A Model of Offensive versus Defensive Technology: When Does Terrorism Occur?

Jin Yeub Kim*

June 11, 2020

Abstract

Offense-defense theory argues that conflict and war are more likely when offense has the advantage over defense. This paper explores how the prime hypotheses in offense-defense theory can be challenged when it comes to making predictions about conflicts or terrorism caused by revisionist states. To do so, I propose a model of a continuous time conflict game between two asymmetric states, in which a status quo state develops defensive technology and a revisionist state develops offensive technology, with incomplete information on the revisionist state's militancy level. The equilibrium analysis shows that the risk of terrorism is maximized when the overall offense-defense balance is relatively even. This paper suggests an alternative framework in explaining asymmetric conflicts in the contemporary context, and provides insights into understanding the likelihood of terrorism.

Keywords: military technology; offense-defense balance; conflict; terrorism.

Word count: 9877

*School of Economics, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul 03722, Republic of Korea (email: shiningjin@gmail.com).

Introduction

In a series of influential articles on international relations, offense-defense theory has become the most powerful and useful neorealist theory on the causes of war, and it has accomplished a lot of explanatory and theoretical work for decades.[1] Many neorealist scholars have noted the likelihood of war is powerfully affected by changes in technology. Military technologies can influence decisions for or against war, and depending on whether the 'offense' or 'defense' has the edge, states will choose to use forces against the status quo.

Offense-defense theory conceptualizes the *offense-defense balance* by examining the comparative statics of changes in the relative efficacy of offensive and defensive fighting technologies in models of conflict. More formally, advantage refers to a difference between the relative expected value or utility of attacking versus defending, holding other factors constant. Any technological factor that tends to increase the difference shifts the offense-defense balance towards the offense. In international relations scholarship and foreign policy analysis, the offense-defense balance has been used pervasively to explain international outcomes, especially the occurrence of violent conflict and the likelihood of interstate war.

Among prominent conjectures that offense-defense theorists contend about the causes of violent conflict and war, one of the far-reaching claims that Jervis (1978) initiated and Van Evera (1998) explicated is that offensive advantages exacerbate the security dilemma, making wars more likely in periods when the offense is dominant than in other periods.[2] In another exposition of offense-defense theory, Fearon (1997) disputed the claim, arguing instead that greater offensive advantages may have countervailing effects on the probability of war and thus, yield the possibility of peace.

The vast bulk of the literature on conflicts still tends to focus on the historical and empirical assessments of specific wars and patterns of warfare in history between great powers. However, conventional interstate wars between great powers ended in the aftermath of the Cold War, and many new forms of conflicts beyond the full control of nation-states have emerged. In particular, terrorism, whether it is fed by religious fundamentalism or not, has become a serious threat in international society.[3] The hallmark of terrorism is the effective wielding of force by a weak contender against an opponent that is far superior in conventional military power. But the fact that terrorists do not fight on clear fronts and do not play according to the rules of war makes struggling with terrorism extremely difficult.

Scholars have tried to employ offense-defense theory to explain a wide array of political relations in the post-Cold War era, including nuclear strategy, arms races, deterrence, and ethnic conflict. Yet, the literature lacks rigorous formal theory, such as the perspective of game theory or rational choice analysis, to explain terrorism caused by *revisionist states*.[4] Therefore, a comprehensive theoretical analysis of the conflicts between a revisionist state and a status quo state is needed. My interest is in understanding the conflicts between these asymmetric states and the likelihood of terrorism through a game-theoretic model of the wedge between offensive versus defensive technology; and in exploring how the conventional hypotheses in offense-defense theory can be challenged when applied to such asymmetric conflicts in the contemporary context.[5]

In particular, I study the offense-defense balance in a continuous time model of conflict between a large state and a small state. The large state (A) is a status quo power and at each instant in time can either attack or continue. The small state (B) is a provocateur with private information on its militancy level, and at each instant can attack, back down, or continue. The game ends in war by either side's attack, or ends in capitulation of B by B's back down. If both states continue, the conflict game continues in stalemate. The B's continue strategy can represent threats of terrorism by provocative military actions without an explicit attack.[6]

The offense-defense balance is understood as the technological advantages of attacking versus being attacked, conditional on a war occurring. I allow for differences in these technologies across the two states: The A develops defensive technology and B develops offensive technology over time. This assumption is the key source for asymmetry between states. In particular, A can be thought of as a pacifist, limiting the use of military technology to defense, while the provocateur B actively uses force to change situations of which it disapproves.[7]

I obtain two main theoretical results on the equilibria set of my model: (1) In any equilibrium, there exists a unique finite *duration of stalemate*–a level of continuation after which B is not expected to back down; and (2) in equilibrium, A chooses to attack right after the duration, while B will continue up to the duration and attack if and only if her privately known militancy level is sufficiently large, and will back down on or before the duration with some probability distribution if and only if her privately known militancy level is relatively low.

The results of my model suggest surprising reversals of the classical hypotheses produced by offense-defense theory that war is more likely in offense-dominant eras, as well as the argument of Fearon (1997) that greater offensive advantages may make states favor peace more. First, in my setting, terrorism is most likely in periods when the system has relatively even offense-defense balance than in other periods. More precisely, offensive or defensive advantages are not decisive in determining the likelihood of war over stalemate or peace. Second, greater offensive advantages never increase a state's inclination toward peace.

The existing literature, focused as it is on models of conflict between symmetric states, has never suggested that an even offense-defense balance might maximize the risk of war. As my model offers an interesting departure that yields new conjectures, this paper thereby calls for a theoretical refinement of offense-defense theory in terms of understanding the contests between large status quo and small revisionist states that today fuel terrorism and stalemate.

The Model

My model presented here is a variant of Fearon's (1995) model; hence some terminologies and derivations are adopted from his paper. Consider a conflict game in which two asymmetric states, a status quo A ('he') and a revisionist B ('she'), are involved in a conflict.[8] The conflict occurs in continuous time, starting at $t = 0$. For every finite $t \geq 0$, A can choose whether to attack or continue, and B can choose to attack, back down, or continue. The conflict ends when A attacks, B attacks or B backs down. When the conflict ends, it is the end of strategic interactions. The conflict is said to be in stalemate when both sides continue.

At every instant of time $t \geq 0$, A has defensive capability of size $m_A(t)$ and B has offensive capability of size $m_B(t)$. Functional forms of $m_i(t)$, $i = A, B$, are common and public knowledge. Also, $m_A(t)$ and $m_B(t)$ are continuous and strictly increasing functions of the amount of continuation with $m_i(0) = 0$.[9] I consider the linear cases, $m_A(t) = \pi t$ and $m_B(t) = \mu t$, where π and μ are parameters indicating how rapidly continuation strengthens military capabilities for each state. Let $\pi > 0$ be A's defensiveness, or value for peace, which is common knowledge. Let $\mu > 0$ be B's offensiveness, or militancy level, which is private information.

At the outset of the game, μ is chosen by nature and B learns her own true militancy level. The A knows only the distribution of B's militancy level. He believes that μ is drawn from a nonatomic distribution over the interval $M = [0, \bar{\mu}]$, $\bar{\mu} > 0$, according to a cumulative distribution function F that has continuous and strictly positive density f. The density f represents A's prior beliefs about B's militancy level.

Given the above parameters and information structure, increasing capability imposes a technological cost of being attacked to A, and a technological benefit of attacking to B, if there is a continuation of conflict ended by B's attack. That is, when B attacks, A suffers a cost from his defensive capability being destroyed by B's larger offensive capability, as offensive capability caused by stalemate has accumulated more, while B benefits from her offensive capability. Thus, military technology can be interpreted as martial effectiveness, or continuation value (benefit or cost) of stalemate from technology.

