An Analysis of Artillery Shoot-and-Scoot Tactics

Younglak Shim[*], Michael P. Atkinson[†]

July 8, 2018

Abstract

Firing multiple artillery rounds from the same location has two main benefits: a high rate of fire at the enemy and improved accuracy as the shooter's aim adjusts to previous rounds. However, firing repeatedly from the same location carries significant risk that the enemy will detect the artillery's location. Therefore, the shooter may periodically move locations to avoid counter-battery fire. This maneuver is known as the shoot-and-scoot tactic. This paper analyzes the shoot-and-scoot tactic for a time-critical mission using Markov models. We compute optimal move policies and develop heuristics for more complex and realistic settings. Spending a reasonable amount of time firing multiple shots from the same location is often preferable to moving immediately after firing an initial salvo. Moving frequently reduces risk to the artillery, but also limits the artillery's ability to inflict damage on the enemy.

Keywords: combat models, salvo equations, stochastic duels, Markov models, artillery

[*]younglakshim@gmail.com Center for Army Analysis and Simulation, Republic of Korea Army
[†]mpatkins@nps.edu Operations Research Department, Naval Postgraduate School

1 Introduction

Joseph Stalin said that "Artillery is the god of war" (Holmes and Singleton, 2001). Artillery has been a crucial part of warfare from at least as far back as the the thirteenth century (Dastrup, 1992) and is a method to apply *indirect* fire on an enemy. Indirect fire is applied to a target outside of visual range using coordinates supplied by an external source (Army, 2014), such as a forward observer or surveillance UAV (USMC, 1998). The range of modern artillery exceeds 20 kilometers (Dastrup, 1992).

Artillery remains a workhorse of modern militaries, accounting for roughly 25% of the United States Army and Marine Corps ground combat regiments (Army, 1990; USMC, 2016). Artillery is often used in a support role, such as providing suppression fire to allow infantry forces to reach an objective. Other common artillery missions include targeting an enemy's logistics centers and transportation links, or providing counter-fire against the enemy's artillery weapons (Army, 2001). Artillery and air support play similar roles as they both can strike at long ranges and put minimal forces at risk. However, there are many benefits to using artillery over air strikes, including: artillery is far cheaper to use and aircraft can only operate in certain enviromental and threat conditions (Army, 2001; Kopp, 2005; Harris, 2017). Artillery often performs suppression of enemy air defense (SEAD) operations to destroy anti-aircraft weapons to allow aircraft to perform their missions (Army, 2001).

In the last fifteen years, artillery has played a role in many conflicts throughout the world, including in Iraq (Ratliff, 2017), Afghanistan (Bowman, 2016), Syria (Snow, 2018), and the Russian-Ukranian conflict (Fox, 2017). Artillery dictates that tensions in Korea be handled delicately because North Korea has thousands of artillery weapons within range of Seoul (McCarthy, 2017). In recent years, technological improvements have resulted in many countries increasing their investments in artillery capabilities (Kopp, 2005; Miller, 2017).

This paper focuses on the tactics of mobile artillery units. An artillery's mobility is key to survival because an adversary can pinpoint a firing artillery's location with counter-battery radar and return fire (Kopp, 2005). We formulate models to examine when an artillery force

(Blue) should move locations. The United States Army's field manuals recognize the trade-off between survivability and artillery responsiveness when changing locations (see Chapter 2 of Army (1996)). The benefits to Blue firing from the same location are a high firing rate and increasing accuracy as Blue adjusts its aim based on the results of earlier Blue rounds. However, if Blue remains in its original position, the enemy (Red) will eventually counter-fire. As both artillery maneuverability and counter-detection capabilities increased in the second half of the 20th century, so did the importance of Blue artillery moving to avoid counter-fire. This led to the advent of the "shoot-and-scoot" tactic, where artillery moves quickly after firing at a target (Koba, 1996). Until recently shoot-and-scoot was primarily used by unconventional forces as a harassing tactic (Sharp, 2006) and by conventional forces as a defensive measure to limit casualties. However, the newest generation of artillery cannons can now fire from one location and then move to a new location and fire again within minutes (Kopp, 2005; Miller, 2016). This leads to the potential for a new offensive oriented variant of the shoot-and-scoot tactic, where Blue artillery can perform many cycles of shoot/move. In the last year, the United States Army practiced new tactics that use this shoot-and-scoot principle, referred to as "hide and seek" tactics (Tinder, 2017).

In this paper, we generalize the shoot-and-scoot tactic and examine how long Blue should "shoot" in one location before "scooting" to a new location. Moving frequently limits the risk to Blue but also reduces Blue's offensive effectiveness. The United States military recognizes the importances of moving artillery, spending many pages in manuals describing how to choose firing positions and travel routes (Army, 1996). However, very little guidance is given for when the commander should move locations. The United States Army and Marine Corps field manuals suggest the move decision should consider the amount of time at the current location, number of rounds fired, and the enemy threat level (Army, 2001; USMC, 2002). However it is up to the specific commander to weigh the various factors to determine when to move. We incorporate in our models the factors listed in the field manuals, and we believe our analysis can provide artillery commanders with insight to produce more effective

shoot-and-scoot tactics.

Section 2 describes the setting in more detail. We then frame our work within the literature in Section 3. Sections 4–7 present the models and analysis.

2 Setting

We refer to individual artillery weapons as *cannons* throughout this paper. One side (Blue) uses artillery to fire at enemy (Red) targets. Red targets might include bridges, logistics depots, radar systems, or anti-aircraft weapons. Red has its own artillery whose objective is to detect Blue's artillery and return fire. The Red artillery is located at a separate location from the Red targets. Red artillery fires on Blue artillery, but Blue artillery only fires at Red targets, not at the Red artillery. Possible reasons for Blue not firing at Red artillery include: Red artillery is outside the range of Blue artillery, Blue cannot detect the location of Red artillery (e.g., Blue's counter-battery radars have been destroyed), or the Blue commander orders all fire to focus on Red targets. While Blue does not fire at Red artillery, Blue can move quickly to another position to avoid Red counter-fire. Figure 1 illustrates the setting.

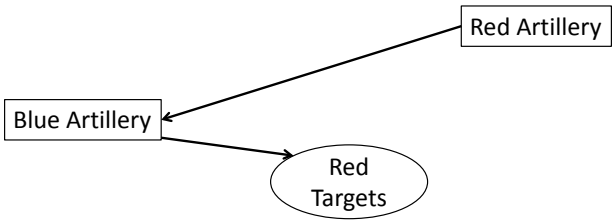


Figure 1: Blue fires only at Red Targets and receives counter-fire from Red Artillery

Blue has an initial estimate of the location of the Red targets and initiates the battle by firing first. Red will quickly determine an estimate of Blue's location and eventually return fire. At some point, Blue may move to another location to avoid Red counter-fire. After moving to a new position, Blue fires one *risk-free* round at the Red targets. We use the term risk-free because Red does not know Blue's new location until after Blue fires the first shot from the new location. Consequently, Blue has the advantage with the first shot

from a new location because Red cannot hit Blue before Blue fires this risk-free shot. After the risk-free shot, both Blue and Red have an estimate of the other's location, and they proceed to repeatedly fire at each other. The probability a shot hits the adversary may be low for both Blue and Red after Blue moves to a new location. As Blue remains in the same location, the accuracy of Blue and Red improves over time as both adjust their aim. Blue makes adjustments based on feedback from surveillance assets about the results of Blue's earlier fire. Red likewise may adjust to its own fire, but Red also adjusts its aim based on counter-battery radar reports on Blue fire; Blue does not obtain any targeting information from observing Red fire as the Red artillery and targets are not co-located. Eventually Blue moves to a new location and the process resets and repeats. Blue maintains an estimate of the location of the Red targets from move to move. However, after every move Blue's accuracy resets to a lower level due to self-location errors, environmental conditions, and calibration perturbations to the cannons during the move.

Blue's objective is to destroy Red targets before Red destroys Blue. With no other structure on the problem, Blue should move immediately after its risk-free shot because this results in no Blue casualties. This is the classic scoot-and-shoot approach. In our setting, we assume a high-priority time-critical mission: Blue must destroy Red within a certain *window of opportunity*. If the time-window closes before Blue destroys Red, then Blue effectively "loses" the battle. An example is Blue plans an air strike in six hours, and Blue artillery must destroy Red's anti-aircraft missiles so Blue aircraft can perform the mission safely. If moving is time-consuming, then frequently changing locations will not be an effective strategy.

We formulate three Markovian models of increasing complexity. Section 4 presents a simple model that only considers the time-window closing. This model does not account for the increase in risk to Blue (and Red) from improved firing accuracy when Blue remains at the same location. Section 5 formulates a model that captures the increased risk faced by both Blue and Red the longer Blue stays in the same location. In Section 5 Blue knows the current time and risk level, but in Section 6 Blue does not know the underlying risk level.

Markov models have strong underlying assumptions and are approximations to reality. Our goal is to gain insight into the problem and evaluate how to more effectively use shoot-and-scoot tactics. In Section 7 we describe a simulation experiment that makes more realistic assumptions. The insights from the simpler models in Sections 4–6 produce robust heuristics that perform well in the more realistic simulation setting.

3 Literature Review

Our model incorporates the firing rate of artillery cannons and how aim improves over time. Alexander (1977) describes volley patterns, aiming procedures, errors, and the likelihood a round will cause damage. Nadler and Eilbott (1971); Barr and Piper (1972) analyze shoot-adjust-shoot strategies where the shooter incorporates feedback from earlier shots to improve his aim. Kwon et al. (1997); Cha and Kim (2010) develop algorithms to schedule when each artillery cannon should fire at specific targets. All these articles are complementary to our work by providing realistic parameter estimates for our model.

Our approach falls under the umbrella of combat modeling and has similarities to Lanchester attrition models, search theory, missile defense, salvo equations, and duels (Washburn and Kress, 2009). The Lanchester Linear Law is often used to model artillery conflict because the attrition rate depends upon attacker fire-power and defensive density, which is the dynamic in many indirect fire scenarios (Lucas and Turkes, 2004; Keane, 2011). The flaming datum problem from search theory (Hohzaki and Washburn, 2001) has similarities to our setting because a submarine flees after an attack, leaving the enemy scrambling to counter.

Artillery and naval battles both have two sides lobbing volleys of fire at each. The Hughes' salvo equations focus on the damage inflicted by each volley of a naval battle (Hughes, 1995; Armstrong, 2004). Most studies explore the relative importance of force size, offensive capabilities, defensive capabilities, and survivability (Armstrong, 2005, 2007, 2014b). Moving to avoid fire is not a relevant tactic, but there is another tactic in the naval setting that has

a trade-off similar to moving in the artillery context. A Blue ship turns off many of its sensors when Blue employs restrictive emissions control, which makes it difficult for Red to detect Blue's position. However without the use of its sensors, Blue's ability to execute offensive and defensive operations are significantly degraded. Perhaps our approach could be adapted to salvo models where Blue tunes its emissions control policy to best balance Blue's vulnerability to Red detection vs. Blue fighting effectiveness.

Our model has similarities with stochastic duels (Williams and Ancker, 1963; Choe, 1983). While we share some modeling overlap with many duel models, no one paper incorporates as many factors as we do. Furthermore, most duel models are descriptive in nature, whereas we take a prescriptive approach. Ancker and Williams (1965); Williams and Ancker (1981) examine a variant that allows one player to defensively dodge fire, and this dodge tactic prevents the player from returning fire. In our model, Blue can avoid casualties by moving, and Blue cannot fire at Red during the move. Another variant by Ancker (1966) imposes a time-limit similar to the one in our model. We allow the hit probability to increase in time as aim improves, which also appears in the duel models of Bhashyam and Singh (1967); Ancker (1975); Barfoot (1989). In our model, Blue has the advantage after moving because Red does not know Blue's new location. Barfoot (1989); McNaught (2002) consider a similar first-shooter advantage, where the second-shooter has to detect and react to the initial shot.

Harari (2008) considers insurgents (Blue) utilizing shoot-and-scoot tactics to fire rockets at government forces (Red). If the government returns fire quickly, the government counter-fire is not very accurate and may cause collateral casualties. However, if the government waits too long to better pinpoint the cannon's location, the insurgents may have already moved. We are not analyzing an insurgent setting where collateral damage is likely; in our model as soon as Red detects Blue, Red counter-fires. Armstrong (2014a) also models defenses against rocket attacks, but evaluates the defender's ability to shoot down rounds in the air. The problem examined by Armstrong (2014a) relates to missile defense, which primarily focuses on knocking out missiles as they approach the intended destination (Brown et al., 2005;

Davis et al., 2017) In our setting, Red targets the source of fire rather than incoming shells. Marshall (1994) develops a model that incorporates both counter-targeting of live missiles and the missile launchers themselves, but the attacker does not move to avoid counter-fire.

Several Naval Postgraduate theses develop models that consider artillery movement, including one by the first author upon which this paper is based (Shim, 2017). Temiz (2016) proposes an algorithm to choose the firing locations and corresponding travel routes based on topological features of the terrain. Finlon (1991) develops a Markov model of artillery combat to evaluate artillery unit organization. The model considers the decrease in effectiveness due to movement but does not include an enemy threat and provides no prescriptive analysis about when to move. The work most similar to ours is the thesis of Guzik (1988), who formulates a Markov model to analyze when an artillery unit should move locations. Guzik (1988) assumes Blue fires multiple shots before Red detects Blue's location. Counter-detection capabilities have improved in the last thirty years; we reasonably assume Red detects Blue after Blue's first shot and that Blue will receive counter-fire if Blue remains in the same location. Guzik's model is predicated upon Blue trying to avoid *all* Red counter-fire, which is consistent with the classic notion of shoot-and-scoot. This emphasis on survivability is appropriate in low priority missions, but other missions may require Blue to expose itself to hostile fire (Army, 1996). Our analysis focuses on how much risky counter-fire Blue should endure before moving. Our model also accounts for a crucial factor missing from other analyses of mobile artillery: Blue benefits from firing from the same location via improving firing rate and accuracy. Focusing on the disadvantages of risky-fire (Red counter-fire) and ignoring the advantages (improving Blue effective fire), will skew results toward moving more frequently. From a methodological standpoint, Guzik (1988) examines steady-state behavior of the system, whereas we consider a time-critical battle and compute the probability Blue wins. Furthermore, we incorporate a time-window, track the current time and risk (and how the battle dynamics evolve as risk changes), consider a partial-information setting where Blue is unaware of the precise current risk level, and present provably optimal move policies.

4 Time-window Model

This section focuses on a simple model. There is only one Red target and one direct hit is sufficient to destroy either side. Blue fires at effective rate α and Red fires at effective rate β. Thus α (β) represents the overall firing rate of Blue (Red) multiplied by the single-shot hit (kill) probability. All random times are independent and have exponential distributions. Therefore the time until Blue (Red) destroys Red (Blue) has an exponential distribution with rate α (β). In Section 5 we incorporate "risk" by allowing the firing rates α and β to increase as Blue continues to fire from the same location.

Blue must destroy Red within a window of opportunity. This window comprises N time periods; each time period is an exponential random variable with rate γ and all time periods are independent. Therefore the time-window has a Gamma distribution with shape parameter N and rate parameter γ. If Blue does not hit Red by the conclusion of the Nth time period, the battle is over and Blue loses. In Appendix I, we examine a time-window of deterministic length, which may be more realistic in some scenarios. We only consider a binary outcome to the battle: Blue wins or Blue loses. Blue loses if either Blue is destroyed by Red fire or the time-window closes. In Section 4.3 we incorporate an additional cost if Red destroys Blue.

Immediately after moving to a new location, Blue fires one risk-free shot, which hits Red with probability p_0. It takes no time to set up and fire the risk-free shot after Blue arrives to the new location. If Blue misses, then Blue's new location is revealed to Red. Blue and Red proceed to fire at each other until one of three events occurs: (1) one side is hit, (2) the time-window closes, or (3) Blue moves. We refer to a situation where Blue and Red exchange fire simultaneously as *risky-fire* because Blue exposes itself to risk from Red's salvos. Blue knows the time period $t \in \{1, 2, 3, \ldots, N\}$, and hence can tailor its movement decision to the current time period. We define λ_t as the rate at which Blue moves. During risky-fire in time period t, Blue moves after an exponential time with rate λ_t. Blue moving randomly allows us to maintain analytic tractability. We examine more realistic move dynamics in Section 7. However within the fog of war, a random aspect to the movement decision may not be that

unreasonable of an approximation. Blue chooses λ_t^* to maximize its probability of winning the engagement. The optimal move policy actually leads to non-random behavior.

The time to travel to a new location is an exponential random variable with rate δ. During travel Blue is safe from Red fire, however Blue cannot hit Red either. To summarize, time is divided into two components: Blue travel to new firing locations when neither side can be destroyed, and a risky-fire portion when both sides fire at each other. The transition between travel and risky-fire is marked by one risk-free Blue shot. Blue wins if Blue hits Red before (1) Red hits Blue and (2) the time-window closes. The model parameters appear below.

- α: rate at which Blue fires fatal shots during risky-fire
- β: rate at which Red fires fatal shots during risky-fire
- γ: rate at which time transitions to next period
- λ_t: rate at which Blue moves; Blue decision variable that depends upon time period t
- δ: rate at which travel completes
- p_0: probability Blue hits Red with its risk-free shot

Given the movement rates $(\lambda_1, \lambda_2, \lambda_3 \ldots, \lambda_N)$, we denote $\mathbf{P}[t]$ as the probability Blue wins starting from risky-fire in period t. We define $\mathbf{P}[N+1] \equiv 0$: Blue loses if the time-window closes. To compute $\mathbf{P}[t]$, we condition on which event happens next:

$$\begin{aligned}\mathbf{P}[t] =& \frac{\alpha}{\alpha+\beta+\gamma+\lambda_t} + \frac{\gamma}{\alpha+\beta+\gamma+\lambda_t} \times \mathbf{P}[t+1] \\ &+ \frac{\lambda_t}{\alpha+\beta+\gamma+\lambda_t} \times \frac{\delta}{\delta+\gamma} \sum_{s=t}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} (p_0 + (1-p_0)\mathbf{P}[s]) \end{aligned} \tag{1}$$

If Red hits Blue first (rate β), then Blue loses. With rate α, Blue hits Red and wins immediately. With rate γ, time increments to the next period and Blue now wins with probability $\mathbf{P}[t+1]$. When Blue moves (rate λ_t), the probability Blue wins moving forward

is more complicated. While traveling between locations one of two events can occur: Blue arrives to the new location (rate δ) or time increases (rate γ). Therefore, the number of periods for Blue to complete its travel is a Geometric random variable (starting at 0) with success probability $\frac{\delta}{\delta+\gamma}$. If Blue initiates a moves in period t, the travel will complete in period $s \geq t$ with probability $\frac{\delta}{\delta+\gamma}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t}$. If Blue's transit takes more than $N-t$ periods, the time-window closes and Blue loses. Otherwise Blue completes the move in time period s, for some $t \leq s \leq N$. Blue then fires a risk-free shot, which destroys Red with probability p_0. If the risk-free shot misses, then Blue has probability $\mathbf{P}[s]$ of winning moving forward.

Rearranging (1), we write $\mathbf{P}[t]$ recursively in terms of $\mathbf{P}[s]$ for $t+1 \leq s \leq N$:

$$\begin{aligned}\mathbf{P}[t] = &\frac{(\delta+\gamma)\left(\alpha+\gamma\mathbf{P}[t+1]\right)}{(\alpha+\beta+\gamma)(\delta+\gamma)+\lambda_t(\gamma+\delta p_0)} \\ &+\lambda_t\frac{\delta p_0+\gamma\frac{\delta}{\delta+\gamma}\sum_{s=t+1}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-1}\left(p_0+(1-p_0)\mathbf{P}[s]\right)}{(\alpha+\beta+\gamma)(\delta+\gamma)+\lambda_t(\gamma+\delta p_0)}\end{aligned} \tag{2}$$

Given that $\mathbf{P}[N+1]=0$, we can solve for $\mathbf{P}[t]$ for any $1 \leq t \leq N$ using backward recursion.

4.1 Optimal Move Policy

We now compute the optimal movement rates $(\lambda_1^*, \lambda_2^*, \lambda_3^* \ldots, \lambda_N^*)$ to maximize the probability Blue wins the battle. For the remainder of this section, $\mathbf{P}[t]$ denotes the maximum probability using the optimal $(\lambda_1^*, \lambda_2^*, \lambda_3^* \ldots, \lambda_N^*)$. We compute λ_t^* using (2) by first solving for λ_N^*, and then solving for λ_{N-1}^*, etc. The first result states there are only two possibilities for λ_t^*.

Proposition 1. $\lambda_t^* \in \{0, \infty\}$. *If* $\lambda_t^* = 0$,

$$\mathbf{P}[t] = \mathbf{F}[t] \equiv \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\mathbf{P}[t+1] \tag{3}$$

If $\lambda_t^* = \infty$,

$$\mathbf{P}[t] = \mathbf{S}[t] \equiv \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0} \frac{\delta}{\delta + \gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta + \gamma} \right)^{s-t-1} (p_0 + (1 - p_0)\mathbf{P}[s]) \tag{4}$$

The proof appears in Appendix A. The optimal policy specifies Blue should either fire until time increments ($\lambda_t^* = 0$), or immediately scoot after firing the risk-free shot ($\lambda_t^* = \infty$). We refer to the $\lambda_t^* = 0$ case as *fight* and the $\lambda_t^* = \infty$ case as *scoot*. If Blue fights, then Blue wins with probability $\mathbf{F}[t]$ in (3), which corresponds to the first two terms of (1). If Blue scoots, then Blue wins with probability $\mathbf{S}[t]$ in (4). With probability $\frac{\delta p_0}{\gamma + \delta p_0}$, Blue hits Red before time increments. Otherwise, if time increments, then the complicated summation term in (4) follows from the same logic as the corresponding term in (1). Proposition 1 implies that $\mathbf{P}[t] = \max(\mathbf{F}[t], \mathbf{S}[t])$. If Blue is indifferent between fighting and scooting ($\mathbf{F}[t] = \mathbf{S}[t]$), Blue fights ($\lambda_t^* = 0$). Before stating the conditions when $\lambda_t^* = 0$ vs. $\lambda_t^* = \infty$, we present an intuitive property of $\mathbf{P}[t]$.

Proposition 2. $\mathbf{P}[t]$ *is monotonically non-increasing in* t.

The proof appears in Appendix B. As the time-window draws nearer to closing, Blue's chances of winning decrease.

As time increments, eventually Blue will remain and fight because Blue does not have enough time to scoot. Determining this time period is a key part of our analysis. We first define the probability Blue wins, starting in period t, if Blue fights for all remaining time periods (i.e., $\lambda_s = 0$ for all $t \leq s \leq N$)

$$\hat{\mathbf{F}}_{\mathbf{all}}[t] \equiv \frac{\alpha}{\alpha + \beta} \left(1 - \left(\frac{\gamma}{\alpha + \beta + \gamma} \right)^{N-t+1} \right). \tag{5}$$

Equation (5) follows by substituting $\lambda_s = 0$ for all $t \leq s \leq N$ into (1). Intuitively (5) is the probability Blue fires a fatal shot before Red, multiplied by the probability the time-window does not close.

We next define $\hat{\mathbf{S}}_{\mathbf{F}}[t]$, which is the probability Blue wins when Blue scoots in period t, but fights for $s > t$. That is $\lambda_t = \infty$ and $\lambda_s = 0$ for all $t+1 \leq s \leq N$.

$$\hat{\mathbf{S}}_{\mathbf{F}}[t] \equiv \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0} \frac{\delta}{\delta + \gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta + \gamma} \right)^{s-t-1} \left(p_0 + (1 - p_0) \hat{\mathbf{F}}_{\mathbf{all}}[s] \right) \tag{6}$$

Equation(6) follows from the same logic that produces $\mathbf{S}[t]$ in (4). To summarize the notation:

- $\mathbf{F}[t]$: probability Blue wins if Blue fights in period t, given Blue behaves optimally for all time periods $s \geq t+1$ (equation (3))
- $\mathbf{S}[t]$: probability Blue wins if Blue scoots in period t, given Blue behaves optimally for all time periods $s \geq t+1$ (equation (4))
- $\hat{\mathbf{F}}_{\mathbf{all}}[t]$: probability Blue wins when Blue fights for all time periods $s \geq t$ (equation (5))
- $\hat{\mathbf{S}}_{\mathbf{F}}[t]$: probability Blue wins when Blue scoots in time period t and fights for all time periods $s \geq t+1$ (equation (6))

With this new notation, we next define the last time period when Blue scoots:

$$\tau = \sup\{s \mid \lambda_s^* = \infty\} = \sup\{s \mid \hat{\mathbf{S}}_{\mathbf{F}}[s] > \hat{\mathbf{F}}_{\mathbf{all}}[s]\} \tag{7}$$

If $\tau = N$, then Blue scoots in the final time period, and if $\tau = -\infty$, then Blue never scoots. By construction $\lambda_t^* = 0$ for all $t > \tau$ and $\lambda_\tau^* = \infty$. To complete the optimal policy, we need λ_t^* for $t < \tau$. The optimal policy is a threshold policy dictated by τ:

Proposition 3. *If $t > \tau$, then $\lambda_t^* = 0$, otherwise $\lambda_t^* = \infty$ for $1 \leq t \leq \tau$, where τ is defined by (7).*

The proof appears in Appendix C. A threshold policy is intuitive, however it is not trivial to prove. Early in the battle, Blue has time to take the safe approach by scooting and avoiding Red fire. Near the end of the time-window Blue must fight because a move risks expiration of the time-window before Blue arrives to the new location.

