, \mu(m))[1 - m - \mu(m)] + 1\right].\end{aligned}$$

If $\mu(m) > 0$, then in a neighborhood of m, Equation 9 must hold (per Lemma 13), and thus

$$\mathrm{RV}'(m) = -2\gamma p(\mu(m), m).$$

By Lemma 14, this expression is strictly increasing in a neighborhood of m, so RV is strictly convex in that neighborhood. Alternatively, suppose $\mu(m) = 0$ in a neighborhood of m. Then we have

$$\begin{aligned}\mathrm{RV}'(m) &= -\gamma p(0,m)\left[\theta p(m,0)[1-m]+1\right]\\ &= -\underbrace{\frac{\gamma}{1+\exp(\theta m)}}_{(a)}\underbrace{\left[\frac{\theta\exp(\theta m)}{1+\exp(\theta m)}[1-m]+1\right]}_{(b)}\end{aligned}$$

Term (a) is obviously strictly decreasing in m. For term (b), recall that the first-order condition for $\mu(m) = 0$ is $\theta p(m,0)[1-m] \leq 1$, via Equation 9. We therefore have

$$\begin{aligned}\frac{d}{dm}\left[\frac{\exp(\theta m)}{1+\exp(\theta m)}[1-m]\right] &= \frac{\theta\exp(\theta m)}{[1+\exp(\theta m)]^2}[1-m] - \frac{\exp(\theta m)}{1+\exp(\theta m)}\\ &= p(m,0)\left[\theta p(0,m)[1-m]-1\right]\\ &< p(m,0)\left[\theta p(m,0)[1-m]-1\right]\\ &\leq 0,\end{aligned}$$

so term (b) is strictly decreasing in m as well. Altogether this proves $\mathrm{RV}'(m)$ is strictly increasing in a neighborhood of m, so RV is strictly convex. □

The inefficiency of uneven coercive investments in the symmetric model with economies of scale is almost immediate from the previous lemma.

Proposition 8. *Assume $y_1 = y_2 = \frac{1}{2}$ and $\theta_1 = \theta_2$ in the model with economies of scale. For any institution that sustains peace with $m_1^* \neq m_2^*$, there is an institution that sustains peace with strictly less total investment in which $m_1^* = m_2^*$.*

Proof. There is a peaceful equilibrium with $m_1 \neq m_2$ only if $\psi(m_1) + \psi(m_2) \leq 1$. As ψ is strictly convex, per Lemma 15, this implies $2\psi(m-\epsilon) \leq 1$ for sufficiently small $\epsilon > 0$, where $m = (m_1 + m_2)/2$. Therefore, by the same line of argument as in the proof of Proposition 2, there exists a peaceful equilibrium in which each player invests $m - \epsilon$. □

A.9 Proof of Proposition 9

For an extension of this argument to an environment with two-sided private information, see subsection A.12.

In the following proof, we let $W_i(m_i, m_j, \theta_2)$ denote player i's payoff from conflict with the given investments when player 2's type is θ_2

Proposition 9. *In the private information model, if $\gamma > \frac{1}{2}$ and $\Pr(\theta_2 = 1)$ is sufficiently low, then any peaceful equilibrium is* ex ante *inefficient.*

Proof. We begin by showing that a conditional monopoly of violence is sustainable as an equilibrium and that the inefficiency of such an equilibrium vanishes as $\pi \to 0$. Specifically, we claim that the following institution and strategies constitute an equilibrium:

- Player 1 invests $m_1^* = (\frac{\pi}{1+\pi})^2$.
- Player 2 invests $m_{20}^* = 0$ if $\theta_2 = 0$ and invests $m_{21}^* = \frac{\pi}{(1+\pi)^2}$ if $\theta_2 = 1$.
- The institution is defined by

$$V_1(m_1, m_2) = \begin{cases} 1 - m_1^* - m_2 & m_1 < m_1^*, \\ 1 - m_1 - m_2 & m_1 \geq m_1^*. \end{cases}$$