If B attacks at t, the military technology yields a time-contingent benefit $m_B(t)$ to B and a time-contingent cost $m_A(t)$ to A. In the linear case that I focus on, μ indicates how rapidly continuation creates benefit for B and π indicates how rapidly continuation creates cost for A. Each A and B suffers additional fixed military costs of $d_A > 0$ and $d_B > 0$, respectively.[10] If A attacks at t, A suffers a cost $c_A > 0$ and B suffers a cost $c_B > 0$.

Payoffs are realized only when the conflict ends, given as follows. If A attacks at t, two states receive their expected payoffs $-c_A$ for A and $-c_B$ for B. If B backs down at t, payoffs are $-s$ for A and s for B, where $s \geq 0$. This s can be interpreted as A's offer of a surplus to B in return for B backing down and terminating the conflict. If B attacks at t, two states receive expected payoffs $-m_A(t) - d_A$ for A and $m_B(t) - d_B$ for B. Note that the asymmetry plays out only when an attack is instigated by B, who benefits from being an attacker and incurs cost to A for being attacked. The game tree illustrated in Figure 1 depicts the structure of the game, with payoffs indicated in the cases that A attacks first at t' and B attacks or backs down first at t''.

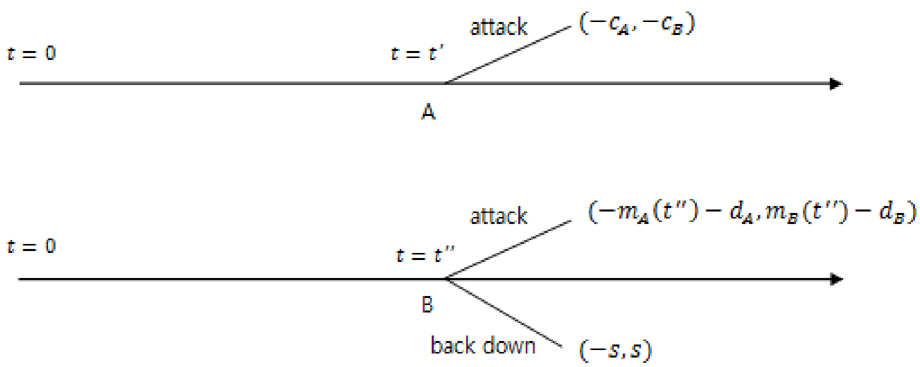


Figure 1: The game tree

For A to find it in his interest to not attack at $t = 0$, it must be $s < c_A$. Also, B will respond to A's offer s only if the offer is credible, so A's offer must not exceed his cost when B attacks, i.e. $s < \pi t + d_A$ for all t; so, $s < d_A$. The restrictions on the fixed costs A and B suffer when A attacks are $c_B > d_B$ and $c_A > d_A$.[11] I take attacking to be inefficient, so that $\pi \geq \bar{\mu}$. With these parameter restrictions, I study perfect Bayesian equilibrium. The formal statement of the solution concept is given in Appendix A.

Main Results

I begin by introducing an informal description of the complete information case, in which A knows B's militancy level with certainty. I have omitted a discount factor to make the analysis more tractable, focusing on the independent effects of the wedge between offensive and defensive capability over time in continuation of a conflict game.

If A knew B's level of militancy, he could look ahead and anticipate what would happen if B were to continue and create a stalemate. That is, A would see that ultimately the benefit for B from increasing offensive capability, regardless of B's militancy level, would be so large that B would in effect commit herself not back down after some point, say t'_μ.[12] Anticipating this at the outset, a rational A would use a strategy of 'continue up to t'_μ and attack for $t > t'_\mu$' to prevent himself from paying the larger costs that would go with continuation.[13] Knowing this, a rational B would back down immediately, because the best she could do is to back down before t'_μ. The game ends in peace, with B backing down and having no incentive to launch a conflict game at the outset, and no conflict would occur.

With complete information, a provocateur will have to settle immediately rather than suffer costs from a status quo state's preventive attack. I can now ask, if B's militancy level is unknown to A and if it were known that B with higher militancy level preferred continuing to ending the conflict, should B with lower militancy level necessarily back down immediately?

Equilibria under Incomplete Information

I now consider equilibria of my model in which B has private information about her militancy level. With preliminary results and all proofs given in Appendix B, I provide the following theoretical results.

First, I establish the existence of a unique finite *duration of stalemate*—a level of continuation after which B is not expected to back down and thus war is certain—in some equilibrium in which a conflict may occur.

Proposition 1. *Let t^* be such that $-\pi t^* - d_A = -c_A$. In any equilibrium where A believes B will attack at some t with positive probability, there exists a unique finite duration of stalemate t^*.*

Proposition 1 shows that if a finite duration of stalemate exists, then it is uniquely defined as $t^* = \frac{c_A - d_A}{\pi}$. This result guarantees that A is willing to attack after $t = t^*$ as a preventive measure to end the conflict.

Proposition 2 characterizes the set of pure strategy equilibria in the incomplete information conflict game in which a finite duration of stalemate occurs.[14]

Proposition 2. *Let the unique finite duration of stalemate be* $t^* = \frac{c_A - d_A}{\pi}$. *The following describes the set of equilibrium strategies for A and equilibrium strategies for B as a function of* μ*:*

(i) A plays $\{attack\}_A^t$ *for* $t > t^*$.

(iia) For $\mu \geq \frac{s+d_B}{t^*}$, *B plays* $\{attack\}_B^t$ *at* $t = t^*$.

(iib) For $\mu < \frac{s+d_B}{t^*}$, *B plays* $\{backdown\}_B^t$ *for any* $t \leq t^*$.

The belief system in equilibrium is given as follows:

(iii) For $t \leq t^*$, *A believes that the probability B will not back down is* $1 - F\left(\frac{s+d_B}{t^*}\right)$.

(iv) For any $t > t^*$ *off-the-equilibrium path, A believes that B has* $\mu > \frac{s+d_B}{t^*}$ *and is distributed according to F truncated at* $\frac{s+d_B}{t^*}$.

Proposition 2 asserts that there is a set of equilibria, where A's optimal strategy would be to attack right after t^* and B's optimal strategy would be to attack at t^* if she has a high enough militancy level, or to back down on or before t^* with some positive probability if she has a lower militancy level. The interpretation of these strategies is as follows. The status quo state will not strike first to disarm a revisionist state prior to some fixed duration of stalemate, because there is a chance that the revisionist state might react by backing down. The status quo state with defensive objectives, however, believes that if he does not attack after that time, his enemy will do so, forcing destructive war upon them on less favorable terms and at higher costs. Thus, he commits himself to attack for sure right after the duration. If the revisionist state is very aggressive (high μ), then she will continue stalemate up to the duration and attack at the very moment the duration is reached; while if the revisionist state is not so aggressive (low μ), she will back down, ending the conflict on or before the duration with positive probability.

Note that in any equilibrium in which conflict occurs, the probability distribution on equilibrium outcomes is indeterminate up to the duration t^*.[15] That is, on or before t^*, B may choose any time to back down, because the payoff structure leaves this open by not incorporating any incentives for either an immediate back down or prolonged stalemate. Hence, up to the duration t^*, stalemate may continue, and once war seems likely (i.e. t^* is reached), there are strong pressures to pre-empt an attack. Each side knows the other sees the situation the same way, thus increasing the perceived danger that the other would attack, and giving each added reasons to precipitate an attack if conditions seem favorable.

