Making Multilateral Peace*

Scott Wolford
The University of Texas at Austin
swolford@austin.utexas.edu

June 5, 2018

Abstract

I analyze a game-theoretic model in which a war-winning coalition's ability to deter postwar revisionist challenges depends on weathering shifts in intra-coalition power. Bargains within and across sides are linked, such that the failure of one presages failure of the other, and the success of one presages success of the other. Under some conditions, individual incentives for preventive conflict interact to make the settlement more, rather than less, stable. But when one member of a war-winning coalition will grow substantially stronger, peace settlements can fail because (a) coalition partners fight over shares of the postwar pie or (b) former enemies attack to exploit failures of collective deterrence. An empirical analysis of peace settlements following coalition victories since 1816 shows that the risk of war between former enemies increases in the predicted risk of war between coalition partners that hope to deter them.

*Thanks to Jeff Carter, Kyle Haynes, Amy Liu, Pat McDonald, Glenn Palmer, Toby Rider, Emily Ritter, Chris Zorn, and audiences at Penn State and Purdue Universities, for helpful comments and suggestions, to Julie Phillips for research assistance, and to Black Star Co-Op, Crown & Anchor Pub, Gabriel's Café, Growler USA, The Park, Pinthouse Pizza, Spider House Café, Taco Flats, and Workhorse Bar in Austin; Hotel Park and Human Fish Brewing in Ljubljana; BEAST Bar & Grill in Kaohsiung; OrgAsmo Curry and Zhang Men Brewing in Taichung; Whiskers in State College; Harry's Chocolate Shop in West Lafayette; and Southland Beer in Los Angeles for excellent work environments.

Seven weeks after the Balkan League of Serbia, Bulgaria, Greece, and Montenegro defeated it in the First Balkan War, the Ottoman Empire ordered its armies back into southeastern Europe on 20 July 1913. The previous war ended on 30 May, when the Treaty of London ratified Ottoman losses in Macedonia and Thrace, including the strategic city of Adrianople; created an independent Albania; and transferred Crete to Greece. Overwhelming defeat and subsequent great power guarantees of its new, reduced borders notwithstanding, the Ottomans regained Adrianople over the course of the Second Balkan War, and the subsequent Treaty of Constantinople gave legal imprimatur to these revisions before the year was out. By almost any standard, the settlement that ended the First Balkan War was a failure, keeping former belligerents from each other's throats for barely seven weeks. The First Balkan War cost tens of thousands of lives, strained national finances, prompted coup attempts in Constantinople, and almost drew in Austria-Hungary, Russia, and Germany. It was bloody, wasteful, and ostensibly decisive, yet the principals fought again less than two months later, over an issue nominally settled by the previous war. Why?

When former enemies return to war, political science often asks whether the settlement was flawed. Did the content of the ceasefire, armistice, or treaty correspond to military realities on the ground (Werner and Yuen, 2005)? Did the settlement provide ample means for monitoring and enforcement (Fortna, 2003, 2004)?[1] Was it imposed (Quackenbush and Venteicher, 2008; Senese and Quackenbush, 2003) or guaranteed (Arena and Pechenkina, 2016) by third parties? Were the terms robust to changing circumstances, from shifting military capabilities and alignments to domestic upheaval,[2] that might encourage one side to

[1]For applications of similar questions to peace after civil conflict, see, *inter alia*, Fortna (2008), Mattes and Savun (2009, 2010) and Hultman et al. (2015).

[2]On the link between changing domestic circumstances and changes in foreign policy, see Licht (2014) and Mattes et al. (2016).

restart the war (Lo et al., 2008; Werner, 1999)? The answers to these questions bode well for a stable peace after the First Balkan War, whose largely uninterrupted tide of battle consistently favored the victors and whose outcome won guarantees from the great powers. So how should we explain former enemies' return to war in 1913? The state of the art would have us look at the wrong postwar bargain; it wasn't the bargain struck between the Balkan League and Turkey, between victors and vanquished, that first proved unsustainable. Rather, the winners' bargain over the terms of their shared victory cracked first, paving the way for the Ottomans to recoup their losses once Bulgarian forces had already fallen upon their onetime Serb and Greek partners. Understanding why peace breaks down after multilateral wars requires that we look at the terms of settlement not only *across* formerly warring sides but also *within* them. Doing so paints a very different picture of the Second Balkan War.

On 29 June, three weeks before Ottoman forces broke out into Thrace and a month after the signing of the Treaty of London, the coalition that won the First Balkan War imploded. The Bulgarian Second and Fourth armies attacked, respectively, Greek and Serbian troops occupying recently-conquered parts of Macedonia. The preceding conflict had lasted seven months, from October 1912 to May 1913, during which the Balkan League secured territory from the region's ancient imperial power. Yet with Serbia and Greece straddling Macedonian lands coveted by Bulgaria and unwilling to either yield them or submit to a previously agreed Russian mediation (Glenny, 2012, pp. 243-248), the winners fought a Second Balkan War over part of the large swathes of territory held by the victors yet left explicitly unaddressed by the Treaty of London. Only then, the victors' collective ability to defend their gains was compromised by an intra-coalition war, did the Ottomans try to overturn the Treaty of London. Whatever the flaws of the bargain between the Empire and the Balkan League that ended the First Balkan War, they didn't precipitate Ottoman participation in the Second.[3] It was the agreement among the the victors that first collapsed, undermining

[3]The absence of a formal agreement might arise as a possible explanation of the League's post-victory collapse (Owsiak et al., 2017; Owsiak and Rider, 2013). But the very anticipa-

the Balkan League's ability to defend the terms of the Treaty of London and the Ottoman commitment to the same, tempting the Empire to strike back.

The outbreak of the Second Balkan War makes clear that peace settlements entail bargains negotiated not only across but also within formerly warring sides. In the simplest case, war-winning coalitions must agree to the division and defense of a postwar status quo defined by the chaotic, destructive, and redistributive process of fighting. War reallocates land, resources, population, wealth, influence, prestige, and privileges, often in unpredictable ways, and in the process of taking from the vanquished, the victors must devise a multilaterally credible postwar bargain lest they fall into war themselves (see Wolford, 2017). Peace is multidimensional; one war produces multiple bargains, each of which must sustain the cooperation of the relevant parties if the settlement writ large is to succeed. Each component bargain depends on the others, in that the failure of any one can trigger another's failure, while the success of one can ensure the success of another. Serbia, Bulgaria, Greece, and Turkey could not all be satisfied simultaneously with their shares of the pie created by the First Balkan War, but it took the violent collapse of the bargain among the victors to cause the violent collapse of the bargain between victors and vanquished. To look only to the politics between former enemies is to risk mischaracterizing some failed settlements as successful or misunderstanding why some settlements, like the one that ended the First Balkan War, fail. Indeed, standard coding rules (see Fortna, 2003; Lo et al., 2008; Quackenbush and Venteicher, 2008; Senese and Quackenbush, 2003; Werner, 1999; Werner and Yuen, 2005) would identify the Ottoman invasion of Thrace rather than the intra-coalition war that preceded it as the collapse of the First Balkan War settlement.

---

tion of conflict over Macedonia explains (a) why Serbia and Bulgaria agreed only informally to Russian mediation before the war and (b) why Bulgaria wouldn't submit to a formal agreement ratifying Serbia's gains in the summer of 1913. Whatever the possible effects of the formal delineation of borders, they were avoided here due to anticipated conflict.

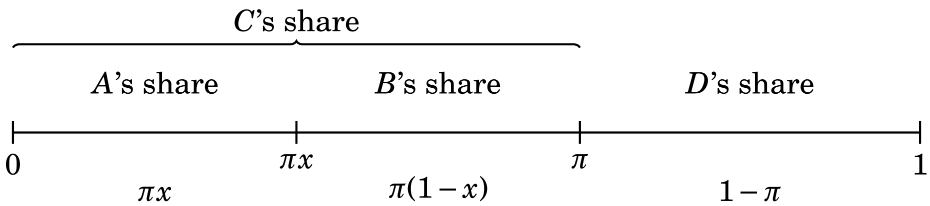
Nearly half (40%) of all interstate wars fought in the last two centuries have seen military coalitions fight on at least one side (Sarkees and Wayman, 2010), and more often than not those coalitions emerge victorious (Morey, 2016). But whether or not they win, coalitions must share in the distribution (Starr, 1972) and defense of a new postwar status quo if they're to sustain the bargain struck with their enemies. Bargains across and within warring sides are strategically interdependent, and a durable settlement requires that both bargains be simultaneously stable. Most work on postwar peace focuses on former belligerents at the expense of peace between former partners, but it strains credulity to claim that a peace settlement remains successful only if former belligerents do not return to war when the war itself also reshapes the terms on which its victors live with one another. In the sections that follow, I analyze a game-theoretic model that focuses on settlements that divide a flow of benefits both *across* and *within* formerly warring sides after a coalition military victory. Settlements survive when no player attacks another over the terms of either bargain, while the rising strength of one coalition member creates fears of relative decline in both a former enemy that would like to recoup its losses and an erstwhile partner hoping to keep a favorable share of the fruits of victory. I show, first, that individual incentives for preventive conflict can interact to make settlements more, rather than less, stable. Then, I show that the same factors related to cooperation failure inside the coalition are related to the breakdown of the bargain between victors and vanquished. First, shifts in the distribution of power can lead to preventive war inside the coalition, though the threat of drawing in former enemies with their own preventive motives can dampen those incentives. Second, even when the winning coalition itself has yet to collapse, a former enemy may attack, confident that shifting intra-coalition power will undermine the collective deterrent threat that sustains the bargain across sides. Solving the commitment problem along one dimension can solve it along the other, but failing to solve it along one can lead to failures along the other. An event-history analysis of the duration of peace following coalition victories in interstate wars from 1816-2007 uncovers the expected positive relationship between the risk of intra-coalition war and the occurrence of war between victorious coalitions and their former enemies. The stability of bargains *within* formerly warring sides thus proves critical to understanding the stability of bargains *across* formerly warring sides.