- After observing an investment m_2, player 1's belief that $\theta_2 = 1$ is

$$\lambda(m_2) = \begin{cases} 1 & m_2 = m_{21}^*, \\ 0 & m_2 \neq m_{21}^*. \end{cases}$$

- After investments (m_1, m_2):
  - Player 1 selects $w_1 = 0$ if and only if $V_1(m_1, m_2) \geq E_{\lambda(m_2)}[W_1(m_1, m_2, \theta_2)]$.
  - If $\theta_2 = 0$, player 2 selects $w_2 = 0$.
  - If $\theta_2 = 1$, player 2 selects $w_2 = 0$ if and only if $V_2(m_1, m_2) \geq W_2(m_1, m_2, 1)$.

Along the path of play, we have peace if $\theta_2 = 0$, as then

$$V_1(m_1^*, 0) = 1 - m_1^* \geq \gamma(1 - m_1^*) = W_1(m_1^*, 0, 0).$$

We have conflict if $\theta_2 = 1$, as then

$$V_2(m_1^*, m_{21}^*) = 0 < W_2(m_1^*, m_{21}^*, 1).$$

To show that this is an equilibrium, first note that the decisions of w_i are best responses by construction, and player 1's beliefs about θ_2 are consistent with Bayes' rule wherever possible. Moving up to the investment stage, there is clearly no deviation by player 1 that would give her a greater payoff from the institution. Let $\xi(m_1)$ denote her expected value from investing m_1 and forcing a conflict, so that

$$\xi(m_1) = \gamma \left[(1 - \pi)(1 - m_1) + \pi \frac{m_1}{m_1 + m_{21}^*} (1 - m_1 - m_{21}^*) \right].$$

To prove that she cannot raise her payoff from deviating to $m_1 \neq m_1^*$ and forcing a conflict, notice that $m_1^* + m_{21}^* = \frac{\pi}{1+\pi}$ and thus

$$\begin{aligned} \xi'(m_1^*) &\propto -(1 - \pi) + \pi \left[\frac{m_{21}^*}{(m_1^* + m_{21}^*)^2} (1 - m_1^* - m_{21}^*) - \frac{m_1^*}{m_1^* + m_{21}^*} \right] \\ &= \pi - 1 + (1 + \pi) \left(\frac{(1 + \pi)(1 - \pi)}{(1 + \pi)^2} \right) \\ &= 0. \end{aligned}$$

For player 2, neither type can unilaterally deviate to raise their payoff from the institution, as $V_2(m_1^*, m_2) = 0$ for all m_2. Additionally, if $\theta_2 = 0$, then player 2 obviously cannot raise her payoff by deviating to conflict. Finally, if $\theta_2 = 1$, notice that

$$\begin{aligned}\frac{\partial W_2(m_1^*, m_2^*, 1)}{\partial m_2} &\propto \frac{m_1^*}{(m_1^* + m_{21}^*)^2}(1 - m_1^* - m_{21}^*) - \frac{m_{21}^*}{m_1^* + m_{21}^*} \\ &= \underbrace{\left(\frac{\pi}{1+\pi}\right)}_{\frac{m_1^*}{m_1^*+m_{21}^*}} \underbrace{\left(\frac{1+\pi}{\pi}\right)}_{\frac{1}{m_1^*+m_{21}^*}} \underbrace{\left(\frac{1}{1+\pi}\right)}_{1-m_1^*-m_{21}^*} - \frac{1}{1+\pi} \\ &= 0,\end{aligned}$$

so this type cannot increase her payoff by deviating to $m_2 \neq m_{21}^*$ and forcing a conflict. Therefore, the strategy profile constitutes an equilibrium under the given institution. Notice that $m_1^* \to 0$ and $m_{21}^* \to 0$ as $\pi \to 0$, as does the equilibrium probability of conflict, confirming the efficiency claim.