Illustrative Example

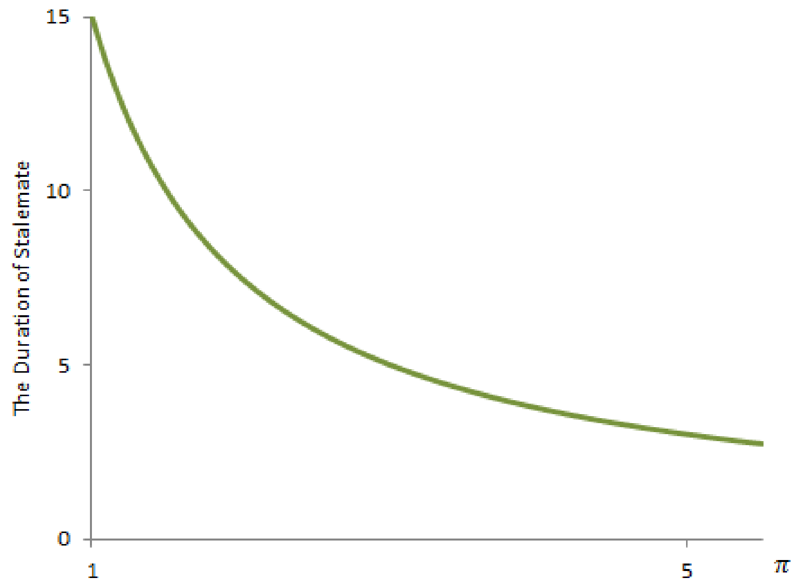
A comparative statics analysis of the equilibrium characterized in Proposition 2 yields theoretical insights into how the slopes of military technology matter in equilibrium behavior and affect conflict outcomes. A striking feature is that player A who possesses lower defensiveness is more likely to endure longer duration of stalemate because t^* proves to rise as π falls. On the other hand, player A with very large defensiveness would commit to strike first early, inducing equilibrium behavior of B to be more likely to back down prior to the duration of stalemate.

Intuition can justify the opposite prediction. One might think that an A who values peace less would fear conflict less and strike first faster. But this argument misses the point that the cost of being attacked by the adversary, generated by increasing defensive technology, rises more slowly. While such an A may not fear attack, if he does choose to continue the conflict, the lesser costs created by lower value for peace mean that he is more likely to be patient with stalemate, conditional on having started a conflict. Therefore, as A endures a longer duration of stalemate, B would then be less inclined to back down prior to the duration. Figure 2 illustrates the duration of stalemate and the prior probabilities of B's equilibrium strategies in terms of A's defensiveness.

As can be seen from Figure 2, as A's defensiveness π increases, he has fewer expected periods of waiting and the risk of war decreases. Formally, the probability that B will back down prior to the duration of stalemate t^* is $F\left(\frac{s+d_B}{t^*}\right)$, which increases in π; and the probability that B will attack at t^* is $1 - F\left(\frac{s+d_B}{t^*}\right)$, which decreases in π. If A's value for peace is lower, then B is less likely to back down than when facing A with higher value for peace. This result holds regardless of the surplus amount and fixed costs of attack, and regardless of A's initial belief about B. Thus, when the status quo state's desire for preserving peace for its own sake is not that great, the revisionist states would back down less often compared to when the revisionist faces a very peace-loving status quo state. Conversely, A with lower value for peace will face higher probability that B will attack at a larger t^*, conditional on a conflict occurring.

The comparative statics of how A's defensiveness matter on the duration of stalemate in the equilibrium has two implications. First, a high value of defensiveness helps the status quo state, by making the revisionist state more likely to shy away from conflicts and more likely to back down once in them. That is, the risk of war given a conflict decreases as π increases above $\bar{\mu}$.

Second, as the expected duration of stalemate decreases with π, conflicts between



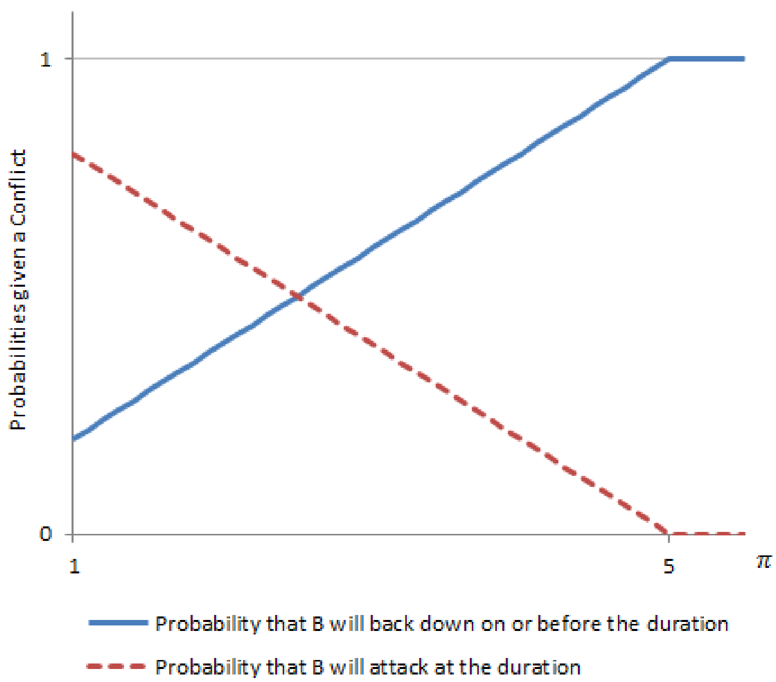


Figure 2: Comparative statics in terms of A's defensiveness

Notes: Parameter values for computation are $s = 1$, $c_A = 20$, $d_A = 5$, and $d_B = 2$. The μ is drawn according to a uniform distribution F over the interval $M = [0, 1]$. By the parameter restriction $\pi \geq \bar{\mu}$, the range of π starts from one.

a revisionist state and a more peace-loving status quo state should see significantly fewer periods of standoff than conflicts between a revisionist state and a less peace-loving status quo state.[16] In the model, A's ex ante expected payoff for continuing up to t^* decreases with its defensiveness π. This result provides a rationale for why, ex ante, a status quo state with higher defensiveness would want to strike first earlier in international conflicts, even though that might intuitively force wars upon them sooner (but with lower probability). That is, A's higher defensiveness implies that A is more stalemate-averse and more determined to pre-empt an attack in order to terminate the conflict in the first place.

The Offense-Defense Balance Revisited

Offense-defense theorists have made arguments about war and peace that remain conventional premises in international relations. Among those, Jervis (1978) and Van Evera (1998) developed the commonly accepted hypothesis that offensive advantages make wars more likely. On the other hand, Fearon (1997) argued that greater offensive advantages may have conflicting effects, one of which works against war.[17] In this section, I provide paradoxical results against these arguments. I begin by defining the overall offense-defense balance in my context.

An Overall Offense-Defense Balance

Jervis (1978) insists that military factors are the primary causes of offense and defense dominance.[18] Then the offense-defense balance can be understood as the military technological advantages of attacking versus being attacked. More formally, *advantage* refers to a difference between the relative expected benefits and costs of attacking versus being attacked, holding other factors constant. Any technological factor that increases the difference tips the offense-defense balance towards the offense.

Then, for a particular state, the offense-defense balance is said to be 'even' when the difference is, or is relatively near, zero; the defense has the advantage, or defense dominance, when the difference is negative; and the offense has the advantage, or offense dominance, when the difference is positive. Also, a larger difference above zero indicates a balance more in favor of offense and the greater aggressor's prospects for value of offense. This measure of balance is defined for a particular state, which Glaser and Kaufmann (1998) call *directional balance*.

Motivated by how the use of military technology differs between large versus small countries, I assumed in my model that two states face asymmetric opportunities: A small revisionist state strengthens offensive technology while a big status quo state strengthens its defensive technology over time. Then, I can gauge the impact of increasing offensive and defensive technology on the direction of change in the offense-defense balance for a particular state, holding all other factors constant. In my model, increasing defensive technology over time for a status quo state increases the difference between the relative expected utility of attacking versus being attacked (defending), which equals $[-c_A] - [-\pi t - d_A]$. The directional offense-defense balance for a status quo state starts from negative, reaches a point where it equals zero, and then increases over zero. This implies that a status quo state has defense dominance at first, an even offense-defense balance at some later point, and then offense dominance over time. Similarly, increasing offensive technology for a revisionist state also increases the difference between the relative expected utility of attacking versus being attacked, which equals $[E(\mu)t - d_B] - [-c_B]$. As the directional offense-defense balance for a revisionist state starts from being positive, a revisionist's offense dominance grows much stronger over time.