Theoretical Model

The theoretical model isolates the specific case of a two-state coalition and a third state that it's just defeated in war, where the coalition must negotiate over the distribution and defense of a new status quo, an obstacle that doesn't confront states that win wars on their own (Wolford, 2017, pp. 703-705). Decisive war outcomes are rare (Kecskemeti, 1958; Slantchev, 2004; Wagner, 2007), but these are equally relevant for stalemates or negotiated settlements, like the Korean War (Reiter, 2009, Ch. 5). Still, this post-victory specification captures the problem in its starkest and most illuminating form. Defeat clarifies the war-ending distribution of power, removing an informational motive for subsequent war (see Werner and Yuen, 2005), but it also tempts the defeated state to launch a revisionist attack. The defeated state can be deterred if the coalition will cooperate, but that cooperation is complicated by the fact that one member, hereafter the *rising* state (see Powell, 1999, Ch. 4), can't promise not to exploit rising strength to extract a larger share of the fruits of victory from its partner who, along with its recent enemy, is also a declining state.[4] The declining states' individual incentives for preventive war are well understood (Powell, 2006), but the present model shows how the interaction of preventive mottives across sides can ex-

[4] I abstract away from the case in which one member of the coalition will see its absolute capabilities fall; though it creates a similar incentive for intra-coalition preventive war, it creates strong incentives for the defeated state not to attack before the shift but to wait. This kind of shifting power may be empirically relevant, but I abstract away from it to isolate one form of interdependence between across and within bargains.

Figure 1: Allocations in the Status Quo Settlement



plain non-obvious collective outcomes, such as how shifting power can *enhance* rather than undermine the credibility of commitments to the settlement.

Suppose that a coalition of two states, $C = \{A, B\}$, has defeated state D in a war over a flow of unit-sized benefits of which each player would, all else equal, prefer a larger share. A *settlement* is a network of bargains across and within formerly warring sides that define who gets what in the postwar status quo. As shown in Figure 1, the settlement ending the previous war between C and D has two dimensions; a bargain *across* sides allocates $\pi \in (0,1)$ to C and $1-\pi$ to D, while a bargain *within* the coalition allocates πx and $\pi(1-x)$, where $x \in (0,1)$, respectively to A and B. Harsher across bargains (large π) encourage D's revisionism but increase the pie to be shared between the victors; lenient across bargains (small π) keep D satisfied yet leave less for to be shared in the within bargain. War can break out, and the settlement collapse, over either dimension. The challenge for the coalition is to cooperate in the distribution and defense of the status quo as one member grows stronger in absolute terms, which can tempt the declining member to attack or abandon its erstwhile partner in a bid to preserve its share of the pie.[5]

In each of a potentially infinite number of periods $t > 0$, the settlement survives if no player attacks any other. Peaceful renegotiations, even if they occur in the shadow of war, don't constitute a failure of the settlement.[6] With a common discount rate $\delta \in (0,1)$, payoffs

[5]See Wolford (2017) for an analysis of the problem of shared victory in isolation.

[6]Successful settlements, like the Concert of Europe, often include means of peaceful rene-

for a stable settlement are

$$u_A = \frac{\pi x_t}{1-\delta}, \quad u_B = \frac{\pi(1-x_t)}{1-\delta}, \quad \text{and} \quad u_D = \frac{1-\pi}{1-\delta},$$

where the coalition's shares of the within bargain are indexed by t to indicate that they may be peacefully renegotiated. State D, however, can renegotiate its share only by going to war. This is unrealistic, but it focuses the analysis on how distributive politics inside the coalition can undermine the across settlement even when D can guarantee itself $1-\pi$ simply by not attacking. Next, as long as it hasn't been defeated in war, B's military capabilities increase between periods t_0 and t_1, influencing the attractiveness of both dimensions of the settlement; D's chances of winning a war with B on the other side will fall, and A may be forced to accept a less favorable within bargain. Finally, to focus on preventive motives for war, I assume that the rising B is the only possible target of attack; the declining states have no direct incentives to attack one another.

Play begins with a three-actor stage game that occurs in every period that doesn't follow a war, which transitions play to a different stage game. First, D chooses whether to attack state B or to pass. If D passes, then A and B have the opportunity to renegotiate the within settlement; A either attacks B over π or proposes a division based on shares $(x_t, 1-x_t)$, which B can accept, implementing the deal for the current period, or reject, which amounts to attacking A. An intra-coalition war eliminates the loser, leaving the winner to face D in a new stage game in which D chooses in every period whether to attack the surviving coalition member or pass. If D attacks B in its first move of the three-player stage game, A then chooses whether to help B or deny assistance. If A denies, then states fight a war of all against all (cf. Gallop, 2017) in which A sides with neither of initial belligerent but secures a chance to win the whole prize for itself.[7] If A helps B, then the coalition cooperates

gotiation and adjustment (Slantchev, 2005).

[7]This treats subsequent processes of war expansion and negotiation in reduced form, but

against D. Should the coalition win, A and B once again bargain, with A proposing shares of $(y_t, 1-y_t)$ before transitioning to the next stage game; while they bargain over only π after D passes, after eliminating D they divide the entire flow of benefits, less the costs of war. In each such two-player stage game, A and B bargain in the shadow of war just as they do at the end of the three-player stage game. If any war leaves only one player left standing, then the game ends with the last player enjoying the remaining pie in perpetuity.

The outcomes of negotiations and wars determine payoffs. First, if D passes in any period during which it has the option, it receives $1-\pi$. Second, war is costly, destroying a fraction $d \in (0,1)$ of the flow of benefits; if the coalition defeats D, the per-period flow is $(1-d)$, and should A and D then fight, only $(1-d)^2$ remains. If B accepts a proposal from A, coalition members receive $(1-d)y_t$ and $(1-d)(1-y_t)$ if they defeated D in the present period and πx_t and $\pi(1-x_t)$ if D passed. The outcomes of wars, which eliminate losing states, depend on relative military capabilities, the efficiency of cooperation between coalition members, and whether war occurs before or after B's military capabilities have increased. If A supports B, the coalition and D win, respectively, with probabilities

$$\frac{c(m_A+m_B+sI)}{c(m_A+m_B+sI)+m_D} \quad \text{and} \quad \frac{m_D}{c(m_A+m_B+sI)+m_D},$$

where $c \geq 1$ is the marginal efficiency of military cooperation (cf. Powell, 1999, Ch. 5), $s > 0$ is the increase in B's capabilities, and I equals 0 before ($t = t_0$) and 1 after ($t \geq t_1$) B's

it ensures that A's payoffs are a function of its own military power relative to the other two players, as would be the case if it were to engage in subsequent bargaining with the winner of a two-party war between B and D. An alternative specification would allow A a period of neutrality before bargaining with the eventual winner of the war, but it would only complicate the analysis without changing the substantive results.

capabilities increase. If A denies help, probabilities of victory in a war of all against all are

$$\frac{m_A}{m_A+m_B+sI+m_D}, \quad \frac{m_B+sI}{m_A+m_B+sI+m_D}, \quad \text{and} \quad \frac{m_D}{m_A+m_B+sI+m_D}, \tag{1}$$

for A, B, and D respectively. State i's probability of winning a bilateral war against $-i$ in any two-player continuation is $m_i/(m_i+m_{-i})$.

An increase in B's military capabilities impacts the distributions of power underlying both within and across bargains, creating *two* declining states, in contrast to bilateral treatments of the commitment problem (cf. Powell, 2006). That shifting power across sides may cause settlements to break down is uncontroversial, but this model captures the strategic linkages between multiple declining states with divergent distributive interests in the preservation of peace. A faces the prospect of negotiating with a stronger partner in the future over the within bargain, but its partner's rising strength limits D's incentives to attack. D faces a closing window in which to attack before its military position declines, an opportunity made all the more tempting should A prove unwilling to help D. Further, by assuming that D can increase its share only by the costly action of attacking, the analysis isolates the cause of the across bargain's breakdown. If the across bargain collapses, it does so not because D fears a renegotiation of the bargain but because shifting intra-coalition power creates incentives for D to attack.

Equilibrium Analysis

I first prove the existence of an equilibrium in which (a) neither declining state attacks the rising state and (b) A helps B in the event of outside attack. In such an equilibrium, the threat posed by a revisionist B encourages A to tolerate B's rise and the attendant future concessions over the terms of the within bargain. Then, I derive conditions sufficient for such a peaceful equilibrium to break down in two ways, when (a) A attacks B in a preventive intra-coalition war that can, under some conditions, cause D to challenge the across bargain as well, or (b) D attacks B because A denies help. The shadow of shifting intra-coalition power creates an endogenous relationship between the stability of across and within bargains; the failure of one can presage the failure of the other, and the success of one can support the success of the other. The solution concept is Markov Perfect Equilibrium (MPE), a refinement of Subgame Perfect Equilibrium that selects only payoff-relevant strategies (Maskin and Tirole, 2001). Markovian strategy profiles entail mutual best responses in each proper subgame, the five payoff relevant states of which entail combinations of B's military strength, the present value of the flow of benefits, and the number of players.[8] I discuss each MPE informally, saving all proofs for the appendix.