Now, to prove the inefficiency of peaceful equilibria, we will show that we cannot construct a sequence of peaceful equilibria such that total investment vanishes as $\pi \to 0$. To this end, let $\{\pi^k\}_{k=1}^{\infty}$ be a sequence such that each $\pi^k > 0$ and $\lim_{k\to\infty} \pi^k = 0$. For each k, take a peaceful equilibrium with investments m_1^k, m_{20}^k (for player 2 when $\theta_2 = 0$), and m_{21}^k (for when $\theta_2 = 1$) supported by an institution V_1^k. In order for player 1 not to prefer to deviate to conflict, we must have

$$(1-\pi^k)V_1^k(m_1^k, m_{20}^k) + \pi^k V_1^k(m_1^k, m_{21}^k) \geq \max_{m_1} \left[\pi^k W_1(m_1, m_{21}^k, 1) + (1-\pi^k)\gamma(1 - m_1 - m_{20}^k)\right].$$

Assume $m_{20}^k \to 0$, as is required for efficiency. Because $\pi^k \to 0$, the above inequality then implies $\liminf_{k\to\infty} V_1^k(m_1^k, m_{20}^k) \geq \gamma$. Similarly, for type $\theta_2 = 1$ of player 2 not to prefer to deviate to conflict,

$$V_2^k(m_1^k, m_{21}^k) \geq \max_{m_2} W_2(m_1^k, m_2, 1) = \gamma\left[1 + m_1^k - 2\sqrt{m_1^k}\right],$$

where the equality follows from Lemma 2 and Lemma 3. Because the equilibrium is peaceful, incentive compatibility for player 2 requires $V_2^k(m_1^k, m_{20}^k) = V_2^k(m_1^k, m_{21}^k)$. Combined with player 2's no-deviation condition, this implies

$$V_1^k(m_1^k, m_{20}^k) + \gamma\left[1 + m_1^k - 2\sqrt{m_1^k}\right] \leq 1 - m_1^k - m_{20}^k.$$

Efficiency requires $m_1^k \to 0$. But then, because $V_1^k(m_1^k, m_{20}^k) \geq \gamma$ in the limit, the above may hold for sufficiently large k only if $\gamma \leq \frac{1}{2}$. □

A.10 N-Player Model

To support the claim of footnote 2, we extend the baseline model to incorporate N players.

A.10.1 Setup

Let $\mathcal{N} = \{1, \ldots, N\}$ denote the set of players in the N-player model. At the start of the game, each player simultaneously chooses a coercive investment, $m_i \in M_i = [0, y_i]$. The players then observe the vector of coercive investments, $m = (m_i)_{i \in \mathcal{N}}$. At this point, each player simultaneously opts for the institution or conflict. If every player opts for the institution, then each player receives $V_i(m)$, where $\sum_{j=1}^{N} V_j(m) = 1 - \sum_{j=1}^{N} m_j$. Otherwise, if any player opts for conflict, then each player receives her expected payoff from conflict,

$$W_i(m) = \gamma \times p_i(m) \times \left[1 - \sum_{j=1}^{N} m_j\right],$$

where the N-player contest success function p_i is the N-player analogue to Equation 1,

$$p_i(m) = \frac{\theta_i m_i}{\sum_{j=1}^{N} \theta_j m_j}. \tag{10}$$

For ease of exposition, we assume the players are *ex ante* identical, with $\theta_i = 1$ and $y_i = 1/N$ for all $i \in \mathcal{N}$.

Before moving on to the results of the extension, we briefly remark on two features of the N-player model. First is the requirement that every player opt for the institution in order for peace to prevail. This is in line with the voluntary agreements assumption in the literature on mechanism design and conflict, which restricts attention to equilibria in which no player receives less than her expected payoff from fighting (Fey and Ramsay 2011). Moreover, our focus on institutional arrangements that can garner unanimous consent is in line with our broader focus on the conditions that sustain peace.