Combining two directional offense-defense balances for each state, I define, in a two-state system, an *overall offense-defense balance* to be an arithmetic product of the two balances.[19] Then, when the defense has the advantage for the status quo state, the product is less than zero, and thus, the overall offense-defense balance favors the defense. That is, the defense is dominant system-wide. In this case, I say that the revisionist state has relatively *mild* offense dominance over the course of time. When the offense-defense balance is even for the status quo state, the product is equal to zero, and thus, the overall offense-defense balance is even system-wide and the revisionist state's offense dominance is referred to as being *moderate*. If the product is greater than zero, i.e. when the offense is dominant for both states, then the offense has greater system-wide advantage, which corresponds to *strong* offense dominance for the revisionist state.

Equilibrium Analysis

Based on the definition of the overall balance given above, I now elaborate on the likelihoods of stalemate and peace, and the risk of war in equilibrium.

In equilibrium, war only occurs at t^* with probability of $1 - F\left(\frac{s+d_B}{t^*}\right)$, the ex ante pre-conflict probability that the revisionist state will choose to attack. If war actually occurs, then the war occurs for sure only at $t^* = \frac{c_A - d_A}{\pi}$. That is, when the war occurs, the offense-defense balance is exactly even for the status quo state, while the revisionist has a mild offensive advantage.[20] Also, in equilibrium, peace occurs on or before t^* with positive probability of $F\left(\frac{s+d_B}{t^*}\right)$, while peace is impossible after t^*. That is, when peace occurs, it is only when the defense has the advantage for the status quo state and offensive advantage for the revisionist state is mild. In addition, before t^* when the defense dominates for the status quo state, stalemates are more surely likely than wars.

In other words, in any equilibrium, the risk of war is positive only when the overall offense-defense balance is even, while the likelihood of peace, equivalently the likelihood of stalemate, is positive only when the overall offense-defense balance favors the defense. Thus, this result implies that war by a revisionist state, or terrorism, is most likely at a point when the overall offense-defense balance is relatively even than in any other periods; when the defense is system-wide dominant before t^*, stalemates may happen; and, greater offensive advantages never make states favor peace more. Figure 3 shows the likelihood of stalemate/peace and war over time given t^* as the duration in equilibrium.

New Conjectures

The equilibrium analysis implies new conjectures contradicting the salient predictions about war and peace in offense-defense theory. Figure 4 associates the likelihoods of peace, stalemate, and war to the offense-defense balance in continuation of a conflict between asymmetric states.

First, an *even* overall offense-defense balance makes wars by revisionist states most likely. The revisionist state develops offensive capabilities because they have aggressive aims unrelated to their security requirements. These capabilities raise the risk of war, but when wars actually occur, they do not stem principally from offense dominance. My first conjecture is a surprising reversal of the basic prediction that war is far more common in periods when the offense-defense balance favors the offense. More precisely, when wars are more likely, I cannot conclude whether the offense is dominant or the defense is dominant. The fact that I cannot find a significant association between the offense-defense balance and the likelihood of war

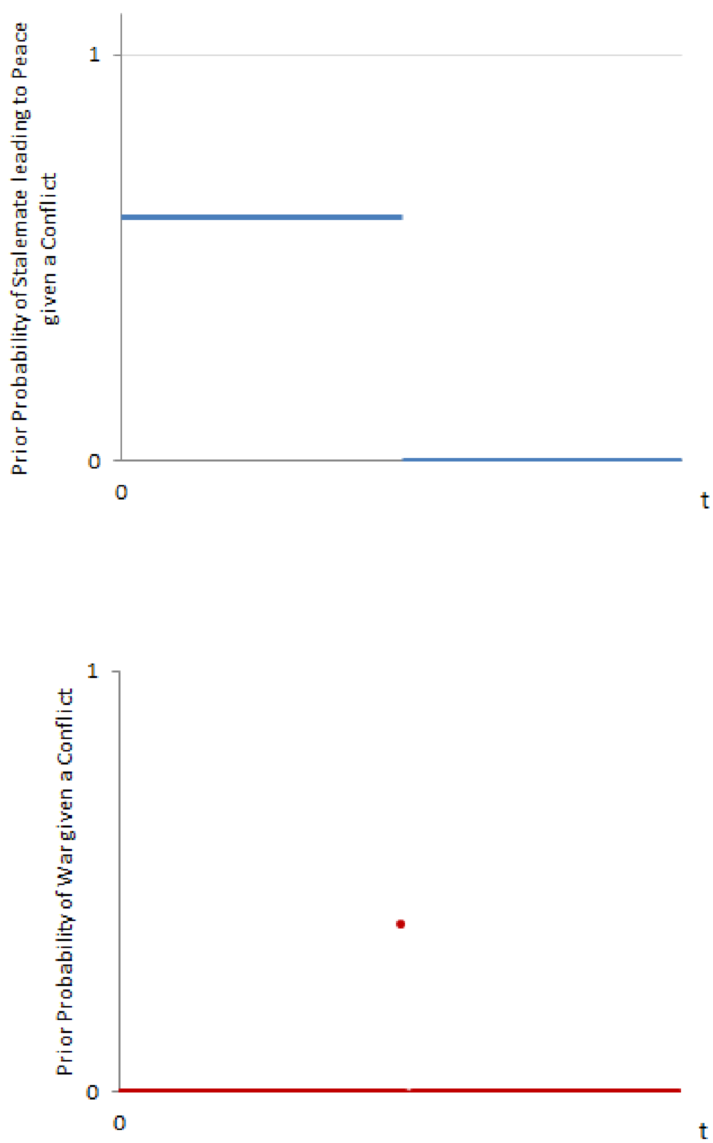


Figure 3: The likelihoods of stalemate/peace and war in equilibrium

Notes: Parameter values for computation are the same as in Figure 2.

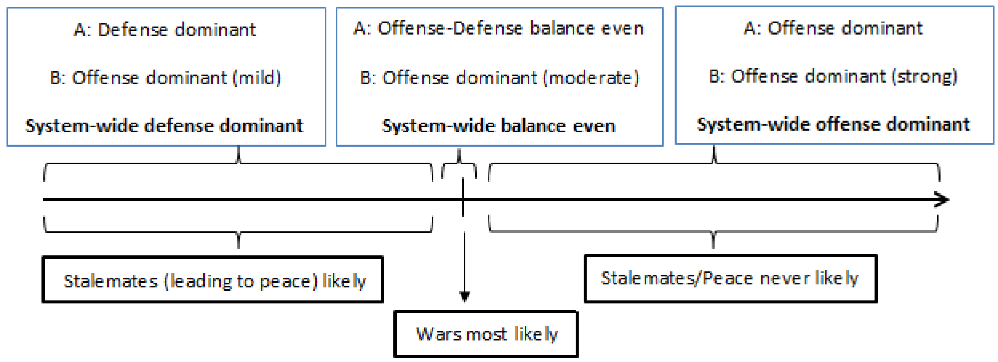


Figure 4: The offense-defense balance and the likelihoods of peace, stalemate, and war

runs counter to offense-defense theorists' common-sense assumption.

Second, greater offensive advantages over time never increase states' inclination toward peace. In other words, increasing offensive advantages over time do not have conflicting effects on the risk of war. There is a positive likelihood of peace preceded by a finite duration of stalemate only within the range of periods when the overall offense-defense balance favors the defense. My second conjecture refutes Fearon (1997)'s argument that offensive advantages may have conflicting effects on the likelihood of war and thus greater offensive advantages may make states more receptive to peace.