Durable Settlements

Many MPE with durable settlements exist, but most are uninteresting—e.g., no state attacks because war is too costly or D has no incentive to attack because it's too weak. In these equilibria, peace obtains by default.[9] But in each period of Proposition 1's *successful deterrence* equilibrium, D passes, deterred by the credibility of the A's commitment to help B in meeting an attack, and A chooses not to launch a preventive war against B, in part because D will attack an A that stands alone after defeating B. D's temptation to attack the coalition diminishes once B's rise is complete, which encourages A to forego its own attack before B's rise, but for A the stability of the settlement comes at the price of a renegotiation of the within bargain once B's rise is complete. The settlements that survive in this equilibrium avert two kinds of wars, both of which declining states find themselves tempted to

[8]See the proof of Proposition 1 for the game's states and transition rules between them.

[9]The existence of such equilibria is easy to prove; if war is too costly, then the status quo is sure to be more attractive than fighting, and if D can't gain from overturning the settlement—if it's satisfied—then it can't be induced to attack.

launch before B's increase in power: a revisionist attack on the across settlement launched by D and a preventive war over the within settlement that soon spreads to the across settlement as well. Proposition 1 states that this equilibrium exists when B's growth in power (s) isn't too great, when D retains sufficient capabilities (m_D) after its defeat, and when the across bargain (π) is neither too harsh nor too lenient.

Proposition 1 (Successful Deterrence)**.** *Let $\delta \to 1$. When $\underline{\pi} \le \pi \le \overline{\pi}$, $s \le \hat{s}$, and $m_D > \hat{m}_D$, the following strategy profiles constitute a Markov Perfect Equilibrium. In the three-player stage game, D passes in all periods; A helps, proposes $y_t^* = \overline{y}_t$, and proposes $x_t^* = \overline{x}_t$; B accepts iff $y_t \le \overline{y}_t$ and iff $x_t \le \overline{x}_t$. In the relevant two-player continuations, D attacks A but not B; A proposes $z^* = \overline{z}$, and B accepts iff $z \le \overline{z}$.*

That peace survives when the distribution of power is stable isn't surprising (Powell, 2006; Werner, 1999; Werner and Yuen, 2005). B's rise is small enough ($s \le \hat{s}$) that A finds it tolerable, and D's prospects won't decline so far that it's impelled to attack before it loses its chance. But these two processes are interdependent, linked by the terms of the across bargain and the extent of D's revisionist threat: A tolerates the concessions that follow B's rise because it needs B to ensure deterrence, and D doesn't take advantage of a closing window of opportunity because of A's commitment to bolster B's capabilities before power shifts. A's commitment to help B reduces the size of the effective shift in power against D; a given value of s has a smaller effect on D's probability of victory the more the pre-shift distribution of power already favors B's side. Next, sustainable military cooperation inside the coalition can deter B from attacking, but it is the prospect of drawing in a powerful D, like Bulgaria did after attacking Greece and Serbia in 1913, that convinces A not to launch its own preventive war. When D retains sufficient military capabilities,

$$m_D > m_A^2 \frac{(1-d)(1-d\pi)}{\pi(m_A + m_B d)} - m_A = \hat{m}_D, \tag{2}$$

its threat to the across settlement solves the commitment problem that threatens the within settlement, reducing the prospects that A can emerge a winner from either abandoning or launching its own war against its rising partner. Likewise, when the across bargain isn't too harsh, such that

$$\pi \leq \frac{c(m_A + m_B) + dm_D}{c(m_A + m_B) + m_D} = \overline{\pi},$$

the coalition's threat to cooperate deters D. The bargain must, however, still be harsh enough ($\pi \geq \underline{\pi}$) to tempt D into a war against A alone in order for A to fear the consequences of intra-coalition war.

When across bargains are neither too harsh nor too lenient, when defeated states can credibly threaten to return to war under some conditions, and when intra-coalition power isn't shifting too far, then war-winning coalitions can weather shifting power and deter revisionist challenges. Proposition 1 shows when these relationships emerge in equilibrium, but it also shows that peace settlements can survive in multilateral settings under conditions that would cause them to break down in bilateral settings. Launching its own preventive war against B in the successful deterrence equilibrium ensures that A must also fight D afterwards if it wins, which acts as a check on A's intra-coalition behavior. And if it didn't share an interest in defending the within settlement with its rising partner, A would fight to prevent B's rise for even smaller shifts in power. Likewise, without the support of its coalition partner before its military power rises, B might face attack from a revisionist D eager to recoup losses before the window of opportunity closes. Declining states may be tempted into preventive war in isolation, but the effect in this multilateral equilibrium is to deter one another from attacking. When a defeated state is neither stripped of too many military capabilities nor denied a sufficient stake in the postwar settlement, war-winning coalitions can sustain otherwise impossible cooperation over the fruits of victory. Some states that would otherwise fall victim to preventive war in a dyadic setting may be avoid them because they share a recently defeated but still dangerous former enemy. This reveals an incentive after coalition wars not to disarm one's opponent, to leave a buffer in place that may suffer when times are bad (see Fazal, 2007) but that can help keep the peace between other states when it survives (Wagner, 2007, Ch. 4).

Failed Settlements

When do multilateral settlements fail? This section clarifies when and how shifting power inside the war-winning coalition can lead to both intra-coalition wars over the within bargain and revisionist wars over the across bargain prompted by anticipated failures of collective deterrence. First, Proposition 2 characterizes sufficient conditions for any MPE to entail A attacking B over the within settlement in an intra-coalition war that may or may not expand to include a revisionist D. Next, Proposition 3 identifies conditions sufficient for any MPE to entail D launching a revisionist war because A will refuse to help a rising B. After discussing each path to settlement failure, I show that similar conditions—large increases in B's military capabilities—support each set of violent equilibria.

First, intra-coalition war can break out when power is shifting so far in B's favor that fighting to prevent that shift can recoup A's costs of war, whether those costs entail the simple destruction of fighting or subsequent war against its former enemy.

Proposition 2 (Intra-Coalition War)**.** *Let $\delta \to 1$. When d is sufficiently small and*

$$s > \max\left\{ \frac{(m_A + m_B)\left(m_A\left(CV_A^D - \pi\right) - d\pi m_B\right)}{d\pi m_B - m_A\left(CV_A^D - d\pi\right)}, \frac{d(m_A + m_B)^2}{m_A(1-2d) - dm_B} \right\} \tag{3}$$

all MPE entail strategy profiles in which A attacks B in period t_0.

Proposition 2 characterizes the minimum shift in power necessary to ensure that any MPE entails a strategy profile in which A attacks B if D passes. The left term inside the brackets in Inequality (3), where CV_A^D is A's continuation value for war against D, is the minimum shift necessary to ensure that A attacks even at the cost of future war with D over the entire flow of benefits (e.g., in Proposition 1's successful deterrence equilibrium);

the right term is the minimum shift necessary to tempt A to attack B when D won't attack a lone, victorious A. When s passes both thresholds, then no pre-shift bargain with B, even a maximally generous deal of $x_t = 1$ that allocates all of π to A, can convince A not to attack. A risen B can secure a substantial revision of the within bargain, and rather than see its postwar position negotiated away in the future, A opts for a preventive war. Whether the right or left term binds depends on the across bargain (π), which shapes how much A stands to gain from a subsequent war with D over the entire flow. When

$$\pi < \frac{m_A}{(1+d)m_A + m_D} = \pi^{\dagger},$$

A launches preventive war for smaller shifts in power when the war won't expand to include D, but when $\pi \geq \pi^{\dagger}$, A launches preventive wars for smaller shifts when subsequent war with D is certain. Thus, there is no consistent bivariate relationship between D's revisionist threat and the outbreak of intra-coalition war. Rather, an interactive relationship between shifting power and the terms of the across bargain shapes the relative risk of war across cases in which D does and doesn't pose a revisionist threat.

The intra-coalition war equilibrium offers a compelling explanation for the fragility of the settlement that ended the First Balkan War. The Ottoman Empire attacked its former Bulgarian enemy opportunistically, not because Bulgaria was unable to honor a commitment not to push farther into Turkey-in-Asia, but because Bulgaria couldn't trust Serbia to honor a commitment to negotiate over the status of Macedonia. The within bargain proved unstable, and once the winners of the previous war fell upon themselves, the across bargain was ripe for revision. Had Serbia and Bulgaria been able to negotiate the distribution of Macedonian lands peacefully and sustain military cooperation—that is, had the Second Balkan War not already broken out—it's unlikely that the Ottoman Empire would have been willing to launch a war to regain eastern Thrace. Thus, failing to account for the link between within and across bargains risks both misdiagnosing the causes of their failure and mis-specifying empirical models of war recurrence.

Preventive war isn't the only path by which intra-coalition bargaining frictions can undermine the across settlement. Shifting power can also weaken the credibility of coalition partners' commitments to fight alongside one another, and Proposition 3 characterizes the conditions under which collective deterrence is sufficiently compromised to ensure that D attacks an unreliable coalition in every MPE.

Proposition 3 (Failed Deterrence)**.** *Let $\delta \to 1$. When d and c are sufficiently small,*

$$s > \frac{d(m_A + m_B)^2}{m_A(1-2d) - dm_B}, \quad \textit{and} \quad \pi > \max\left\{\frac{m_A + m_B + dm_D}{m_A + m_B + m_D}, \frac{m_B}{d(m_A + m_B + m_D)}\right\},$$

all MPE entail strategy profiles in which D attacks B and A denies help.

