Arming and Arms Races

James D. Fearon Department of Political Science Stanford University DRAFT

March 24, 2011

Abstract

I consider a model of arming in which states choose in successive periods whether to build weapons and whether to attack, and in which arms accumulate across periods. Surprisingly, despite a long history of formal models of arms races, the models we have in the literature lack one or more these features – no choices (Richardson models), no option to use the weapons (Prisoners' Dilemma models), or no accumulation of arms over multiple periods (Powell 1993, Jackson and Morelli 2009). In a Markov Perfect Equilibrium with two symmetric states, the states either (a) build up to stability at the first arms level such that a one-period "break out" does not give enough advantage to make attack worthwhile; or (b) fight at the outset because they anticipate that building to stability would not be worth the cost. (a) is more likely when the military technology favors defense and when the states have less "expansionist" preferences, in terms of value for territory. When the states differ enough in some respect – capacity to build, preferences for conquest, or size of economy, for example – mixed strategies appear on the equilibrium path. War or peace may result from an arms race between asymmetric rivals, with perhaps some tendency for races to be more dangerous as they proceed.

1 Introduction

Some of the first formal models developed in political science were models of arms races. Lewis Richardson (1919) represented an arms competition with a pair of differential equations. States were assumed to build at a rate determined by their perception of threat, which was in turn a function of the other side's arms level and its own population's "resistance" to arms expenditures. In these models states react automatically. There are no decisions or choices, whether about arms levels or attacking. With developments in game theory after World War II and the U.S.-Soviet arms race, the Prisoners' Dilemma game was commonly used as a model of the problem, and repeated

PD as a model of arms racing with cooperative equilibria representing the possibility of an arms control agreement (Snyder, 1971; Axelrod, 1984; Downs and Rocke, 1990; Downs, 1991).

Given this long tradition, it is surprising that there is still no model of arming and arms races in the literature that has what one might think are the minimal features needed to analyze the strategic situation properly. Namely, there are no models in the literature in which

1. states can choose whether to build arms in successive periods;

- 2. arms accumulate across (potentially many) periods; and
- 3. states can choose to attack the other in any period if they wish.

Richardsonian models of arming do not involve any choices or decision-making. Repeated Prisoners' Dilemmas do not represent either the accumulation of arms or the possibility of attack and war, which introduce a non-stationary aspect to the situation and so argue against any standard repeated game as a good model.¹

Powell (1993) studies a model in which states choose in successive periods how much to build (or buy) and whether to attack the other state, but in his treatment arms do not accumulate across periods, so there is no possibility of arms racing in the usual sense. Jackson and Morelli (2009) study a similar game but in which states choose arms levels in each period simultaneously before deciding whether to attack. Like Powell, they assume that arms do not accumulate across periods. The Powell and Jackson and Morelli models might thus be thought of as models of the decision of how many soldiers to employ each year, rather than models of the accumulation of military capital.²

¹Skaperdas (1992), Hirshleifer (1995) and others analyze "contest models" in which states simultaneously choose levels of arms, but there is no choice whether to attack. Instead, either war is assumed to occur for sure, or transfers are made automatically as a function of the balance of arms (perhaps representing an unmodelled, efficient bargaining process). I am not aware of papers in this literature that have studied a multi-period model with the possibility of accumulation of arms.

²One can imagine an "arms race" in the sense of escalation of the size of standing forces in two states, but this possibility does not arise in the Powell and Jackson and Morelli models. I believe the main reason is the assumption that the only constraint on how much a state can increase military spending above the previous period's levels is the

Brito and Intriligator (1985), Kydd (1997, 2000, 2005), and Slantchev (2005) consider models that have both arming and attack decisions, but with only one period (or round) of arming. As such they are not well-suited for exploring the dynamics or stability conditions of armament decisions. This is not the focus of these articles in any event. Instead, especially in Kydd's work and in Slantchev (2005), the main concern is whether and how arming decisions signal private information about a state's costs for war or preferences over taking territory, and the overall implications for the risk of war. These are important issues, at the heart of the cluster of informal arguments known as the "security dilemma," the "spiral model," and the "deterrence model," which are in turn at the heart of international relations theory (Jervis, 1976, 1978; Glaser, 2000, 2010). The model considered in this paper, which has complete information about states' types, does not get to considerations about whether and how arming might signal preference for peace or conquest (although I briefly consider a simple version with private information about types in section 5). It is, hopefully, a step towards a more thorough analysis of that problem.

I consider the simplest possible model that incorporates all three features mentioned above. Two states choose in successive periods whether to build a unit of arms and, after observing the arms decisions, whether to attack the other. In the first version, considered in the next section, an attack is successful and the loser is eliminated if the attacker has at least m > 1 times as many arms as the defender. I show that in a Markov Perfect Equilibrium, if the states put enough value on controlling the other's territory relative to the costs of building and if they are sufficiently patient, then they build up to the first level such that neither could attack and win by building one more time when the other does not. The equilibrium level of armament is increasing in the advantage for offense and in the amount they can build without a reaction by the other side.