My conjectures imply that offense-defense theory may overstate the association between the offense-defense balance and the occurrence of violent conflicts in making predictions about conflicts involving transnational terrorist groups. On that account, I insist that the offense-defense balance is not the best predictor of the risk of terrorism. Revisionist states that are not satisfied with an existing international order may not resort to war, and status quo states that are truly interested in preserving peace may well accept stalemate. Thus, the most common usage of the concept of the offense-defense balance by itself does not determine the optimal strategic choices of asymmetric states and the equilibrium outcome as to whether terrorism occurs or not. Therefore, offense-defense theory demands refinements to incorporate the contemporary realities of international conflicts. Understanding the limitations of offense-defense theory will help international politics scholarship develop better causal explanations and analytical leverage in describing the relationship between technological change and asymmetric conflicts.

Concluding Remarks

In this paper, I develop a game-theoretic model of increasing offensive and defensive technology, and explore the dynamics of threats posed by revisionist states or terrorist organizations. The equilibrium analysis shows that offense dominance in conflicts between two asymmetric states is not associated with war. In this sense, my paper provides an alternative framework to the formal conflict literature, highlighting the qualitative difference in the structure of international conflicts between asymmetric states from that of conflicts between equally matched states.

A few extensions are proposed for future research. First, the definition of the overall offense-defense balance may seem stylized; a generalization of the concept is needed for broader applications to conflicts associated with revisionist states. Second, my model assumes that the surplus s that A offers to B is exogenously given, which may not be self-enforcing. One solution that would fix this problem would be to make B receive interest payment streams of s forever at every instant of time after her concession; in which case B would not have an incentive to restart the game after backing down and thus, a payment of s becomes self-enforcing. Further, because the arbitrariness of s's path may limit its applications, it would be interesting to endogenize s so that A has an option of choosing the amount $s(t)$ for every finite

$t \geq 0$. In this way, I can study the profound effects of the role of credibility and the learning in information revelation process on the duration of stalemate and on what terms conflict can be ended. Such analyses may guide many states' decision making in creating rational incentives to achieve international cooperation, even when the offense is dominant, as well as finding ways to end stalemate.

Notes

[1]In attempting to explain the causes of international conflict, theorists of power transition attribute shifts in the distribution of power as the important source. Several neorealists argue that the distribution of power in the international system is the key variable. For example, Waltz (1979), one of the founders of neorealism, defines the likelihood of war in terms of power differentials between states and groups of states. Offense-defense theory is a variant of neorealist theory that diverges from standard power based structural realist analyses in that the theory converts power into military capabilities, explaining many aspects of international politics that the distribution of power does not account for. Betts (2005, p. 357) points out that 'Military capabilities — and political leaders' understanding or misunderstanding of them — can exert an independent influence on decisions for or against war. Depending on which side in a conflict has the edge, the military factors make it easier or harder for a government to risk resorting to forces against the status quo.'

[2]The crucial works and developments on offense-defense theory are Jervis (1978), Lynn-Jones (1995), Van Evera (1998), and Glaser and Kaufmann (1998).

[3]The term terrorism has been described variously as a tactic, a reaction to oppression, and a crime. The United Nations defines terrorism as 'an anxiety-inspiring method of repeated violent action, employed by clandestine individual, group, or state actors, for idiosyncratic, criminal or political reasons, where in contrast to assassination the direct targets of violence are not the main targets.'

[4]Revisionist state is a term from Power Transition Theory within the wider field of international relations. It is used to describe states, that unlike status quo states, see an inherent injustice in the international system of states. Powerful and influential nations in international relations such as the United States, Britain, France and other nations like Japan who have benefited from western

liberalism, are likely to fall under the category of status quo states while North Korea, Iran, and other nations dissatisfied with their place on the international spectrum are often considered Revisionist States.

[5]See Mearsheimer (1983), Lieber (2000), Gortzak, Haftel and Sweeney (2005), and Nilsson (2012) for the critics of offense-defense theory that have challenged the theory's main predictions concerning the frequency and the expected duration of wars on both theoretical and empirical grounds.

[6]Provocative military actions may include troop mobilization, accumulation of weapons, arms proliferation, executing missile tests, launching nuclear programs, or acquiring nuclear arsenal. 'In virtually all international crises, we observe states using military preparations and diplomatic statements' as costly signaling strategies (Fearon, 1992, p. 144). Then, these actions may be called signaling by explicit threats, or 'brinkmanship.' For a discussion of 'brinkmanship' arguments, see Nalebuff (1986), Powell (1990), and Schelling (1966).

[7]I assume that A possesses the ability to strike first but only as a war limitation measure in order to disarm offensive actions by an aggressor state B.

[8]I use male pronouns to refer A and female pronouns to refer B, for the sake of simplicity.

[9]Note that there is no upper or lower bound on $m_i(t)$. If I consider $m_i(t)$ as 'physical technology,' then it is reasonable to have a technological constraint. However, I regard $m_i(t)$ as some kind of martial effectiveness that A and B accumulates as time goes.

[10]Attacks or wars are associated with fixed costs suffered by dealing with the use of arms—costs of fighting and destruction.

[11]Because we observe a state attacking first suffering less in the real world situation, it might be more realistic to assume $c_A < d_A$, i.e. it's less costly for A to be the first to 'attack' than being attacked by B at the outset of the game. However, I assume that A does not want to be too much of an aggressor; rather it is an extreme pacifist, who would use preventive attacking only as war limitation measure. That is, A would rather prefer being attacked by B to attacking in the early stage of the game when the cost from technology that B can impose on A have not increased that much with continuation.

[12]For each μ, there exists some $t^{'}_{\mu} \equiv t^{'}(\mu)$ such that $\mu t^{'}_{\mu} - d_B = s$. After this point, B would strictly prefer attacking to backing down.

[13]I assume that $-\pi t^{'}_{\bar{\mu}} - d_A \leq -c_A$ for A to find it in his interest to attack after $t^{'}(\mu)$ for all μ.

[14]For all values of μ, there is one other equilibrium outcome (which does not involve a finite

duration of stalemate) obtainable in a subgame perfect equilibrium: A plays $\{attack\}_A^t$ for $t > 0$; and B plays $\{backdown\}_B^0$. If B expects A to attack immediately after $t = 0$, she has no incentive to deviate. Also, if A expects B to back down immediately at $t = 0$, he has no incentive to deviate. This equilibrium disappears in an alternating-move version of the game. Thus, I focus on the set of equilibria discussed in the text.

[15]A formal proof for this result is available upon request.

[16]In an extreme case where A's value for peace increases indefinitely, $\pi \to \infty$, the duration t^* converges to zero. Then, the prior probability that B will choose a strategy involving attack approaches zero, and thus, the equilibrium yields optimal strategy of all types of B to back down immediately.

[17]Finding Jervis (1978)'s arguments to be partial and incomplete, Fearon (1997) develops a brief assessment of how the standard 'offensive-advantages-cause-war' hypothesis fares against the record of interstate war in Europe since 1648. He claims that the historical record suggests that war appears to have been less frequent when offensive advantages have been greater. He argues that offense dominance tends also to increase the variance of military outcomes, making war more risky for the attacker, and thus, greater offensive advantages might make states favor peace rather than war and be more reluctant to resort to war.

[18]For discussions of other factors, such as the cumulativity of resources, nationalism, force size, alliance behavior, regime popularity, military doctrine, etc., see Jervis (1978), Lynn-Jones (1995), Van Evera (1998), and Glaser and Kaufmann (1998).

[19]The definition differs from that of 'compound balance' introduced in Glaser and Kaufmann (1998, p. 58). They argue that the offense-defense balance is defined 'as the ratio of the cost of the forces the attacker requires to take territory to the cost of the forces the defender has deployed' (Glaser and Kaufmann, 1998, p. 50).