Proposition 3 shows that D launches a revisionist war over the across bargain when power will shift sufficiently to guarantee that A won't help B. In any such MPE, A prefers a war of all against all at B's initial capabilities to the compounded costs and risks of a coalition war against D that, if won, still leads A to preventive war against B. If A won't attack B after defeating D, then it's sure to help in the event of a revisionist attack; but if power is shifting sufficiently, the costs of war aren't prohibitively large, and military cooperation isn't too efficient ($c < 1/(1-d)$), then A refuses help once D attacks B in its own preventive war. Faced with a deterrent threat of questionable reliability (cf. Leeds, 2003; Smith, 1995), D attacks when sufficiently dissatisfied with the across bargain—i.e., when π is large enough. Therefore, war-winning coalitions beset by shifting power and overseeing harsh postwar settlements should be particularly vulnerable to deterrence failures, their own struggles over the within bargain leading to the violent breakdown of the across bargain.

Finally, the conditions supporting the violent equilibria of Propositions 2 and 3 overlap: many parameter values predicting failures of the within bargain also predict failures of the across bargain, either of which is sufficient to ensure that the settlement fails. When power will shift too far in B's favor, A is tempted both to attack B and to abandon it should D

attack. This leads to a clear empirical implication for the stability of across settlements, which depends on the credibility of commitments to the within bargain:

> **Hypothesis.** The risk of war between former enemies after coalition victories increases in the risk of war between members of the war-winning coalition.

Whether war with former enemies occurs before (Proposition 3) or after (Proposition 2) war breaks out within the coalition depends on D's willingness to risk a war of all against all or to wait and face A alone. In contrast to Proposition 1's successful deterrence MPE, where increases in D's military capabilities (m_D) bolster settlements, D's willingness to attack in Proposition 3 increases in those same capabilities.[10] A powerful, revisionist D can enhance cooperation between coalition partners when power isn't shifting too far, but when shifting power undermines cooperation between A and B, then war over the across settlement is more likely when D is more powerful. Defeated states' capabilities—from Germany in 1919, Germany and Japan in 1945, to Iraq in 1991—are frequent subjects of peace settlements, but whether leaving an enemy powerful or weak sustains or undermines postwar peace depends on shifting power inside the winning coalition. The stability of within and across settlements are linked; the success of one can ensure the success of the other, while the failure of one can cause the failure of the other. Individual incentives for preventive war can be offset in equilibrium with the addition of more declining states into the strategic system, discouraging rather than encouraging the outbreak of preventive war.

Research Design

The theoretical model anticipates a positive relationship between the risks of war over the within and across bargains created by coalition military victories, and the empirical analysis

[10] In addition to the successful deterrence equilibrium existing when $m_D > \hat{m}_D$, the minimum values of π above which D attacks both decrease in m_D.

focuses on a specific element of that relationship: how frictions over the within bargain can undermine deterrence of challenges to the across bargain. Most work disaggregates peace settlements into multiple component dyads (Fortna, 2003; Lo et al., 2008; Quackenbush and Venteicher, 2008; Senese and Quackenbush, 2003; Werner, 1999; Werner and Yuen, 2005), which accounts for the differential timing with which states leave wars. That approach obscures the effects of settlement-level variables and inflates the importance of settlements (and the wars that produce them) with a large number of parties (Wolford, 2017, p. 707). Further, its focus on dyads can credit settlements with survival when they've already failed, with dyads remaining at peace after others have already gone to war over either dimension of the settlement. Finally, the theoretical model focuses on the deterrence of revisionist attacks after coalition victories, so the proper unit of analysis is the bargain produced by a coalition victory, and the proper outcome variable is the breakdown of the across bargain: any two former enemies going to war with one another after the previous war. The bargaining problem is multilateral, entailing collectively credible commitments across and within formerly warring sides, so the unit of analysis and outcome variable are multilateral as well (see Poast, 2016). I focus on settlement failure through a particular causal path—one that leads from cooperative problems over the within bargain to the breakdown of the across bargain—so the research design clarifies which bargains, within or across, fail first. The settlement that ended the First Balkan War, for example, collapses not with the Ottoman attack on Bulgaria but with the Bulgarian attack on Serbia and Greece.

I model the survival of settlements in an event history framework, using a Cox Proportional Hazards (PH) model to estimate the probability that a settlement fails at time t given survival until that point.[11] I identify relevant multilateral settlements (i.e., the subjects of

[11]The Cox PH model assumes that covariates have a constant, time-invariant effect on the underlying hazard rate, though it's sufficiently flexible to account for non-proportional hazards (see Box-Steffensmeier and Jones, 2004). However, as discussed below, the present models don't violate the proportional hazards assumption, allowing for the estimation of

the analysis) with multilateral military victories in the Correlates of War Project's list of interstate wars from 1816-2007 (Sarkees and Wayman, 2010).[12] After recording the days on which wars won by coalitions end as the latest ending day for any state on the winning side, I identify the outcome variable *Across War* as the first day on which any two parties to the across settlement go to war with one another. Consistent with the theoretical model, this marks the collapse (in event-history terms, the failure) of that settlement. Settlements can fail short of war, but armed conflict rising to the level of war is an unambiguous indicator of failure that can apply to every settlement in the sample. Other indicators, like the abrogation of alliances or the suspension of cooperation in other areas, don't apply across either particular settlements or time periods (Wolford, 2017, p. 708). And lower-level conflicts, like Militarized Interstate Disputes (Palmer et al., 2015), may indicate successful adjustments of the settlement when parties avert war. Finally, if the within bargain collapses before the across bargain can fail, the observation is right-censored; the survival of the across bargain to that point is informative, contributing to the likelihood function, but failure of the within bargain implicates Proposition 2's processes of war expansion, not Proposition 3's failed deterrence. This results in a sample of 23 settlements that don't see the winning coalition collapse before (a) the across bargain collapses or (b) the observation period ends in 2007, only 8 of which experience uncensored failures of the across bargain.[13]

Cox models without any adjustments.

[12]The present sample doesn't include the Napoleonic Wars or the Iraq and Afghanistan wars in Wolford's (2017) sample. It does include France and Spain's 1958 victory in the Ifni War, which Reiter et al. (2016) delete from their revised list of COW wars, but the statistical results are robust to excluding that conflict from the sample.

[13]Four multilateral settlements—the War of Italian Unification (COW #28), the First Balkan War (COW #100), and the two World Wars (COW #106, 139)—see the winning coali-

The theoretical variable is the predicted risk that war breaks out between two states inside the war-winning coalition. Propositions 2 and 3 show that this risk is related to A's willingness to abandon D and, as a result, the failure of collective deterrence—i.e., the outcome variable. To generate this variable, I re-estimate Wolford's (2017) model of the duration of peace inside war-winning coalitions, which focuses on publicly observable features of coalitions that render them more or less susceptible to wars caused by shifting power: the coalition's size and the density of its alliance commitments raise the risk, while the presence of a great power smooths out otherwise dangerous shifts in power (p. 710, Table 1). Table 1 presents the results on a sample of 26 war-winning coalitions, including COW coalition victories and the end of the Napoleonic Wars in 1915, from which I generate predicted risks of war over each winning coalition's within bargain. The resulting variable, *Within Risk*, is the linear prediction for each coalition generated from the Cox PH model, where higher values indicate higher risks of intra-coalition war.[14] *Within Risk* is time-invariant, such that the empirical model estimates the extent to which generally at-risk within settlements are linked to generally at-risk across settlements.

Figure 2 presents a histogram of *Within Risk*, where the mean and median are both

tion implode before the across bargain can fail, censoring them out of the present sample before the ultimate failures of their across bargain.

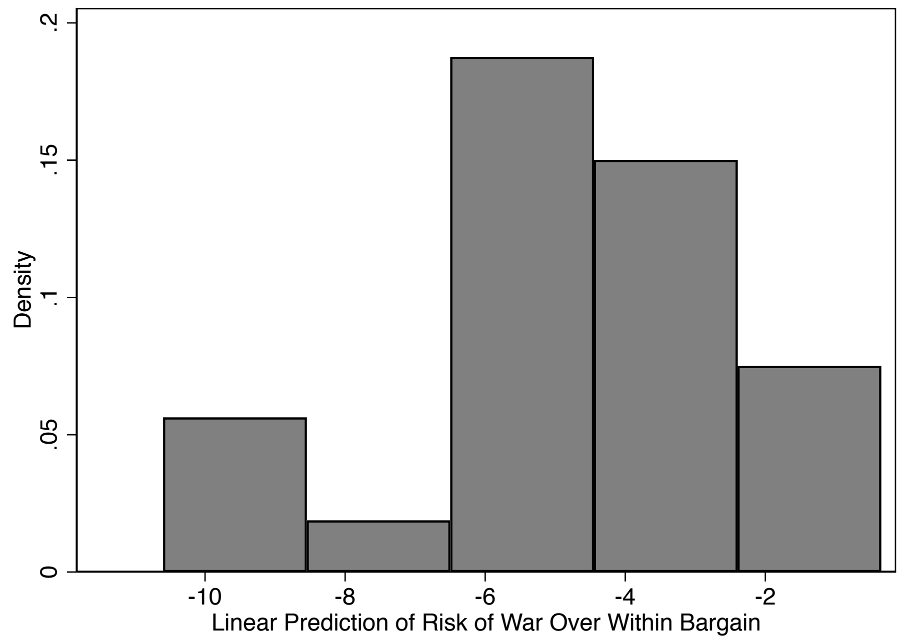
[14]Results are substantively identical if I instead take the natural log of the predicted hazard ratio. The hazard ratio is problematic as an explanatory variable, precisely because it's a ratio; 0.01 is radically different compared to a value of 1 than is 1.01. High hazard ratios exercise outsized influence on the outcome variable in such a setting, so the linear prediction provides an easier, more parsimonious way of capturing substantive differences in the predicted fragility of coalitions without requiring subsequent exponentiation and log-transformation.