These comparative statics have been suggested in informal treatments (e.g., Hoag 1961, Glaser 2000). The formal analysis characterizes in addition the "off the path" strategies that give rise to the "on the path" behavior. Here, we find that a somewhat complicated set of mixed strate-

total amount of resources available. Acemoglu et al. (2011) consider a dynamic model of resource competition in which one state decides how much to arm and whether to attack each period; arms do not accumulate.

gies obtain if either player deviates. Thus armed conflict occurs with positive probability off the path, and it turns out that deviations are more dangerous (in the sense of greater risk of conflict) the earlier they occur in the race.

In the third section I consider a richer version of the game in which the probability of winning a war can be a continuous function of current arms levels, although I continue to assume symmetric states (that is, they have the same preferences for conquest, costs for arming, and building capacities). This modification seems minor, but the new game proves more difficult to analyze. As before, the problem is non-stationary in that today's actions can change the situation tomorrow by changing the baseline military balance. Thus optimal behavior at arms levels (a_1, a_2) depends on what would happen at levels $(a_1+i, a_2+j), i, j \in \{0, 1\}$, which in turn depends on what would happen at higher levels of arms. With a natural diminishing returns assumption about the "contest success function" $p(a_1, a_2)$, it can be shown that if at (a_1, a_2) a state prefers attack to stability with no further building, then in any MPE the state attacks at that point. Along with the conjecture that an arms level is stable if neither state gains from building once and attacking, this result makes it possible to pin down MPE strategies for all other arms levels.

The main results for the equilibrium path are similar to the "threshold" case, though we get a richer set of comparative statics. There is a smallest level of arms such that if both states reach this point, neither wants to "break out" and attack. This first stable arms level is higher the more "greedy" the states are about taking new territory, the lower the costs of war, the more the offensedefense balance favors offense, and the worse their ability to monitor "break out." Comparative statics on the discount rate and the incremental cost of arms are indeterminate. Both states building up to the first stable point is the equilibrium path provided that the discounted costs of the build up are not so large that fighting with a one half chance of winning is preferable in the first period. Thus, immediate conflict to avoid the costs of an arms race to stability can also be the equilibrium outcome; this is more likely the more offense is favored and the more greedy the states (relative to the costs of war), both of which imply higher costs to arms race to stability.

Off the path, mixed strategies are less common than in the case of the threshold contest

success function, but can still appear.

In section 4, I provide a partial analysis of two asymmetric cases. In the first, one state is a "security seeker" that has no value for taking new territory (or has very high costs for war), while the other is a "greedy state" that values conquest per se.³ Markov perfect equilibrium now involves mixed strategies on the equilibrium path, as both states mix on "build" and "not build" until either (a) the security seeker arms and the greedy state does not, in which case the race ends peacefully, (b) the greedy state arms and the security seeker does not, in which case the greedy state attacks (or extorts), or (c) mutual building leads to a stable armament level. The intuition for mixed strategies is that the security seeker wants to build to deter attack or coercion by the greedy state, but the greedy state has no reason to build a large army if it will not be able to use it to gain anything. Comparative statics on when arms races are "dangerous" in the sense of generating a risk of war are then possible, although the results mainly suggest that war would not be a very likely outcome.

The other asymmetric case examined in section 4 involves one state that builds arms faster than the other. The results are similar in that there must be mixing on the equilibrium path (if the advantaged state is not so advantaged that it wants to attack right away). Here, however, both states may build for sure at first, so that the mixing, and thus a chance of war, start later in the race. Also, whereas in the "greedy state vs. security seeker" case the security seeker arms more in expectation, when the asymmetry is in building rates the advantaged state ends up with more weapons (or more effective weapons).

Section 5 informally interprets a number of the main results, contrasting them with related models, and briefly discusses some natural extensions: "carrying costs" for arms, and private information about values for changing the status quo relative to war costs.

³This specific language is due to Glaser (1992), but see also Jervis (1976).

2 An arming game with a 'threshold' military technology

There are two states, 1 and 2, that interact in successive periods t = 0, 1, 2, ... At the start of each period each state receives resources v > 0. They next choose simultaneously whether to "build" or "not build," where building costs $c \in (0, v)$. After observing these actions, they choose simultaneously whether to attack the other state. If neither state attacks then play proceeds to the next period.

Arms accumulate across periods and there is no depreciation, so that if state *i* has built a_i times and has not lost a war, it has a_i units of arms. Consumption occurs prior to the decisions about whether to attack, so that a state gets *v* in a period when it has not built, and v - c in a period when it has built.