[20]At t^*, the difference between the relative expected utility of attacking versus being attacked is zero for the status quo state whereas the difference is positive ($\mu t^* - d_B + c_B \geq s + c_B > 0$) for the revisionist state.

Acknowledgments

This work was initiated while I was studying at the University of Chicago, and was conducted further at Yonsei University. This work was supported by the Yonsei University Future-Leading Research Initiative under Grant 2019-22-0187. I am deeply grateful to Roger Myerson, Ethan Bueno de Mesquita, and Lars Stole for their valuable advice and guidance on this paper during my doctoral studies. I greatly appreciate Scott Ashworth, Sandeep Baliga, Tiberiu Dragu, Thomas Dolan, Alex Frankel, Mehdi Shadmehr, Richard van Weelden, and participants at the 71st and 72nd Midwest Political Science Association Conferences for helpful suggestions and discussions. I also thank Jin Young Kim, Charles Lipson, and John Mearsheimer for helpful insights on international politics.

Disclosure statement

There is no potential conflict of interest.

References

Betts, Richard K. 2005. *Conflict after the Cold War: Arguments on Causes of War and Peace*. New York: Longman.

Fearon, James D. 1992. Threats to Use Force: Costly Signals and Bargaining in International Crises PhD thesis University of California, Berkeley.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Oranization* 49(3):379–414.

Fearon, James D. 1997. "The Offense-Defense Balance and War since 1648." A Revised Version of a Paper Presented at the Annual Meetings of the International Studies Association, Chicago, February 21-25, 1995.

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

Gortzak, Yoav, Yoram Z. Haftel and Kevin Sweeney. 2005. "Offense-Defense Theory: An Empirical Assessment." *The Journal of Conflict Resolution* 49(1):67–89.

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

Lieber, Keir A. 2000. "Grasping the Technological Peace: The Offense-Defense Balance and International Security." *International Security* 25(1):71–104.

Lynn-Jones, Sean M. 1995. "Offense-Defense Theory and Its Critics." *Security Studies* 4(4):660–691.

Mearsheimer, John J. 1983. *Conventional Deterrence*. Ithaca, NY: Cornell University Press.

Nalebuff, Barry. 1986. "Brinkmanship and Nuclear Deterrence: The Neutrality of Escalation." *Conflict Management and Peace Science* 9:19–30.

Nilsson, Marco. 2012. "Offense-Defense Balance, War Duration, and the Security Dilemma." *The Journal of Conflict Resolution* 56(3):467–48.

Powell, Robert L. 1990. *Nuclear Deterrence Theory: The Search of Credibility*. Cambridge: Cambridge University Press.

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

Van Evera, Stephen. 1998. "Offense, Defense, and the Causes of War." *International Security* 22(4):5–43.

Waltz, Kenneth N. 1979. *Theory of International Politics.* New York: McGraw-Hill.

Appendix A: The Solution Concept

The definitions and derivations in this appendix closely follow Fearon (1995). First, I define a Bayesian Nash equilibrium for the normal form version of the conflict game, denoted Γ. To avoid complications, I restrict attention to pure strategy equilibria. Here, a pure strategy for A is a rule s_A that specifies $t \in \mathbb{R}^+$ to $\{attack\}$, where $\mathbb{R}^+$ is the set of nonnegative reals. Also, a pure strategy for B is a map $s_B : M \to \mathbb{R}^+ \times \{attack, backdown\}$. I write $\{attack\}_A^t$ for the subgame strategy 'continue up to t, then attack' for A, and $\{attack\}_B^t$ for the subgame strategy 'continue up to t, then attack' and $\{backdown\}_B^t$ for 'continue up to t, then back down' for B. For every t, let $R(t)$ be a unique cumulative probability that A attacks on or before t in some equilibrium if A follows s_A. Using F, s_B induces a unique pair of cumulative distributions $P(t)$ and $Q(t)$, which are the probabilities that B attacks on or before t, or backs down on or before t in some equilibrium if B follows s_B. By the properties of cumulative distribution functions, $P(t)$ and $Q(t)$ are increasing and have well-defined left-hand limits for all t. Let $P^-(t) \equiv \lim_{s \to t^-} P(s)$ and $Q^-(t) \equiv \lim_{s \to t^-} Q(s)$. Note that I assume $R(t)$ to be increasing, left-continuous, and have well-defined right-hand limits for all t.

Payoffs have been defined except for simultaneous $\{attack\}_A^t$ and $\{attack\}_B^t$ at t, or $\{attack\}_A^t$ and $\{backdown\}_B^t$. When A attacks and B attacks at the same instant time, then B suffers from A's preventive attack and A suffers from B's attack, so that they each suffers costs that occur by the adversary's action. If A attacks while B chooses to back down at the same time t, A's preventive attack dominates the withdrawal of B. So I assume the following. However, these payoffs do not matter in the sequel.

Assumption 1. If A chooses to attack at t and B chooses to attack at the same time, they receive $(-m_A(t) - d_A, -c_B)$.

Assumption 2. If A chooses to attack at t and B chooses to back down at the same time, they receive $(-c_A, -c_B)$.

Given Assumptions 1 and 2, A's expected payoff for $\{attack\}_A^t$, given s_B, is

$$\begin{aligned} U_A^{attack}(t) \equiv{} & P^-(t)\left[-\pi t - d_A\right] + \big(P(t) - P^-(t)\big)\left[-\pi t - d_A\right] \\ & + Q^-(t)\left[-s\right] + \big(Q(t) - Q^-(t)\big)\left[-c_A\right] + (1 - P(t) - Q(t))\left[-c_A\right] \\ ={} & P(t)\left[-\pi t - d_A\right] + Q^-(t)\left[-s\right] + \big(1 - P(t) - Q^-(t)\big)\left[-c_A\right]. \end{aligned}$$

Similarly, B's expected payoff for $\{attack\}_B^t$, given s_A, is

$$U_B^{attack}(t), \equiv R(t)\left[-c_B\right] + (1 - R(t))\left[\mu t - d_B\right], \tag{A.1}$$

and B's expected payoff for $\{backdown\}_B^t$, given s_A, is

$$U_B^{backdown}(t) \equiv R(t)\left[-c_B\right] + (1 - R(t))\left[s\right]. \tag{A.2}$$

Definition 1. $\{attack\}_A^{t'}$ is a best reply for A given s_B if

$$t' \in \arg\max_t U_A^{attack}(t).$$

Definition 2. $\{attack\}_B^{t'}$ and $\{backdown\}_B^{t'}$ are best replies for B with μ given s_A if, respectively,

$$t^{'} \in \arg\max_{t} U_B^{attack}(t) \text{ and } U_B^{attack}(t^{'}) \geq \max_{t} U_B^{backdown}(t),$$

$$t^{'} \in \arg\max_{t} U_B^{backdown}(t) \text{ and } U_B^{backdown}(t^{'}) \geq \max_{t} U_B^{attack}(t).$$

Definition 3. (s_A, s_B) is a Bayesian Nash equilibrium for the normal form version of Γ if under s_i, $i = A, B$, every i chooses a best reply, given s_j, $j \neq i$, where s_A induces $R(t)$ and $\{F, s_B\}$ induces $\{P(t), Q(t)\}$.

In the dynamic setting, a complete pure strategy in Γ for A is a rule $s_A : \mathbb{R}^+ \to \mathbb{R}^+ \times \{attack\}$, with the restriction that if $s_A(t^{'}) = \{attack\}_A^t$, then $t^{'} \leq t$. Also, a complete pure strategy for B is a map $s_B : \mathbb{R}^+ \times M \to \mathbb{R}^+ \times \{attack, backdown\}$, with the restriction that if $s_B(t^{'}, \sigma) = \{attack\}_B^t$ or $\{backdown\}_B^t$, then $t^{'} \leq t$.