Table 1: Cox PH model of the risk of intra-coalition war, 1816-2007

Replication of Wolford (2017)	
Size	4.66**
	(2.23)
Percent Allied	0.04**
	(0.02)
Great Power	-4.69**
	(1.95)
CINC Winner	0.08
	(0.06)
CINC Loser	-0.07
	(0.11)
Pre WWII	6.27**
	(2.98)
Territory	-1.12
	(1.63)
Revisionist	2.23
	(1.51)
War Duration	-1.22***
	(0.44)
Model Statistics	
Subjects/failures	26/12
Log Likelihood	-21.05

Significance levels : * : 10% ** : 5% *** : 1%

Figure 2: Histogram of *Within Risk*



approximately -4.7 with low skewness, requiring no transformations before inclusion in the empirical models. According to this estimate, the most vulnerable war-winning coalitions that either don't collapse or collapse only after the across bargain fails are (1) the Entente and its Associated Powers after the First World War (COW #106), (2) NATO after the 1999 Kosovo War (COW #221), (3) the Franco-British-Israeli coalition after the 1956 Sinai War (COW #155), (4) Czechoslovakia and Romania after the War of Hungarian Adversaries in 1919 (COW #112), and (5) Cuba and Ethiopia after the second phase of the Second Ogaden War in 1978 (COW #187). Those coalitions that experience intramural war before the across bargain fails contribute information the likelihood function, but I right-censor them once the within bargain fails. Finally, the inclusion of an estimate on the right hand side is problematic in that *Within Risk* enters the model with its own error structure, so all models are estimated with bootstrapped standard errors (1000 reps).

I also include four control variables whose absence might confound the relationship between *Within Risk* and *Across War*. First, the more parties to the settlement, the more opportunities for war over both within and across settlements. Therefore, *Total Parties* counts the number of parties to the war-ending settlement as identified by COW (Sarkees and Wayman, 2010). Second, longer wars tend to be costlier than shorter wars, which depresses the risk of conflict between both former belligerents (Werner and Yuen, 2005) and former partners (Wolford, 2017). Thus, *War Duration* is the length of the war in days beginning with the earliest recorded entry dates and latest exit dates for war participants (Sarkees and Wayman, 2010).[15] Finally, stronger coalitions may be resistant to shifts in power that would compromise weaker coalitions (Wolford, 2017, pp.706-707), making collective deterrence easier, while stronger defeated sides, whether coalitions or singletons, can both enhance cooperation over the within settlement (recall Proposition 1) and have greater incentives to revisit the last war. Thus, *Winner CINC* and *Loser CINC* are the summed values of the Composite Index of National Capabilities for each member of the winning and losing sides in the previous war (Singer, 1987).

I estimate the following Cox PH model,

$$h(t|X) = h(t)\exp[\beta_1(\text{Within Risk}) + \mathbf{X}\beta],$$

where $h(t|X)$ is the probability that at least two states on opposite sides of the last war return to war over the across bargain on day t, given peace until day t, as a function of a set of regressors and their coefficients. Coefficient β_1 is estimated on the theoretical variable, and the vector $\mathbf{X}\beta$ includes control variables and their coefficients. The sample is not left-censored, and right-censored observations still contribute to the likelihood function until they leave the sample. The Cox PH model is also semi-parametric, estimating the base-

[15]Advances in battlefield medicine have reduced combat deaths substantially over the sample period, rendering them a problematic measure of the costs of war (Fazal, 2014). The length of the war is thus a superior indicator of war's cost across the whole sample period.

Table 2: Cox PH models of the risk of settlement collapse, 1816-2007

	Model 1 All Across	**Model 2** Multilateral	**Model 3** Multi. w/Risk
Theoretical Variable			
Within Risk	–	–	0.75**
			(0.33)
Control Variables			
Total Parties	0.14**	0.155	0.20*
	(0.06)	(0.10)	(0.11)
War Duration	-0.44***	-0.83**	-0.88*
	(0.15)	(0.39)	(0.47)
Winner CINC	-1.01	-2.08	-8.53*
	(2.08)	(3.56)	(4.42)
Loser CINC	12.17***	15.34*	37.83**
	(3.85)	(8.87)	(16.38)
Model Statistics			
Subjects/failures	81/33	23/8	23/8
Log Likelihood	-118.98	-17.05	-13.64

Significance levels : $*$: 10% $**$: 5% $***$: 1%

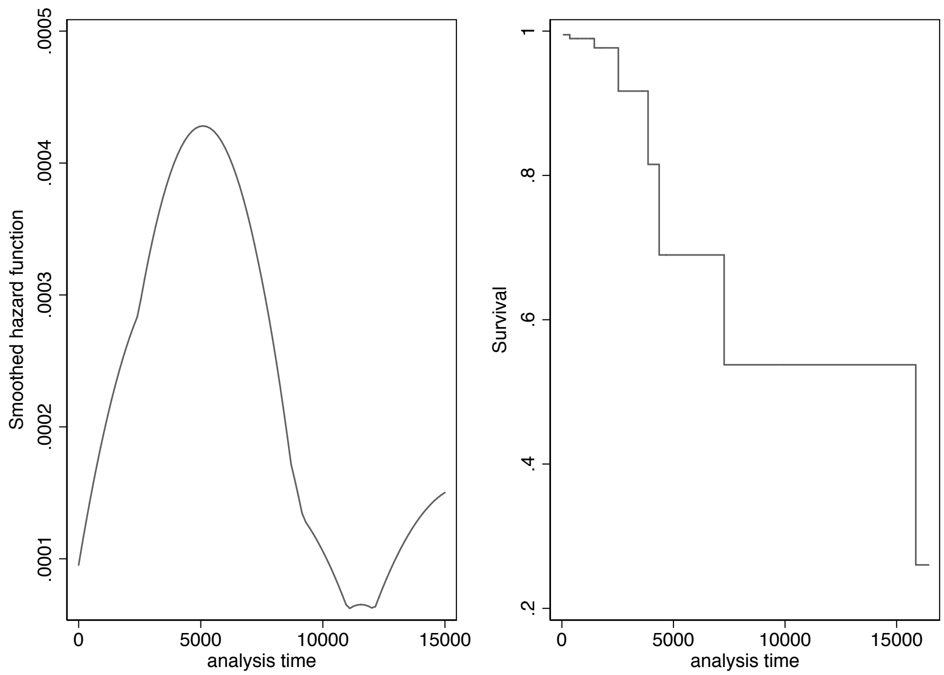
line hazard, though it assumes that the variables have a proportional, i.e. time-constant, effect on the hazard. Grambsch and Therneau's (1994) test reveals no violations of the PH assumption in either individual variables or the full model, so I make no adjustments for non-proportional hazards.

Results

Table 2 presents three Cox PH models, where positive hazard coefficients indicate an increase and negative coefficients a decrease in the risk of settlement collapse. Model 1 estimates the effect of the control variables on 81 war-ending settlements that followed military victories according to COW,[16] both bilateral and multilateral, showing that the number of

[16]Other categories include compromise/tied, stalemate, and transformations into other types of war (extra-state, intra-state, non-state, etc.), and conflict continuing at sub-war

Figure 3: Estimated baseline hazard and survival functions

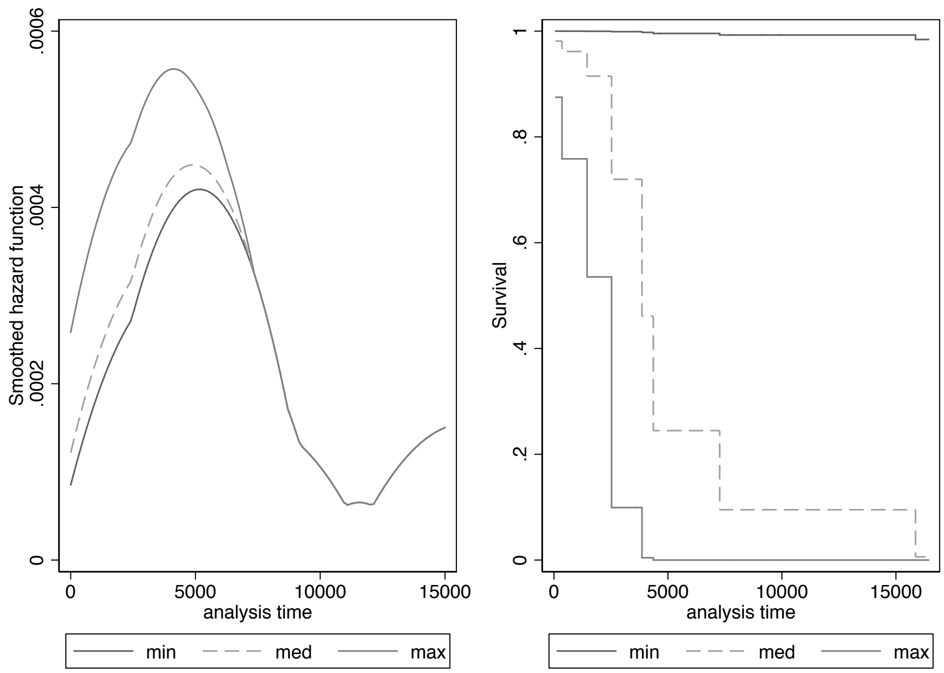


parties involved and the duration of the war have the expected relationships with the stability of settlements even in a highly aggregated unit of observation. Model 2 limits the sample to only those 23 settlements built on coalition victories, and Model 3 includes bioth theoretical and control variables. Figure 3 presents the estimated baseline hazard (left panel) and survivor (right panel) functions. Across settlements are at their greatest risk about 16 years (6000 days) after the end of the previous war; the hazard function, which plots the daily risk of war, peaks at that point, and the survivor function, which plots the probability of survival through day t conditional on having reached day t, levels off. But the risk of war over the within bargain can change raise and lower these risks substantially.