Time-based threshold policies suggested by Proposition 3 often produce near optimal results for the more complicated models we examine in Sections 5–7.

4.2 Comparative Statics

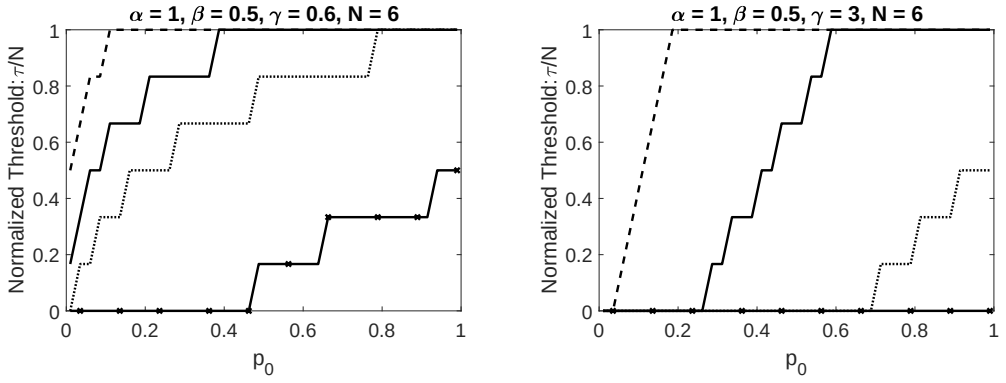
We next illustrate how the threshold τ and Blue win-probability vary for different parameters. We formalize the comparative statics below the figures in Proposition 4. Figure 2 plots $\frac{\tau}{N}$ against p_0 for four values of δ. Each subfigure fixes a different combination of β and γ. For all figures $\alpha = 1$ and $N = 6$. Recall that Blue scoots for all values less than or equal to τ. Consequently, curves closer to the northwest produce more scenarios where Blue scoots. Blue is more likely to scoot for smaller γ (Figure 2a and 2c) because that corresponds to a longer time-window. Smaller δ corresponds to longer travel times, which lessens the likelihood that Blue will scoot. In the extreme when $\delta = 0.2$ in Figures 2b and 2d, Blue never scoots and fights to the finish from the start. Larger values of β result in more scooting because Blue wants to avoid Red's more effective fire.

Figure 3 presents the Blue win-probability at the beginning of the battle, $\mathbf{P}[1]$, in a similar format to Figure 2. The always-fight policy provides a lower bound for $\mathbf{P}[1]$. This lower bound corresponds to $\hat{\mathbf{F}}_{\mathbf{all}}[1]$ from (5) and is achieved in the figures as $p_0 \to 0$. The relationships follow what we expect. The probability is increasing in p_0 and δ and decreasing in β and γ. The probability can vary by a substantial amount as we change the inputs.

We summarize the relationships illustrated in Figures 2–3.

Proposition 4. $\mathbf{P}[t]$ *and* τ *vary with the input parameters in the following manner:*

- $\mathbf{P}[t]$ *is non-decreasing in* α*.*
- $\mathbf{P}[t]$ *is non-increasing in* β *and* τ *is non-decreasing in* β*.*
- $\mathbf{P}[t]$ *and* τ *are non-decreasing in* δ*.*
- $\mathbf{P}[t]$ *is non-increasing in* γ*.*



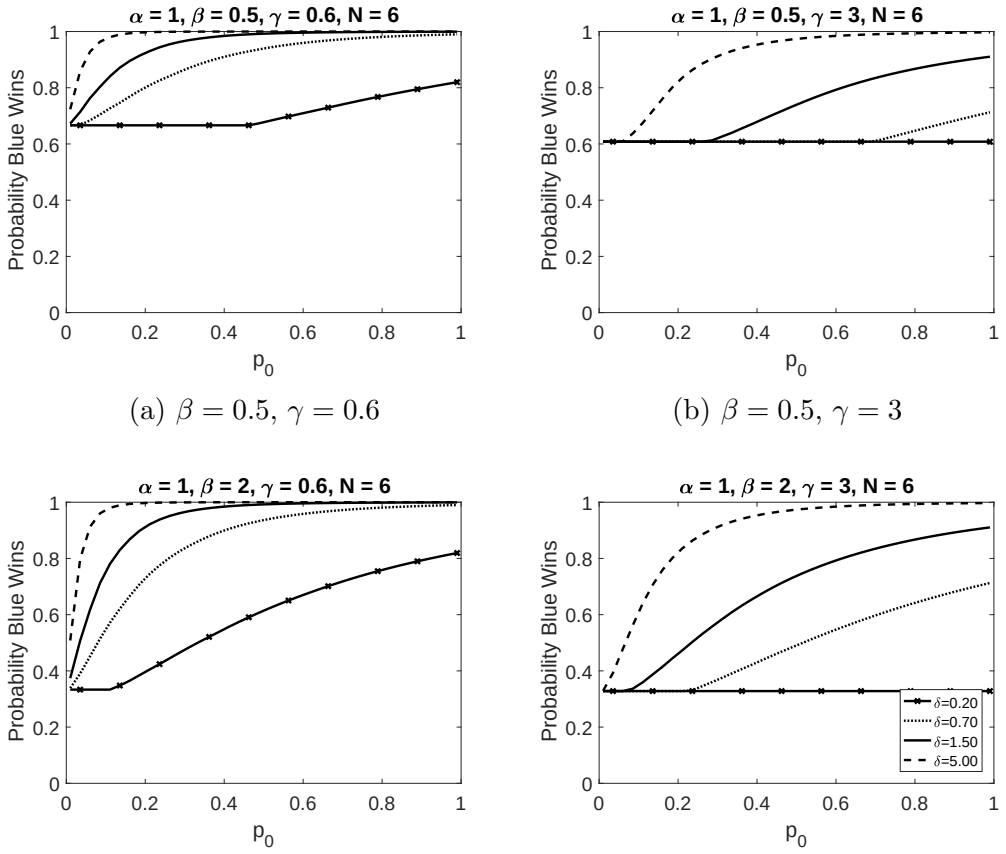
(a) $\beta = 0.5$, $\gamma = 0.6$

(b) $\beta = 0.5$, $\gamma = 3$

(c) $\beta = 2$, $\gamma = 0.6$

(d) $\beta = 2$, $\gamma = 3$

Figure 2: Normalized threshold $\frac{\tau}{N}$ for $\alpha = 1$, $N = 6$ and various values of β, γ, δ.



(a) $\beta = 0.5$, $\gamma = 0.6$

(b) $\beta = 0.5$, $\gamma = 3$

(c) $\beta = 2$, $\gamma = 0.6$

(d) $\beta = 2$, $\gamma = 3$

Figure 3: Probability Blue wins starting in period 1 ($\mathbf{P}[1]$) for $\alpha = 1$, $N = 6$ and various values of β, γ, δ.

- $\mathbf{P}[t]$ *and* τ *are non-decreasing in* p_0.

The proof appears in Appendix D. Most of the relationships in Proposition 4 are intuitive. However as α increases, the threshold τ may increase or decrease. If Blue is more effective during risky-fire, Blue may exploit that by fighting earlier (which leads to a smaller τ). However a larger α may allow Blue to delay fighting and rely on more risk-free shots, knowing that when Blue finally switches to fight, Blue will have a better chance of winning.

4.3 Additional Cost if Blue Destroyed

In the base model, Blue is indifferent between being destroyed and the time-window closing. In reality the Blue commander would prefer that Blue not suffer attrition. We now incorporate the additional cost of Blue casualties and frame the problem from an expected value perspective. If Blue wins, Blue receives reward 1, if the time-window closes Blue receives reward 0, and if Blue is destroyed Blue receives reward $-C$ for some $C > 0$. We slightly modify equation (1), including replacing $\mathbf{P}[t]$ with $\mathbf{V}[t]$, to highlight the expected value focus:

$$\begin{aligned}\mathbf{V}[t] =& \frac{\alpha}{\alpha+\beta+\gamma+\lambda_t} - \frac{\beta}{\alpha+\beta+\gamma+\lambda_t}C + \frac{\gamma}{\alpha+\beta+\gamma+\lambda_t} \times \mathbf{V}[t+1] \\ &+ \frac{\lambda_t}{\alpha+\beta+\gamma+\lambda_t} \times \frac{\delta}{\delta+\gamma}\sum_{s=t}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t}(p_0 + (1-p_0)\mathbf{V}[s])\,. \end{aligned} \tag{8}$$

Our results generalize in the natural way, which we summarize in the following corollary.

Corollary 1. *If* $t > \tau^C$*, then* $\lambda_t^* = 0$*, otherwise* $\lambda_t^* = \infty$ *for* $1 \le t \le \tau^C$*, where*

$$\tau^C = \sup\{j \mid \hat{\mathbf{S}}_{\mathbf{F}}^C[j] > \hat{\mathbf{F}}_{\mathbf{all}}^C[j]\},$$

and

$$\hat{\mathbf{F}}^{C}_{\mathbf{all}}[t] \equiv \frac{\alpha - \beta C}{\alpha + \beta} \left(1 - \left(\frac{\gamma}{\alpha + \beta + \gamma} \right)^{N-t+1} \right)$$

$$.\hat{\mathbf{S}}^{C}_{\mathbf{F}}[t] \equiv \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0} \frac{\delta}{\delta + \gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta + \gamma} \right)^{s-t-1} \left(p_0 + (1 - p_0) \hat{\mathbf{F}}^{C}_{\mathbf{all}}[s] \right).$$

The proof appears in Appendix E. Blue scoots in more periods compared to the base model ($\tau_C \geq \tau$) because of the additional cost of being destroyed. Appendix I contains another extension to the model where we examine a time-window of deterministic length.

5 Risk Model: Full Information

We now incorporate the concept of *risk*. As Blue continues to fire from the same location, the firing accuracy will increase for both Blue and Red. Recall that one operational way for Blue's accuracy to improve is that Blue adjusts its aim after incorporating feedback from surveillance assets about the results of Blue's earlier rounds. Risk represents the increase in the effective firing rate for both sides as Blue remains in the same location. After Blue changes locations, the effective firing rate, and hence risk, resets to its lowest level. Even though Blue maintains an estimate of Red's location throughout the battle, Blue's effective firing rate resets because errors are introduced into Blue's firing calculus after a move. These errors are caused by uncertainty in self-location, changing environmental conditions, and mechanical perturbations incurred during transit.

In addition to tracking the time period t, we now track the risk level r. This approach assumes one index appropriately captures the risk to both Blue and Red, when in reality the risk may be asymmetric. It is theoretically straightforward to incorporate a separate risk level for each side by tracking both a Blue and Red risk, but for simplicity we do not pursue that here. The model in Section 7 does treat the risk to Blue and Red separately.

As in Section 4, Blue fires a risk-free shot immediately after changing locations. After

every risk-free shot from Blue, the battle enters risk level $r = 1$, where both Blue and Red begin firing at each. As Blue fires from the same location, the risk to Blue increases over time as Red obtains better information about Blue's location. This information may come from counter-battery radar that tracks Blue's fire or surveillance assets, such as UAVs. The time until the risk increases from level r to level $r+1$ has an exponential distribution with rate μ_r, which may depend upon the current risk level r. There are M risk levels, and by definition $\mu_M = 0$. The time until the risk increases includes intelligence processing to determine an updated aimpoint and the time to recalibrate and aim the artillery with the new aimpoint. Similarly, the risk to Red also increases as Blue fires from the same location because Blue improves its accuracy based on feedback from earlier shots. As discussed earlier, the risk is symmetric. Currently there is no explicit connection between shots fired by either side and increasing risk. A more realistic model would track the number of shots fired by each side and then the risk rate μ would depend upon shots fired. The simulation model in Section 7 does explicitly connect risk to shots fired by each side.

The battle begins at time $t = 1$ and risk level $r = 1$; that is immediately after Blue's first (presumed missed) risk-free shot. As in Section 4, time advances to the next period according to an exponential distribution with rate γ. The effective firing rates α_r and β_r now also depend upon the risk level r. These firing rates should increase in r as Blue and Red better localize the other's position.

Moving has one important advantage for Blue: it generates a risk-free shot opportunity. Depending upon the values of α_r and β_r, moving may have an additional benefit to Blue in that it puts Blue in a relative firing advantage (e.g., if $\alpha_1 \gg \beta_1$). In this section we assume Blue has perfect information about the risk level and thus tailors its movement decision based on time t and risk level r. That is Blue chooses its movement rate $\lambda_{t,r}$ for all time periods t and risk levels r to maximize its win-probability. In Section 6, we consider an imperfect information scenario where Blue knows the current time period t, but not the risk level r, and thus chooses a move-rate that depends only upon the time period t. As in Section 4,

the travel time to the new location has an exponential distribution with rate δ. Once Blue arrives to the new location, the risk resets to $r = 1$ after the ensuing risk-free shot.

With this new notation, we update the win-probability equation from (1), which is now a function of the tuple (t, r).

$$\begin{aligned}\mathbf{P}[t,r] =& \frac{\alpha_r}{D(t,r)} + \frac{\gamma}{D(t,r)} \times \mathbf{P}[t+1,r] + \frac{\mu_r}{D(t,r)} \times \mathbf{P}[t,r+1] \\ &+ \frac{\lambda_{t,r}}{D(t,r)} \times \frac{\delta}{\delta+\gamma} \sum_{s=t}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} (p_0 + (1-p_0)\mathbf{P}[s,1])\end{aligned} \tag{9}$$

where the denominator $D(t, r)$ is the sum of all the rates

$$D(t,r) = \alpha_r + \beta_r + \gamma + \mu_r + \lambda_{t,r} \tag{10}$$

The first, second, and last terms in (9) correspond directly to terms in (1). Blue hits Red with rate α_r and wins. Time increments to period $t+1$ with rate γ, and when time increments the risk remains at its current level r. Blue moves with rate $\lambda_{t,r}$ and arrives to the new location at some period in the future. If Blue arrives prior to the closing of the time-window (period N or earlier), than Blue takes one risk-free shot with success probability p_0. If Blue misses, then the risk resets to level 1, and hence the $\mathbf{P}[s, 1]$ term in the summation. The new term in (9) relative to (1) is the third term: the risk level increases at rate μ_r. When risk increases, the time remains at period t.

5.1 Results

As in the time-window model of Section 4, the optimal move policy is deterministic. Upon entering a given time-period/risk-level tuple (t, r), Blue should either scoot immediately or remain and fight until either the battle ends or the time period or risk level increments. The following proposition generalizes Proposition 1.

Proposition 5. $\lambda^*_{t,r} \in \{0, \infty\}$. *If* $\lambda^*_{t,r} = 0$,

$$\mathbf{P}[t,r] = \frac{\alpha_r}{\alpha_r + \beta_r + \gamma + \mu_r} + \frac{\gamma}{\alpha_r + \beta_r + \gamma + \mu_r} \times \mathbf{P}[t+1,r] + \frac{\mu_r}{\alpha_r + \beta_r + \gamma + \mu_r} \times \mathbf{P}[t,r+1] \quad (11)$$

If $\lambda^*_{t,r} = \infty$,

$$\mathbf{P}[t,r] = \frac{\delta}{\delta+\gamma} \sum_{s=t}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} (p_0 + (1-p_0)\mathbf{P}[s,1]) \quad (12)$$

The proof appears in Appendix F. Appendix F.2 describes an $O(NM^2)$ algorithm to determine whether $\lambda^*_{t,r}$ is 0 or ∞. In Section 4 one could compute the optimal λ^*_t via backward induction. We take a similar approach with the model described by (9). Starting at $t = N$, we compute $\lambda^*_{t,r}$ for all $1 \le r \le M$, and then we proceed to $t-1$ and repeat. While we can use backward induction in t, we cannot perform backward induction in r because $\mathbf{P}[t,r]$ depends upon $\mathbf{P}[t,1]$ (see the summation in the last term of (9)). Therefore for each fixed t, we first solve for $\mathbf{P}[t,1]$ and then we can solve for $\mathbf{P}[t,r]$, $r > 1$, using backward induction in r. See Appendix F for more details.

We illustrate Proposition 5 with an example with three time periods ($N = 3$) and three risk levels ($M = 3$). The expected length of the time-window is 75 minutes, which corresponds to $\gamma = \frac{1}{25}$. The average travel time is 12 minutes ($\frac{1}{\delta}$). Blue hits Red on a risk-free shot with probability $p_0 = 0.1$. The risk rate is constant: $\mu_r = \mu = \frac{1}{15}$. This implies it takes on average 30 minutes to reach the maximum risk level after Blue's risk-free shot. Blue's effective firing rate is also constant: $\alpha_r = \alpha = \frac{1}{20}$. Finally, Red's effective firing rate β_r increases with risk: $\beta_1 = \frac{1}{80}, \beta_2 = \frac{1}{19}, \beta_3 = \frac{1}{5}$. Table 1 summarizes the parameters for this scenario.

$\alpha = \frac{1}{20}$	$\beta_r = \left(\frac{1}{80}, \frac{1}{19}, \frac{1}{5}\right)$	$\mu = \frac{1}{15}$	$\gamma = \frac{1}{25}$	$\delta = \frac{1}{12}$	$p_0 = 0.1$

Table 1: Example 1 parameter values

Table 2 presents the optimal move-rate $\lambda^*_{t,r}$ for the scenario defined by Table 1. We solve for $\lambda^*_{t,r}$ using the approach described in Appendix F.2. If $\lambda^*_{t,r} = 0$, Blue fires from the same

location until either one side is hit, the risk level increases (r increases), or time increments (t increases). If time or risk increases, Blue reevaluates with the new value of $\lambda^*_{t,r}$ using the updated indices. If $\lambda^*_{t,r} = \infty$, then Blue moves immediately upon entering this state. In this scenario Blue benefits little from the risk-free shot as p_0 is small. However, the benefit to scooting is that the risk resets to $r = 1$, where Blue has a significant firing advantage: $\alpha_1 = \frac{1}{20}$ vs. $\beta_1 = \frac{1}{80}$. Blue exploits this firing advantage by fighting at risk level $r = 1$ for all time periods ($\lambda^*_{t,1} = 0, \forall t$). For higher risk levels, Blue primarily scoots. For risk level $r = 3$, Blue is at a firing disadvantage ($\alpha_3 = \frac{1}{20}$ vs. $\beta_3 = \frac{1}{5}$) and so Blue scoots for all time periods ($\lambda^*_{t,3} = \infty,\ \forall t$). At risk level $r = 2$, Blue and Red are close to parity in terms of their firing rates. For early time periods ($t \in \{1, 2\}$), if Blue moves it is unlikely the time-window will close before Blue finishes traveling (travel-rate $\delta = \frac{1}{12}$ vs. time-rate $\gamma = \frac{1}{25}$), and thus Blue prefers to scoot and reset the risk to the favorable level $r = 1$. However, in the last time period Blue chooses to fight in risk level $r = 2$ because there is a reasonable chance the time-window will close before Blue finishes traveling. If Blue uses the optimal strategy in Table 2, Blue has a 0.707 probability of winning the battle starting in $t = 1$, $r = 1$.

	$r = 1$	$r = 2$	$r = 3$
$t = 1$	0	∞	∞
$t = 2$	0	∞	∞
$t = 3$	0	0	∞

Table 2: $\lambda^*_{t,r}$ for the parameters in Table 1

In the time-window model of Section 4, the optimal policy is a time-threshold policy (see Proposition 3). As Table 2 illustrates, a time-threshold policy will not be optimal once we include risk. There may be some risk levels where Blue has a firing advantage and other risk levels where Blue is at a significant firing disadvantage. A pure time-threshold cannot account for the differences across risk. The optimal time-threshold using the parameters in Table 1 has Blue scoot for $t = 1$ and fight for $t \in \{2, 3\}$. This time-threshold policy produces a win-probability of 0.582, which is a reasonable amount less than the optimal of 0.707.

While the optimal policy may not take the form of a pure time-threshold policy, the solution in Table 2 suggests a risk-based time-threshold might perform well. Such a generalized policy would still have Blue scooting in early periods and fighting in later periods, but the exact crossover time may depend upon the risk level. Unfortunately as the following example illustrates, the optimal policy is not necessarily a risk-based time-threshold. This counter-example has $N = 6$ time periods and $M = 3$ risk levels. The remaining parameters appear in Table 3 and the optimal policy appears in Table 4.

$\alpha_r = (0.1, 2.7, 10)$	$\beta_r = (0.6, 0.7, 0.9)$	$\mu_r = (6.5, 7.8, 0)$	$\gamma = 8.2$	$\delta = 7.6$	$p_0 = 0.29$

Table 3: Example 2 parameter values

	$r = 1$	$r = 2$	$r = 3$
$t = 1$	∞	0	0
$t = 2$	∞	0	0
$t = 3$	0	0	0
$t = 4$	0	0	0
$t = 5$	∞	0	0
$t = 6$	∞	0	0

Table 4: $\lambda^*_{t,r}$ for the parameters in Table 3

Table 4 shows the surprising result that Blue oscillates between scooting and fighting in risk level $r = 1$. In this scenario Blue has a significant firing advantage in risk level $r = 3$ ($\alpha_3 = 10$ vs. $\beta_3 = 0.9$) and $r = 2$ ($\alpha_2 = 2.7$ vs. $\beta_2 = 0.7$), but Red holds the advantage in risk level $r = 1$ ($\alpha_1 = 0.1$ vs. $\beta_1 = 0.6$). To reach the favorable risk levels, Blue needs to fight Red in risk level $r = 1$ and hope to survive. For early time periods Blue scoots because it has enough time remaining to try to get lucky with a risk-free shot as the travel-rate δ is comparable to the time-rate γ. For time periods close to the end, if Blue fights in $r = 1$ it is unlikely Blue fires a fatal shot before the time-window closes, and thus Blue is better off scooting in the hope of hitting a risk-free shot. The interesting middle time periods ($t \in \{3, 4\}$), provide a sweet spot for Blue to attempt to reach the favorable higher risk levels

by fighting in $r = 1$. There is a reasonable chance Blue reaches a higher risk level with enough time remaining to effectively exploit the firing advantage. The Blue win-probability using $\lambda^*_{t,r}$ in Table 4 is 0.762. If instead Blue scoots in risk level $r = 1$ for all time ($\lambda_{t,1} = \infty$, $\forall t$), the win-probability drops slightly to 0.760. While an interesting theoretical result that the optimal policy may oscillate, in practice the difference between the optimal policy and a threshold-type policy may not be operationally meaningful.

5.2 Heuristic Performance

We conclude this section by examining the performance of the two heuristics discussed earlier: a pure time-threshold policy and a more robust risk-based time-threshold. Formally we define these two heuristics as:

- *Time-Threshold*: $\lambda_{t,r} = \infty$ for $t \leq k$ for some $1 \leq k \leq N$; otherwise $\lambda_{t,r} = 0$. This heuristic uses a time threshold similar to Section 4. Blue scoots during early time periods and fights in later time periods, without consideration for the risk level.

- *Risk-Time-Threshold*: $\lambda_{t,r} = \infty$ for $t \leq k_r$ for some $1 \leq k_r \leq N$; otherwise $\lambda_{t,r} = 0$. The time threshold k_r depends upon the risk level.

We perform a simulation experiment to compare the win-probability using the optimal $\lambda^*_{t,r}$ to the win-probability achieved by the two heuristics. We generate the number of time periods N as a uniform integer between 2 and 15 and the number of risk levels M a uniform integer between 2 and 5. The rate parameters $\alpha_r, \beta_r, \mu_r, \gamma, \delta$ are uniformly distributed between 0 and 10 and the risk-free hit probability p_0 has a uniform distribution between 0 and 1.

Table 5 contains the results for the two heuristics. To compute the optimal thresholds for the heuristics we use brute force enumeration. This is computationally easy for the *Time-Threshold* heuristic, but for *Risk-Time-Threshold* the brute force method can be burdensome for larger values of N and M. Our metric of interest is relative suboptimality of the win-probability starting in $t = 1$ and $r = 1$: $\frac{optimal - heuristic}{optimal}$. Table 5 presents the 95th percentile of

this metric over 9000 replications. We repeat the experiment, but limit ourselves to scenarios that satisfy two conditions: the optimal win-probability lies in the interval (0.15,0.85) and the optimal policy both scoots and fights. When one of these conditions does not hold the heuristics often produce near optimal results. By focusing on these more "challenging" scenarios, there is more variability in heuristic performance. Both heuristics perform very well, with *Risk-Time-Threshold* being optimal at the 95th percentile.