If state *i* has at least m > 1 times as many arms as state *j*, then if there is a war state *i* takes over *j* and the strategic interaction ends. In this event *i* receives $v(1 + \beta)$ in all future periods, where $\beta \in (0, 1)$. β can be interpreted in three main ways: first, as the the share of the other state's resources that the winning side can appropriate; second, as the value that the winning state places on controlling the losing state, or changing its regime; and third, as a reflection of the costs of fighting (higher β means lower costs). The losing side receives 0 in all future periods. If a state attacks another when it does not have enough to win, suppose that it is taken over by the other state.⁴ Both players discount future payoffs by the factor $\delta \in (0, 1)$ per period.

The parameter m represents one aspect of what international relations scholars call the "offense-defense balance," and economists refer to as "decisiveness" when describing contest success functions.⁵ Military strategists talk about the force ratio needed for successful attack. So, for example, if m = 3 then the attacker needs a three-to-one advantage in force size to prevail. This

⁴If both states attack in a period when neither can win, assume that both are eliminated; this is for the simplicity and convenience of not having a continuation game after a state chooses attack. In the more continuous formulation for the chances of winning in section 3, war always results in one side being defeated if both attack.

⁵On the offense-defense balance, see for example Jervis (1978) and Glaser and Kaufmann (1998). The other main aspect of the offense-defense balance in typical informal discussions in political science is the advantage accruing to the side that strikes first; in the present model, there are no first-strike advantages.

would represent a situation in which defense has a moderate advantage, or in which conflict is not highly "decisive."

Let $a^t = (a_1, a_2)$ represent the arms level, or balance of forces, in period t, where a_i is the number of times state i has built prior to period t. A complete strategy in the stage game is an element of $\{0, 1\} \times \{0, 1\}^4$, where the first component represents the state's decision whether to build and the next four say whether the state would attack as a function of the four possible outcomes of the building choices.

I will focus on Markov Perfect Equilibria, defined as follows. Since each of the four "attack" subgames of the stage game starts a proper subgame of the whole game, it seems natural to require that the states make the same build decisions and the same attack decisions whenever arms level is the same. Let $\sigma_i(a^t) \in \{0, 1\}$ represent state *i*'s choice of whether to build or not whenever the arms level is a^t , and let $\tau_i(a^t) \in \{0, 1\}$ be *i*'s choice of whether to attack or not given arms level a^t . In an MPE, for each point $a' = (a'_1, a'_2)$, the choices $\sigma_i(a')$ and $\tau_i(a')$ maximize *i*'s discounted expected payoffs given $\sigma(a) = (\sigma_1(a), \sigma_2(a))$ and $\tau(a') = (\tau_1(a), \tau_2(a))$ for all other points $a \neq a'$. I will also restrict attention to equilibria in which states do not play weakly dominated strategies, which rules out mutual attack equilibria when one or both would be better off not attacking.

The simple form of the "threshold" military technology makes it is easy to pin down the states' attack decisions in any attack subgame.

Proposition 1. If $a_i \ge ma_j$, state *i* attacks, $\tau_i(a_1, a_2) = 1$. At all other points, *i* does not attack, $\tau_i(a_1, a_2) = 0$.

To prove this, note that when $a_i \ge ma_j$, *i* gets $v(1+\beta)/(1-\delta)$ by attacking, which is greater than any payoff *i* could possibly get by not attacking. Further, attacking is weakly dominated when $a_i < ma_j$, which rules out mutual attacks when neither can win.⁶ For the case of m = 1.5, Figure 1 illustrates the arms levels at which state 1 attacks by marking them in red.

The harder problem is to characterize Markov perfect building decisions, σ . Since everything is symmetric, we can focus on points where $a_1 \ge a_2$ without losing generality. It is immediate that

⁶In fact, these arguments imply that the claim holds in any subgame perfect equilibrium.

if the balance of forces is so skewed in state 1's favor that it can attack and win even if state 2 builds while state 1 does not, then it must be that neither state builds and state 1 then attacks. This is because state 2 is surely going to be attacked and defeated, so neither has any reason to waste current resources on another unit of arms. Such points are indicated in Figure 1 by the squares.

To make further progress, consider whether we can construct an MPE in which an arms level $a = (a_1, a_2)$ is *stable*, meaning that neither state builds, if neither state could attack successfully by building once when the other does not: thus $a_i + 1 < ma_j$ for both states. Such points are indicated with circles in Figure 1. Proposition 1 confirms that this is indeed possible; what remains is to specify strategies for arms levels that are neither stable nor that involve certain attack by the advantaged state. Unfortunately most such points prove to require mixed strategies on building and not building, and there can be a number of distinct equilibrium mixed strategy pairs for different arms levels.