Most important for the hypothesis is the positive and statistically discernible ($p < 0.05$) coefficient on *Within Risk* in Model 3, which indicates that the risk of war over the across

levels (Sarkees, 2010).

Figure 4: Estimated hazard and survival functions by risk of intra-coalition war



bargain increases in the predicted risk of war over the within bargain. When defeated states anticipate cooperative problems over the within bargain inside the coalitions that defeated them, deterrence fails, and they're more inclined to go to war over the across bargain. The hazard ratio for *Within Risk* is 2.12; each one-unit increase in the variable more than doubles the risk that the across bargain collapses in war. The range of *Within Risk* is about ten units, so if we think in terms of deciles, each 10% increase along the spectrum from no risk of war between former coalition partners to certain war between former coalition partners more than doubles the risk of war between former enemies. Put differently, coalitions at the greatest risk of internal breakdown experience a risk of war with former opponents 1024 times greater than those at no risk of intra-coalition war.

Figure 4 represents this relationship graphically. The left panel plots the hazard function, i.e. the daily risk of war between former enemies, at the minimum, median, and maximum values of *Within Risk* over time. The right panel plots the survival function, or the probability that a settlement endures until a given day, as a function of the same values of *Within Risk*. The estimated hazards of minimal, median, and maximal risks of intra-coalition war converge at around 20 years after the end of the previous war. Before that convergence, the risk of the across bargain collapsing increases dramatically in the estimated risk of war inside the war-winning coalition, which Propositions 2 and 3 show are both related to shifting power inside war-winning coalitions. The survivor function tells the same story in different form, with the most stable war-winning coalitions seeing high probabilities of across-bargain survival even 27 years after the end of the previous war. Median and maximal values of *Within Risk*, however, see dramatic drops in their rate of survival before 13 years (5000 days on the x-axis); the most stable war-winning coalitions don't fall below a 90% chance of survival even thirty years after victory, while the most fragile war-winning coalitions see their hard-won across bargains survive at a rate below 50% before ten years have passed after victory.

Conclusion

Peace settlements are multidimensional, their success predicated on the simultaneous viability of bargains negotiated both across and within formerly warring sides. That very complexity can help sustain cooperation under multilateral conditions that would surely lead to war in bilateral settings. The threat of a revisionist war can encourage declining states inside victorious coalitions to tolerate otherwise intolerable shifts in military capabilities, allowing rising states to avoid war with both former partners and former enemies, when in isolation former partners and enemies would both launch preventive wars. But these settlements can fail, collapsing in renewed war with a former enemy, when the winning coalition itself struggles to cooperate in the face of one member's rise. Defeated states, like the Ottoman Empire, can return to war because the within bargain has already collapsed in violence, the fate that befell the Balkan League in the summer of 1913, or because they anticipate a weakened commitment to collective defense also prompted by shifting intra-coalition power. The same factors lead to both intra-coalition war and failures of collective deterrence, and this relationship between the anticipated risk of war between members of war-winning coalitions and the occurrence of war with former enemies emerges in a sample of 23 settlements produced by coalition victories from 1816-2007. Durable peace settlements therefore rest on simultaneously credible commitments to bargains struck both *across* and *within* formerly warring sides.

Few studies of the duration of peace settlements consider the problem of shared victory. Werner (1999, p. 925) comes closest, finding no relationship between the number of parties involved in a war and failures of the across bargain.[17] The present analysis makes sense of this apparently null relationship; multilateral victories may produce settlements no more or less fragile than unilateral victories, but the politics of the winning coalition play a heretofore unappreciated role in the success of those settlements. Shifting power between D and B, for example, may not lead to war when D's present capabilities are already bolstered by a partner willing to come to the rising state's aid. Observed shifts in power appear to have an inconsistent relationship with recurrent war both within and across dyadic studies (see Lo et al., 2008; Werner and Yuen, 2005), but that inconsistency may be a function of misspecified empirical models that (a) treat settlement-dyads as independent and (b) fail to account for the coalition politics of sharing in and defending the status quo. The standard approach is to calculate robust standard errors clustered on former-enemy dyads produced by the same war (see Fortna, 2003; Lo et al., 2008; Quackenbush and Venteicher, 2008; Senese and Quackenbush, 2003; Werner and Yuen, 2005), but a correction for heteroskedasticity isn't a correction for model misspecification. Dyadic models are appropriate for some questions about the duration of settlements, but not questions that implicate the problems of shared

[17]The results in Lo et al.'s (2008) study are also robust to the inclusion of a variable indicating whether the preceding war was multilateral.

victory and collective deterrence.

The politics of shared victory isolates a particular strategic problem by which coalition members must agree over the distribution and defense of a new status quo, but multidimensional settlements aren't limited to coalition victories. The Korean War, for example, ended in a stalemate that created one across settlement and two within settlements, one apiece for the coalitions on each side in the war. In 1953, the Korean Peninsula—and by extension East Asia—was divided between two great power coalitions, with North Korea, China, and the Soviet Union on one side, South Korea, the United States, and their myriad partners on the other, each dividing political influence over their respective shares of the settlement and the burdens of deterring attempts by the other side to overturn the outcome of the previous war (see Reiter, 2009). The theory and evidence presented above extend readily to this more complex strategic environment. Should cracks in either rival coalition begin to show, should one side anticipate that collective deterrence will fail due to disagreements within the other coalition, then the overall settlement may be in jeopardy. Collective deterrence, as well as intra-coalition restraint of two dissatisfied Korean governments (Benson, 2012; Fang et al., 2014), was key to the settlement that ended the Korean War and remains central to its survival. The war produced a "strong" ceasefire agreement (Fortna, 2003, pp. 343, 346, 350), tying great powers to its enforcement and creating a Demilitarized Zone (DMZ) that helps control accidents. But mutual deterrence may fail, the obstacle of the DMZ notwithstanding, if the coalitions that guarantee the settlement appear fragile, even if the issues over which they disagree don't directly implicate the across settlement. More complicated models with coalitions on both sides and without a stark attack-or-pass choice for the defeated state may yield insights into a wider range of war outcomes and settlements, but the basic tension of sharing in the division and defense of the status quo is important for any coalition that finds itself party to a peace settlement—even those on the losing side.

The politics of multilateral settlements also figure prominently in questions of global order. To the extent that the balance of power is supported by a status quo coalition expected to come to its defense in a crisis, that balance should become more fragile when potential challengers anticipate that the bargain struck between the winning coalition from the last great power war is fragile. If we take the theoretical model as representing collective deterrence in general, then the successful deterrence equilibrium offers some insights into the United States-dominated postwar order. First, a strong opponent helps smooth out shifting power inside the status quo coalition. Against a formidable Soviet Union, cooperation in the Western bloc was relatively easy compared to the fissures that emerged after the Soviet collapse, when Russian military and economic power bottomed out, allowing a relatively weak Russia to seize parts of Georgia in 2008 and Ukraine in 2014, safe in the knowledge that a coordinated Western response wasn't forthcoming. Second, collective deterrence also rests on a stable distribution of power—at a minimum, one to which states can peacefully adjust the distribution of benefits—inside the status quo coalition. From Korea to the South and East China Seas to Eastern Europe, the durability of the international status quo, forged in 1945 and adjusted substantially after the Korean War and the collapse of the Soviet Union, depends as much on Chinese and Russian revisionism as it does on the continued survival of the coalition that has, thus far, presented a credible deterrent threat in the status quo's defense. Substantial changes in relative power inside the status quo coalition, especially to the extent that they undermine the credibility of commitments to fight together, may presage future challenges and even, at the extreme, great power war.

Appendix

Proof of Proposition 1. There are five payoff-relevant states between which states transition as a function of stochastic war outcomes and B's exogenous increase in military power:

1. Post-shift, A faces B after defeating D

2. D faces A after the latter has defeated B

3. Post-shift, D faces B after the latter has defeated A

4. Post-shift, three players, no previous war

5. Pre-shift, thee players, no previous war

To prove the existence of the proposed MPE, I demonstrate that strategy profiles entail mutual best responses in subgames characterized by these states.

State 1. In this two-player continuation, A and B (which has grown stronger) negotiate over a flow reduced to $(1-d)$. In equilibrium, B accepts in each period iff $z \le \bar{z}$, and A proposes $z^* = \bar{z}$. Begin with B's acceptance rule, such that it accepts when

$$\frac{(1-z)(1-d)}{1-\delta} \ge \frac{m_B + s}{m_A + m_B + s}\left(\frac{(1-d)^2}{1-\delta}\right),$$

or when

$$z \le \frac{m_A + d(m_B + s)}{m_A + m_B + s} = \bar{z}.$$

A will never propose any $z < \bar{z}$, because doing so leaves A strictly worse off than meeting B's acceptance constraint at equality. Therefore, it chooses between proposing $z = \bar{z}$ and attacking, opting for the former when

$$\frac{\bar{z}(1-d)}{1-\delta} \ge \frac{m_A}{m_A + m_B + s}\left(\frac{(1-d)^2}{1-\delta}\right),$$

which is sure to be true given the costs of war.