The second assumption worth noting is the nature of the N-player contest. Some readers might wonder about the possibility of coalitions forming, thereby altering the conflict payoffs. Even with coalitions, the condition we characterize would remain necessary (though potentially insufficient) for a peaceful equilibrium. We show below that this necessary condition, at least, is harder to meet with a monopoly of violence than under other arrangements. Second, if one assumes that a coalition would divide its spoils in proportion to its members' coercive investments, then expected utilities from a conflict with coalitions would still be given by Equation 10 (Skaperdas 1998, 31).

A.10.2 Results

Define $W : \mathbb{R}_+^2 \to \mathbb{R}$ as follows:

$$\begin{aligned} W(m, T) &= \gamma \times \frac{m}{m + T} \times [1 - m - T] \\ &= \gamma m \left[\frac{1}{m + T} - 1\right]. \end{aligned}$$

Under our symmetry assumption, $W_i(m) = W(m_i, \sum_{j \neq i} m_j)$; i.e., a player's expected payoff from conflict depends only on her own investment and the total investment of all other players. We may thus define reservation values in the symmetric N-player game as

$$\mathrm{RV}(T) = \sup_{m \in [0,1/N]} W(m,T).$$

As in the baseline model, to prove our efficiency result it will be important to demonstrate that the reservation value function is strictly convex.

Lemma 16. *In the symmetric N-player game,* RV *is strictly convex.*

Proof. For any $T > 0$, $W(m,T)$ is strictly concave in m and therefore has a unique maximizer μ on $[0, 1/N]$. The envelope theorem then implies

$$\mathrm{RV}'(T) = \frac{\partial W(\mu, T)}{\partial T} = -\frac{\gamma \mu}{(\mu + T)^2}.$$

This expression is obviously strictly increasing in T, so RV is strictly convex. □

The inefficiency of any unbalanced scheme of coercive investments follows almost immediately from Lemma 16.

Proposition 10. *In the symmetric N-player game, for any peaceful equilibrium such that $m_i \neq m_j$, there is a peaceful equilibrium with strictly less total investment in which every $m_i = m$.*

Proof. Suppose there is a peaceful equilibrium with allocations m' such that $m'_i \neq m'_j$, and let $T' = \sum_{i=1}^N m'_i > 0$. Under the necessary conditions for an equilibrium,

$$\sum_{i=1}^N \mathrm{RV}(T' - m'_i) \leq 1 - T'.$$

Because RV is strictly convex, Jensen's inequality then implies

$$N\,\mathrm{RV}\left(\frac{(N-1)T'}{N}\right) < \sum_{i=1}^N \mathrm{RV}(T' - m'_i) \leq 1 - T'.$$

Therefore, for sufficiently small $\epsilon > 0$, we meet the necessary condition for an equilibrium in which each player invests $m = T'/N - \epsilon$:

$$N\,\mathrm{RV}\left(\frac{(N-1)m}{N}\right) \leq 1 - Nm.$$

We now construct the institutional mechanism necessary to support such an equilibrium. Define the functions $V_1, \ldots, V_N$ as follows:[17]

- If every $m_i = m$, then every $V_i(m_1, \ldots, m_N) = \frac{1}{N} - m$.

[17] As in the proof of Lemma 10, the specification of V_i in case $m_i \neq m$ and $m_j \neq m$ ($i \neq j$) is immaterial.

- If $m_i < m$ and each other $m_j = m$, then $V_i(m_1, \ldots, m_N) = \frac{1}{N} - m$ and each other $V_j(m_1, \ldots, m_N) = \frac{1}{N} - m + \frac{m-m_i}{N-1}$.
- If $m_i > m$ and each other $m_j = m$, then $V_i(m_1, \ldots, m_N) = \min\{\frac{1}{N} - m, 1 - \sum_k m_k\}$ and each other $V_j(m_1, \ldots, m_N) = \max\{\frac{1}{N} - m - \frac{m-m_i}{N-1}, 0\}$.
- If $m_i \neq m$ and $m_j \neq m$ for distinct i, j, then each $V_k(m_1, \ldots, m_N) = \frac{1}{N}(1 - \sum_\ell m_\ell)$.