	Time-Threshold	*Risk-Time-Threshold*
All replications	0.020	0
Optimal in $(0.15, 0.85)$	0.056	0

Table 5: 95th percentile of the relative suboptimality in the Blue win-probability

We divide the replications into 9 categories of 1000 replications each depending upon the values of α_r, β_r, μ_r. We consider three groupings based on whether the ratio $\frac{\alpha_r}{\beta_r}$ is increasing, decreasing, or non-monotonic. If that ratio increases then Blue has the advantage as risk increases, and Red has the advantage if the ratio decreases. Similarly we consider three groupings based on whether μ_r is increasing, decreasing, or non-monotonic. When $\frac{\alpha_r}{\beta_r}$ is increasing, then *Time-Threshold* is optimal at the 95th percentile. If it is optimal to fight in risk level $r = 1$ for a fixed time t, then it is also likely optimal for Blue to fight in higher risk levels where Blue has a stronger firing advantage. *Time-Threshold* performs the worst when $\frac{\alpha_r}{\beta_r}$ and μ_r are both decreasing, which leads to a 95th percentile suboptimality of 0.095. These cases are similar to the scenario in Tables 1–2: Blue fights for lower risk levels and scoots for higher risk levels, which is not compatible with *Time-Threshold*. The *Risk-Time-Threshold* heuristic performs excellently. The worst relative suboptimality across all 9000 replications for *Risk-Time-Threshold* is 0.043, but the actual difference in win-probabilities is only 0.005.

To summarize: *Risk-Time-Threshold*, while not technically optimal, generates essentially optimal results. The much simpler *Time-Threshold* also performs very well, coming within 10% of the optimal at the 95th percentile even with worst-case parameter relationhips. We also ran a separate experiment where the parameters take values across more realistic ranges (e.g., similar to those in Table 1) and the heuristics perform even better in these scenarios.

6 Risk Model: Partial Information

In Section 5 Blue knows both the current time period t and the risk level r. In some scenarios the Blue commander may not be able to accurately determine the true risk level in real-time during the chaos battle. In this section Blue knows the current time period t, but not the risk level r. More precisely Blue only knows the risk level immediately after the risk-free shot following a location change when the risk resets to level $r = 1$. As Blue continues to fire from the same location, the risk level will increase, but Blue will not know the exact risk level.

In this setting the move-rate λ cannot depend upon the unknown risk level r. We use $\lambda_t^{(j)}$ to denote Blue's move-rate in time period t given the last risk-free shot occurred in period $j \leq t$. Immediately after each move, Blue knows the current time period j and risk level $r = 1$, and thus uses a new family of move-rates $\lambda_t^{(j)}$ moving forward. Because the move-rate now depends upon the last period j a risk-free shot occurred, we modify the definition of our win-probability in (9) to $\mathbf{P}^{(j)}[t, r]$ to also depend upon j

$$\begin{aligned}\mathbf{P}^{(j)}[t,r] =& \frac{\alpha_r}{D^{(j)}(t,r)} + \frac{\gamma}{D^{(j)}(t,r)} \times \mathbf{P}^{(j)}[t+1,r] + \frac{\mu_r}{D^{(j)}(t,r)} \times \mathbf{P}^{(j)}[t,r+1] \\ &+ \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)} \times \frac{\delta}{\delta+\gamma} \sum_{s=t}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} \left(p_0 + (1-p_0)\mathbf{P}^{(s)}[s,1]\right) \end{aligned} \tag{13}$$

where the denominator $D^{(j)}(t, r)$ is the sum of all the rates

$$D^{(j)}(t,r) = \alpha_r + \beta_r + \gamma + \mu_r + \lambda_t^{(j)} \tag{14}$$

Note how the superscript index j varies for $\mathbf{P}^{(j)}[t, r]$ on the right-hand side of (13). If time (rate γ) or risk (rate μ) increment, then the superscript remains j because Blue remains in the same location. However, if Blue moves (rate $\lambda_t^{(j)}$), then Blue finishes the transit at period $s \geq t$, fires a risk-free shot, and then proceeds to use new move-rates $\lambda_t^{(s)}$, which generate the $\mathbf{P}^{(s)}[s, 1]$ win-probabilities in the summation of (13).

After completing a move at time j, Blue's win-probability moving forward after the risk-

free shot is $\mathbf{P}^{(j)}[j,1]$. Blue chooses $\lambda_t^{(j)}$ to maximize $\mathbf{P}^{(j)}[j,1]$ for each j. To solve for the optimal $\lambda_t^{(j)*}$, we can proceed with backward induction in j, however we cannot use backward induction in t. To determine an effective $\lambda_t^{(j)}$, it is helpful to have an estimate of the risk level in time period t. However, the risk level in period t depends upon the move-rates in *earlier* periods $\lambda_s^{(j)}$ for $j \leq s < t$. Because the decisions are now coupled across time, we cannot use backward induction in t. Furthermore the binary nature (scoot or fight) of the optimal decision from the full-information settings in Sections 4–5 (see Propositions 1 and 5) no longer holds.

6.1 Results

We discuss the technical details behind solving for $\lambda_t^{(j)*}$ in Appendix G, which involves standard numerical optimization techniques. Figure 4 illustrates how the Blue win-probability varies as we change $\lambda_t^{(j)}$. The win-probability is not a concave function of $\lambda_t^{(j)*}$, which makes it difficult to derive properties of the optimal solution. In Figure 4 we assume one time period ($N = 1$) and two risk levels ($M = 2$); the other baseline parameters appear in Table 6. There is only one move-rate $\lambda_1^{(1)}$, which we denote simply λ in Figure 4. The three curves correspond to three different combinations of α_r and β_r. In the bottom curve Red has a significant advantage in risk level $r = 2$, and so Blue sets $\lambda^* = \infty$ to avoid the higher risk level. In the middle dashed curve, Blue has a significant advantage in risk level $r = 2$, and so Blue sets $\lambda^* = 0$ to attempt to reach that risk level. The more interesting case appears in the top curve where Blue has the advantage in risk level $r = 1$ and Red has the advantage in risk level $r = 2$. Blue wants to spend some time firing to benefit from the advantage in the lower risk level but knows that if Blue fires for too long the risk will increase and tip the scales in favor of Red; therefore the optimal move-rate is $\lambda^* = 3.73$ and not one of the extremes. If we extend the x-axis in Figure 4 indefinitely, all three plots converge to the same value: $\frac{\delta p_0}{\delta p_0 + \gamma}$. In Appendix G.2 we present two more numerical illustrations, including a scenario with a local maximum that is not the global maximum.

$\mu_r = (5, 0)$	$\gamma = 2$	$\delta = 3$	$p_0 = 0.55$

Table 6: Parameter values associated with Figure 4

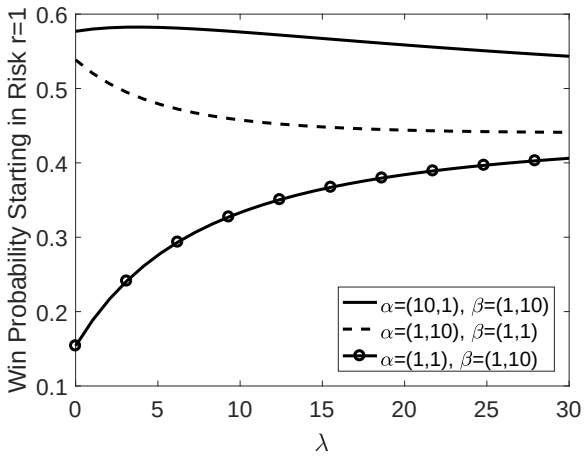


Figure 4: $N = 1$, $R = 2$, for various combination of α_r and β_r. Other parameters appear in Table 6

The top curve in Figure 4 is similar to the scenario defined in Table 1 in Section 5 because Blue has the firing advantage in lower risk levels and Red has the advantage in higher risk levels. Table 7 presents $\lambda_t^{(j)*}$ for the scenario defined in Table 1 in the partial-information setting. We must take care in comparing the partial-information move-rates $\lambda_t^{(j)*}$ in Table 7 with the full-information move-rates $\lambda_{t,r}^*$ in Table 2. Both tables have a row for each time period t. However Table 2 has a column for each risk level r, whereas Table 7 has a column for each time period. The time period j along the columns in Table 7 corresponds to the last period Blue fired a risk-free shot. The diagonal terms $\lambda_t^{(t)*}$ represent the move-rates when Blue enters risk level $r = 1$ after a risk-free shot in period t, and hence are most comparable to the first column of Table 2. It is interesting to see how $\lambda_t^{(j)*}$ changes in Table 7 as we vary j for a fixed t. For $t = 2$, $\lambda_2^{(2)*}$ is smaller than $\lambda_2^{(1)*}$ because Blue uses $\lambda_2^{(2)*}$ immediately after entering risk level $r = 1$ and wants to take advantage of the favorable relative firing rate in $r = 1$. However, Blue uses $\lambda_2^{(1)*}$ only after some time has elapsed since the last risk-free shot in the previous time period; thus it is more likely that the risk level has changed to a less favorable firing condition for Blue.

	$j=1$	$j=2$	$j=3$
$t=1$	0.29	—	—
$t=2$	0.11	0.05	—
$t=3$	0	0	0

Table 7: $\lambda_t^{(j)*}$ for the parameters in Table 1

For the scenario defined in Table 1, the Blue win-probability in the full-information setting using $\lambda_{t,r}^*$ in Table 2 is 0.707. In the partial-information setting using $\lambda_t^{(j)*}$ in Table 7, the win-probability is 0.614, which is a substantial decrease from the full-information setting. We conclude this section with a simulation experiment to assess how much worse Blue does in the partial-information setting compared to full-information. We also examine how well the *Time-Threshold* heuristic described in Section 5.1 performs in the partial-information setting. Table 8 contains the results for the same 9000 scenarios as in Table 5 in Section 5 and in a similar format. The metric of interest is performance relative to the optimal partial-information win-probability. Therefore the metric is positive for the *Time-Threshold* heuristic and negative for the full-information optimal policy.

	Time-Threshold	Full-Information Optimal
All replications	0.0003	-0.020
Optimal in $(0.15, 0.85)$	0.006	-0.053

Table 8: 95th percentile performance relative to the optimal partial-information Blue win-probability

The *Time-Threshold* heuristic generates nearly optimal results. Furthermore, there is only a small difference between the full-information solution and the partial-information solution. Situations where there is a larger performance gap between both *Time-Threshold* and partial-information and full-information and partial-information are similar to the scenario in Table 1. In these scenarios Blue has the firing advantage for early risk levels and Red has the firing advantage for later risk levels, which implies greater benefits if Blue can tailor the move strategy to the specific risk level. For the Table 1 scenario the full-information win-probability is 0.707, partial-information is 0.614, and *Time-Threshold* is 0.582.

Throughout Sections 4–6 we require only one hit to destroy either side. In some scenarios destruction may require multiple hits. Blue may consist of a battery with six cannons, and so Red must hit Blue six times to destroy Blue. One could approximate a multiple hit requirement by decreasing the effective fire rates α and β. Another approach would be to explicitly track the *health* of both Blue and Red during the battle (e.g., number of remaining Blue cannons). This would increase the state space from two dimensions (time and risk) to four with the addition of Blue and Red health. In the next section we incorporate the health level into a more general simulation model.

7 Simulation Model

The models in Sections 4–6 are approximations of reality. Perhaps the most controversial part of our model is that the move decision is random. This may not be unreasonable because the actual departure time may vary substantially from the planned departure time due to the chaos on the battlefield. Furthermore, the optimal move decisions are often deterministic: Blue either scoots immediately ($\lambda^* = \infty$) or fights until the state changes ($\lambda^* = 0$). The deterministic *Time-Threshold* heuristic also performs well in Section 6 when the optimal move time is non-deterministic. Even if our model is a simplified approximation, the results can provide insight about more realistic settings.

We could extend the models in Sections 4–6 to incorporate more realism. As mentioned at the end of the previous section, we could track the health of Blue and Red. Blue may consist of several cannons and Red may have several targets (e.g., bridges, refineries, logistic centers). Therefore to destroy either side requires multiple hits. Another way to add realism to the model would be to have a separate Blue risk and Red risk and count the number of shots fired by each side to better capture how risk increases. Incorporating these aspects into our modeling framework from Sections 4–6 is conceptually straightforward and requires expanding our state space to track additional factors. Rather than continue with this ap-

proach, we instead develop a simulation model. In the interest of space, we defer most of the implementation details to Appendix H. We no longer restrict random variables to follow exponential distributions and for simplicity we primarily use uniform distributions for most of the random variables in the simulation. The travel time and the inter-shot times for both Blue and Red are all uniform random variables. The time-window is deterministic. We also track Blue and Red health, which corresponds to the number of remaining hits until the side is destroyed. We explicitly connect an increase in risk to improved firing accuracy from processing intelligence. Specifically, after every Blue and Red shot, Red processes the shot information (e.g., radar signature) to improve firing accuracy. The time to process the shot information is a uniform random variable. Blue's accuracy improves only after adjusting to a previous Blue shot. Because Blue does not fire at the Red artillery, Red artillery trajectory provides no information on the location of the desired Red targets.

Blue's move decision now depends upon five factors

- Blue's health
- Number of Blue shots fired
- Number of Red shots fired
- Current time
- Time since Blue's risk-free shot at current location

The move decision does not depend upon Red's health because we assume that Blue has imperfect battle damage assessment and thus does not know Red's health in real time. Blue maps the above five factors to the time when Blue moves to a new location. If the current time is less than the move time, Blue stays at the current location, otherwise Blue moves. Even with a coarse discretization of the two time variables listed above (e.g., 5 to 10 levels), the number of move decision variables can easily exceed 1000. Optimizing over that many variables embedded in a simulation is difficult. We use a genetic algorithm approach

(see Whitley (1994); Mitchell (1998)) to solve for the move variables using Matlab's built-in genetic algorithm functionality. Unfortunately, there is no guarantee that the solution generated by the genetic algorithm is optimal or even near optimal. We compare the results from the genetic algorithm to simpler heuristics based on the results in Sections 4–6 in the hopes that the heuristics provide effective move policies. Based on the effectiveness of a pure time-threshold policy in Sections 4–6, we focus on two heuristics

- *Pure-Time-Threshold*: Blue scoots for all time less than some threshold τ and fights from one location thereafter. This corresponds to the optimal policy in Section 4 and the *Time-Threshold* heuristic in Sections 5–6.

- *Health-Based-Time-Threshold*: If Blue's current health level is h_B, then Blue scoots for all time less than threshold $\tau(h_B)$ and fights thereafter until Blue's health decreases.

For the *Pure-Time-Threshold* heuristic, we estimate the optimal threshold by enumeration. For the *Health-Based-Time-Threshold* we must estimate the optimal threshold for each Blue health level. We utilize a genetic algorithm to do this, which performs well because we only need to optimize over at most six decision variables for the scenarios we consider.

We run each simulation 10000 times to estimate Blue's win-probability. We consider 750 different scenarios. Both heuristics are within 0.05, in absolute terms, of the genetic algorithm solution in 0.91 of the scenarios. The strong performance of the heuristics is partly the result of the difficulty of finding optimal solutions with genetic algorithms. Future work could improve the optimization approach to generate solutions closer to the optimum. However, it is reassuring that the intuitive time-threshold policy suggested by the simple model in Section 4 performs nearly as well as a more complicated and much more time-intensive algorithm. The scenarios where the heuristics perform poorly relative to the genetic algorithm had large increases in Blue firing accuracy as Blue observes more Blue shots. Blue is apt to stay in the same location for a few shots in order to fire high accuracy rounds. The heuristics do not account for this information and thus have a larger performance gap.

8 Conclusion

We formulate several models of increasing complexity to analyze shoot-and-scoot tactics for artillery forces. Most commanders are risk averse and want to avoid casualties and thus tend to move frequently to avoid the enemy's counter-fire. Moving frequently reduces risk to Blue, but limits Blue's ability to inflict damage on Red. Fighting from the same location improves accuracy and increases the overall firing rate. When the mission is time-critical, frequent scooting may no longer be an effective tactic. In some situations, Blue should spend a reasonable amount of time firing at Red from the same location. These include situations where the time-window is close to expiring.

We primarily focus on Markov models to gain insight and develop simple heuristics. The heuristics are variants of time-threshold policies, where Blue scoots in early periods and fights in later periods. The pure time-threshold policy is optimal for the model in Section 4. These heuristics perform well when applied to more complex and realistic settings via a simulation experiment.

Future work could more explicitly model where Blue fires from after a move. For example a larger travel-rate δ corresponds to moving to a closer location. If Blue moves to a location too close to the previous location, Red may hit Blue before Blue fires the risk-free shot. Another future effort could consider a setting where Blue artillery and Red artillery fire directly at each other. We could develop a sequential game theory model with two mobile players. Each round a player chooses to either move or fire at an enemy location. Some locations are hidden from the enemy until a player fires from the location for the first time.

Acknowledgment

This paper is based on the Master's Thesis of Younglak Shim (first author). We would like to thankful Captain Alex Ryan, USMC, for very informative discussions about the nature of artillery engagements.

References

Alexander, Robert Michael. 1977. An analysis of aggregated effectiveness for indirect artillery fire on fixed targets. Master's thesis, Georgia Institute of Technology.

Ancker, C. J. 1966. Stochastic duels of limited time duration. *Canadian Operational Research Society Journal* **4**(2) 59–81.

Ancker, C. J. 1975. Stochastic duels with round-dependent hit probabilities. *Naval Research Logistics* **22**(3) 575–583.

Ancker, C. J., Trevor Williams. 1965. Some discrete processes in the theory of stochastic duels. *Operations Research* **13**(2) 202–216.

Armstrong, Michael J. 2004. Effects of lethality in naval combat models. *Naval Research Logistics* **51**(1) 28–43.

Armstrong, Michael J. 2005. A stochastic salvo model for naval surface combat. *Operations Research* **53**(5) 830–841.

Armstrong, Michael J. 2007. Effective attacks in the salvo combat model: Salvo sizes and quantities of targets. *Naval Research Logistics* **54**(1) 66–77.

Armstrong, Michael J. 2014a. Modeling short-range ballistic missile defense and Israel's Iron Dome system. *Operations Research* **62**(5) 1028–1039.

Armstrong, Michael J. 2014b. The salvo combat model with a sequential exchange of fire. *Journal of the Operational Research Society* **65**(10) 1593–1601.

Barfoot, C Bernard. 1989. Continuous-time Markov duels: Theory and application. *Naval Research Logistics* **36**(3) 243–253.

Barr, Donald R, Larry D Piper. 1972. A model for analyzing artillery registration procedures. *Operations Research* **20**(5) 1033–1043.

Bhashyam, N, Naunihal Singh. 1967. Stochastic duels with varying single-shot kill probabilities. *Operations Research* **15**(2) 233–244.

Bowman, Tom. 2016. U.S. military trainers teach Afghan troops to wield artillery. July 5, 2016. National Public Radio.

Brown, Gerald, Matthew Carlyle, Douglas Diehl, Jeffrey Kline, Kevin Wood. 2005. A two-sided optimization for theater ballistic missile defense. *Operations Research* **53**(5) 745–763.

Cha, Young-Ho, Yeong-Dae Kim. 2010. Fire scheduling for planned artillery attack operations under time-dependent destruction probabilities. *Omega* **38**(5) 383–392.

Choe, Jum Soo. 1983. Some stochastic-duel models of combat. Master's thesis, Naval Postgraduate School.

Dastrup, Boyd L. 1992. King of battle: a branch history of the US Army's field artillery. Army Training And Doctrine Command, Fort Monroe, VA.

Davis, Michael T, Matthew J Robbins, Brian J Lunday. 2017. Approximate dynamic programming for missile defense interceptor fire control. *European Journal of Operational Research* **259**(3) 873–886.

Deelstra, Griselda, Guillaume Plantin. 2013. *Risk theory and reinsurance*. Springer.

Finlon, Matthew A. 1991. Analysis of the field artillery battalion organization using a Markov chain. Master's thesis, Naval Postgraduate School.

Fox, Major Amos. 2017. Understanding modern Russian war: Ubiquitous rocket, artillery to enable bttlefield swarming, siege warfare. *Fires Bulletin* **107**.

Guzik, Dennis M. 1988. Markov model for measuring artillery fire support effectiveness. Master's thesis, Naval Postgraduate School.

Harari, Lior. 2008. Defense against rocket attacks in the presence of false cues. Master's thesis, Naval Postgraduate School.

Harris, Michael C. 2017. Is tube artillery a viable fire support platform for the United States military on the battlefields of the future. Master's thesis, Air War College, Air University Maxwell AFB United States.

Hohzaki, R, A Washburn. 2001. The diesel submarine flaming datum problem. *Military Operations Research* **6**(4) 19–30.

Holmes, Richard, Charles Singleton. 2001. *The Oxford companion to military history*. Oxford University Press, Oxford.

Hughes, Wayne P. 1995. A salvo model of warships in missile combat used to evaluate their staying power. *Naval Research Logistics* **42**(2) 267–289.

Keane, Therese. 2011. Combat modelling with partial differential equations. *Applied Mathematical Modelling* **35**(6) 2723–2735.

Koba, Michael. 1996. Artillery strike force. School of Advanced Military Studies, United States Army Command and General Staff College.

Kopp, Carlo. 2005. Artillery for the Army: Precision fire with mobility. *Defence Today* **4**(3).

Kwon, Ojeong, Kyungsik Lee, Sungsoo Park. 1997. Targeting and scheduling problem for field artillery. *Computers & Industrial Engineering* **33**(3-4) 693–696.

Lucas, Thomas W, Turker Turkes. 2004. Fitting lanchester equations to the battles of kursk and ardennes. *Naval Research Logistics* **51**(1) 95–116.

Marshall, Kneale T. 1994. Quantifying counterforce and active defense in countering theater ballistic missiles. *Military Operations Research* **1** 35–48.

McCarthy, Niall. 2017. Why the North Korean artillery factor makes military action extremely risky. October 2, 2017. *Forbes* .

McNaught, Ken R. 2002. Markovian models of three-on-one combat involving a hidden defender. *Naval Research Logistics* **49**(7) 627–646.

Miller, Stephen W. 2016. Shoot and scoot. *Armada International* **1**.

Miller, Stephen W. 2017. Asia's artillery rennaissance. *Asian Military Review* **2**.

Mitchell, Melanie. 1998. *An introduction to genetic algorithms*. MIT press.

Nadler, Jack, Joan Eilbott. 1971. Optimal sequential aim corrections for attacking a stationary point target. *Operations Research* **19**(3) 685–697.

Ratliff, Timothy M. 2017. Field artillery and fire support at the operational level: An analysis of operation desert storm and operation iraqi freedom. Master's thesis, US Army School for Advanced Military Studies.

Sharp, Jeremy. 2006. Lebanon: The Israel-Hamas-Hezbollah conflict. Congressional Research Service.

Shim, Younglak. 2017. An analysis of shoot-and-scoot tactics. Master's thesis, Naval Postgraduate School.

Snow, Shawn. 2018. These marines in Syria fired more artillery than any battalion since Vietnam. February 6, 2018. *Marine Corps Times* .

Temiz, Yusuf. 2016. Artillery survivability model. Master's thesis, Naval Postgraduate School.

The MathWorks, Inc. 2016. Matlab 2016b. URL `https://www.mathworks.com/products/matlab.html`.

Tinder, Brett. 2017. Artillery units embrace 'hide and seek' tactics to gain decisive edge. August 2, 2017. *Army.mil* .

United States Army. 1990. Army regulation 600-82: The U.S. Army regimental system. Headquarters, Department of the Army.

United States Army. 1996. FM 6-50: Tactics, techniques, and procedures for the field artillery cannon battery. Headquarters, Department of the Army.

United States Army. 2001. FM 3-09.22: Tactics, techniques, and procedures for corps artillery, division artillery, and field artillery brigade operations. Headquarters, Department of the Army.

United States Army. 2014. FM 3-09: Field artillery operations and fire support. Headquarters, Department of the Army.

United States Marine Corps. 1998. MCWP 3-16.6: Supporting arms observer, spotter and controller. Marine Corps Combat Development Command.

United States Marine Corps. 2002. MCWP 3-16.1: Artillery operations. Marine Corps Combat Development Command.

United States Marine Corps. 2016. Organization of the United States Marine Corps. Headquarters, United States Marine Corps.

Washburn, Alan R, Moshe Kress. 2009. *Combat modeling*. Springer, Heidelberg.

Whitley, Darrell. 1994. A genetic algorithm tutorial. *Statistics and computing* **4**(2) 65–85.

Williams, G Trevor, C. J. Ancker. 1981. Stochastic duels with displacements (suppression). *Naval Research Logistics* **28**(3) 519–524.

Williams, Trevor, Clinton J. Ancker. 1963. Stochastic duels. *Operations Research* **11**(5) 803–817.