State 2. In this two-player continuation, D chooses whether to attack A, where the present flow of benefits is $\pi(1-d)+1-\pi$, since only the within bargain was at stake in the previous war. A has no moves. In equilibrium, D attacks in the first play of the stage game. This is a best response when

$$\frac{m_D}{m_A + m_D}\left(\frac{(\pi(1-d)+1-\pi)(1-d)}{1-\delta}\right) > 1-\pi+\delta\left(\frac{(\pi(1-d)+1-\pi)(1-d)}{1-\delta}\right),$$

or when

$$\pi > \frac{m_A + dm_D}{m_A + (1-d(1-d))m_D}.$$

State 3. In this two-player continuation, D chooses whether to attack B after the latter has grown stronger, where the present flow of benefits is $\pi(1-d)+1-\pi$, since only the within bargain was at stake in the previous war. D has no moves. In equilibrium, D never attacks. This is a best response when

$$\frac{1-\pi}{1-\delta} \geq \frac{m_D}{m_B + s + m_D}\left(\frac{(\pi(1-d)+1-\pi)(1-d)}{1-\delta}\right),$$

or when

$$\pi \leq \frac{m_B + s + dm_D}{m_B + s + (1-d(1-d))m_D}.$$

State 4. In this three-player continuation, B has grown stronger and the none of the pie has been destroyed. In equilibrium, D passes; A helps, proposes $y_1^* = \overline{y}_1$, and proposes $x_t^* = \overline{x}_1$; B accepts iff $y_1 \leq \overline{y}_1$ and iff $x_t \leq \overline{x}_1$. Begin with B's acceptance rule after D passes. If B rejects x_1, then if it wins the game transitions to State 3, where D never attacks. B thus accepts when

$$\frac{(1-x_1)\pi}{1-\delta} \geq \frac{m_B + s}{m_A + m_B + s}\left(\frac{\pi(1-d)}{1-\delta}\right),$$

or when

$$x_1 \leq \frac{m_A + d(m_B + s)}{m_A + m_B + s} = \overline{x}_1.$$

If A wishes to induce acceptance, it will not propose any $x_1 < \overline{x}_1$, so it chooses between $x_1^* = \overline{x}_1$ and attacking, which transitions the game to State 2, where D attacks in the first play of that continuation. Therefore, A proposes $x_1^* = \overline{x}_1$ when

$$\frac{\overline{x}_1 \pi}{1-\delta} \geq \frac{m_A}{m_A + m_B + s}\left(\pi(1-d) + \delta \frac{m_A}{m_A + m_D}\left(\frac{(\pi(1-d)+1-\pi)(1-d)}{1-\delta}\right)\right),$$

or when, as $\delta \to 1$,

$$\pi \geq \frac{m_A^2(1-d)}{m_A^2(1+d(1-d)) + d m_D(m_B + s) + m_A(s(m_B+s) + m_D)}.$$

A's and B's incentives after defeating D are identical to State 1, to which the game transitions are a single exchange of y_1, so A proposes $y_1^* = \overline{y}_1$, where

$$\overline{y}_1 = \frac{m_A + d(m_B + s)}{m_A + m_B + s} = \overline{z},$$

and B accepts iff $y_1 \leq \overline{y}_1$. Next, A helps B, setting up an exchange over y_1 and transition to State 1 after victory, rather than deny when

$$\frac{c(m_A + m_B + s)}{c(m_A + m_B + s) + m_D}\left(\frac{\overline{y}_t(1-d)}{1-\delta}\right) \geq \frac{m_A}{m_A + m_B + s + m_D}\left(\frac{1-d}{1-\delta}\right),$$

which is true for all parameters. Finally, D passes when

$$\frac{1-\pi}{1-\delta} \geq \frac{m_D}{c(m_A + m_B + s) + m_D}\left(\frac{1-d}{1-\delta}\right),$$

or when

$$\pi \leq \frac{c(m_A + m_B + s) + d m_D}{c(m_A + m_B + s) + m_D}.$$

State 5. In this three-player continuation, B has yet to grow stronger, and none of the pie has been destroyed. In equilibrium, D passes; A helps, proposes $y_0^* = \overline{y}_0$, and proposes $x_0^* = \overline{x}_0$; B accepts iff $y_0 \leq \overline{y}_0$ and iff $x_0 \leq \overline{x}_0$. Begin with B's acceptance rule after D passes.

If B rejects x_t, then if it wins the game transitions to State 3, where D never attacks; if B accepts, the game transitions to State 4. B thus accepts when

$$\pi(1-x_0)+\delta\frac{\pi(1-\overline{x}_t)}{1-\delta} \geq \frac{m_B}{m_A+m_B}\left(\frac{pi(1-d)}{1-\delta}\right),$$

or when

$$x_0 \leq d+\frac{m_A(1-d)}{(m_A+m_B)(1-\delta)}-\frac{\delta m_A(1-d)}{(1-\delta)(m_A+m_B+s)}=\overline{x}_0.$$

If A wishes to induce acceptance, it meets B's acceptance constraint at equality, so A proposes $x_0^*=\overline{x}_0$, rather than attack and if victorious transition to State 2, when

$$\overline{x}_0\pi+\delta\frac{\overline{x}_1\pi}{1-\delta} \geq \frac{m_A}{m_A+m_B}\left(\pi(1-d)+\delta\frac{m_A}{m_A+m_D}\left(\frac{(\pi(1-d)+1-\pi)(1-d)}{1-\delta}\right)\right),$$

or, as $\delta \to 1$, when

$$\pi \geq \frac{m_A^2(1-d)}{m_A^2(1+d(1-d))+dm_Bm_D+m_A(dm_B+m_D)}.$$

and when

$$s \leq \frac{(m_A+m_B)\left(\pi m_A(dm_B+m_D)+m_A^2(d(-d\pi+\pi+1)+\pi-1)+d\pi m_Bm_D\right)}{-d\pi m_A(m_B+m_D)+m_A^2((d-2)d\pi-d+1)-d\pi m_Bm_D}=\overline{s}_1 \qquad (4)$$

and

$$m_D > m_A\left(\frac{(d-1)m_A(d\pi-1)}{\pi(m_A+dm_B)}-1\right)=\hat{m}_D, \qquad (5)$$

where the last two constraints jointly ensure that s is not so large that A will forego any bargain, even $x_0=1$, and attack instead. When the coalition bargains after defeating D, B's acceptance transitions the game to State 1; war transitions the game to a trivial continuation in which the victor controls the now twice-reduced flow of benefits. B accepts y_0 when

$$(1-y_0)(1-d)+\delta\frac{(1-y_1)(1-d)}{1-\delta} \geq \frac{m_B}{m_A+m_B}\left(\frac{(1-d)^2}{1-\delta}\right),$$

or when

$$y_0 \leq d+\frac{m_A(1-d)}{(1-\delta)(m_A+m_B)}-\frac{\delta m_A(1-d)}{(1-\delta)(m_A+m_B+s)}=\overline{y}_0.$$

If A wishes to induce acceptance, it meets B's acceptance constraint at equality by proposing $y_0^*=\overline{y}_0$, which it is sure to do as long as $\delta \to 1$ and

$$s \leq \frac{d(m_A+m_B)^2}{1-2d)m_A-dm_B}=\overline{s}_2, \tag{6}$$

which again ensures that A is willing to make a proposal acceptable to B rather than choose attack over even the most generous proposal. Anticipating a peaceful bargain with B over y_0, A helps when

$$\frac{c(m_A+m_B)}{c(m_A+m_B)+m_D}\left(\overline{y}_0(1-d)+\delta\frac{\overline{y}_1(1-d)}{1-\delta}\right) \geq \frac{m_A}{m_A+m_B+m_D}\left(\frac{1-d}{1-\delta}\right),$$

which is sure to be true since A is willing to propose $y_0^*=\overline{y}_0$. Finally, D passes in each period when

$$\frac{1-\pi}{1-\delta} \geq \frac{m_D}{c(m_A+m_B)+m_D}\left(\frac{1-d}{1-\delta}\right),$$

or when

$$\pi \leq \frac{c(m_A+m_B)+dm_D}{c(m_A+m_B)+m_D}.$$

Summarizing the constraints, $\hat{m}_D$ is as defined in (5), lower and upper bounds on π are

$$\underline{\pi} = \max\left\{\frac{m_A + dm_D}{m_A + (1-(1-d)d)m_D}, \frac{(d-1)m_A^2}{-m_A(dm_B + m_D) + ((d-1)d-1)m_A^2 - dm_B m_D}\right\}$$

and

$$\overline{\pi} = \min\left\{\frac{m_B + dm_D + s}{m_B + (1-(1-d)d)m_D + s}, \frac{c(m_A + m_B) + dm_D}{c(m_A + m_B) + m_D}\right\},$$

while $\hat{s} = \min\{\overline{s}_1, \overline{s}_2\}$ as defined by Equations (4) and (6). □

Proof of Proposition 2. To prove the claim, it is sufficient to show that A prefers attacking in period t_0 to striking the most favorable possible bargain, $x_0 = 1$, before transitioning to state 1 when D will and won't attack in the first period.