Notice that $\sum_i V_i(m_1, \ldots, m_N) = 1 - \sum_i m_i$, as required. Following along the lines of the proof of Lemma 10, we now construct a strategy profile that constitutes a peaceful equilibrium:

- In the investment stage, each player invests $m_i = m$.
- In the institutional stage, each player opts for the institution ($w_i = 0$) if and only if $V_i(m_1, \ldots, m_N) \geq W_i(m_1, \ldots, m_N)$.

The strategies in the institutional stage are mutual best responses by construction, so we need only check that the initial coercive investments are optimal. There is no incentive for any player to force a conflict, as each $V_i(m, \ldots, m) = \frac{1}{N} - m \geq \mathrm{RV}(\frac{(N-1)m}{N})$. So we must only check that no player can receive more from the institution by deviating—as is obvious from the construction of V_i. Therefore, the constructed strategy profile is an equilibrium. Finally, observe that peace prevails on the equilibrium path, as each $V_i(m, \ldots, m) \geq \mathrm{RV}(\frac{(N-1)m}{N}) \geq W_i(m, \ldots, m)$. □

A.11 Baseline Model with Boundary Conditions

In the main text, we assumed each $y_i > \bar{\mu}_i$ so that boundary conditions would never bind best responses. We now briefly discuss the consequences of relaxing this assumption.

If initial wealth is disproportionately concentrated in one player, then the possibilities for equilibrium with a monopoly of violence might expand or contract. Figure 8 illustrates these findings. Even if $\gamma \approx 1$, it is possible to sustain a monopoly of violence by either player—even one that is weaker in terms of coercive effectiveness—if that player possesses a large enough share of initial wealth. On the other hand, even if $\gamma < \hat{\gamma}_i$, a monopoly of violence by player i might be unsustainable if that player possesses too little of the initial wealth.

An equilibrium with distributed coercion will continue to exist even when the boundary constraints might bind. To be clear, some such equilibria that would exist when each $y_i > \bar{\mu}_i$ might no longer be sustainable, but at least one will be. To see why, remember that the crucial condition for the existence of an equilibrium with distributed coercion is that there exist m_1, m_2 such that

$$\psi_1(m_2) + \psi_2(m_1) \leq 1.$$

Per Lemma 7, ρ_1, ρ_2 are the values at which the LHS of this expression is minimized in the unconstrained case. But suppose this is infeasible for one player: $y_i < \rho_i$. Then to sustain an equilibrium we will need to find m_j such that $\psi_i(m_j) + \psi_j(y_i) \leq 1$.[18] At first this might appear

[18] The wealth constraint can only ever bind for one player: $y_i < \rho_i$ implies $y_i < \bar{\mu}_i < 1/2$ and thus $y_j > 1/2 > \bar{\mu}_j$.