APPENDIX

A Proof of Proposition 1: $\lambda_t^* \in \{0, \infty\}$

Simplifying the notation of equation (2) yields

$$\mathbf{P}[t] = \frac{a}{b + \lambda_t c} + \lambda_t \frac{d}{b + \lambda_t c}. \tag{A.1}$$

It is a straightforward exercise via differentiation to show that $\mathbf{P}[t]$ is a monotonic function of λ_t. It increases if $\frac{d}{c} > \frac{a}{b}$. Substituting in the specific values of a, b, c, d from (2), reveals that $\lambda_t^* = 0$ if

$$\mathbf{F}[t] \geq \mathbf{S}[t],$$

otherwise $\lambda_t^* = \infty$. $\mathbf{F}[t]$ and $\mathbf{S}[t]$ are the fighting costs and scooting costs, respectively, defined in Proposition 1.

B Proof of Proposition 2: Monotonicity of $\mathbf{P}[t]$

We will proceed via induction. First we prove the base case: $\mathbf{P}[N-1] \geq \mathbf{P}[N]$. By inspection of (3)–(4)

$$\mathbf{F}[N-1] \geq \mathbf{F}[N] \tag{B.1}$$

$$\mathbf{S}[N-1] \geq \mathbf{S}[N]. \tag{B.2}$$

This follows because $\mathbf{P}[N+1] = 0$. Because $\mathbf{P}[t] = \max(\mathbf{F}[t], \mathbf{S}[t])$, conditions (B.1)–(B.2) imply that

$$\mathbf{P}[N-1] \geq \mathbf{F}[N-1] \geq \mathbf{F}[N] \tag{B.3}$$

$$\mathbf{P}[N-1] \geq \mathbf{S}[N-1] \geq \mathbf{S}[N]. \tag{B.4}$$

Conditions (B.3)–(B.4) imply that $\mathbf{P}[N-1] \geq \max(\mathbf{F}[N], \mathbf{S}[N]) = \mathbf{P}[N]$, which completes the base case.

The general case follows similar logic. First we assume that $\mathbf{P}[s] \geq \mathbf{P}[s+1]$ for all $t < s \leq N$. This induction assumption implies

$$\begin{aligned} \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\mathbf{P}[t+1] &\geq \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\mathbf{P}[t+2] \\ &\rightarrow \mathbf{F}[t] \geq \mathbf{F}[t+1] \end{aligned} \tag{B.5}$$

Similarly, $\mathbf{S}[t] \geq \mathbf{S}[t+1]$. To see this:

$$\begin{aligned} \mathbf{S}[t] - \mathbf{S}[t+1] = &\frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\delta+\gamma}\sum_{s=t+2}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-2}(1-p_0)(\mathbf{P}[s-1]-\mathbf{P}[s]) \\ &+ \frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\delta+\gamma}\left(\frac{\gamma}{\delta+\gamma}\right)^{N-t-1}(p_0+(1-p_0)\mathbf{P}[N]) \\ \geq &\, 0 \end{aligned} \tag{B.6}$$

Condition (B.6) follows because $\mathbf{P}[s-1] - \mathbf{P}[s] \geq 0$ by the induction assumption. We combine conditions (B.5) and (B.6) in the same fashion as Conditions (B.3)–(B.4), which yields $\mathbf{P}[t] \geq \mathbf{P}[t+1]$. This completes the proof.

C Proof of Proposition 3: Optimality of Threshold Policy

We start with several edge cases before turning to the general case. The first Lemma handles the case when $\tau = N$: Blue scoots in the final time period.

Lemma 1. *If $\tau = N$, then $\lambda_t^* = \infty$ for all time periods $1 \leq t \leq N$.*

The proof appears in Appendix C.1.

If $\tau = -\infty$ (Blue never scoots) or $\tau = 1$ (Blue scoots in first period, and then fights for the remainder of the battle), then Proposition 3 follows immediately by definition of τ. For the remainder of this section, we focus on the general case where $1 < \tau < N$. We must show that $\lambda_t = \infty$ for all $1 \leq t \leq \tau$. The following Lemma states that if Blue optimally scoots in two consecutive periods, Blue should optimally scoot in all previous periods.

Lemma 2. *If $\lambda_t^* = \infty$ and $\lambda_{t-1}^* = \infty$, then $\lambda_s^* = \infty$ for all $1 \leq s \leq t-1$.*

The proof of Lemma 2 appears in Appendix C.2. By definition of τ, $\lambda_\tau = \infty$. By Lemma 2, we can complete the proof if we show that $\lambda_{\tau-1} = \infty$. The following Lemma provides this final piece

Lemma 3. *For $1 < \tau < N$, $\lambda_{\tau-1}^* = \infty$.*

The proof of Lemma 3 appears in Appendix C.3.

C.1 Proof of Lemma 1: Scoot in All Time Periods

If Blue scoots in period N ($\lambda_N^* = \infty$), then by definition $\hat{\mathbf{S}}_{\mathbf{F}}[N] > \hat{\mathbf{F}}_{\mathbf{all}}[N]$. This implies, by (5) and (6), that

$$\frac{\delta p_0}{\gamma + \delta p_0} > \frac{\alpha}{\alpha + \beta + \gamma}. \tag{C.1}$$

To show Blue should scoot in all periods ($\lambda_t^* = \infty$ for all $1 \leq t \leq N$), we follow an inductive approach. We assume that there exists some $s \leq N$, such that it is optimal to scoot for all $s \leq k \leq N$: $\lambda_k^* = \infty$. The base case $s = N$ is one such s by assumption. We can explicitly compute $\mathbf{P}[s]$ when Blue scoots for all $s \leq k \leq N$. We denote $\hat{\mathbf{S}}_{\mathbf{all}}[s]$ as the probability Blue wins if it uses a scoot-only policy for all time periods from s onward:

$$\mathbf{P}[s] = \hat{\mathbf{S}}_{\mathbf{all}}[s] \equiv 1 - \left(\frac{\gamma}{\gamma + \delta p_0}\right)^{N-s+1}. \tag{C.2}$$

Using this scoot-only policy, Blue wins unless the time-window closes. One can formally show that (2) simplifies to (C.2) when $\lambda_k = \infty$ for all $s \leq k \leq N$ via induction.

To determine the optimal action in period $s-1$, Blue compares

$$\mathbf{P}[s-1] = \max\left(\frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{S}}_{\mathbf{all}}[s], \hat{\mathbf{S}}_{\mathbf{all}}[s-1]\right) \tag{C.3}$$

Substituting in for $\hat{\mathbf{S}}_{\mathbf{all}}[s]$ from (C.2) into (C.3) and rearranging terms, we see that Blue will also scoot in state $s-1$ if

$$\begin{aligned}\left(1 - \left(\frac{\gamma}{\gamma+\delta p_0}\right)^{N-s+2}\right) &> \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\left(1 - \left(\frac{\gamma}{\gamma+\delta p_0}\right)^{N-s+1}\right) \\ \rightarrow \frac{\beta}{\alpha+\beta+\gamma} &> \left(\frac{\gamma}{\gamma+\delta p_0}\right)^{N-s+1}\left(\frac{\gamma}{\gamma+\delta p_0} - \frac{\gamma}{\alpha+\beta+\gamma}\right)\end{aligned} \tag{C.4}$$

If the right-hand side in (C.4) is negative, then condition (C.4) holds for all s, and thus Blue should scoot in period $s-1$ and inductively for all periods. If the right-hand side in (C.4) is positive, then the right-hand side of (C.4) is an increasing function of s for $s \leq N+1$. Condition (C.4) holds for $s = N+1$ because condition (C.4) simplifies to condition (C.1) for $s = N+1$. By the monotonicity of the right-hand side of (C.4), condition (C.4) will thus also hold for all $s < N+1$, and hence Blue should scoot in all time periods. This completes the proof.

C.2 Proof of Lemma 2: Two Consecutive Scoot Periods

If $\tau = N$, then the result follows immediately by Lemma 1. For the remainder of this proof, we assume $\tau < N$, which holds if and only if

$$\frac{\delta p_0}{\gamma + \delta p_0} \leq \frac{\alpha}{\alpha + \beta + \gamma} \tag{C.5}$$

If $\lambda_t^* = \infty$, then the expression for $\mathbf{S}[t-1]$ simplifies greatly from (4). We present the following as a Lemma because we use it for other results.

Lemma 4. *If $\lambda_t^* = \infty$, then*

$$\mathbf{S}[t-1] = \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\mathbf{P}[t]. \tag{C.6}$$

The proof of Lemma 4 appears in Appendix C.2.1. If, in addition, $\lambda_{t-1}^* = \infty$, then

$$\begin{aligned} &\mathbf{S}[t-1] > \mathbf{F}[t-1] \\ \rightarrow &\frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\mathbf{P}[t] > \frac{\alpha}{\alpha + \beta + \gamma} + \frac{\gamma}{\alpha + \beta + \gamma}\mathbf{P}[t]. \end{aligned} \tag{C.7}$$

We use (C.6) and (3) to produce (C.7). Condition (C.7) is a comparison of two linear functions where the independent variable is $\mathbf{P}[t]$. By (C.5), the "slope" of the left-hand side of (C.7) is larger than the slope of the right-hand side of (C.7):

$$\frac{\gamma}{\gamma + \delta p_0} \geq \frac{\gamma}{\alpha + \beta + \gamma}$$

Thus condition (C.7) will hold if we replace $\mathbf{P}[t]$ with any $x \geq \mathbf{P}[t]$. Namely, by the monotonicity of $\mathbf{P}[t]$ (Proposition 2), condition (C.7) will also hold for $\mathbf{P}[t-1]$. Hence, Blue should scoot in period $t-2$ ($\lambda_{t-2}^* = \infty$). Iterating this logic, condition (C.7) holds for all $1 \leq s \leq t-1$ and thus $\lambda_s^* = \infty$.

C.2.1 Proof of Lemma 4: Win-Probability if Next Period Optimally Scoots

The expressions for $\mathbf{S}[t]$ and $\mathbf{S}[t-1]$ only differ by one term in the summation of (4). Manipulating (4) yields $\mathbf{S}[t-1]$ as a function of $\mathbf{S}[t]$

$$\mathbf{S}[t-1] = \left(\mathbf{S}[t] - \frac{\delta p_0}{\gamma + \delta p_0}\right)\frac{\gamma}{\gamma+\delta} + \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\gamma+\delta}\left(p_0 + (1-p_0)\mathbf{P}[t]\right) \quad \text{(C.8)}$$

Because we assume $\lambda_t^* = \infty$, we can substitute $\mathbf{S}[t] = \mathbf{P}[t]$ in (C.8). Further algebra transforms (C.8) to (C.6).

C.3 Proof of Lemma 3: Scoot in Period $\tau - 1$

By definition of τ in equation (7), we have

$$\hat{\mathbf{F}}_{\mathbf{all}}[t] \geq \hat{\mathbf{S}}_{\mathbf{F}}[t] \quad \forall \tau + 1 \leq t \leq N \quad \text{(C.9)}$$

$$\hat{\mathbf{F}}_{\mathbf{all}}[\tau] < \hat{\mathbf{S}}_{\mathbf{F}}[\tau] \quad \text{(C.10)}$$

To complete the proof, we must show that $\lambda_{\tau-1}^* = \infty$, which occurs if and only if (see Proposition 1)

$$\mathbf{F}[\tau-1] < \mathbf{S}[\tau-1]$$

$$\rightarrow \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\mathbf{P}[\tau] < \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\mathbf{P}[\tau] \quad \text{(C.11)}$$

$$\rightarrow \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{S}}_{\mathbf{F}}[\tau] < \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\hat{\mathbf{S}}_{\mathbf{F}}[\tau] \quad \text{(C.12)}$$

The left-hand side of (C.11) follows directly from (3) and the right-hand side of (C.11) is a result of Lemma 4 because $\lambda_\tau^* = \infty$. Finally we replace $\mathbf{P}[\tau]$ in (C.11) with $\hat{\mathbf{S}}_{\mathbf{F}}[\tau]$ in (C.12) by (C.9)–(C.10).

We will derive a contradiction by assuming that (C.12) does not hold. That is we assume

$$\frac{\alpha}{\alpha+\beta+\gamma}+\frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{S}}_{\mathbf{F}}[\tau]\geq\frac{\delta p_0}{\gamma+\delta p_0}+\frac{\gamma}{\gamma+\delta p_0}\hat{\mathbf{S}}_{\mathbf{F}}[\tau]. \tag{C.13}$$

If condition (C.13) holds, then we will show that condition (C.10) will be violated, which produces a contradiction. Combining (C.9), (C.10), and Proposition 2 yields

$$\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]=\mathbf{P}[\tau+1]\leq\mathbf{P}[\tau]=\hat{\mathbf{S}}_{\mathbf{F}}[\tau] \tag{C.14}$$

By Proposition 1 if $\lambda_N^*=0$ (Lemma assumption) then

$$\frac{\alpha}{\alpha+\beta+\gamma}\geq\frac{\delta p_0}{\gamma+\delta p_0} \tag{C.15}$$

Condition (C.15) also implies

$$\frac{\gamma}{\alpha+\beta+\gamma}\leq\frac{\gamma}{\gamma+\delta p_0} \tag{C.16}$$

Conditions (C.15) and (C.16) imply that if (C.13) holds, then (C.13) will also hold if we replace $\hat{\mathbf{S}}_{\mathbf{F}}[\tau]$ in (C.13) with any value $x\leq\hat{\mathbf{S}}_{\mathbf{F}}[\tau]$. This follows because (C.13) is just the comparison of two linear functions. We use similar logic in Appendix C.2. Consequently (C.13) holds for $\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]\leq\hat{\mathbf{S}}_{\mathbf{F}}[\tau]$, by (C.14):

$$\begin{aligned}\frac{\alpha}{\alpha+\beta+\gamma}+\frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]&\geq\frac{\delta p_0}{\gamma+\delta p_0}+\frac{\gamma}{\gamma+\delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]\\ \rightarrow\hat{\mathbf{F}}_{\mathbf{all}}[\tau]&\geq\frac{\delta p_0}{\gamma+\delta p_0}+\frac{\gamma}{\gamma+\delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1].\end{aligned} \tag{C.17}$$

The second line of (C.17) follows from the first line by (3) and (C.9). The final step is to

show that

$$\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] \geq \frac{\delta}{\delta+\gamma} \sum_{s=\tau+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-\tau-1} \left(p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[s]\right). \tag{C.18}$$

If we substitute condition (C.18) into the right-hand side of (C.17) we produce $\hat{\mathbf{S}}_{\mathbf{F}}[\tau]$ (see (6)). Consequently, condition (C.17) would imply $\hat{\mathbf{F}}_{\mathbf{all}}[\tau] \geq \hat{\mathbf{S}}_{\mathbf{F}}[\tau]$, which is the desired contradiction of (C.10). We show (C.18) by brute force. Manipulating the right-hand side (C.18)

$$\begin{aligned}
&\frac{\delta}{\delta+\gamma} \sum_{s=\tau+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-\tau-1} \left(p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[s]\right) \\
&= \frac{\delta(1-p_0)}{\delta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] + \left(\frac{\delta p_0}{\delta+\gamma} + \frac{\gamma}{\delta+\gamma}\frac{\delta}{\delta+\gamma} \sum_{s=\tau+2}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-\tau-2} \left(p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[s]\right)\right).
\end{aligned} \tag{C.19}$$

Inspection of (6) reveals that the large term in parentheses in (C.19) is a constant multiple of $\hat{\mathbf{S}}_{\mathbf{F}}[\tau+1]$, which simplifies (C.19) to

$$\begin{aligned}
\frac{\delta}{\delta+\gamma} \sum_{s=\tau+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-\tau-1} \left(p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[s]\right) &= \frac{\delta(1-p_0)}{\delta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] + \frac{\gamma+\delta p_0}{\delta+\gamma}\hat{\mathbf{S}}_{\mathbf{F}}[\tau+1] \\
&\leq \hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]
\end{aligned} \tag{C.20}$$

The second line in (C.20) follows from the first by (C.9). Condition (C.20) produces inequality (C.18) and completes the proof.

D Proof of Proposition 4: Comparative Statics

D.1 Blue firing rate α

We introduce the notation $\mathbf{P}[t;\alpha]$, $\mathbf{S}[t;\alpha]$, $\tau(\alpha)$, $\hat{\mathbf{F}}_{\mathbf{all}}[t;\alpha]$, and $\hat{\mathbf{S}}_{\mathbf{F}}[t;\alpha]$ to explicitly connect these quantities to α. By inspection of equation (5) $\hat{\mathbf{F}}_{\mathbf{all}}[t;\alpha]$ is an increasing function of α

$$\frac{\partial}{\partial\alpha}\hat{\mathbf{F}}_{\mathbf{all}}[t;\alpha] \geq 0, \qquad \forall\alpha > 0, 1 \leq t \leq N \tag{D.1}$$

To show that $\mathbf{P}[t;\alpha]$ is also non-decreasing in α, we consider $\alpha_1 < \alpha_2$. We show the result by considering two cases for the value of t. For $t > \max(\tau(\alpha_1), \tau(\alpha_2))$, it is optimal to fight for both α_1 and α_2. By (D.1)

$$\mathbf{P}[t;\alpha_1] = \hat{\mathbf{F}}_{\mathbf{all}}[t;\alpha_1] \leq \hat{\mathbf{F}}_{\mathbf{all}}[t;\alpha_2] = \mathbf{P}[t;\alpha_2] \qquad \forall t > \max(\tau(\alpha_1), \tau(\alpha_2)) \tag{D.2}$$

Next we consider $1 < t \leq \max(\tau(\alpha_1), \tau(\alpha_2))$. We argue by induction that

$$\mathbf{F}[t;\alpha_1] \leq \mathbf{F}[t;\alpha_2]$$
$$\mathbf{S}[t;\alpha_1] \leq \mathbf{S}[t;\alpha_2],$$

where $\mathbf{F}[t;\alpha]$ and $\mathbf{S}[t;\alpha]$ are defined in (3)–(4). Since by definition $\mathbf{P}[t;\alpha] = \max(\mathbf{F}[t;\alpha], \mathbf{S}[t;\alpha])$, this completes the proof.

The base case is $t = \max(\tau(\alpha_1), \tau(\alpha_2))$. $\mathbf{S}[t;\alpha]$ is a linear combination of $\mathbf{P}[t;\alpha]$ for $t > \max(\tau(\alpha_1), \tau(\alpha_2))$. Therefore $\mathbf{S}[t;\alpha_1] \leq \mathbf{S}[t;\alpha_2]$ follows from (D.2). Similar logic applies to $\mathbf{F}[t;\alpha]$. Furthermore $\mathbf{F}[t;\alpha]$ is a weighted combination of 1 and $\mathbf{P}[t;\alpha] \leq 1$, and the weight on 1 increases with α. Consequently $\mathbf{F}[t;\alpha_1] \leq \mathbf{F}[t;\alpha_2]$ holds and the base case is complete.

The general case follows by the same logic because for a general t the induction assumption

implies

$$\mathbf{P}[s;\alpha_1] = \max(\mathbf{F}[s;\alpha_1], \mathbf{S}[s;\alpha_1]) \leq \max(\mathbf{F}[s;\alpha_2], \mathbf{S}[s;\alpha_2]) = \mathbf{P}[s;\alpha_2] \quad \forall s > t \tag{D.3}$$

This completes the proof.

D.2 Red firing rate β

We introduce the notation $\mathbf{P}[t;\beta]$, $\tau(\beta)$, $\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]$, and $\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta]$ to explicitly connect these quantities to β. We first present three Lemmas that provide the pieces for our proof.

Lemma 5.

$$\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] \leq 0, \qquad \forall \beta > 0, 1 \leq t \leq N$$

The proof of Lemma 5 and Lemma 6 appears in Appendix D.2.1.

Lemma 6. *For all* $1 \leq t \leq N$,

$$\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]\Big|_{\beta=\beta_0} \leq \frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[s;\beta]\Big|_{\beta=\beta_0}, \qquad \forall \beta_0 > 0, t < s \leq N$$

Lemma 7. *If* $\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta_1] > \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_1]$, *then* $\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta_2] > \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_2]$ *for any* $\beta_2 > \beta_1$

The proof appears in Appendix D.2.2.

Lemma 7 immediately implies that $\tau(\beta)$ is non-decreasing in β by the definition of τ in equation (7).

To show that $\mathbf{P}[t;\beta]$ is non-increasing in β we consider $\beta_1 < \beta_2$. We assume Blue uses threshold $\tau(\beta_2) \geq \tau(\beta_1)$ when $\beta = \beta_1$, instead of the optimal $\tau(\beta_1)$. This produces the suboptimal win-probability of $\bar{\mathbf{P}}[t;\beta_1]$:

$$\bar{\mathbf{P}}[t;\beta_1] = \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\delta+\gamma}\sum_{s=t+1}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-1}\left(p_0+(1-p_0)\bar{\mathbf{P}}[s;\beta_1]\right) \qquad 1 \le t \le \tau(\beta_2) \tag{D.4}$$

$$\bar{\mathbf{P}}[t;\beta_1] = \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_1] = \mathbf{P}[t;\beta_1] \qquad t > \tau(\beta_2) \tag{D.5}$$

The second equality in (D.5) follows because $\tau(\beta_2) \ge \tau(\beta_1)$ (see discussion below Lemma 7).

Clearly $\bar{\mathbf{P}}[t;\beta_1] \le \mathbf{P}[t;\beta_1]$ since Blue does not use the optimal threshold $\tau(\beta_1)$. Lemma 5 implies that

$$\bar{\mathbf{P}}[t;\beta_1] = \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_1] \ge \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_2] = \mathbf{P}[t;\beta_2] \qquad \text{for } t > \tau(\beta_2). \tag{D.6}$$

Condition (D.6) combined with comparison of (4) with (D.4) yields that $\bar{\mathbf{P}}[t;\beta_1] \ge \mathbf{P}[t;\beta_2]$ for $1 \le t \le \tau(\beta_2)$. Thus $\bar{\mathbf{P}}[t;\beta_1] \ge \mathbf{P}[t;\beta_2]$ for all $1 \le t \le \tau(\beta_2)$ and hence $\mathbf{P}[t;\beta_1] \ge \mathbf{P}[t;\beta_2]$ and the proof is complete.

D.2.1 Proof of Lemma 5 and Lemma 6

Differentiating (5) yields

$$\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] = \frac{\alpha}{(\alpha+\beta)^2}\left(-1+\left(\frac{\gamma}{\alpha+\beta+\gamma}\right)^{N-t+1}\left(1+(N-t+1)\frac{\alpha+\beta}{\alpha+\beta+\gamma}\right)\right)$$

Ignoring the initial constant and replacing $x = N-t+1$ and $p = \frac{\gamma}{\alpha+\beta+\gamma}$ produces

$$\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] = -1 + p^x\left(1+x(1-p)\right) \tag{D.7}$$

To complete the proof for Lemma 6, we show that the right-hand side of (D.7) decreases in x. Since $x = N-t+1$, this implies the non-decreasing relationship in Lemma 6. Taking

the derivative of (D.7) with respect to x yields

$$p^x \left(\log p + (1-p)(x \log p + 1)\right) \tag{D.8}$$

We show (D.8) is always negative and hence $\frac{\partial}{\partial \beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]$ is non-decreasing in t. We ignore the outer p^x factor and generate the following inequality by setting $x = 0$ in the interior

$$\log p + (1-p)(x \log p + 1) \leq \log p + (1-p) \tag{D.9}$$

Condition (D.9) follow because the $x \log p$ on the left-hand side of (D.9) is non-positive. Thus if

$$\log p + (1-p) \leq 0, \tag{D.10}$$

we have proven Lemma 6. The left-hand side of (D.10) is 0 if $p = 1$ and it increases over $p \in [0,1]$. Therefore, the left-hand side of (D.9) must be non-positive for any $p \in [0,1]$.

To complete the proof for Lemma 5, we note that by Lemma 6, the largest derivative occurs at $t = N$, so to complete the proof it suffices to show $\frac{\partial}{\partial \beta}\hat{\mathbf{F}}_{\mathbf{all}}[N;\beta] \leq 0$:

$$\frac{\partial}{\partial \beta}\hat{\mathbf{F}}_{\mathbf{all}}[N;\beta] = \frac{\alpha}{(\alpha+\beta)^2}\left(-1 + \left(\frac{\gamma}{\alpha+\beta+\gamma}\right)\left(1 + \frac{\alpha+\beta}{\alpha+\beta+\gamma}\right)\right)$$

By inspection the above is non-positive.