First, A prefers attacking B at t_0 to $x_0 = 1$ when

$$\frac{m_A}{m_A + m_B}\left(\pi(1-d) + \frac{\delta}{1-\delta} CV_A^D\right),$$

where

$$CV_A^D = \frac{m_A}{m_A + m_B}(\pi(1-d) + 1 - \pi)(1-d).$$

if D attacks in state 2. This inequality is satisfied when, as $\delta \to 1$,

$$s > \frac{(m_A + m_B)\left(m_A\left(CV_A^D - \pi\right) - d\pi m_B\right)}{d\pi m_B - m_A\left(CV_A^D - d\pi\right)} \quad \text{and} \quad d < \frac{m_A CV_A^D}{\pi(m_A + m_B)},$$

where the first term is equal to $\overline{s}_1$ in Equation (4). Next, if D won't attack in state 1, then $CV_A^D = \pi(1-d)$, and A is sure to attack B when, as $\delta \to 1$,

$$s > \frac{d(m_A + m_B)^2}{(1-2d)m_A + dm_B} \quad \text{and} \quad d < \frac{m_A}{2m_A + m_B}.$$

Therefore, whether D attacks a victorious A or passes in state 2, A's strategy profile entails attacking B at the first opportunity in t_0 if D passes. □

Proof of Proposition 3. To prove the claim, it is sufficient to identify sufficient conditions at time t_0 for A to deny B help and for D to attack. First, as demonstrated in the proofs of Propositions 1 and 2, A is sure to help B unless it will attack B after defeating D. Next, A prefers to attack B rather than strike a bargain at $y_0 = 1$ when

$$\frac{m_A}{m_A + m_B}\left(\frac{(1-d)^2}{1-\delta}\right) > (1-d) + \delta\frac{\overline{y}_1(1-d)}{1-\delta}$$

or when

$$s > \frac{d(m_A + m_B)^2}{(1-2d)m_A + dm_B} \quad \text{and} \quad d < \frac{m_A}{2m_A + m_B}.$$

Then, A refuses to help when

$$\frac{m_A}{m_A + m_B + m_D}\left(\frac{1-d}{1-\delta}\right) > \frac{c(m_A + m_B)}{c(m_A + m_B) + m_D}\left(\frac{m_A}{m_A + m_B}\left(\frac{(1-d)^2}{1-\delta}\right)\right)$$

or when either $c \leq 1/(1-d)$ or

$$c > \frac{1}{1-d} \quad \text{and} \quad m_D < \frac{cd(m_A + m_B)}{c(1-d) - 1}.$$

Finally, D attacks when

$$\frac{m_D}{m_A + m_B + m_D}\left(\frac{1-d}{1-\delta}\right) > 1 - \pi + \frac{\delta}{1-\delta}CV_D^C,$$

where CV_D^C is D's continuation value for post-shift states that follow passing at time t_0. If

A and B will strike a deal at t_0 and fight together for $t \geq t_1$, such that

$$CV_D^C = \frac{m_D}{m_A + m_B + m_D}(1-d),$$

then D attacks at t_0 rather than pass when

$$\pi > \frac{m_A + m_B + dm_D}{m_A + m_B + m_D}.$$

If A will win a war against B at t_0, such that D attacks a lone A in the relevant state,

$$CV_D^C = \frac{(1-d)((1-d)\pi - \pi + 1)m_D}{m_A + m_D},$$

and D attacks when

$$\pi > \frac{m_B}{d(m_A + m_B + m_D)}.$$

If B will win at war, leading D to pass in the relevant state, then $CV_D^C = 1 - \pi$, and D attacks when

$$\pi > \frac{m_A + m_B + dm_D}{m_A + m_B + m_D}.$$

Therefore, when A denies B help, a sufficient condition for D to attack is

$$\pi > \max\left\{\frac{m_A + m_B + dm_D}{m_A + m_B + m_D}, \frac{m_B}{d(m_A + m_B + m_D)}\right\}.$$

Thus, when d and c are sufficiently small and s and π are sufficiently large, all MPE entail D attacking B and A denying help. $\square$

References

Arena, P. and A. Pechenkina (2016). External subsidies and lasting peace. *Journal of Conflict Resolution 60*(7), 1278–1311.

Benson, B. V. (2012). *Constructing International Security: Alliances, Deterrence, and Moral Hazard*. Cambridge: Cambridge University Press.

Box-Steffensmeier, J. M. and B. S. Jones (2004). *Event History Modeling: A Guide for Social Scientists*, Volume New York. Cambridge University Press.

Fang, S., J. C. Johnson, and B. A. Leeds (2014). To concede or resist? the restraining effect of military alliances. *International Organization 68*(4), 775–809.

Fazal, T. M. (2007). *State Death: The Politics and Geography of Conquest, Annexation, and Occupation*. Princeton: Princeton University Press.

Fazal, T. M. (2014). Dead wrong? battle deaths, military medicine, and exaggerated reports of war's demise. *International Security 39*(1), 95–125.

Fortna, V. P. (2003). Scraps of paper? agreements and the durability of peace. *International Organization 57*(2), 337–372.

Fortna, V. P. (2004). Interstate peacekeeping: Causal mechanisms and empirical effects. *World Politics 56*(4), 481–519.

Fortna, V. P. (2008). *Does Peacekeeping Work: Shaping Belligerents' Choices After Civil War*. Princeton: Princeton University Press.

Gallop, M. (2017). More dangerous than dyads: How a third party enables rationalist explanations for war. *Journal of Theoretical Politics 29*(3), 353–381.

Glenny, M. (2012). *The Balkans: Nationalism, War, and the Great Powers 1804-2011*. New York: Penguin.

Grambsch, P. M. and T. M. Therneau (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika 81*(3), 515–526.

Hultman, L., J. D. Kathman, and M. Shannon (2015). United nations peacekeeping dynamics and the duration of post-civil conflict peace. *Conflict Management and Peace Science 33*(3), 231–249.

Kecskemeti, P. (1958). *Strategic Surrender: The Politics of Victory and Defeat*. Stanford: Stanford University Press.

Leeds, B. A. (2003). Do alliances deter aggression? the influence of military alliances on the initiation of militarized interstate disputes. *American Journal of Political Science 47*(3), 427–439.

Licht, A. (2014). Regular turnover details. Avaliable at https://sites.google.com/site/amandalichtphd/research/regular-turnover-details.

Lo, N., B. Hashimoto, and D. Reiter (2008). Ensuring peace: Foreign-imposed regime change and postwar peace duration, 1914-2001. *International Organization 62*(4), 717–736.

Maskin, E. and J. Tirole (2001). Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory 100*(2), 191–219.

Mattes, M., B. A. Leeds, and N. Matsumura (2016). Measuring change in source of leader support: The chisols dataset. *Journal of Peace Research 53*(2), 259–267.

Mattes, M. and B. Savun (2009). Fostering peace after civil war: Commitment problems and agreement design. *International Studies Quarterly 53*(3), 737–759.

Mattes, M. and B. Savun (2010). Information, agreement design, and the durability of civil war settlements. *American Journal of Political Science 54*(2), 511–524.

Morey, D. S. (2016). Military coalitions and the outcome of interstate wars. *Foreign Policy Analysis 12*(4), 533–551.

Owsiak, A. P., P. F. Diehl, and G. Goertz (2017). Border settlement and the movement toward and from negative peace. *Conflict Management and Peace Science 34*(2), 176–193.

Owsiak, A. P. and T. J. Rider (2013). Clearing the hurdle: Border settlement and rivalry termination. *Journal of Politics 75*(3), 757–772.

Palmer, G., V. D'Orazio, M. Kenwick, and M. Lane (2015). The mid4 dataset, 2002-2010: Procedures, coding rules and description. *Conflict Management and Peace Science 32*(2), 222–242.

Poast, P. (2016). Dyads are dead, long live dyads! the limits of dyadic designs in international relations research. *International Studies Quarterly 60*(2), 369–374.

Powell, R. (1999). *In the Shadow of Power*. Princeton: Princeton University Press.

Powell, R. (2006). War as a commitment problem. *International Organization 60*(1), 169–203.

Quackenbush, S. L. and J. F. Venteicher (2008). Settlements, outcomes, and the recurrence of conflict. *Journal of Peace Research 45*(6), 723–742.

Reiter, D. (2009). *How Wars End*. Princeton: Princeton University Press.

Reiter, D., A. C. Stam, and M. C. Horowitz (2016). A revised look at interstate wars, 1816-2007. *Journal of Conflict Resolution 60*(5), 956–976.

Sarkees, M. R. (2010). Inter-state wars (version 4.0): Definitions and variables. Available at http://www.correlatesofwar.org/data-sets/COW-war, Accessed 9 April 2018.

Sarkees, M. R. and F. Wayman (2010). *Resort to War: 1816-2007*. Washington: CQ Press.

Senese, P. D. and S. L. Quackenbush (2003). Sowing the seeds of conflict: The effect of dispute settlements on durations of peace. *Journal of Politics 65*(3), 696–717.

Singer, J. D. (1987). Reconstructing the correlates of war dataset on material capabilities of states, 1816-1985. *International Interactions 14*(2), 115–132.

Slantchev, B. (2004). How initiators end their wars: The duration of warfare and the terms of peace. *American Journal of Political Science 48*(4), 813–829.

Slantchev, B. (2005). Territory and commitment: The concert of europe as a self-enforcing equilibrium. *Security Studies 14*(4), 565–606.

Smith, A. (1995). Alliance formation and war. *International Studies Quarterly 39*(4), 405–425.

Starr, H. (1972). *War Coalitions: The Distributions of Payoffs and Losses*. Lexington, MA: Lexington Books.

Wagner, R. H. (2007). *War and the State: The Theory of International Politics*. Ann Arbor: The University of Michigan Press.

Werner, S. (1999). The precarious nature of peace: Resolving the issues, enforcing the settlement, and renegotiating the terms. *American Journal of Political Science 43*(3), 912–934.

Werner, S. and A. Yuen (2005). Making and keeping peace. *International Organization 59*(2), 261–292.

Wolford, S. (2017). The problem of shared victory: War-winning coalitions and postwar peace. *Journal of Politics 79*(2), 702–716.