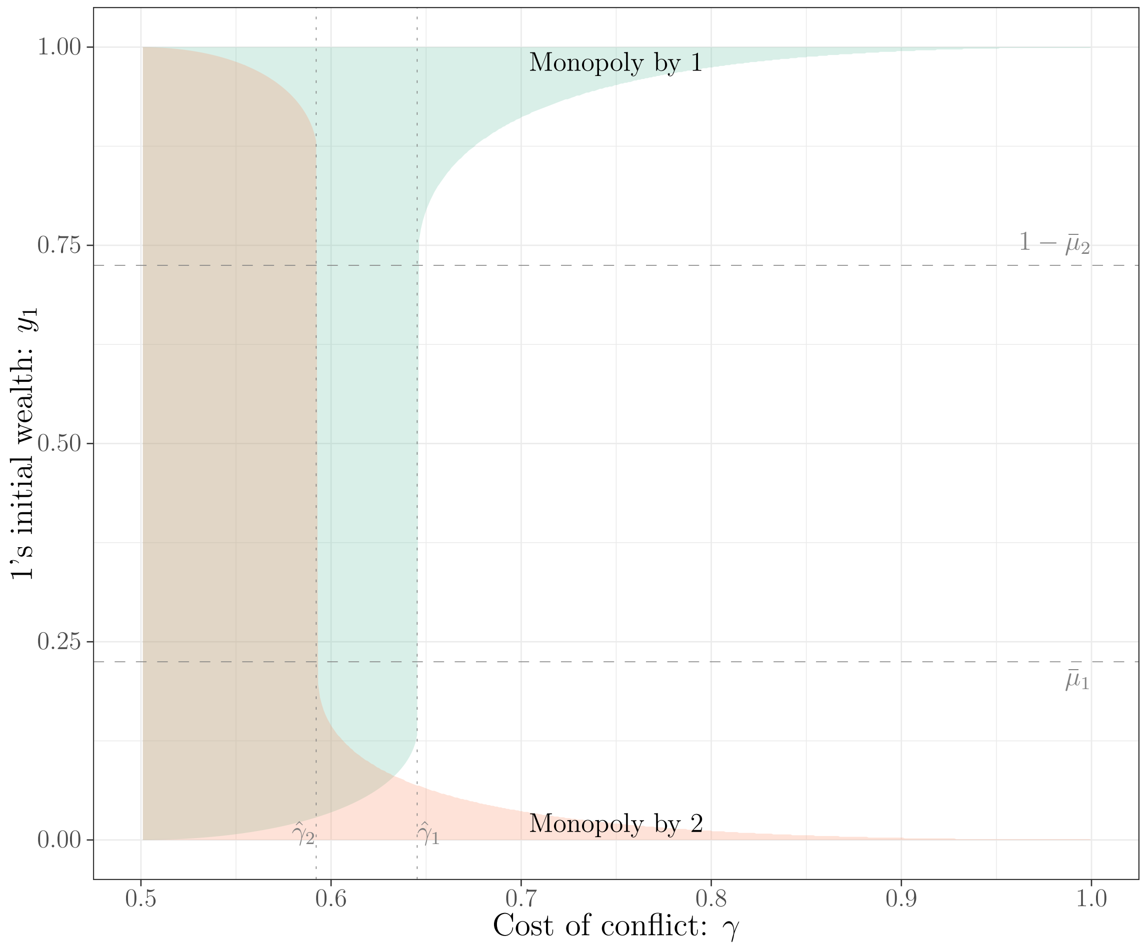


Figure 8. Existence of a monopoly of violence when boundary conditions might bind. (Parameters: $\theta_1 = 1.5$, $\theta_2 = 1$.)

to be a more difficult problem than in the unconstrained case, as $\psi_j(y_i) > \psi_j(\rho_i)$. However, this is offset by the fact that RV_i (and thus ψ_i) is lower than in the unconstrained case, as now the wealth constraint sometimes binds i's optimal response to j's coercive investment.

A.12 Equilibria with Conflict — Two-Sided Private Information

The goal of this section is to show how there are conditions such that it is better to allow for positive probability of conflict when both players have private information. We compare two institutions with a focus on efficiency, one in which there is conflict with probability 0, and one in which conflict is allowed with positive probability. We assume that $y_1 = y_2 = \frac{1}{2}$ and $\gamma > \frac{1}{2}$. Both players are privately informed about their coercive effectiveness $\theta_i \in \{0, 1\}$, where $\pi = \Pr(\theta_i = 1) \in (0, 1)$. We first look for the most efficient equilibrium without conflict, and then compare it to the most efficient equilibrium with conflict.