D.2.2 Proof of Lemma 7

By the fundamental theorem of calculus we have

$$\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_2] = \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_1] + \int_{\beta_1}^{\beta_2} \frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]d\beta,$$
$$\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta_2] = \hat{\mathbf{S}}_{\mathbf{F}}[t;\beta_1] + \int_{\beta_1}^{\beta_2} \frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta]d\beta.$$

Since by assumption $\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta_1] > \hat{\mathbf{F}}_{\mathbf{all}}[t;\beta_1]$, to complete the proof it suffices to show

$$\int_{\beta_1}^{\beta_2} \frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta]d\beta \geq \int_{\beta_1}^{\beta_2} \frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]d\beta.$$

Since integration preserves inequalities, we will prove

$$\frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta] \geq \frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] \qquad \forall\beta$$

To do this we define $\frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta]$ in terms of $\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta]$. Differentiating (6)

$$\frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta] = \frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\delta+\gamma}(1-p_0)\sum_{s=t+1}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-1}\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[s;\beta] \leq 0 \tag{D.11}$$

The inequality in (D.11) follows by Lemma 5. By Lemma 6

$$\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[s;\beta] \geq \frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] \qquad t < s \leq N \tag{D.12}$$

Combining (D.11) and (D.12) yields

$$\frac{\partial}{\partial\beta}\hat{\mathbf{S}}_{\mathbf{F}}[t;\beta] \geq \kappa\frac{\partial}{\partial\beta}\hat{\mathbf{F}}_{\mathbf{all}}[t;\beta] \tag{D.13}$$

where

$$1 \geq \kappa = \frac{\gamma}{\gamma + \delta p_0} \frac{\delta}{\delta + \gamma} (1 - p_0) \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta + \gamma} \right)^{s-t-1} \tag{D.14}$$

$\frac{\partial}{\partial \beta} \hat{\mathbf{F}}_{\mathbf{all}}[t; \beta]$ and $\frac{\partial}{\partial \beta} \hat{\mathbf{S}}_{\mathbf{F}}[t; \beta]$ are both non-positive by Lemma 5 and (D.11). This non-positivity combined with $\kappa \in [0, 1]$ and (D.13) implies

$$\frac{\partial}{\partial \beta} \hat{\mathbf{S}}_{\mathbf{F}}[t; \beta] \geq \frac{\partial}{\partial \beta} \hat{\mathbf{F}}_{\mathbf{all}}[t; \beta],$$

which is the desired condition and completes the proof.

D.3 Blue move-rate δ

We introduce the notation $\mathbf{P}[t; \delta]$, $\tau(\delta)$, and $\hat{\mathbf{S}}_{\mathbf{F}}[t; \delta]$ to explicitly connect these quantities to δ. Note that $\hat{\mathbf{F}}_{\mathbf{all}}[t]$ does not depend upon δ. We use the following Lemma in our proof

Lemma 8.

$$\frac{\partial}{\partial \delta} \hat{\mathbf{S}}_{\mathbf{F}}[t; \delta] \geq 0, \qquad \forall \delta > 0, 1 \leq t \leq N$$

The proof for Lemma 8 appears in Appendix D.3.1.

Because $\hat{\mathbf{F}}_{\mathbf{all}}[t]$ does not depend upon δ, Lemma 8 implies

$$\begin{aligned} &\text{if } \hat{\mathbf{S}}_{\mathbf{F}}[t; \delta_1] > \hat{\mathbf{F}}_{\mathbf{all}}[t] \\ &\text{then } \hat{\mathbf{S}}_{\mathbf{F}}[t; \delta_2] > \hat{\mathbf{F}}_{\mathbf{all}}[t] \quad \text{for any } \delta_2 > \delta_1 \end{aligned} \tag{D.15}$$

Condition (D.15) implies that $\tau(\delta)$ is non-decreasing in δ by the definition of τ in equation (7).

To show that $\mathbf{P}[t; \delta]$ is non-decreasing in δ, we consider $\delta_1 < \delta_2$. We now assume that when $\delta = \delta_2$ Blue uses threshold $\tau(\delta_1) \leq \tau(\delta_2)$ instead of the optimal $\tau(\delta_2)$. This produces

the suboptimal win-probability of $\bar{\mathbf{P}}[t;\delta_2]$:

$$\bar{\mathbf{P}}[t;\delta_2] = \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\frac{\delta}{\delta+\gamma}\sum_{s=t+1}^{N}\left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-1}\left(p_0+(1-p_0)\bar{\mathbf{P}}[s;\delta_2]\right) \qquad 1 \le t \le \tau(\delta_1) \tag{D.16}$$

$$\bar{\mathbf{P}}[t;\delta_2] = \hat{\mathbf{F}}_{\mathbf{all}}[t] \qquad t > \tau(\delta_1) \tag{D.17}$$

Clearly $\bar{\mathbf{P}}[t;\delta_2] \le \mathbf{P}[t;\delta_2]$ since Blue does not use the optimal threshold $\tau(\delta_2)$. By definition $\mathbf{P}[t;\delta_1] = \hat{\mathbf{F}}_{\mathbf{all}}[t]$ for $t > \tau(\delta_1)$, and thus by (D.17) $\bar{\mathbf{P}}[t;\delta_2] = \mathbf{P}[t;\delta_1]$ for $t > \tau(\delta_1)$. The final piece is to show that $\bar{\mathbf{P}}[t;\delta_2] \ge \mathbf{P}[t;\delta_1]$ for $1 \le t \le \tau(\delta_1)$. This follows using the same logic as the proof for Lemma 8 in Appendix D.3.1. We define $\bar{\mathbf{P}}[t;\delta_2]$ as a weighted sum of 1 and an expectation of a function of a geometric random variable (see (D.18)). Appendix D.3.1 contains the details for this approach.

D.3.1 Proof of Lemma 8

First let us write rewrite $\hat{\mathbf{S}}_{\mathbf{F}}[t;\delta]$ in terms of a weighted sum of expected values

$$\hat{\mathbf{S}}_{\mathbf{F}}[t;\delta] = w(\delta) \times 1 + (1-w(\delta)) \times \mathbf{E}[g(Z(\delta))] \tag{D.18}$$

Where $w(\delta) = \frac{\delta p_0}{\gamma+\delta p_0}$, $Z(\delta)$ is a geometric random variable with success probability $\frac{\delta}{\delta+\gamma}$, and the function $g(\cdot)$

$$g(k) = \begin{cases} p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[t+k] & \text{if } 1 \le k \le N-t \\ 0 & \text{if } k > N-t \end{cases}$$

Note that $g(k) \le 1$ for all k and hence $\mathbf{E}[g(Z(\delta))] \le 1$ for all δ.

By inspection of (5), $\hat{\mathbf{F}}_{\mathbf{all}}[t]$ decreases in t and hence $g(k)$ is non-increasing in k. If $\delta_1 < \delta_2$, then $Z(\delta_1)$ has first-order stochastic dominance over $Z(\delta_2)$. See Chapter 1.1.2i of Deelstra

and Plantin (2013) for a discussion of stochastic dominance. Since $g(k)$ is a non-increasing function, first-order stochastic dominance implies that $\mathbf{E}[g(Z(\delta_2))] \geq \mathbf{E}[g(Z(\delta_1))]$.

For any $\delta_1 < \delta_2$ we have $w(\delta_1) \leq w(\delta_2)$ and $\mathbf{E}[g(Z(\delta_1))] \leq \mathbf{E}[g(Z(\delta_2))] \leq 1$. Using this inequalities in conjunction with (D.18) yields $\hat{\mathbf{S}}_{\mathbf{F}}[t;\delta_1] \leq \hat{\mathbf{S}}_{\mathbf{F}}[t;\delta_2]$. Since this holds for any $\delta_1 < \delta_2$, $\hat{\mathbf{S}}_{\mathbf{F}}[t;\delta]$ is an non-decreasing function, which completes the proof.

D.4 Time-rate γ

The proof is nearly identical to the corresponding proof for α in Appendix D.1, so we just sketch the approach here. We introduce the notation $\mathbf{P}[t;\gamma]$, $\mathbf{S}[t;\gamma]$, $\tau(\gamma)$, $\hat{\mathbf{F}}_{\mathbf{all}}[t;\gamma]$, and $\hat{\mathbf{S}}_{\mathbf{F}}[t;\gamma]$ to explicitly connect these quantities to γ. By inspection of equation (5) $\hat{\mathbf{F}}_{\mathbf{all}}[t;\gamma]$ is a non-increasing function of γ

$$\frac{\partial}{\partial\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[t;\gamma] \leq 0, \qquad \forall\gamma > 0, 1 \leq t \leq N \tag{D.19}$$

To show that $\mathbf{P}[t;\gamma]$ is non-increasing in γ, we consider $\gamma_1 < \gamma_2$. We show the result by considering two cases for the value of t. For $t > \max(\tau(\gamma_1), \tau(\gamma_2))$, it is optimal to fight for both γ_1 and γ_2. By (D.19)

$$\mathbf{P}[t;\gamma_1] = \hat{\mathbf{F}}_{\mathbf{all}}[t;\gamma_1] \geq \hat{\mathbf{F}}_{\mathbf{all}}[t;\gamma_2] = \mathbf{P}[t;\gamma_2] \qquad \forall t > \max(\tau(\gamma_1), \tau(\gamma_2)) \tag{D.20}$$

Next we consider $1 < t \leq \max(\tau(\gamma_1), \tau(\gamma_2))$. We argue by induction that

$$\mathbf{F}[t;\gamma_1] \geq \mathbf{F}[t;\gamma_2] \tag{D.21}$$

$$\mathbf{S}[t;\gamma_1] \geq \mathbf{S}[t;\gamma_2], \tag{D.22}$$

where $\mathbf{F}[t;\gamma]$ and $\mathbf{S}[t;\gamma]$ are defined in (3)-(4). The remainder of the proof follows the same logic as in in Appendix D.1.

D.5 Risk-free hit probability p_0

The proof is nearly identical to the corresponding proof for δ in Appendix D.3, so we just sketch the approach here. We introduce the notation $\mathbf{P}[t; p_0]$, $\tau(p_0)$, and $\hat{\mathbf{S}}_{\mathbf{F}}[t; p_0]$ to explicitly connect these quantities to p_0. Note that $\hat{\mathbf{F}}_{\mathbf{all}}[t]$ does not depend upon p_0. We use the following Lemma in our proof

Lemma 9.

$$\frac{\partial}{\partial p_0}\hat{\mathbf{S}}_{\mathbf{F}}[t; p_0] \geq 0, \qquad \forall p_0 \in [0, 1), 1 \leq t \leq N$$

The proof for Lemma 9 follows the same logic as the analogous proof for Lemma 8 in Appendix D.3.1. Because $\hat{\mathbf{F}}_{\mathbf{all}}[t]$ does not depend upon p_0, Lemma 9 implies

$$\begin{aligned} &\text{if } \hat{\mathbf{S}}_{\mathbf{F}}[t; p_0^{(1)}] > \hat{\mathbf{F}}_{\mathbf{all}}[t] \\ &\text{then } \hat{\mathbf{S}}_{\mathbf{F}}[t; p_0^{(2)}] > \hat{\mathbf{F}}_{\mathbf{all}}[t] \quad \text{for any } p_0^{(2)} > p_0^{(1)} \end{aligned} \tag{D.23}$$

Condition (D.23) implies that $\tau(p_0)$ is non-decreasing in p_0 by the definition of τ in equation (7).

To show that $\mathbf{P}[t; p_0]$ is non-decreasing in p_0, we consider $p_0^{(1)} < p_0^{(2)}$. We then proceed using the same logic as in the proof for δ in Appendix D.3.

E Proof of Corollary 1: Cost if Blue Destroyed

Adding the cost C requires replacing α with $\alpha - \beta C$ in the numerator of several of the terms in equations (2)–(6). Equation (2) becomes

$$\mathbf{V}[t] = \frac{(\delta+\gamma)\left(\alpha - \beta C + \gamma \mathbf{V}[t+1]\right)}{(\alpha+\beta+\gamma)(\delta+\gamma) + \lambda_t(\gamma + \delta p_0)} + \lambda_t \frac{\delta p_0 + \gamma \frac{\delta}{\delta+\gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t-1} (p_0 + (1-p_0)\mathbf{V}[s])}{(\alpha+\beta+\gamma)(\delta+\gamma) + \lambda_t(\gamma + \delta p_0)}.$$

Equation (3) becomes

$$\mathbf{V}[t] = \mathbf{F}[t] \equiv \frac{\alpha - \beta C}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\mathbf{V}[t+1]$$

and (4) remains unchanged except for replacing $\mathbf{P}[s]$ with $\mathbf{V}[s]$. Theorems 1 – 3 and the proofs and results in Appendices A – C all still hold after making these $\alpha \to \alpha - \beta C$ substitutions.

F Proof of Proposition 5 : $\lambda^*_{t,r} \in \{0, \infty\}$ for Full-Information Risk Model

We provide the proof in Appendix F.1 and present an algorithm to compute $\lambda^*_{t,r}$ in Appendix F.2.

F.1 Proof

We first introduce notation in Appendix F.1.1. They key to the proof is optimizing the first risk level, $\mathbf{P}[t, 1]$, which we focus on in Appendix F.1.2. We tie up the loose ends for higher risk levels in Appendix F.1.3

F.1.1 Notation

We first define several intermediate values to simplify notation. We separate out the $s = t$ term from the summation in (9) and define the remaining summation starting at $t + 1$ as $G[t + 1]$

$$G[t+1] = \frac{\delta}{\delta+\gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} (p_0 + (1-p_0)\mathbf{P}[s,1]) \tag{F.1}$$

Note that $G[t + 1]$ does not depend upon the current risk level r. Using $G[t + 1]$, we rewrite $\mathbf{P}[t, r]$ from (9) by separating out the $\mathbf{P}[t, 1]$ term:

$$\begin{aligned}\mathbf{P}[t,r] =& \frac{\alpha_r}{D(t,r)} + \frac{\gamma}{D(t,r)}\mathbf{P}[t+1,r] + \frac{\mu_r}{D(t,r)}\mathbf{P}[t,r+1] + \frac{\lambda_{t,r}}{D(t,r)}G[t+1] \\ &+ \frac{\lambda_{t,r}}{D(t,r)}\frac{\delta}{\delta+\gamma}p_0 + \frac{\lambda_{t,r}}{D(t,r)}\frac{\delta}{\delta+\gamma}(1-p_0)\mathbf{P}[t,1]\end{aligned} \tag{F.2}$$

We next define $\mathbf{Q}[t, r]$ as all the terms of $\mathbf{P}[t, r]$ in (F.2) except the μ_r and $\mathbf{P}[t, 1]$ terms

$$\mathbf{Q}[t,r] = \frac{\alpha_r}{D(t,r)} + \frac{\gamma}{D(t,r)}\mathbf{P}[t+1,r] + \frac{\lambda_{t,r}}{D(t,r)}G[t+1] + \frac{\lambda_{t,r}}{D(t,r)}\frac{\delta}{\delta+\gamma}p_0 \tag{F.3}$$

Combining (F.2) with (F.3) results in

$$\mathbf{P}[t,r] = \left(\mathbf{Q}[t,r] + \frac{\lambda_{t,r}}{D(t,r)}\frac{\delta}{\delta+\gamma}(1-p_0)\mathbf{P}[t,1]\right) + \frac{\mu_r}{D(t,r)}\mathbf{P}[t,r+1] \tag{F.4}$$

Examining (F.2)–(F.4), we see that $\mathbf{P}[t, r]$ depends upon $\mathbf{P}[i, j]$ for $i \geq t$ and $j \geq r$, *except* for the $\mathbf{P}[t, 1]$ term in (F.4). Consequently if we knew $\mathbf{P}[t, 1]$ for all $1 \leq t \leq N$, then we could solve for $\mathbf{P}[t, r]$ and the corresponding optimal policy $\lambda^*_{t,r}$ via backward induction in both t and r. For a fixed t, we would solve for $\mathbf{P}[t, r]$ working backward from $r = M$ to $r = 2$, and then we would proceed to time index $t - 1$ and repeat. Therefore, the key is determining $\mathbf{P}[t, 1]$.

F.1.2 Optimizing $\mathbf{P}[t,1]$

If we iterate out F.4, we can rewrite $\mathbf{P}[t,r]$ only in terms of $\mathbf{Q}[t,\cdot]$ and $\mathbf{P}[t,1]$

$$\mathbf{P}[t,r] = \sum_{k=r}^{M} \left(\prod_{i=r}^{k-1} \frac{\mu_i}{D(t,i)} \right) \left(\mathbf{Q}[t,k] + \frac{\lambda_{t,k}}{D(t,k)} \frac{\delta}{\delta+\gamma} (1-p_0) \mathbf{P}[t,1] \right) \tag{F.5}$$

Substituting in $r = 1$ in (F.5) allows us to compute $\mathbf{P}[t,1]$ as a function of the move-rates $\lambda_{t,r}$:

$$\mathbf{P}[t,1] = \frac{\sum_{k=1}^{M} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \mathbf{Q}[t,k]}{1 - \frac{\delta}{\delta+\gamma}(1-p_0) \sum_{k=1}^{M} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \frac{\lambda_{t,k}}{D(t,k)}} \tag{F.6}$$

We can now choose the optimal move-rates $\lambda^*_{t,r}$ to maximize $\mathbf{P}[t,1]$ in (F.6). $\mathbf{P}[t,1]$ depends upon $\mathbf{Q}[t,k]$, which is a function of the current time period t only through the decision variables $\lambda_{t,r}$. Because we compute $\mathbf{P}[t,\cdot]$ and $\lambda^*_{t,r}$ in reverse chronological order, we can view $\mathbf{P}[t+1,r]$ and $G[t+1]$ in as (F.3) as known constants in our optimization of $\mathbf{P}[t,1]$.

To show that $\lambda^*_{t,r} \in \{0, \infty\}$, we pick an arbitrary risk index r and multiply the numerator and denominator of $\mathbf{P}[t,1]$ in (F.6) by $D(t,r)$. We then break up the summations in (F.6) into three categories: indices less than r, indices greater than r, and the index exactly equal to r:

$$\mathbf{P}[t,1] = \frac{AD(t,r) + C + B\left(\alpha_r + \gamma \mathbf{P}[t+1,r] + \lambda_{t,r}\left(G[t+1] + \frac{\delta}{\delta+\gamma} p_0\right)\right)}{\tilde{A}D(t,r) - \tilde{C} - \tilde{B}\lambda_{t,r}} \tag{F.7}$$

where

$$
\begin{aligned}
A =& \sum_{k=1}^{r-1} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \mathbf{Q}[t,k] \\
B =& \left(\prod_{i=1}^{r-1} \frac{\mu_i}{D(t,i)} \right) \\
C =& \mu_r \left(\prod_{i=1}^{r-1} \frac{\mu_i}{D(t,i)} \right) \sum_{k=r+1}^{M} \left(\prod_{i=r+1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \mathbf{Q}[t,k] \\
\tilde{A} =& 1 - \frac{\delta}{\delta+\gamma}(1-p_0) \sum_{k=1}^{r-1} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \frac{\lambda_{t,k}}{D(t,k)} \\
\tilde{B} =& \frac{\delta}{\delta+\gamma}(1-p_0) \left(\prod_{i=1}^{r-1} \frac{\mu_i}{D(t,i)} \right) \\
\tilde{C} =& \frac{\delta}{\delta+\gamma}(1-p_0)\mu_r \left(\prod_{i=1}^{r-1} \frac{\mu_i}{D(t,i)} \right) \sum_{k=r+1}^{M} \left(\prod_{i=r+1}^{k-1} \frac{\mu_i}{D(t,i)} \right) \frac{\lambda_{t,k}}{D(t,k)}
\end{aligned}
$$

Note that the constants $A, B, C, \tilde{A}, \tilde{B}, \tilde{C}$ do not depend upon $\lambda_{t,r}$. Of course they depend upon $\lambda_{t,k}$ for risk level $k \neq r$. Because $D(t,r)$ is a linear function of $\lambda_{t,r}$ (see (10)), $\mathbf{P}[t,1]$ in (F.7) is just the ratio of two linear functions of $\lambda_{t,r}$. As discussed in the proof for Proposition 1 in Appendix A, a ratio of linear functions leads to the desired $\lambda^*_{t,r} \in \{0, \infty\}$.

When computing the optimal solution $\lambda^*_{t,r}$ to maximize $\mathbf{P}[t,1]$ in (F.6), the optimal solution takes the form of $\lambda^*_{t,r} = 0$ for all $1 \leq r \leq m^* - 1$ and $\lambda^*_{t,m^*} = \infty$ for some $1 \leq m^* \leq M+1$. That is Blue fights for all risk levels $1 \leq r < m^*$ and scoots at risk level m^*. If $m^* = M+1$, Blue fights for all risk levels at time t. The values of $\lambda_{t,r}$ for $r > m^*$ do not impact $\mathbf{P}[t,1]$ as the summation terms are 0 in (F.6) for $k > m^*$. If Blue uses this risk threshold policy with risk threshold m, then (F.6) can be written as:

$$
\mathbf{P}[t,1;m] = \frac{\left(\prod_{i=1}^{m-1} \frac{\mu_i}{\alpha_i+\beta_i+\gamma+\mu_i} \right) \left(G[t+1] + \frac{\delta}{\delta+\gamma} p_0 \right) + \sum_{k=1}^{m-1} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{\alpha_i+\beta_i+\gamma+\mu_i} \right) \frac{\alpha_k + \gamma \mathbf{P}[t+1,k]}{\alpha_k+\beta_k+\gamma+\mu_k}}{1 - \frac{\delta}{\delta+\gamma}(1-p_0) \left(\prod_{i=1}^{m-1} \frac{\mu_i}{\alpha_i+\beta_i+\gamma+\mu_i} \right)} \tag{F.8}
$$

We can determine m^* by enumeration: compute $\mathbf{P}[t, 1; m]$ for all $1 \leq m \leq M$ using (F.8). As in Section 4.1, we assume that if Blue is indifferent between fighting and scooting, Blue fights. Therefore, if there are multiple thresholds that produce the maximum $\mathbf{P}[t, 1; m]$, we choose the largest threshold as m^*.

F.1.3 Optimizing $\mathbf{P}[t, 1]$ for $r > 1$

We still have two outstanding issues. One is that when determining $\lambda^*_{t,r}$ by maximizing $\mathbf{P}[t, 1]$ as described in Appendix F.1.2, we only determine $\lambda^*_{t,r}$ for $1 \leq r \leq m^*$. To determine the $\lambda^*_{t,r}$ for $r > m^*$, we perform backward induction on (F.2) for both t and r. Examining (F.2), $\mathbf{P}[t, 1]$ is now known and both $\mathbf{P}[t + 1, r]$ and $G[t + 1]$ have been computed at previous time iterations. Finally, we compute $\mathbf{P}[t, r + 1]$ on the previous step as we work backward in risk. The computation for $\mathbf{P}[t, r]$ in (F.2) is therefore just a weighted sum of constants, where $\lambda_{t,r}$ controls the weights. To maximize a weighted sum, we put all the weight on the largest value; we achieve this by setting $\lambda^*_{t,r} \in \{0, \infty\}$. If Blue is indifferent, we set $\lambda^*_{t,r} = 0$.

The above describes two methods to compute $\lambda^*_{t,r}$. The first method maximizes $\mathbf{P}[t, 1]$ by solving for m^* using (F.8) and setting $\lambda^*_{t,r} = 0$ for all $1 \leq r \leq m^* - 1$ and $\lambda^*_{t,m^*} = \infty$. We can also compute $\lambda^*_{t,r}$ for $2 \leq r \leq M$ by maximizing $\mathbf{P}[t, r]$ via backward induction on (F.2) as discussed in the previous paragraph. The last issue we need to resolve is to ensure these two methods for computing $\lambda^*_{t,r}$ produce the same results for $1 \leq r \leq m^*$. This consistency follows from a contradiction argument. We denote $\lambda^*_{t,r}$ as the move-rates computed to maximize $\mathbf{P}[t, 1]$ using risk threshold m^*. We denote $\tilde{\lambda}_{t,r}$ as the move-rates to optimize each $\mathbf{P}[t, r]$ separately via backward induction on (F.2). Finally we denote $\mathbf{P}[t, r]$ as the probabilities associated with $\lambda^*_{t,r}$ and $\tilde{\mathbf{P}}[t, r]$ as the probabilities associated with $\tilde{\lambda}_{t,r}$. By construction $\mathbf{P}[t, 1] \geq \tilde{\mathbf{P}}[t, 1]$ and $\mathbf{P}[t, r] \leq \tilde{\mathbf{P}}[t, r]$ for $r > 1$. Assume there exists some $1 \leq j \leq m^*$ such that $\lambda^*_{t,j} \neq \tilde{\lambda}_{t,j}$, which implies $\mathbf{P}[t, j] < \tilde{\mathbf{P}}[t, j]$. By (F.2)–(F.5), we have

$$\mathbf{P}[t,1] = \left(\prod_{i=1}^{j-1} \frac{\mu_i}{D(t,i)}\right) \mathbf{P}[t,j] + \sum_{k=1}^{j-1} \left(\prod_{i=1}^{k-1} \frac{\mu_i}{D(t,i)}\right) \mathbf{Q}[t,k] \tag{F.9}$$

To move from (F.4) to (F.9) we use $\lambda^*_{t,k} = 0$ for $1 \leq k \leq j-1 < m^*$. $\mathbf{P}[t,1]$ is a direct function of $\mathbf{P}[t,j]$ in (F.9). Consequently, if we use move-rate $\tilde{\lambda}_{t,i}$ instead of $\lambda^*_{t,i}$ for $i \geq j$, we could replace $\mathbf{P}[t,j]$ in the right-hand side of (F.9) with $\tilde{\mathbf{P}}[t,j] > \mathbf{P}[t,j]$. However, this substitution would imply that using $\tilde{\lambda}_{t,i}$ would produce a strictly greater value for $\mathbf{P}[t,1]$ than using $\lambda^*_{t,i}$, which yields the desired contradiction. Therefore, the two methods for computing $\lambda^*_{t,r}$ are consistent and we have addressed the remaining outstanding issue. The proof is complete: $\lambda^*_{t,r} \in \{0, \infty\}$

F.2 Algorithm to Compute $\lambda^*_{t,r}$

We conclude with a description for computing the optimal $\lambda^*_{t,r}$

1. Start at time period $t = N$

    2. Compute $\mathbf{P}[t,1;m]$ for each $1 \leq m \leq M$ using (F.8) and determine m^* such that $\mathbf{P}[t,1;m^*] \geq \mathbf{P}[t,1;m]$ for each $1 \leq m \leq m^*$ and $\mathbf{P}[t,1;m^*] > \mathbf{P}[t,1;m]$ for each $m^* < m \leq M$

    3. Set $\lambda^*_{t,r} = 0$ for all $1 \leq r \leq m^* - 1$ and $\lambda^*_{t,m^*} = \infty$

    4. Define $\mathbf{P}[t,1] = \mathbf{P}[t,1;m^*]$

    5. Start at risk level $r = M$

        6. Compute the right-hand side of (F.2) for $\lambda_{t,r} \in \{0, \infty\}$ and denote $\lambda^*_{t,r}$ as the maximizer and $\mathbf{P}[t,r]$ as the maximum

        7. Decrement $r \leftarrow r - 1$

        8. If $r > m^*$ return to step 6, otherwise proceed to step 9

9. Decrement $t \leftarrow t-1$

10. If $t > 1$ return to step 2, otherwise algorithm complete

To compute m^* in Step 2 is $O(M)$ and we have to work backward in time and risk to compute the remaining values of $\lambda^*_{t,r}$ and $\mathbf{P}[t,r]$. Therefore the complexity of the entire algorithm is $O(NM^2)$.