For all $t^{'} \geq 0$, define the continuation game $\Gamma(t^{'})$ as follows: (i) payoffs are as in Γ, beginning at $t^{'}$; and (ii) initial beliefs for A are given by a cumulative distribution function $F(\cdot; t^{'})$ on M. A strategy s_i implies a strategy for state i in every continuation game $\Gamma(t^{'})$ called $s_i|t^{'}$. Further, for every $t^{'}$, let $R(t|t^{'})$ be a unique conditional cumulative distribution over $s_A|t^{'}$, analogous to $R(t)$. Also, using $F(\cdot; t^{'})$, $s_B|t^{'}$ induces a unique pair of cumulative distributions $P(t|t^{'})$ and $Q(t|t^{'})$, defined analogously to $P(t)$ and $Q(t)$. Expected payoff functions for $\Gamma(t^{'})$, $U_A^{attack}(t|t^{'})$, $U_B^{attack}(t|t^{'})$, and $U_B^{backdown}(t|t^{'})$ are defined as before.

To rule out some optimistic off-equilibrium-path beliefs on surprises, I give Condition 1. For example, if B continues unexpectedly at t, then A will think that B is of low militancy level and maintain this optimistic belief as conflict continues. Because this would allow B with higher militancy level to continue accumulating more offensive capability and attack before A, we need the following criterion to impose restrictions on how A would interpret completely unexpected behavior by B and thus, refine the set of equilibria.

Condition 1. For all $t > 0$ such that $P(t) + Q(t) = 1$, $F\left(\frac{s+d_B}{t}; t\right) = 0$.

In other words, if B continues beyond t when it was expected to have attacked or backed down prior to t, then A believes that B's payoff for attacking is at least as great as B's payoff for backing down at time t.

Definition 4. $\{(s_A, s_B), F(\cdot;\cdot)\}$ is an *equilibrium* for Γ if (i) (s_A, s_B) induces a Bayesian Nash equilibrium in Γ and for all $t \geq 0$, $(s_A|t, s_B|t)$ induces a Bayesian Nash equilibrium in $\Gamma(t)$ using $F(\cdot; t)$; (ii) for all t such that t is reached with positive probability under s_A and s_B (i.e. $R(t) < 1$ and $P(t) + Q(t) < 1$), $F(\cdot; t)$ is $F(\cdot)$ updated using Bayes' Rule and s_i, $i = A, B$; and (iii) Condition 1.

Appendix B: Preliminaries and Proofs

In this appendix, I give the formal definition of a *duration of stalemate*, and develop some lemmas that will build up to proving the main results.

Definition 5. τ is a duration of stalemate for Γ in some equilibrium if in this equilibrium, $\tau = \inf\{t | Q(t) = Q(\infty) \text{ and } t \neq 0\}$.

Lemma 1. *In an equilibrium of* Γ*, if* $Q(t)$ *is atomic at* t'*, then,* $\{attack\}_A^{t'}$ *is never a best reply for A and is chosen with zero probability in equilibrium.*

Proof. Suppose to the contrary that in some equilibrium A chooses $\{attack\}_A^{t'}$ where $Q(t') > Q^-(t')$. The A then receives an ex ante expected payoff of

$$U_A^{attack}(t') = \quad P(t')\left[-\pi t' - d_A\right] + Q^-(t')\left[-s\right] + \left(1 - P(t') - Q^-(t')\right)\left[-c_A\right]. \quad \text{(B.1)}$$

By right continuity of $Q(t)$, the deviation $\{attack\}_A^{t'+\varepsilon}$ for $\varepsilon > 0$, yields an expected payoff arbitrarily close to

$$P(t')\left[-\pi t' - d_A\right] + Q(t')\,[-s] + \left(1 - P(t') - Q(t')\right)[-c_A] \tag{B.2}$$

as $\varepsilon \to 0$. Note that (B.2) is strictly greater than (B.1). Thus, $\{attack\}_A^{t'}$ cannot be a best reply for A and so, it will be chosen with zero probability in equilibrium. □

I can now write A's equilibrium ex ante expected payoff for $\{attack\}_A^t$ as

$$U_A^{attack}(t) = \quad P(t)\,[-\pi t - d_A] + Q(t)\,[-s] + (1 - P(t) - Q(t))\,[-c_A]. \tag{B.3}$$

Recall that B's equilibrium ex ante expected payoff for $\{attack\}_B^t$ and $\{backdown\}_B^t$ are given in (A.1) and (A.2) respectively.

Lemma 2. *In any equilibrium of* Γ*, if there exists a* $t' > t$ *such that* $Q(t') > Q(t)$*, then A will never choose* $\{attack\}_A^t$ *and it must be* $R(t) = 0$*.*

Proof. A's equilibrium ex ante expected payoff function $U_A^{attack}(t)$ in (B.3) increases with $Q(t)$ since

$$\frac{\partial U_A^{attack}(t)}{\partial Q(t)} = -s + c_A > 0,$$

by the parameter restriction $s < c_A$. Thus, it is straightforward that A will never choose $\{attack\}_A^t$, and it must be $R(t) = 0$ for all $t < t'$ such that $Q(t') > Q(t)$. □

Remark 1. Suppose $\tau > 0$ is a finite duration of stalemate in some equilibrium of Γ. (1) By definition, for all $\varepsilon > 0$, B backs down with positive probability in the interval $[\tau - \varepsilon, \tau]$. (2) By Lemma 2 and (1), $R(t) = 0$ for $t < \tau$. Thus, $\{attack\}_B^{t>\tau}$ yields an ex ante expected payoff of $U_B^{attack}(t) = R(t)\,[-c_B] + (1 - R(t))\,[\mu t - d_B]$, while $\{backdown\}_B^{t<\tau}$ yields $U_B^{backdown}(t) = s$.

Lemma 3. *In an equilibrium of* Γ *where A believes B will attack at some* t *with positive probability, there must exist a finite duration* $\tau < \infty$*.*

Proof. Suppose to the contrary that there exists an equilibrium of Γ in which continuation may occur, where A believes B will attack at some t with positive probability, and in which $Q(t)$ is strictly increasing for all t. By Lemma 2, $R(t) = 0$ for all $t \geq 0$ since for all $t \geq 0$ there exists a $t^{'} > t$ such that $Q(t^{'}) > Q(t)$. Then, for B, $\{attack\}_B^t$ for any $t \geq 0$ yields a payoff of $U_B^{attack}(t) = \mu t - d_B$ while $\{backdown\}_B^t$ for any $t \geq 0$ yields a payoff of $U_B^{backdown}(t) = s$. Then, for each μ, there exists some $t^{''} \equiv t^{''}(\mu)$ such that $\mu t^{''} - d_B = s$ and so $\mu t - d_B > s$ for all $t > t^{''}$. Also, $\frac{\partial U_B^{attack}(t)}{\partial t} > 0$ for all $t > t^{''}(\mu)$ for all μ, which implies that $\{attack\}_B^t$ is never a best response for all t for any μ, and thus, $P(t) = 0$ for all $t \geq 0$, contradicting that B will attack at some t with positive probability. Therefore, it must be $Q(t) = Q(\hat{t})$ for all $t \geq \hat{t}$ for some $\hat{t}$ and by definition there exists a finite duration. $\square$