Equilibria without conflict. We look for symmetric equilibria with the following structure, where player i invests m_1^* if he is of type $\theta_i = 1$, and m_0^* if he is of type 0. First, consider investments $m_1^* \neq m_0^*$ in a most efficient equilibrium without conflict. In this equilibrium, welfare is characterized as $W^{NC} = 1 - (1-\pi)m_0^* - \pi m_1^*$. This equilibrium exists if reservation values for conflict are not too high. In particular, the more stringent condition is placed on the high type ($\theta_i = 1$), which implies that the condition under which this equilibrium exists is

$$RV_1(m_0^*, m_1^*, \theta_1 = 1) + RV_2(m_0^*, m_1^*, \theta_2 = 1) \leq 1 - 2((1-\pi)m_0^* + \pi m_1^*), \tag{11}$$

as both players of the high type have to get more from the institution in expectation than from the best possible deviation. Recall that $RV_i(\cdot)$ is strictly convex if the wealth condition does not bind. Given that $y_1 = y_2 = \frac{1}{2}$, this result applies. Hence, we can consider an alternative equilibrium which is pooling with investments

$$m^* = (1-\pi)m_0^* + \pi m_1^* - \epsilon, \tag{12}$$

where $\epsilon > 0$ is small. Due to the strict convexity of the $RV_i(\cdot)$ function, the condition of Equation 11 also holds and such an equilibrium exists and is more efficient. Thus, the most efficient equilibrium without conflict is a pooling equilibrium. The goal is now to find the lowest m^* such that conflict is prevented with probability 1. The reservation value for type $\theta_i = 1$ equals

$$\begin{aligned} RV_i(m^*, m^*, \theta_i = 1) &= \max \gamma \left[(1-\pi)(1 - m_1 - m^*) + \pi \frac{m_1}{m_1 + m^*}(1 - m_1 - m^*) \right] \\ &= \gamma(1 + \pi m^* - 2\sqrt{\pi m^*}). \end{aligned}$$

As a result, this equilibrium exists if

$$2\gamma(1 + \pi m^* - 2\sqrt{\pi m^*}) \leq 1 - 2m^*.$$

The lowest m^* for which this equation holds equals

$$m' = \frac{1 - \gamma(2 - \pi) + 2g^2\pi)}{2(1 + \pi\gamma)^2} - \sqrt{2}\sqrt{\frac{\gamma^2(1 - \gamma(2 - \pi))\pi}{(1 + \gamma\pi)^4}}, \tag{13}$$

and welfare without conflict is $W^{NC} = 1 - 2m'$.

Equilibria with conflict. Now consider an equilibrium where the high types engage in conflict when both of them are of the high type. Let both players invest $m = 0$ if they are type 0, and $m^* > 0$ otherwise. Holding fixed this strategy profile, the reservation value for both players of type 1 equals

$$RV_i(0, m_1^*, 1) = (1 - \pi)\gamma(1 - m_1) + \pi\gamma\frac{m_1}{m_1 + m^*}(1 - m_1 - m^*),$$

which is maximized by $m' = \sqrt{m^*\pi} - m^*$. Doing this for both players yields

$$\begin{aligned} m_1^* &= \sqrt{m_2^*\pi} - m_2^*, \\ m_2^* &= \sqrt{m_1^*\pi} - m_1^*, \end{aligned}$$

and substitution gives $m_1^* = m_2^* = \frac{\pi}{4}$. As a result, total welfare with conflict would be

$$\begin{aligned} W^C &= 1 - \pi^2\frac{\pi}{2} - 2\pi(1 - \pi)\frac{\pi}{4} - \pi^2(1 - \gamma) \\ &= 1 - \left(\frac{3}{2} - \gamma\right)\pi^2. \end{aligned}$$

Finally, after comparing W^C and W^{NC}, we show that if $\gamma \in (\frac{1}{2}, \frac{3}{4})$, there exists $\bar{\pi}$ such that if $\pi \in (\frac{2\gamma-1}{\gamma}, \bar{\pi})$, welfare with conflict is greater.