G Optimal Solution for Partial-Information Setting

In Appendix G.1 we define the optimization problem to compute the optimal move-rates. We provide numerical examples in Appendix G.2 to illustrate the properties of the optimization problem. Appendix G.3 presents an algorithm to compute $\lambda_t^{(j)*}$

G.1 Optimization Framework

First we define the analogous notation as in Appendix F.

$$G[t+1] = \frac{\delta}{\delta+\gamma} \sum_{s=t+1}^{N} \left(\frac{\gamma}{\delta+\gamma}\right)^{s-t} \left(p_0 + (1-p_0)\mathbf{P}^{(s)}[s,1]\right) \tag{G.1}$$

As in Appendix F, $G[t+1]$ does not depend upon the risk level. $G[t+1]$ also only depends upon the win-probabilities $\mathbf{P}^{(j)}[j,1]$ for $j > t$. Therefore we can view $G[t+1]$ as a constant from an optimization point of view as it can be calculated with terms derived from previous iterations of our backward induction in j. Next rewrite $\mathbf{P}^{(j)}[t,r]$ in terms of $G[t+1]$

$$\begin{aligned}\mathbf{P}^{(j)}[t,r] =& \frac{\alpha_r}{D^{(j)}(t,r)} + \frac{\gamma}{D^{(j)}(t,r)}\mathbf{P}^{(j)}[t+1,r] + \frac{\mu_r}{D^{(j)}(t,r)}\mathbf{P}^{(j)}[t,r+1] + \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)}G[t+1] \\ &+ \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)}\frac{\delta}{\delta+\gamma}p_0 + \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)}\frac{\delta}{\delta+\gamma}(1-p_0)\mathbf{P}^{(t)}[t,1]\end{aligned} \tag{G.2}$$

Finally we define $\mathbf{Q}^{(j)}[t,r]$.

$$\mathbf{Q}^{(j)}[t,r] = \frac{\alpha_r}{D^{(j)}(t,r)} + \frac{\gamma}{D^{(j)}(t,r)}\mathbf{P}^{(j)}[t+1,r] + \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)}G[t+1] + \frac{\lambda_t^{(j)}}{D^{(j)}(t,r)}\frac{\delta}{\delta+\gamma}p_0 \quad \text{(G.3)}$$

In the full-information setting in Appendix F we could assume $\mathbf{P}[t+1,r]$ in (F.2) was known as it was computed on a previous iteration of backward induction in t and had an associated $\lambda^*_{t+1,r}$. We cannot treat $\mathbf{P}^{(j)}[t+1,r]$ as a known constant in (G.2)–(G.3) because when the system transitions to state $(t+1,r)$, Blue knows the time t but not risk r. The move-rate Blue uses in state $(t+1,r)$ is $\lambda_{t+1}^{(j)}$, which is chosen at time j, not time $t+1$ like in the full-information setting. This coupling implies we cannot perform backward induction in t for a fixed j.

Performing the same steps as in Appendix F yields.

$$\mathbf{P}^{(j)}[t,r] = \sum_{k=r}^{M}\left(\prod_{i=r}^{k-1}\frac{\mu_i}{D^{(j)}(t,i)}\right)\left(\mathbf{Q}^{(j)}[t,k] + \frac{\lambda_t^{(j)}}{D^{(j)}(t,k)}\frac{\delta}{\delta+\gamma}(1-p_0)\mathbf{P}^{(t)}[t,1]\right) \quad \text{(G.4)}$$

In the full-information setting for any state (t,r), we could choose the optimal moves rates to maximize $\mathbf{P}[t,r]$. In the partial-information setting, Blue can only choose the move-rates after a risk-free shot because those are the only times Blue has full-information. Thus Blue chooses $\lambda_t^{(j)*}$ for $j \le t \le N$ to maximize $\mathbf{P}^{(j)}[j,1]$. As in Appendix F we can solve for $\mathbf{P}^{(j)}[j,1]$ from (G.4)

$$\mathbf{P}^{(j)}[j,1] = \frac{\sum_{k=1}^{M}\left(\prod_{i=1}^{k-1}\frac{\mu_i}{D^{(j)}(j,i)}\right)\mathbf{Q}^{(j)}[j,k]}{1-\frac{\delta}{\delta+\gamma}(1-p_0)\sum_{k=1}^{M}\left(\prod_{i=1}^{k-1}\frac{\mu_i}{D^{(j)}(j,i)}\right)\frac{\lambda_j^{(j)}}{D^{(j)}(j,k)}} \quad \text{(G.5)}$$

Blue solves $\max_{\lambda_t^{(j)} \text{ for } j\le t\le N}\mathbf{P}^{(j)}[j,1]$. In the corresponding expression in (F.6) in Appendix F there was a different $\lambda_{t,i}$ for each term in the product expressions. However in (G.5) only one move-rate, $\lambda_j^{(j)}$, appears in the product terms. Therefore if we multiply the numerator and denominator of (G.5) by $\prod_{i=1}^{M}D^{(j)}(j,i)$, then $\mathbf{P}^{(j)}[j,1]$ transforms into the ratio of two

polynomials of order M in $\lambda_j^{(j)}$. Recall in Appendix F that $\mathbf{P}[t,1]$ could be expressed as the ratio of linear expressions for any particular $\lambda_{t,r}$, which led to the binary nature of the optimal solution: $\lambda_{t,r}^* \in \{0,\infty\}$. Unfortunately, once our expression is a ratio of polynomials, we no longer have the optimal solution necessarily being an extreme. The $\mathbf{Q}^{(j)}[j,k]$ in the numerator (G.5) contains $G[j+1]$ (see (G.3)) , which has been computed at a previous iteration of the backward induction on j and we can view as a constant for our optimization purposes. However, $\mathbf{Q}^{(j)}[j,k]$ also contains $\mathbf{P}^{(j)}[j+1,r]$, which as discussed earlier is not known unlike in the full-information setting. The $\mathbf{P}^{(j)}[j+1,r]$ term in $\mathbf{Q}^{(j)}[j,k]$ depends upon $\lambda_t^{(j)}$ for $t > j$ and causes the problem to become a multivariate optimization problem.

G.2 Numerical Illustrations

To illustrate the optimization more concretely we examine the results in the last period ($j = N$) in Appendix G.2.1 and the penultimate period in in Appendix G.2.2.

G.2.1 One Period Optimization

We first consider the simplest setting where $j = N$ and we must determine $\lambda_N^{(N)*}$ to maximize $\mathbf{P}^{(N)}[N,1]$. Since we only have one decision variable, we denote it below as merely λ for notational convenience. In this case, substituting $D^{(N)}(N,i)$ from (14) and $\mathbf{Q}^{(N)}[N,k]$ from (G.3) into (G.5) yields

$$\mathbf{P}^{(N)}[N,1] = \frac{\sum_{k=1}^{M}\left(\prod_{i=1}^{k-1}\frac{\mu_i}{\alpha_i+\beta_i+\gamma+\mu_i+\lambda}\right)\frac{\alpha_k+\lambda\frac{\delta}{\delta+\gamma}p_0}{\alpha_k+\beta_k+\gamma+\mu_k+\lambda}}{1-\frac{\delta}{\delta+\gamma}(1-p_0)\sum_{k=1}^{M}\left(\prod_{i=1}^{k-1}\frac{\mu_i}{\alpha_i+\beta_i+\gamma+\mu_i+\lambda}\right)\frac{\lambda}{\alpha_k+\beta_k+\gamma+\mu_k+\lambda}} \tag{G.6}$$

Unfortunately the win-probability in (G.6) is not a concave function so we cannot derive nice properties about the optimal λ^*. As the example in Table 9 and Figure 5 illustrates, the win-probability in (G.6) may have a local maximum that is not a global maximum. The optimal policy is for Blue to scoot $\lambda = \infty$, which produces a win-probability of $\frac{\delta p_0}{\delta p_0+\gamma} = 0.6$. Risk levels $r = 1$ and $r = 4$ provide a significant firing advantage to Red and risk level $r = 3$

slightly favors Red. Blue has a significant advantage only in risk level $r = 2$. The worst choice for Blue is $\lambda \approx 40$ because while Blue only fires for a short time before moving, Blue is likely to spend most of this time firing in risk level $r = 1$ which favors Red. Fighting ($\lambda = 0$) is not a good option for Blue either as there is a nontrivial probability that the exchange will reach risk level $r = 3$ or $r = 4$, which also favor Red. Blue has a sweet spot of $\lambda \approx 6$, which gives Blue a reasonable chance of reaching risk level $r = 2$, where Blue has the advantage, but it is much less likely to reach the higher risk levels where Blue is at a disadvantage. All of these options are worse, however, than Blue scooting.

$R = 4$	$\alpha_r = (1, 10, 2, 1)$	$\beta_r = (5, 1, 3, 10)$	$\mu_r = (20, 5, 15, 0)$	$\gamma = 1$	$\delta = 10$	$p_0 = 0.15$

Table 9: Parameter values associated with Figure 5

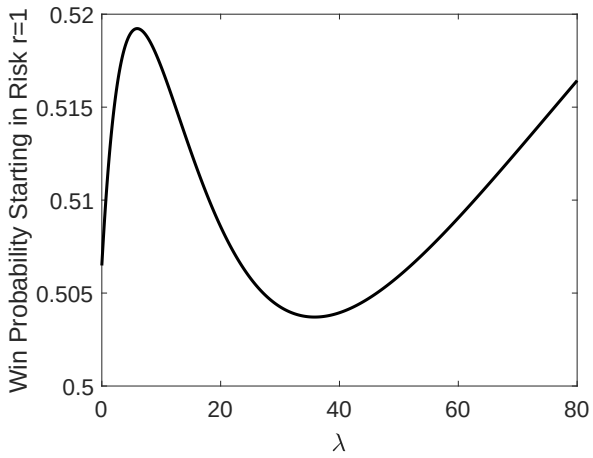


Figure 5: Optimal win-probability $\mathbf{P}^{(N)}[N, 1]$ for parameters in Table 10

To write out $\mathbf{P}^{(j)}[j, 1]$ in (G.5) explicitly for $j < N$ is tedious. However, for fixed $\lambda_t^{(j)}$ it is straightforward to compute $\mathbf{P}^{(j)}[j, 1]$ numerically by iterating over $\mathbf{P}^{(j)}[t, r]$ in (G.2) for $t > j$ to calculate the $\mathbf{Q}^{(j)}[j, k]$ terms in the numerator of $\mathbf{P}^{(j)}[j, 1]$. Therefore we can apply standard numerical nonlinear optimization techniques to compute $\lambda_t^{(j)*}$. We use the built-in optimization functionality in Matlab (The MathWorks, 2016) to compute $\lambda_t^{(j)*}$. As Figure 5 illustrates the problem is non-concave and may have local maximizers that are not the global maximum. Therefore we run the optimization with 20 different initial values, but it is possible

our final solution will not be the global optimum. Based on our experimentation it is rare to have situations with local maxima differing from the global, and often in those situations suboptimal solutions are close the global maximum. Our purpose with this analysis is not to produce a guaranteed optimal solution. We want to generate insight into the nature of the move policy by examining near optimal policies and comparing the results with simple heuristics.

G.2.2 Two Period Optimization

We next consider the decision in the penultimate period ($j = N-1$). For notational simplicity we assume there are only two time periods ($N = 2$). Figure 6 illustrates a scenario with two risk levels ($M = 2$); the other baseline parameters appear in Table 10. We first compute the optimal move-rate when the risk-free shot occurs in the last period $j = 2$: $\lambda_2^{(2)*} = 12.7$. The logic to optimally compute $\lambda_2^{(2)*}$ follows from the discussion in Appendix G.2.1. Given this value of $\lambda_2^{(2)*}$, we plot the Blue win-probability starting in $t = 1$ and $r = 1$ as a function of $\lambda_1^{(1)}$ for three different values of $\lambda_2^{(1)}$. The move-rate at time period $t = 1$, $\lambda_1^{(1)}$, has a much greater impact on the win-probability than the move-rate at time period $t = 2$, $\lambda_2^{(1)}$, because the system may not reach time period 2 before one side is destroyed or Blue moves. The optimal solution, $\lambda_1^{(1)*} = 32.1$ and $\lambda_2^{(1)*} = 21.7$, is an intermediate solution because Blue has the advantage in risk level $r = 1$ and Red has the advantage in risk level $r = 2$.

$\alpha_r = (4, 1)$	$\beta_r = (1, 5)$	$\mu_r = (10, 0)$	$\gamma = 1$	$\delta = 8$	$p_0 = 0.01$

Table 10: Parameter values associated with Figure 6

G.3 Algorithm to Compute $\lambda_t^{(j)*}$

We conclude with a description for computing the optimal $\lambda_t^{(j)*}$

1. Start at time period $j = N$

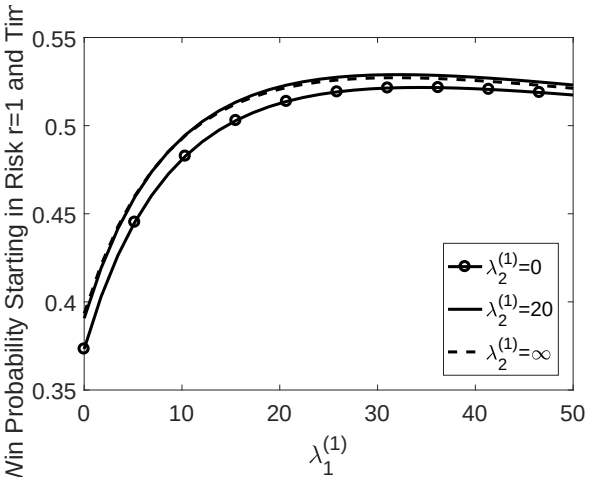


Figure 6: $N = 1$, $R = 2$, for three values of $\lambda_2^{(1)}$. Other parameters appear in Table 10

2. Use numerical optimization techniques to determine $\lambda_t^{(j)*}$ to maximize $\mathbf{P}^{(j)}[j, 1]$ in (G.5) for $j \leq t \leq N$.
3. Given $\lambda_t^{(s)*}$ and $\mathbf{P}^{(s)}[s, 1]$ for all $j \leq s \leq N$, compute $\mathbf{P}^{(j)}[t, r]$ for all $j \leq t \leq N$ and $1 \leq r \leq M$ using (G.2)
4. Decrement $j \leftarrow j - 1$
5. If $j > 1$ return to step 2, otherwise algorithm complete

H Description of Simulation Model

We provide a detailed description of the simulation in Section H.1. Section H.2 contains a list of all the parameters of the simulation model and their corresponding distributions.

H.1 Overview

At a high level, the simulation model is very similar to the risk models in Sections 5–6. One key addition is the inclusion of both Blue and Red health, which corresponds to the number of hits until either side is destroyed. For concreteness we assume one level of Blue health corresponds to one Blue cannon and one unit of Red health corresponds to one Red target

(e.g., refinery, bridge, logistics center). Blue must destroy Red before the time-window closes. In the simulation model the length of the time-window is deterministic. After Blue moves to a new location, there is a risk-free period where Blue prepares for its risk-free shot. After the risk-free shot Red and Blue engage in risky-fire where both sides simultaneously shoot at each. Risk increases as Blue stays in the same location, which corresponds to increased firing accuracy for both Blue and Red. In the simulation model, we now connect risk more directly to the shots fired by each side. After a Blue shot, Red increases its firing accuracy because it better knows Blue's location (e.g., because of counter-battery radar). However, Blue's accuracy also increases after it observes the results of its own shot and recalibrates the cannon based on potential drift errors. Blue's accuracy does not improve after a Red shot because Red artillery fires from a different location than the Red target The time to incorporate the intelligence associated with each shot, which increases firing accuracy, is a uniform random variable. Eventually Blue moves to a new location and the process repeats.

The main additions the simulation model has over the risk models in Sections 5–6 are:

1. Requiring multiple hits to destroy either Red or Blue by tracking Blue and Red health
2. Non-exponential distributions
3. More explicitly connecting risk to improved accuracy from incorporating intelligence about shots fired

The following subsections provide additional detail.

H.1.1 Risk-free Firing

Immediately after moving to a new location, Blue sets up to fire its risk-free shot. The time to fire the risk-free shot is a uniform random variable. The risk-free shot can potentially hit multiple Red targets, which reduces Red's health by more than one level, unlike in the Sections 5–6. This is reasonable if the Blue health corresponds to the number of Blue cannons. After a move, Blue will fire all of its available cannons simultaneously during the risk-free

shot, and thus it is possible Blue will hit multiple Red targets with this risk-free volley. We model the number of Red targets hit with the risk-free shot as a discrete random variable that depends upon the current health of Red and Blue and the firing accuracy of Blue on its risk-free shot. See Section H.2.1 for more details on the distribution for number of Red targets hit on the Blue risk-free shot.

Immediately after the risk-free shot, the battle enters the risky stage (see Section H.1.2), where both Red and Blue fire at each other. Red and Blue now track the shots fired by each side, because that improves future firing accuracy. We count the risk-free shot as just one Blue shot even though multiple cannons might shoot at the same time. We generate a uniform random variable for Blue (Red) to determine when Blue (Red) incorporates the intelligence from the risk-free shot into future shots.

We do not allow Red to detect Blue while Blue is setting up for a risk-free shot immediately after Blue arrives to a new location, although that would be straightforward to incorporate.

H.1.2 Risky Firing

In this stage Blue and Red repeatedly fire at each other. Unlike for the risk-free shot in Section H.1.1, during risky-fire each Blue (Red) shot can only decrease Red's (Blue's) health by one level. After each Blue (Red) shot, we generate a uniform random variable for the time until the next Blue (Red) shot. Red's accuracy improves as it observes and processes shots from either side. For example adjusting aim based on the results of a previous Red shot or better pinpointing Blue's location after processing a radar signal from a Blue shot. Blue's accuracy only improves as it processes its own shots because Red artillery is located at a different location than the Red targets. After each shot we generate a uniform random variable for each side that corresponds to the time when that side incorporates the information from the shot into their firing calculus. The accuracy of Blue (Red) depends upon how many shots by each side Blue (Red) has processed. For simplicity we assume that Blue (Red) incorporates at most 3 shots from each side into its fire system. Even though both sides know how

many shots have been fired, Blue does not know how many shots Red has processed. This corresponds to the partial risk setting in Section 6 where Blue does not know the exact risk level. Blue and Red continue firing and incorporating the corresponding information into their situational awareness picture until either (1) one side is destroyed, (2) the time-window closes, or (3) Blue moves to a new location.

During this stage, Blue is constantly evaluating when it should move. This depends not just on the time, but also the situation on the battlefield: Blue's health and the number of shots fired by both Blue and Red. We provide more details on the move logic in Section H.1.3. When Blue decides to move, we generate a uniform random variable for the travel time and transition into the risk-free period (see Section H.1.1). The risk and situational awareness reset to the lowest levels after a move.

H.1.3 Move Decision

Blue accounts for five factors when determining when it should move: Blue health, Blue shots fired, Red shots fired, current time, time spent in risky-fire at current location. We discretize the current time and risky time into discrete bins; t^{bin} and t^{bin}_{risky} correspond to discrete levels for those factors. This allows us to only define a finite number of combinations for the move decision. We define the threshold $\tau(h_B, s_B, s_R, t^{bin}, t^{bin}_{risky})$ that specifies when Blue should move. The variables h_B, s_B, s_R correspond to Blue health, number of Blue shots fired, and number of Red shots fired, respectively. If the current time $t < \tau(h_B, s_B, s_R, t^{bin}, t^{bin}_{risky})$ then Blue should remain and fire, otherwise Blue should move. Each time one of those five factors change during the risky-fire, Blue checks the updated τ to determine whether Blue should move. Immediately after firing the risk free shot, Blue checks the threshold $\tau(h_B, 1, 0, t^{bin}, 1)$, which only varies with Blue health h_B and current time level t^{bin}.

For a known threshold function $\tau(h_B, s_B, s_R, t^{bin}, t^{bin}_{risky})$, we run the simulation and compute Blue's win-probability. To optimize $\tau(h_B, s_B, s_R, t^{bin}, t^{bin}_{risky})$, we use a genetic algorithm approach. Due to the randomness of the simulation output and the large number of τ com-

binations, the genetic algorithm is slow and provides no guarantees on the optimality gap of the solution generated.

H.2 Simulation Parameters

We now define all the parameters and their corresponding probability distributions where appropriate. We generate 750 scenarios and for each scenario we compute the win-probability by simulating the battle 10000 times. We optimize the move decision and also compute the win-probability against several heuristic. For each scenario there are many parameters that are randomly generated before running the scenario, and then those parameters are fixed for the duration of the scenario. These include the initial health of Blue and Red and the probability Blue or Red hits the other with a risky shot. We describe these parameters in Section H.2.1. Within a particular scenario, we randomly generate several values many times during the duration of the battle. These include the time for Blue to move locations, and the time until the next Red shot during risky-fire. We define these random variables in Section H.2.2

H.2.1 Scenario Input Parameters

Table 11 lists all the parameters. Most of the distributions for the paramters are uniforms, which we denote $U[a, b]$. We describe a few parameters in more detail below the table.

Description	Distribution
Max number of Blue shots used to improve firing accuracy	$U[1,3]$, integer
Max number of Red shots used to improve firing accuracy	$U[1,3]$, integer
Initial Blue health	$U[4,6]$, integer
Initial Red health	$U[4,6]$, integer
Time window	$U[30,240]$ minutes
$Trav_L$: lower bound on Blue travel	$U[2,12]$ minutes
$Trav_U$: upper bound on Blue travel	$Trav_L + U[0,20]$ minutes
$RiskFreeBlue_L$: lower bound on time until Blue fires risk-free shot after arriving to new location	$U[1,6]$ minutes
$RiskFreeBlue_U$: upper bound on time until Blue fires risk-free shot after arriving to new location	$RiskFreeBlue_L + U[0,5]$ min
$RiskyBlue_L$: lower bound on Blue interfire time during risky-fire	$U[0.2,3.2]$ minutes
$RiskyBlue_U$: upper bound on Blue interfire time during risky-fire	$RiskyBlue_L + U[0,3]$ minutes
$RiskyRed_L$: lower bound on Red interfire time during risky-fire	$U[0.1,2]$ minutes
$RiskyRed_U$: upper bound on Red interfire time during risky-fire	$RiskyRed_L + U[0,2]$ minutes
$BlueProc_L$: lower bound on how long it takes Blue to process a shot and improve its firing accuracy	$U[0.5,3.5]$ minutes
$BlueProc_U$: upper bound on how long it takes Blue to process a shot and improve its firing accuracy	$BlueProc_L + U[0,5]$ minutes
$RedProc_L$: lower bound on how long it takes Red to process a shot improve its firing accuracy	$U[0.5,3.5]$ minutes
$RedProc_U$: upper bound on how long it takes Red to process a shot and improve its firing accuracy	$RedProc_L + U[0,5]$ minutes
Distribution for number of Red targets hit by Blue risk-free shot	see below explanation
Probability Blue hits Red during risky-fire	see below explanation
Probability Red hits Blue during risky-fire	see below explanation

Table 11: Inputs for each Scenario of the Simulation Model

We now provide more detail on the three parameters at the bottom of the Table 11.