Proof of Proposition 1. First I show that it cannot be $\tau < t^*$. Suppose to the contrary that it were. Then, $-\pi\tau - d_A > -c_A$. Then, regardless of B's action, A will never choose attack at t, where $\tau \leq t < t^*$, so that $R(t) = 0$ for $t < t^*$. Then, for B, $\{attack\}_B^t$ for $\tau \leq t < t^*$ yields an ex ante expected payoff of $U_B^{attack}(t) = [\mu t - d_B]$ which is strictly increasing in t. This implies that it must be $P(t) = P(t^*)$ for all $\tau \leq t < t^*$. By the definition of the duration, it also must be $Q(t) = Q(\tau)$ for all $t \geq \tau$. This in turn implies that $U_A^{attack}(t)$ strictly decreases in t for $\tau \leq t < t^*$, contradicting $R(t) = 0$ for $t < t^*$. Now I show that it cannot be $\tau > t^*$. Suppose to the contrary that it were. Then, $-\pi\tau - d_A < -c_A$. Then, regardless of B's action, A will choose to attack before τ with positive probability, again contradicting $R(t) = 0$ for $t < \tau$. Thus, it must be $\tau = t^*$. It is straightforward that the duration of stalemate τ is unique. $\square$

The intuition to the proof of Proposition 1 is as follows: If a duration of stalemate τ is such that $\tau > t^*$, then A would prefer to attack before the duration with positive probability, contradicting $R(t) = 0$ for $t < \tau$. On the other hand, if a duration of stalemate is smaller than t^*, then either A would have incentives not to attack at τ and continue for some more time or B would have incentives to continue stalemate beyond τ, in which case B's signaling threats by continuation would not be informative enough to A. Thus, it would make equilibrium not supported at τ. Lemma 4 characterizes the equilibrium behavior of states.

Lemma 4. *In an equilibrium of Γ with finite t^* as the duration of stalemate, (1) if A chooses $\{attack\}_A^t$, it must be the case that $t \geq t^*$; (2) if B chooses $\{attack\}_B^t$, it must be the case that $t \geq t^*$; and (3) B will choose $\{attack\}_B^t$ where $t \geq t^*$ if $m_B(t^*) - d_B > s \leftrightarrow \mu > \frac{s+d_B}{t^*}$ and only if $m_B(t^*) - d_B \geq s$.*

Proof. Part (1) follows immediately from Lemma 2 and Remark 1. Since $R(t) = 0$ for $t < t^*$, A never chooses to attack before t^*. Part (2): Since $R(t) = 0$ for $t < t^*$, $\{attack\}_B^t$ gives $U_B^{attack}(t) = [\mu t - d_B]$ and $\{backdown\}_B^t$ gives $U_B^{backdown}(t) = s$. Since $U_B^{attack}(t)$ is strictly increasing in t, B will never choose $\{attack\}_B^t$ when $t < t^*$, so part (2) follows. Part (3): Fix an equilibrium in which there exists a finite duration t^*. B's ex ante expected payoff for $\{attack\}_B^{t>t^*}$ is

$$U_B^{attack}(t) = R(t)\,[-c_B] + (1 - R(t))\,[\mu t - d_B]\,,$$

and, B's ex ante expected payoff for $\{backdown\}_B^{t>t^*}$ is

$$U_B^{backdown}(t) = R(t)\,[-c_B] + (1 - R(t))\,[s]\,,$$

This implies that B does better to choose $\{attack\}_B^{t\geq t^*}$ if $\mu t^* - d_B > s$ and only if

$\mu t^* - d_B \geq s$. $\square$

Lemma 4 implies that the ex ante pre-conflict probability that B will choose a strategy involving attack is $1 - F\left(\frac{s+d_B}{t^*}\right)$, the prior probability that $\mu t^* - d_B \geq s$. Thus, I can write A's ex ante expected utility for continuing up to a duration t^* as

$$\begin{aligned} u_A(t^*) &= F\left(\frac{s+d_B}{t^*}\right)[-s] + \left(1 - F\left(\frac{s+d_B}{t^*}\right)\right)[-\pi t^* - d_A] \\ &= F\left(\frac{s+d_B}{t^*}\right)[-s] + \left(1 - F\left(\frac{s+d_B}{t^*}\right)\right)[-c_A]. \end{aligned} \tag{B.4}$$

Proof of Proposition 2. Given s_B in (ii), s_A in (i) is an optimal strategy for A: Note that if A plays $\{attack\}_A^{t^*}$, then B would want to attack an infinitesimal time just before t^*, which in turn makes A to play attack before B, contradicting Lemma 4. For A not to deviate from his equilibrium strategy on the interval $t \in [0, t^*]$, B with $\mu < \frac{s+d_B}{t^*}$ needs to make A (weakly) prefer continue to attack for any $t \leq t^*$. By Remark 1 and Lemma 4, $Q(t) > 0$ and $P(t) = 0$ for $t < t^*$. So, A's ex ante expected payoff for $t < t^*$ is $F\left(\frac{s+d_B}{t^*}\right)[-s]$, which is strictly greater than (B.4). Also, A would be indifferent between being attacked and attacking at t^*. Therefore A's optimal strategy is to play *plays* $\{attack\}_A^t$ *for* $t > t^*$. *such that* $R(t) = 0$ *if* $t \leq t^*$ *and* $R(t) = 1$ *if* $t > t^*$.Given s_A in (i), s_B (ii) is a best response strategy for B. This follows immediately from Lemma 4. Note that since A attacks immediately after t^*, if any B deviates from her optimal strategy, then she will receive $-c_B$ which is strictly worse than attacking or backing down at t^*.[21] Therefore, B does better to choose $\{attack\}_B^t$ at t^* for $\mu \geq \frac{s+d_B}{t^*}$ and to choose $\{backdown\}_B^t$ for any $t \leq t^*$ for $\mu < \frac{s+d_B}{t^*}$. [22] Thus, following from Lemma 4, strategies specified in (i) and (ii) of Proposition 2 constitute a Bayesian Nash equilibrium in Γ, given A's belief and opponent's strategy. For the continuation games $\Gamma(t')$, $t' \geq 0$, Bayes' Rule implies that if $\{F, s_B\} \Rightarrow \{P(t), Q(t)\}$, then, $P(t|t^{'}) = \frac{P(t)-P(t^{'})}{1-P(t^{'})-Q(t^{'})}$ and $Q(t|t^{'}) = \frac{Q(t)-Q(t^{'})}{1-Q(t^{'})-P(t^{'})}$ for $t > t^{'}$ such that $P(t^{'}) + Q(t^{'}) < 1$. Then, if $\{backdown\}_B^t$ or $\{attack\}_B^t$is a best response for B given s_A, then it remains so in all continuation games $\Gamma(t^{'} \leq t)$ for all $t^{'}$ such that $t^{'}$ is reached with positive probability, i.e. $P(t^{'})+Q(t^{'}) < 1$ under s_A and s_B. Thus, the strategies specified in (i) and (ii) of Proposition 2 induce a Bayesian Nash equilibria in every continuation game $\Gamma(t^{'})$ up to t^* where $P(t^*) + Q(t^*) = 1$, given beliefs and opponent's strategy. The condition (i) of Definition 4 for an equilibrium of the conflict game Γ is satisfied. For $t \leq t^*$, provided that $P(t)+Q(t) < 1$, A's belief on $F(\cdot; t)$ is $F(\cdot)$ updated using Bayes' Rule. Since, the payoff structure leaves B to choose any time to back down on or before t^*, A's belief on the probability that B will not back down conditional on having not backed down would be $\Pr\left(\mu \geq \frac{s+d_B}{t^*}|t\right) = 1 - F\left(\frac{s+d_B}{t^*}\right)$. Since A's beliefs are consistent with the strategy profile along the equilibrium path using Bayesian updating for all t such that t is reached with positive probability under s_A and s_B, the belief system given the strategy profile, specified in (iii) of Proposition 2 satisfy (ii) of Definition 4 for an equilibrium of the conflict game Γ. Beliefs off the equilibrium path given in (iv) is consistent with Condition 1. Thus, the specified set of equilibrium strategies and the belief system along with the off-the-equilibrium path beliefs constitute the set of equilibria in the incomplete information conflict game. $\square$