Number of Red Targets Hit By Risk-Free Shot

After moving positions, Blue simultaneously fires all of its cannons. During risky-fire, Blue fires one shot at a time as each cannon will have random perturbations in its actual firing

time. Thus during risky-fire, each Blue shot can hit at most one Red target (i.e., decrease Red health by 1). However, Blue's risk-free volley can hit multiple targets. If we denote h_B and h_R as Blue and Red health remaining, then the maximum number of Red targets hit on a risk-free volley is $\min(h_B, h_R)$ because each Blue cannon can hit at most one Red target. If we define $maxHB$ as the starting Blue health (3rd parameter listed in Table 11) , and $maxHB$ as the starting Red health (4th parameter listed in Table 11), then the probability distribution for the number of Red targets hit on a risk-free shot is

$$\begin{aligned}&\mathbf{P}[k \text{ Red targets hit on risk-free volley } ; h_B, h_R] \\ &\quad = a \times b^{maxHB-h_B} \times c^{maxHR-h_R} \times d^{k-1}, \quad 1 \leq k \leq \min(h_B, h_R)\end{aligned} \tag{H.7}$$

For $a, b, c, d \in [0, 1]$. We define the probability that $k = 0$ Red targets are hit on the risk-free shot to normalize the probability distribution to 1. If both Blue and Red are at the maximum health, then the probability that Blue hits exactly 1 Red target equals a. As Blue health decrements, this hit probability degrades by factor b, as there are fewer cannons available to fire at Red. Similarly as Red health decrements the probability degrades by c as there are fewer targets that a round might hit. This is consistent with the area fire nature of artillery. Finally the d^{k-1} term gives a geometric flavor to the distribution: the probability monotonically decreases from $k = 1$ to $k = \min(h_B, h_R)$. For the simulation, we generate a, b, c, d as follows

$$a \sim U[0.5, 0.4] \qquad b \sim U[0.3, 0.95] \qquad c \sim U[0.3, 0.95] \qquad d \sim U[0.5, 0.45]$$

Probability One Risky Blue Shot Hits Red

During risky-fire the probability one Blue shot hits Red depends upon the number of Blue shots fired since Blue arrived to the current location *that Blue has processed into its firing system.* This includes the risk-free shot. We denote the quantity of Blue shots processed

by Blue as $s_B^{(B)}$. Blue improves its aim based on feedback from forward observers about the results of earlier Blue shots. Because Red artillery is firing from a different location than the desired Red target, the number of Red shots fired does not improve Blue's accuracy for aiming at Red targets. If we denote $maxSB$ as the maximum number of Blue shots that Blue can incorporate before reaching its maximum situational awareness (1st parameter listed in Table 11), then

$$\mathbf{P}[\text{Blue hits Red with risky shot}; s_B^{(B)}] = e + \frac{\sum_{i=1}^{s_B^{(B)}} \frac{1}{i}}{\sum_{i=1}^{maxSB} \frac{1}{i}}(f - e) \tag{H.8}$$

For $f, e \in [0, 1]$. e represents the probability of hitting Red without any situational awareness: $s_B^{(B)} = 0$. f represents the hit probability at maximum situational awareness $s_B^{(B)} = maxSB$. The inverse weighting formulation in (H.8) provides a decreasing marginal impact of shots. For example if $maxSB$ is 2, then the first shot provides $\frac{2}{3}$ of the improvement from e to f, and the second shot provides the remaining $\frac{1}{3}$ increase. We generate e, f as follows

$$e = a \sum_{i=0}^{maxHR-1} d^i \qquad f \sim U[0.5, 1]$$

Where a and d are defined above in (H.7). In this case e corresponds to the probability Blue hits at least one Red target during the risk-free shot when both Blue and Red are at their maximum health levels. If $f < e$, then we continue generating f until $f > e$.

Probability One Risky Red Shot Hits Blue

The probability that Red hits Blue with a risky shot follows similar logic, however Red incorporates both Red shots and Blue shots into its situational awareness. Observing Blue shots via radar can help Red pinpoint Blue's location. Observing the results from Red's earlier shots (e.g., from a forward observer or surveillance UAV) allows Red to adjust its aim to improve accuracy. In this case we denote $s_B^{(R)}$ as the number of Blue shots processed by Red, $s_R^{(R)}$ as the number of Blue shots processed by Red, and $maxSR$ as the maximum

number of Red shots required to reach its maximum situational awareness (2nd parameter listed in Table 11). We assume that Red also uses $maxSB$ for the number of Blue shots to reach Red's maximum situational awareness. The parameter w_B represents a weighting for how much emphasis Red puts on Blue shots relative to Red shots. For example a w_B near 1 might occurs if Red's radar system to track Blue shots is very accurate relative to the human observer tracking Red shots. Our final hit probability is a generalization of (H.8)

$$\begin{aligned}&\mathbf{P}[\text{Blue hits Red with risky shot}; s_B^{(R)}, s_R^{(R)}] \\ &= g + \left(w_B \frac{\sum_{i=1}^{s_B^{(R)}} \frac{1}{i}}{\sum_{i=1}^{maxSB} \frac{1}{i}} + (1 - w_B) \frac{\sum_{i=1}^{s_R^{(R)}} \frac{1}{i}}{\sum_{i=1}^{maxSR} \frac{1}{i}} \right) (h - g) \end{aligned} \tag{H.9}$$

g represents the probability of hitting Red without any situational awareness: $s_B^{(R)}, s_R^{(R)} = 0$. h represents the hit probability at maximum situational awareness $s_B^{(R)} = maxSB, s_R^{(R)} = maxSR$. Equation (H.9) has a similar decreasing marginal impact of shots as (H.8), but (H.9) has to track two variables. The relevant parameters have the following distributions

$$g = (0.5 + U[0,1])e \qquad h \sim U[0.5, 1] \qquad w_B \sim U[0,1]$$

Where e is defined above in (H.8). If $h < g$, then we continue generating h and g until $h > g$.

H.2.2 Random Variables within a Scenario

Many of the random variables listed in Table 12 depend upon the input parameters from Table 11.

Note the difference between the interfire time of Blue during risky-fire (3rd random variable listed in Table 12) and Red's interfire time (4th random variable listed in Table 12). Recall from Section H.2.1 that h_B represents the current health (i.e., remaining cannons) of Blue. For larger values of h_B the overall fire rate should be greater because there are more functional cannons; consequently why Blue's interfire time depends inversely on h_B. Because

Description	Distribution
Time for Blue to travel to new location after move	$U[Trav_L, Trav_U]$
Time until Blue fires risk-free shot after arriving to a new location	$U[RiskFreeBlue_L, RiskFreeBlue_U]$
Time between Blue shots during risky-fire	$U\left[\frac{RiskyBlue_L}{h_B}, \frac{RiskyBlue_U}{h_B}\right]$
Time between Red shots during risky-fire	$U[RiskyRed_L, RiskyRed_U]$
Time until Blue processes last Blue shot and improves firing accuracy	$U[BlueProc_L, BlueProc_U]$
Time until Red processes last Blue or Red shot and improves firing accuracy	$U[RedProc_L, RedProc_U]$

Table 12: Random Variables Generated during One Scenario

Red artillery is not subject to attrition, Red's interfire time distribution does not vary over the course of the battle. When the processing times complete (5th and 6th random variables listed in Table 12) then the appropriate value of $s_B^{(B)}, s_B^{(R)}, s_R^{(R)}$ increments (these variables are defined in Section H.2.1)

I Deterministic Window of Opportunity

We now derive the optimal policy when Blue faces a deterministic time-window. As the proof is extremely long and the result is a natural generalization of the main results in Section 4, we place the analysis here in the Appendix. As we make limiting arguments in this Appendix, we use n (rather than N) to represent the number of periods. If we denote the time-window as W_n when we have n periods, then $W_n \sim Gamma(n, \gamma)$. To avoid having to consider unrealistic edge cases in our analysis, we assume the input rate parameters are positive, $\alpha, \beta, \gamma, \delta > 0$, and the risk-free probability $p_0 \in (0, 1)$. Furthermore in this Appendix we use the variables s, t to represent continuous time and the integers i, j to denote discrete time periods.

In this Appendix we define $\gamma = \frac{n}{T}$ and examine the limiting behavior of W_n and the corresponding optimal move policy for $n \to \infty$. In this situation, $\mathbf{E}[W_n] = T$ for all n and

$\mathbf{Var}[W_n] = \frac{T^2}{n}$. We prove that $W_n \to T$ almost surely and derive the following optimal policy for the special case where the time-window is deterministic.

Proposition 6. *If the length of the window of opportunity is almost surely the constant T, the optimal policy for Blue is as follows*

1. *If $\delta p_0 \geq \alpha$, Blue should scoot for all $0 \leq t \leq T$,*

2. *If condition (I.1) holds, Blue should fight for all $0 \leq t \leq T$*

$$\frac{\delta p_0 - (\alpha + \beta p_0)}{\delta - (\alpha + \beta)} \left(1 - \exp(-\delta T)\right) \leq \frac{\alpha}{\alpha + \beta} \frac{\delta p_0 - (\alpha + \beta)}{\delta - (\alpha + \beta)} \left(1 - \exp\left(-(\alpha + \beta)T\right)\right) \quad \text{(I.1)}$$

3. *If the conditions in cases 1–2 do not hold, then Blue should scoot for $0 \leq t \leq x^* T$ and fight for $x^* T < t \leq T$ for the unique $x^* \in (0, 1)$ that satisfies*

$$\frac{\delta p_0 - (\alpha + \beta p_0)}{\delta p_0 - (\alpha + \beta)} \left(1 - \exp(-\delta T(1 - x^*))\right) = \frac{\alpha}{\alpha + \beta} \left(1 - \exp\left(-(\alpha + \beta)T(1 - x^*)\right)\right) \quad \text{(I.2)}$$

The proof appears below. We derive the expressions in equation (I.2) based on the limits of $\hat{\mathbf{S}}_{\mathbf{F}}[i]$ from (6) and $\hat{\mathbf{F}}_{\mathbf{all}}[i]$ from (5).

Proof of Proposition 6: Threshold for Deterministic Time-window

W_n denotes the time-window, where $W_n \sim Gamma(n, \frac{n}{T})$. The proof is a fairly long and tedious exercise in real analysis to ensure that the limiting arguments are rigorous. We present several Lemmas in this section to highlight the key components of the proof. The technical details appear in later subsections. Our first Lemma specifies that examining the limiting behavior of W_n is equivalent to a deterministic time-window

Lemma 10. *$W_n \to T$ almost surely.*

The proof of Lemma 10 appears in Appendix I.1.

We use τ_n from (7) to denote the threshold's dependence on the parameter n. We also define the normalized threshold $x_n = \frac{\tau_n}{n}$. Blue switches from scooting to fighting at the end of period τ_n (equivalently the beginning of period τ_n+1), which is a deterministic integer. We can view this decision point instead as a random time $V_n \sim Gamma(\tau_n, \frac{n}{T})$: Blue switches from scooting to fighting at time V_n. In the special case where Blue always fights, we redefine τ_n from $-\infty$ (see (7)) to $\tau_n = 0$. This allows us to treat $\tau_n + 1$ as a threshold from either perspective: Blue scoots for the last time in period τ_n or Blue fights for the first time in period $\tau_n + 1$. When $\tau_n = 0$ then $x_n = 0$; that is Blue starts fighting immediately.

The remainder of our steps focus on showing that $x_n \to x^*$, where x^* is the optimal threshold defined in Proposition 6. Once we have shown this convergence, then Blue's switching time V_n converges to the deterministic value in Proposition 6 by the next Lemma.

Lemma 11. *If x_n converges to some value $\tilde{x}$, then $V_n \to \tilde{x}T$ almost surely.*

The proof of Lemma 11 appears in Appendix I.2.

We first prove case 1 of Proposition 6, which dictates when Blue should always scoot

Lemma 12. *If $\delta p_0 \geq \alpha$, $x_n \to 1$ and Blue scoots for all $0 \leq t \leq T$.*

The proof of Lemma 12 appears in Appendix I.3. For the remainder of this section we assume $\delta p_0 < \alpha$

To show that $x_n \to x^*$, we first generalize $\hat{\mathbf{F}}_{\mathbf{all}}[i]$ from (5) and $\hat{\mathbf{S}}_{\mathbf{F}}[i]$ from (6) so that they are functions of all real values $x \in [0, 1]$, rather than integers $1 \leq i \leq n$. We denote these generalized functions as $\hat{\mathbf{F}}_{\mathbf{all}}[x; n]$ and $\hat{\mathbf{S}}_{\mathbf{F}}[x; n]$. We introduce an intermediate term $i(x; n) = \lceil xn \rceil$, so that for any n, $\hat{\mathbf{F}}_{\mathbf{all}}[x; n]$ and $\hat{\mathbf{S}}_{\mathbf{F}}[x; n]$ are well defined for all $x \in [0, 1]$

$$\hat{\mathbf{F}}_{\mathbf{all}}[x; n] \equiv \frac{\alpha}{\alpha + \beta} \left(1 - \left(\frac{\frac{n}{T}}{\alpha + \beta + \frac{n}{T}} \right)^{n - i(x;n) + 1} \right) \tag{I.3}$$

$$\hat{\mathbf{S}}_{\mathbf{F}}[x; n] \equiv \frac{\delta p_0}{\frac{n}{T} + \delta p_0} + \frac{\frac{n}{T}}{\frac{n}{T} + \delta p_0} \frac{\delta}{\delta + \frac{n}{T}} \sum_{j = i(x;n)+1}^{n} \left(\frac{\frac{n}{T}}{\delta + \frac{n}{T}} \right)^{j - i(x;n) - 1} \left(p_0 + (1 - p_0) \hat{\mathbf{F}}_{\mathbf{all}}[\frac{j}{n}; n] \right) \tag{I.4}$$

The expressions in (I.2) are related to the limits of $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ of $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$. We present the limits for $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ and $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$, respectively, below.

$$\begin{aligned} w(x) \equiv & \frac{\delta p_0 - (\alpha + \beta p_0)}{\delta - (\alpha + \beta)} \left(1 - \exp(-\delta T(1-x))\right) \\ & + \left(1 - \frac{\delta p_0 - (\alpha + \beta)}{\delta - (\alpha + \beta)}\right) \frac{\alpha}{\alpha + \beta} \left(1 - \exp(-(\alpha + \beta) T(1-x))\right) \end{aligned} \tag{I.5}$$

$$z(x) \equiv \frac{\alpha}{\alpha + \beta} \left(1 - \exp\left(-(\alpha + \beta) T(1-x)\right)\right) \tag{I.6}$$

The left-hand side of (I.2) corresponds to the first line of $w(x)$ in (I.5) (up to a constant factor). The right-hand side of (I.2) corresponds to the difference between $z(x)$ in (I.6) and the second line of $w(x)$ in (I.5) (up to a constant factor). In the discrete setting, we determine the last period where $\hat{\mathbf{S}}_{\mathbf{F}}[i]$ exceeds $\hat{\mathbf{F}}_{\mathbf{all}}[i]$. We consider essentially the same relationship in the limiting case in this Appendix: when does $w(x)$ last exceed $z(x)$?

Our next Lemma formalizes the limiting relationship between $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ ($\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$) and $w(x)$ ($z(x)$):

Lemma 13. *$\hat{\mathbf{S}}_{\mathbf{F}}[\cdot;n]$ converges uniformly to $w(\cdot)$ on the interval $[0,1]$ and $\hat{\mathbf{F}}_{\mathbf{all}}[\cdot;n]$ converges uniformly to $z(\cdot)$ on the interval $[0,1]$.*

The proof of Lemma 13 appears in Appendix I.4. Lemma 13 is a crucial piece that allows us to transform our arguments in the discrete setting to continuous time via limits.

The following two Lemmas present dominance relationships between $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ ($w(x)$) and $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$ ($z(x)$)

Lemma 14. *If $\hat{\mathbf{S}}_{\mathbf{F}}[x;n] > \hat{\mathbf{F}}_{\mathbf{all}}[x;n]$, then $\hat{\mathbf{S}}_{\mathbf{F}}[y;n] > \hat{\mathbf{F}}_{\mathbf{all}}[y;n]$ for all $0 \leq y \leq x$*

Lemma 15. *If $w(x) > z(x)$ for some $x \in [0,1)$, then $w(y) > z(y)$ for all $0 \leq y \leq x$*

The proof of Lemma 14 appears in Appendix I.5, and the proof of Lemma 15 appears in Appendix I.6. These two Lemmas allow us to consider at most one crossover point between $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ ($w(x)$) and $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$ ($z(x)$).

Combining uniform convergence in Lemma 13 with the ordering results in Lemma 14–15 produces the final two cases in Proposition 6:

Lemma 16. *If condition (I.1) holds, Blue should fight for all* $0 \leq t \leq T$

Lemma 17. *If* $\delta p_0 \leq \alpha$ *and condition (I.1) does not hold, then Blue should scoot for* $0 \leq t \leq x^*T$ *and fight for* $x^*T < t \leq T$ *for the unique* $x^* \in (0,1)$ *that satisfies* $w(x^*) = z(x^*)$

The proof for Lemma 16 appears in Appendix I.7, and the proof for Lemma 17 appears in Appendix I.8.

I.1 Proof of Lemma 10 : $W_n \to T$

We use Borel-Cantelli machinery to prove this Lemma . If we can show that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{P}[|W_i - T| > \epsilon] < \infty \qquad \forall \epsilon > 0, \tag{I.7}$$

then by Borel-Cantelli , W_n converges to T almost surely. We use the generalized Chebyshev inequality :

$$\mathbf{P}[|W_i - T| > \epsilon] \leq \frac{\mathbf{E}[(W_i - T)^4]}{\epsilon^4} \tag{I.8}$$

W_i is an Gamma random variable with mean $\mathbf{E}[W_i] = T$, variance $\mathbf{Var}[W_i] = \frac{T^2}{i}$ and kurtosis

$$\frac{\mathbf{E}[(W_i - T)^4]}{(\mathbf{Var}[W_i])^2} = \frac{6}{i} + 3 \tag{I.9}$$

Combining (I.8) and (I.9) produces

$$\mathbf{P}[|W_i - T| > \epsilon] \leq \frac{\mathbf{E}[(W_i - T)^4]}{\epsilon^4} = \frac{T^4}{\epsilon^4}\left(\frac{6}{i^3} + \frac{3}{i^2}\right) \tag{I.10}$$

Substituting (I.10) into the left-hand side of (I.7), we see the infinite sum is finite because $\mathbf{P}[|W_i - T| > \epsilon]$ scales $\sim i^{-2}$. This completes the proof.

I.2 Proof of Lemma 11 : $V_n \to \tilde{x}T$

We only have to make one slight tweak from the proof in Appendix I.1. In Appendix I.1, $\mathbf{E}[W_n] = T$ does not depend upon n, but here $\mathbf{E}[V_n] = x_n T$ does. Therefore we modify the Chebyshev inequality in (I.8)

$$\mathbf{P}[|V_i - \tilde{x}T| > \epsilon] \leq \mathbf{P}[|V_i - x_i T| + T|x_i - \tilde{x}| > \epsilon] = \mathbf{P}[|V_i - x_i T| > \epsilon - T|x_i - \tilde{x}|] \quad \text{(I.11)}$$

Since $x_i \to \tilde{x}$, there exists some N such that for $i > N$, $T|x_i - \tilde{x}| < \frac{\epsilon}{2}$. This relationship implies that

$$\mathbf{P}[|V_i - \tilde{x}T| > \epsilon] \leq \mathbf{P}[|V_i - x_i T| > \frac{\epsilon}{2}], \quad \text{for } i > N \quad \text{(I.12)}$$

At this point, the same logic applies as in Appendix I.1: we combine Chebyshev with the variance and kurtosis of V_i to show that the right-hand side of (I.12) scales $\sim i^{-2}$ for large i.

I.3 Proof of Lemma 12: Blue Always Scoots if $\delta p_0 \geq \alpha$

We start with the strict inequality $\delta p_0 > \alpha$:

$$\begin{aligned} \delta p_0 \frac{n}{T} &> \alpha \frac{n}{T} \\ \to \delta p_0 \beta + \delta p_0 \frac{n}{T} &> \alpha \frac{n}{T} \\ \to \delta p_0 \alpha + \delta p_0 \beta + \delta p_0 \frac{n}{T} &> \alpha \frac{n}{T} + \delta p_0 \alpha \\ \to \frac{\delta p_0}{\frac{n}{T} + \delta p_0} &> \frac{\alpha}{\alpha + \beta + \frac{n}{T}} \end{aligned} \quad \text{(I.13)}$$

Recall that we define $\gamma_n = \frac{n}{T}$, so condition (I.13) implies that Blue should scoot for all periods for any fixed n (see condition (C.1)). That is $\tau_n = n$ for all n and hence $x_n = 1$ for all n. Therefore by Lemma 11, Blue scoots for all $0 \leq t \leq T$. The same logic applies if $\delta p_0 = \alpha$ and $\beta > 0$ as the second step above produces a strict inequality. By assumption $\beta > 0$, so

(I.13) holds for all $\delta p_0 \geq \alpha$ and the proof is complete.

I.4 Proof of Lemma 13: Uniform convergence of $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$ $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$

We present the results in two separate subsections.

I.4.1 $\hat{\mathbf{F}}_{\mathbf{all}}[x;n]$

Recall the definition $i(x;n) = \lceil xn \rceil$, where $i(x;n)$ appears in (I.3)–(I.4). By construction $-1 \leq nx - i(x;n) \leq 0$. Furthermore, we define

$$\kappa(x;n) = \frac{nx - i(x;n) + 1}{n} \leq \frac{1}{n}, \tag{I.14}$$

and rewrite (I.3) as

$$\hat{\mathbf{F}}_{\mathbf{all}}[x;n] = \frac{\alpha}{\alpha+\beta}\left(1 - \left(1 - \frac{\alpha+\beta}{\alpha+\beta+\frac{n}{T}}\right)^{n(1-x+\kappa(x;n))}\right). \tag{I.15}$$

By construction $\kappa(x;n) \to 0$ (see (I.14)). Comparing (I.15) to the desired limit in (I.6), to complete the proof it suffices to show $f(\cdot;n) \overset{unif}{\to} r(\cdot)$, where

$$f(x;n) = \left(1 - \frac{\alpha+\beta}{\alpha+\beta+\frac{n}{T}}\right)^{n(1-x+\kappa(x;n))} \tag{I.16}$$

$$r(x) \equiv \exp\left(-(\alpha+\beta)T(1-x)\right) \tag{I.17}$$

For notational convenience we define

$$h(n) = \frac{\alpha+\beta}{\alpha+\beta+\frac{n}{T}} \tag{I.18}$$

$$g(x;n) = n(1 - x + \kappa(x;n)). \tag{I.19}$$

Thus we want to uniformly bound $f(x;n) = (1-h(n))^{g(x;n)}$ against $r(x)$.

The following steps are adapted from RRL (2015). To show uniform convergence we first note the following inequalities for any $y \in [0, 1]$

$$(1 + y) \leq e^{y} \tag{I.20}$$

$$(1 - y) \leq e^{-y} \tag{I.21}$$

Conditions (I.20)–(I.21) follow by inspection of the Taylor series expansion of the exponential. From (I.18), $h(n) < 1$ and hence condition (I.21) implies

$$\begin{aligned} 0 \leq e^{-h(n)g(x;n)} - f(x;n) &= e^{-h(n)g(x;n)} - (1 - h(n))^{g(x;n)} \\ &= e^{-h(n)g(x;n)} \left(1 - (1 - h(n))^{g(x;n)} e^{h(n)g(x;n)}\right) \end{aligned} \tag{I.22}$$

Combining (I.20) and (I.22) produces

$$\begin{aligned} 0 \leq e^{-h(n)g(x;n)} - f(x;n) &\leq e^{-h(n)g(x;n)} \left(1 - (1 - h(n))^{g(x;n)} (1 + h(n))^{g(x;n)}\right) \\ &= e^{-h(n)g(x;n)} \left(1 - (1 - h^2(n))^{g(x;n)}\right) \end{aligned} \tag{I.23}$$

By Bernoulli's inequality (Weisstein, 2018), (I.23) implies

$$0 \leq e^{-h(n)g(x;n)} - f(x;n) \leq e^{-h(n)g(x;n)} g(x;n) h^2(n) \leq g(0;n) h^2(n) \leq 2n \left(\frac{\alpha + \beta}{\alpha + \beta + \frac{n}{T}}\right)^2 \tag{I.24}$$

Where we have used the fact that $g(x;n)$ achieves its maximum value n at $x = 0$ (see I.19). The bound in (I.24) goes to 0 uniformly in $x \in [0, 1]$.

To complete the proof

$$|r(x) - f(x;n)| \leq |r(x) - e^{-h(n)g(x;n)}| + |e^{-h(n)g(x;n)} - f(x;n)| \tag{I.25}$$

The second term on the right-hand side of (I.25) is uniformly bounded by (I.24). We next turn to the first term. By the mean value theorem

$$\begin{aligned} |r(x) - e^{-h(n)g(x;n)}| &\leq |(\alpha+\beta)T(1-x) - h(n)g(x;n)| \\ &\leq (\alpha+\beta)T \left| 1 - \frac{1}{1+\frac{T}{n}(\alpha+\beta)} \right| + \frac{\alpha+\beta}{\alpha+\beta+\frac{n}{T}} \end{aligned} \tag{I.26}$$

Therefore $|r(x) - e^{-h(n)g(x;n)}|$ uniformly converges to 0 over $x \in [0,1]$ and and so does $|r(x) - f(x;n)|$ in (I.25), which completes the proof.

I.4.2 $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$

$w(x)$ in (I.5) is equivalent to

$$w(x) \equiv \int_x^1 \delta T \exp(-\delta T(y-x)) \left(p_0 + (1-p_0)\frac{\alpha}{\alpha+\beta}\left(1 - \exp\left(-(\alpha+\beta)T(1-y)\right)\right)\right) dy \tag{I.27}$$

In this section we show that $\hat{\mathbf{S}}_{\mathbf{F}}[\cdot;n]$ uniformly converges to the $w(\cdot)$ representation in (I.27).

Comparing $w(x)$ in (I.27) to $\hat{\mathbf{S}}_{\mathbf{F}}[x;n]$ (I.4), we see that to prove $\hat{\mathbf{S}}_{\mathbf{F}}[\cdot;n] \overset{unif}{\to} w(\cdot)$ it is sufficient to prove

$$\frac{1}{n}\sum_{j=i(x;n)+1}^{n} u\left(\frac{j}{n}, \cdot; n\right) \overset{unif}{\to} \int_{\cdot}^{1} v(y,\cdot)dy \tag{I.28}$$

where

$$u(y,x;n) \equiv \left(1 - \frac{\delta}{\delta + \frac{n}{T}}\right)^{n(y-x+\tilde{\kappa}(x,y;n))} \left(p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[y;n]\right) \tag{I.29}$$

$$v(y,x) \equiv \exp(-\delta T(y-x)) \left(p_0 + (1-p_0)\frac{\alpha}{\alpha+\beta}\left(1 - \exp\left(-(\alpha+\beta)T(1-y)\right)\right)\right), \tag{I.30}$$

and in (I.29) we introduce notation $j(y;n)$ and $\tilde{\kappa}(x,y;n)$ similar to $i(x;n)$ and $\kappa(x;n)$ from

Appendix I.4.1

$$j(y;n) = \lceil yn \rceil \tag{I.31}$$

$$\tilde{\kappa}(x,y;n) = \frac{nx - i(x;n) - (ny - j(y;n)) - 1}{n}, \tag{I.32}$$

and by inspection $-\frac{2}{n} \leq \tilde{\kappa}(x,y;n) \leq 0$.

We now argue $u(\cdot,\cdot;n) \overset{unif}{\to} v(\cdot,\cdot)$ for $x \in [0,1]$ and $y \in [x,1]$. Both $u(y,x;n)$ and $v(y,x)$ are products of two exponential terms. Examining the second terms in each, we have

$$p_0 + (1-p_0)\hat{\mathbf{F}}_{\mathbf{all}}[y;n] \overset{unif}{\to} \left(p_0 + (1-p_0)\frac{\alpha}{\alpha+\beta}\left(1 - \exp\left(-(\alpha+\beta)T(1-y)\right)\right)\right), \tag{I.33}$$

which follows immediately from Appendix I.4.1. Since the first terms of $u(y,x;n)$ and $v(y,x)$ are exponential functions, we can use the same logic as in Appendix I.4.1 to show

$$\left(1 - \frac{\delta}{\delta + \frac{n}{T}}\right)^{n(y-x+\tilde{\kappa}(x,y;n))} \overset{unif}{\to} \exp(-\delta T(y-x)), \tag{I.34}$$

Unlike in (I.33), (I.34) depends upon both x and y. Consequently, we have to be careful extending our arguments from Appendix I.4.1, where we only showed uniform convergence of a one-dimensional function. The $\tilde{\kappa}(x,y;n)$ term is uniformly bounded (see discussoin below (I.32)), thus it will uniformly go to zero and we can ignore it for this analysis. The variables x and y only appear as $y-x$, and that difference is bounded between 0 and 1 in our region of interest: $x \in [0,1]$ and $y \in [x,1]$. Therefore we can apply our one-dimensional result from Appendix I.4.1 on $y-x$ to get condition (I.34)

Because the exponential functions in (I.33) and (I.34) are bounded between 0 and 1, the convergence conditions in (I.33) and (I.34) imply that the products also uniformly converge for $x \in [0,1]$ and $y \in [x,1]$: $u(\cdot,\cdot;n) \overset{unif}{\to} v(\cdot,\cdot)$.

We next bound

$$\left| \frac{u(y,x;n)}{n} - \int_{y-\frac{1}{n}}^{y} v(z,x)dz \right| \tag{I.35}$$

To start, for any $\epsilon > 0$ we can find N_1 such that

$$|u(y,x;n) - v(y,x)| < \frac{\epsilon}{2} \qquad \forall n > N_1, x \in [0,1], y \in [x,1] \tag{I.36}$$

Such an N_1 exists by uniform convergence of $u(\cdot,\cdot;n)$ to $v(\cdot,\cdot)$. We next can find an N_2 such that

$$|v(y,x) - v(y_1,x)| < \frac{\epsilon}{2} \qquad \text{if } |y-y_1| < \frac{1}{N_2} \quad \forall x \in [0,1], y \in [x,1] \tag{I.37}$$

This follows by the uniform continuity of $v(y,x)$ over $x \in [0,1], y \in [x,1]$. $v(y,x)$ is uniformly continuous because it is the product of two exponential functions, each of which is bounded between 0 and 1 in our region of interest. It is straightforward to show an exponential function is uniformly continuous by a mean value theorem argument (see discussion around (I.26)).

If we define $N^* = \max(N_1, N_2)$, then

$$\begin{aligned} \left| \frac{u(y,x;n)}{n} - \int_{y-\frac{1}{n}}^{y} v(z,x)dz \right| &\leq \left| \frac{u(y,x;n)}{n} - \frac{v(y,x)}{n} \right| + \left| \frac{v(y,x)}{n} - \int_{y-\frac{1}{n}}^{y} v(z,x)dz \right| \\ &\leq \frac{\epsilon}{2n} + \frac{\epsilon}{2n} \qquad \forall n > N^*, \quad x \in [0,1], y \in [x,1] \end{aligned} \tag{I.38}$$

The first term in (I.38) follows from (I.36) and the second term follows from (I.37).

Condition (I.38) allows us to make a triangle inequality argument to show (I.28):

$$\left| \frac{1}{n} \sum_{j=i(x;n)+1}^{n} u\left(\frac{j}{n}, x; n\right) - \int_{x}^{1} v(y,x)dy \right| \leq \epsilon \qquad \forall n > N^*, \quad x \in [0,1] \tag{I.39}$$

I.5 Proof of Lemma 14: $\hat{\mathbf{S}}_{\mathbf{F}}[x;n] > \hat{\mathbf{F}}_{\mathbf{all}}[x;n]$ implies $\hat{\mathbf{S}}_{\mathbf{F}}[y;n] > \hat{\mathbf{F}}_{\mathbf{all}}[y;n]$ for $y \in [0,x]$

By Lemma 12 we only need to consider the $\delta p_0 < \alpha$ cases. If $\delta p_0 < \alpha$, then there exists an N^* such that

$$\frac{\delta p_0}{\frac{n}{T} + \delta p_0} \leq \frac{\alpha}{\alpha + \beta + \frac{n}{T}} \qquad \forall n > N^* \tag{I.40}$$

Condition (I.40) implies $\tau_n < n$ for all $n > N^*$ (see (C.5)). For the remainder of this section we assume $n > N^*$ and thus condition (I.40) holds and $\tau_n < n$.

Because we focus on a specific fixed value of n, we drop the notational dependence on n and use integers i and j instead of real values x and y. We also use γ as our time rate rather than T/n. We now proceed to show

$$\hat{\mathbf{S}}_{\mathbf{F}}[i] > \hat{\mathbf{F}}_{\mathbf{all}}[i] \to \hat{\mathbf{S}}_{\mathbf{F}}[j] > \hat{\mathbf{F}}_{\mathbf{all}}[j] \quad \forall 1 \leq j \leq i \tag{I.41}$$

We first consider $\tau = 0$ or $\tau = 1$ (that is Blue always fights, or only scoots in the first period). In these two cases condition (I.41) holds trivially. For $\tau > 1$, we rewrite $\hat{\mathbf{F}}_{\mathbf{all}}[i]$ and $\hat{\mathbf{S}}_{\mathbf{F}}[i]$ from (5) and (6)

$$\hat{\mathbf{S}}_{\mathbf{F}}[i] = \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\left(\frac{\gamma + \delta p_0}{\gamma + \delta}\hat{\mathbf{S}}_{\mathbf{F}}[i+1] + \frac{\delta(1-p_0)}{\gamma + \delta}\hat{\mathbf{F}}_{\mathbf{all}}[i+1]\right) \tag{I.42}$$

$$\hat{\mathbf{F}}_{\mathbf{all}}[i] = \frac{\alpha}{\alpha + \beta + \gamma} + \frac{\gamma}{\alpha + \beta + \gamma}\hat{\mathbf{F}}_{\mathbf{all}}[i+1] \tag{I.43}$$

We next move to the general $1 < \tau < n$ case; recall we do not need to consider the $\tau = n$ case (Blue always scoots), as Lemma 12 accounts for this special case.

The steps to show (I.41) are similar to the step in Appendix C.2 to prove Lemma 2.

Based on definition of τ, we have

$$\hat{\mathbf{S}}_{\mathbf{F}}[\tau] > \hat{\mathbf{F}}_{\mathbf{all}}[\tau] \tag{I.44}$$

$$\hat{\mathbf{S}}_{\mathbf{F}}[\tau+1] \leq \hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] \tag{I.45}$$

Substituting $i = \tau$ into (I.42) and using condition (I.45) produces

$$\hat{\mathbf{S}}_{\mathbf{F}}[\tau] \leq \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] \tag{I.46}$$

Combining (I.46) and (I.44) yields

$$\hat{\mathbf{F}}_{\mathbf{all}}[\tau] < \hat{\mathbf{S}}_{\mathbf{F}}[\tau] \leq \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] \tag{I.47}$$

$$\rightarrow \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1] < \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1], \tag{I.48}$$

where (I.48) follows by substituting (I.43) into (I.47). Condition (I.48) is just comparison of two linear functions.

Because we assume $\tau < n$ condition (C.5) holds which implies the intercept on the left-hand side of (I.48) is larger than the intercept on the right-hand side and the slope on the left-hand side of (I.48) is smaller than the slope on the right-hand side. Consequently (I.48) holds if we replace $\hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]$ with any value $z \geq \hat{\mathbf{F}}_{\mathbf{all}}[\tau+1]$, specifically for any $\hat{\mathbf{F}}_{\mathbf{all}}[j]$ where $j \leq \tau + 1$:

$$\frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[j] < \frac{\delta p_0}{\gamma + \delta p_0} + \frac{\gamma}{\gamma + \delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[j], \quad \forall j \leq \tau + 1 \tag{I.49}$$

We now show $\hat{\mathbf{S}}_{\mathbf{F}}[\tau-1] > \hat{\mathbf{F}}_{\mathbf{all}}[\tau-1]$:

$$\hat{\mathbf{S}}_{\mathbf{F}}[\tau-1] = \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\left(\frac{\gamma+\delta p_0}{\gamma+\delta}\hat{\mathbf{S}}_{\mathbf{F}}[\tau] + \frac{\delta(1-p_0)}{\gamma+\delta}\hat{\mathbf{F}}_{\mathbf{all}}[\tau]\right) \tag{I.50}$$

$$\geq \frac{\delta p_0}{\gamma+\delta p_0} + \frac{\gamma}{\gamma+\delta p_0}\hat{\mathbf{F}}_{\mathbf{all}}[\tau] \tag{I.51}$$

$$> \frac{\alpha}{\alpha+\beta+\gamma} + \frac{\gamma}{\alpha+\beta+\gamma}\hat{\mathbf{F}}_{\mathbf{all}}[\tau] \tag{I.52}$$

$$= \hat{\mathbf{F}}_{\mathbf{all}}[\tau-1] \tag{I.53}$$

Where (I.51) follows by inequality (I.44), (I.52) follows from (I.49) and (I.53) follows from (I.43). Using the same steps in (I.50)–(I.52), we can next show $\hat{\mathbf{S}}_{\mathbf{F}}[\tau-2] > \hat{\mathbf{F}}_{\mathbf{all}}[\tau-2]$, and continuing to iterate we have the desired result: $\hat{\mathbf{S}}_{\mathbf{F}}[i] > \hat{\mathbf{F}}_{\mathbf{all}}[i]$ for all $1 \leq i \leq \tau$, which implies (I.41) and the proof is complete.

I.6 Proof of Lemma 15: $w(x) > z(x)$ implies $w(y) > z(y)$ for $y \in [0, x]$

Let us rewrite $w(x)$ from (I.5) in a slightly different form

$$w(x) \equiv (1-\exp(-\delta T(1-x)))\left(p_0 + (1-p_0)\frac{\alpha}{\alpha+\beta}\left(1 - \frac{\delta}{\delta-(\alpha+\beta)}\right)\right)$$
$$+ (1-p_0)\frac{\alpha}{\alpha+\beta}\frac{\delta}{\delta-(\alpha+\beta)}\left(1-\exp(-(\alpha+\beta)T(1-x))\right) \tag{I.54}$$

Note that $w(1) = z(1) = 0$ ($z(x)$ is defined in (I.6)) and

$$\lim_{x\to-\infty} w(x) = p_0 + (1-p_0)\frac{\alpha}{\alpha+\beta} \tag{I.55}$$

$$\lim_{x\to-\infty} z(x) = \frac{\alpha}{\alpha+\beta} \tag{I.56}$$

We next define $v(x)$ as the difference between $w(x)$ and $z(x)$

$$v(x) = w(x) - z(x). \tag{I.57}$$

We have $v(1) = 0$. We next examine the derivative of $v(x)$ to show that $v(x)$ can have at most one local maximum or minimum on $(-\infty, 1)$:

$$v'(x) = \delta T A \exp(-\delta T(1-x)) + (\alpha+\beta) T B \exp(-(\alpha+\beta)T(1-x)) \tag{I.58}$$

for some constants A and B. The only possibility for $v'(x) = 0$ for all x is if both constants $A = B = 0$. This can only occur if $\beta = 0$. Because we assume $\beta > 0$, we can ignore this case. Attempting to solve $v'(x) = 0$ in (I.58), we see there is at most one real solution to $v'(x) = 0$, and therefore $v(x)$ has at most one local extreme point for $x \in (-\infty, \infty)$.

The condition in the Lemma follows immediately from the above argument. If $v(x) > 0$, than $v(y) > 0$ for all $y < x < 1$. Otherwise if there exists some $y < x$ such that $v(y) \leq 0$, then there exists a local maximum in $(y, 1]$ because $v(y) < v(x)$ and $v(x) > v(1)$. Furthermore since

$$\lim_{x \to -\infty} v(x) = \lim_{x \to -\infty} w(x) - z(x) = p_0 \frac{\beta}{\alpha+\beta} > 0 \tag{I.59}$$

there must also be a local minimum on $(-\infty, y]$. This implies $v(y)$ has at least two extreme points, which is a contradiction. Replacing $v(x) = w(x) - z(x)$, we have the condition of the Lemma.

I.7 Proof of Lemma 16: Case 2 of Proposition 6

We use the same notation as in Appendix I.6. Namely $v(x) = w(x) - z(x)$ according to (I.57). We know that $v(1) = 0$. Condition (I.1) in case 2 of Proposition 6 is equivalent to $v(0) \leq 0$. Similar logic to the proof of Lemma 15 in Appendix I.6 produces

$$v(0) \leq 0 \to v(x) < 0 \quad \forall x \in (0, 1) \tag{I.60}$$

Condition (I.60) follows by the one-local-extreme-point argument from Appendix I.6. If $v(0) \leq 0$ then by (I.59) and the Intermediate Value Theorem there exists a $\tilde{x} \leq 0$ such that $v(\tilde{x}) = 0$. We assume (I.60) does not hold and generate a contradiction. If there exists an $x \in (0, 1)$ such that $v(x) \geq 0$, then this implies again by the Intermediate Value Theorem that there is an $\hat{x} \in (0, 1)$ such that $v(\hat{x}) = 0$. However, this yields three roots for $v(\cdot)$, $v(\tilde{x}) = v(\hat{x}) = v(1) = 0$, which implies at least two local extreme points and a contradiction. Therefore (I.60) holds and $w(x) \leq z(x)$ for all $x \in [0, 1)$

If $v(x) < 0$ for $x \in [0, 1)$ and $v(1) = 0$, then $v'(1) > 0$. Before proceeding we eliminate the possibility that $v'(1) = 0$. If $v'(1) = 0$ then, we have one local extreme point at $x = 1$. The local extreme point cannot be a maximum, because this would imply a local minimum on $(-\infty, 1)$ by (I.59), which contradicts our one-local-extreme-point condition. However, if we have a local minimum at $x = 1$, then $v(x) > 0$ for all $x \in (-\infty, 1)$ by Lemma 15, which contradicts the condition of case 2 of Proposition 6: $v(0) \leq 0$. Therefore $v'(1) > 0$. Taking the derivative of $v(\cdot)$ and evaluating at 1 yields

$$v'(1) = -\delta p_0 T + \alpha T \tag{I.61}$$

By inspection of (I.61) $v'(1) > 0$ if and only if $\alpha > \delta p_0$, which implies cases 1 and 2 of Proposition 6 are mutually exclusive.

We next use uniform convergence to show that $x_n \to 0$, which completes the proof. For any $0 < \theta < 1$, $v(\theta) < 0$ by (I.60). Define

$$\epsilon \equiv -v(\theta) \tag{I.62}$$

By uniform convergence in Lemma 13, we next argue that there exists some N^* such that

$$\hat{\mathbf{S}}_{\mathbf{F}}[\theta; n] < \hat{\mathbf{F}}_{\mathbf{all}}[\theta; n] \quad \forall n > N^* \tag{I.63}$$

By Lemma 14, condition (I.63) implies that $x_n < \theta$. As discussed in the outline of Appendix I, we associate $x_n = 0$ with $\tau_n = -\infty$: both represent Blue firing immediately at the start of the battle. Consequently $|x_n| < \theta$ for all $n > N^*$, which proves convergence $x_n \to 0$.

To derive (I.63), we note that uniform convergence implies there exists N_1 and N_2 such that

$$|w(x) - \hat{\mathbf{S}}_{\mathbf{F}}[x; n]| < \frac{\epsilon}{3} \qquad \forall x \in [0,1], \forall n > N_1 \tag{I.64}$$

$$|z(x) - \hat{\mathbf{F}}_{\mathbf{all}}[x; n]| < \frac{\epsilon}{3} \qquad \forall x \in [0,1], \forall n > N_2 \tag{I.65}$$

Conditions (I.64)–(I.65) imply that for $n > N^* = \max(N_1, N_2)$

$$\begin{aligned} \hat{\mathbf{F}}_{\mathbf{all}}[\theta; n] - \hat{\mathbf{S}}_{\mathbf{F}}[\theta; n] &\geq (z(\theta) - \frac{\epsilon}{3}) - (w(\theta) + \frac{\epsilon}{3}) \\ &= -v(\theta) - \frac{2\epsilon}{3} = \frac{\epsilon}{3} > 0, \end{aligned} \tag{I.66}$$

where the last step in (I.66) follows by (I.62). Condition (I.66) yields (I.63). Therefore under condition (I.1) in case 2 of Proposition 6, the limiting policy is fight for all $0 \leq t \leq T$

I.8 Proof of Lemma 17: Case 3 of Proposition 6

We first prove that for this case there is a unique solution $x^* \in (0,1)$ such that $w(x^*) = z(x^*)$.

We use the same notation as in Appendix I.6 and I.7. Namely $v(x) = w(x) - z(x)$ according to (I.57). We know that $v(1) = 0$. If condition (I.1) in case 2 of Proposition 6 does not hold, then by construction $v(0) > 0$ (i.e., $w(0) > z(0)$). If the condition in case 1 of Proposition 6 does not hold, $\delta p_0 < \alpha$, then $v'(1) < 0$ (see (I.61).

The previous paragraph provides three conditions: $v(0) > 0$, $v(1) = 0$, $v'(1) < 0$. These three conditions imply by the Intermediate Value Theorem that there exist an $x^* \in (0,1)$ such that $v(x^*) = 0$. These three conditions also imply that $v(\cdot)$ has a local minimum on $(x^*, 1)$. To prove that x^* is the unique root, we use a contradiction argument. Assume there

exists another $\tilde{x} \in (0,1)$ such that $\tilde{x} \neq x^*$ and $v(\tilde{x}) = 0$. Without loss of generality we assume $\tilde{x} < x^*$ (otherwise we flip the labeling of x^* and $\tilde{x}$). However if $v(\tilde{x}) = v(x^*) = v(1) = 0$, with $0 < \tilde{x} < x^* < 1$, then there must exist at least two local extreme points on $(\tilde{x}, 1)$, which is a contradiction of our one-local-extreme-point result from Appendix I.6. Therefore x^* is the unique solution in $(0,1)$ such that $w(x^*) = z(x^*)$.

We next use uniform convergence to show that $x_n = \frac{\tau_n}{n} \to x^*$, which completes the proof. The steps are similar to the uniform convergence arguments in Appendix I.7. For any $0 < \theta < \min(x^*, 1 - x^*)$, $v(x^* - \theta) > 0$ and $v(x^* + \theta) < 0$ follow from the discussion in the previous paragraph. Define

$$\epsilon \equiv \min(v(x^* - \theta), -v(x^* + \theta)) \tag{I.67}$$

By uniform convergence in Lemma 13, we next argue that there exists some N^* such that

$$\hat{\mathbf{S}}_{\mathbf{F}}[x^* - \theta; n] > \hat{\mathbf{F}}_{\mathbf{all}}[x^* - \theta; n] \quad \forall n > N^* \tag{I.68}$$

$$\hat{\mathbf{S}}_{\mathbf{F}}[x^* + \theta; n] < \hat{\mathbf{F}}_{\mathbf{all}}[x^* + \theta; n] \quad \forall n > N^* \tag{I.69}$$

By Lemma 14, conditions (I.68)–(I.69) imply that $x^* - \theta < x_n < x^* + \theta$. That is $|x^* - x_n| < \theta$, for all $n > N^*$, which proves convergence $x_n \to x^*$.

To derive (I.68)–(I.69), we note that uniform convergence implies there exists N_1 and N_2 such that

$$|w(x) - \hat{\mathbf{S}}_{\mathbf{F}}[x; n]| < \frac{\epsilon}{3} \qquad \forall x \in [0,1], \forall n > N_1 \tag{I.70}$$

$$|z(x) - \hat{\mathbf{F}}_{\mathbf{all}}[x; n]| < \frac{\epsilon}{3} \qquad \forall x \in [0,1], \forall n > N_2 \tag{I.71}$$

Conditions (I.70)–(I.71) imply that for $n > N^* = \max(N_1, N_2)$

$$\begin{aligned}\hat{\mathbf{S}}_{\mathbf{F}}[x^* - \theta; n] - \hat{\mathbf{F}}_{\mathbf{all}}[x^* - \theta; n] &\geq (w(x^* - \theta) - \frac{\epsilon}{3}) - (z(x^* - \theta) + \frac{\epsilon}{3}) \\ &= v(x^* - \theta) - \frac{2\epsilon}{3} \geq \frac{\epsilon}{3} > 0, \end{aligned} \tag{I.72}$$

where the last step in (I.72) follows by (I.67). Condition (I.72) yields (I.68). A similar argument produces (I.69) and completes the proof

$$\begin{aligned}\hat{\mathbf{F}}_{\mathbf{all}}[x^* + \theta; n] - \hat{\mathbf{S}}_{\mathbf{F}}[x^* + \theta; n] &\geq (z(x^* + \delta) - \frac{\epsilon}{3}) - (w(x^* + \theta) + \frac{\epsilon}{3}) \\ &= -v(x^* + \theta) - \frac{2\epsilon}{3} \geq \frac{\epsilon}{3} > 0. \end{aligned} \tag{I.73}$$

I.9 Additional References

RRL. 2015. Uniform Convergence to the Exponential Function over a Compact Interval. *StackExchange*. January 18, 2015. `https://math.stackexchange.com/questions/1108581/uniform-convergence-to-the-exponential-function-over-a-compact-interval` Accessed on April 18, 2018.

Weisstein, Eric W. 2018. Bernoulli Inequality. *MathWorld*. `http://mathworld.wolfram.com/BernoulliInequality.html` Accessed on April 18, 2018.