Proposition 2. If the gains from attack exceed the costs of building once when victory is assured (that is, when $\delta\beta v > c(1 - \delta)$) and if $\delta v > c$, then the following is an MPE of the game described above. Consider (a_1, a_2) such that $a_1 \ge a_2$. (For $a_2 \ge a_1$, flip the subscripts in the conditions below.) State 1 attacks if and only if $a_1 \ge ma_2$. For $m \in (1, 2)$, building decisions are as follows:⁷

- *1.* If $a_1 + 1 < ma_2$ then neither state builds. Neither state attacks at these points.
- 2. If $a_1 = a_2$ and $a_1 + 1 \ge ma_2$ then both states build and neither then attacks.
- 3. If $ma_2 1 \le a_1 < ma_2 + \min\{0, m 2\}$, the states mix on build and not build with build probabilities

$$b_1^* = \frac{c(1-\delta)^2}{\delta v - c\delta(1-\delta)}$$
 and $b_2^* = 1 - \frac{c(1-\delta)}{\delta v\beta}$

Attack by state 1 follows if 1 builds and 2 does not; otherwise neither attacks.

4. If $ma_2 + m - 2 \le a_1 < ma_2$, then the states mix on build and not build. Let $w = v/(1 - \sqrt{7}$ For clarity I also give the attack decisions that follow in equilibrium at each point.

 δ) – c(2 – δ). Build probabilities are

$$b_1^* = \frac{w - \sqrt{w^2 - 4c^2(1 - \delta)}}{2\delta c}$$
 and $b_2^* = 1 - \frac{c(1 - \delta)}{\delta v \beta}$

Attack by state 1 follows if 1 builds and 2 does not; otherwise neither attacks.

- 5. If $ma_2 \le a_1 < ma_2 + m 1$, then state 2 builds and state 1 does not, and neither then attacks.
- 6. If $ma_2 + m 1 \le a_1 < ma_2 + \min\{m, 2(m-1)\}$, then the state mix with build probabilities

$$b_1^* = \frac{\delta v - c(1 - \delta^2)}{\delta v - c\delta(1 - \delta)}$$
 and $b_2^* = \frac{c(1 - \delta)}{\delta \beta v}$.

Neither attacks if 2 built and 1 did not; otherwise state 1 attacks.

7. If $ma_2 + 2(m-1) \le a_1 < ma_2 + m$, then the states mix. Let $V_2^{case4} = v - c + \delta(1 - b_1^*)v/(1-\delta)$ be 2's continuation payoff in case 4 above, using the b_1^* from that case. Then the build probabilities in this case are

$$b_1^* = 1 - \frac{c}{\delta V_2^{case4}}$$
 and $b_2^* = \frac{c(1-\delta)}{\delta \beta v}$.

Neither attacks if 2 built and 1 did not; otherwise state 1 attacks.

- 8. If $a_1 \ge m(a_2 + 1)$, then neither builds and state 1 then attacks.
- 9. If $a_1 = a_2 + 1$ and $a_1 \ge ma_2$, then the states mix. Let a^* be the smallest integer strictly greater than 1/(m-1), and let $z = a^* a^1$. The build probabilities are

$$b_1^* = \frac{\delta v - c(1 - \delta^{z+1})}{\delta v - c\delta(1 - \delta^z)}$$
 and $b_2^* = \frac{c(1 - \delta)}{\delta \beta v - c(1 - \delta^z)}$

Neither attacks if 2 built and 1 did not; otherwise state 1 attacks.

When m > 2, then only cases 1, 2, 3, 5, 6, and 8 apply.

The proposition implies that on the equilibrium path, both states build in each period until they reach the first point at which neither would be able to attack and win if it built one more time while the other did not. At (0,0), case 2 is always satisfied so both states always build, and will

continue building up to the smallest level a^* such that $a^* + 1 < ma^*$. So the long-run equilibrium arms level a^* is the smallest integer greater than 1/(m-1).

The greater the advantage to defense – or equivalently the smaller the "jump" one can get before the other side can observe that you are building – the less time it takes to reach this point. For example, if m = 2, then both states build in the first two periods up to $a^3 = (2, 2)$, at which point one additional unit does not provide enough an advantage to make successful attack possible (3/2 < 2). By contrast, if m = 3, the states build in the first period and then quit, since building one more unit when the other does not gives only a two-to-one advantage (2/1 < 3).

As m approaches 1 from above, the equilibrium stable arms level increases, and in the limit, at m = 1, the equilibrium has both states building in every period and never attacking on the equilibrium path – a permanent arms race. Obviously, states' payoffs are increasing in the size of the defensive advantage m. At m = 1 the average per-period payoff is v - c while for m > 2 it is $v - c(1 - \delta)$.

The model's equilibrium thus captures the usual "if you want peace, prepare for war" maxim, and also generates natural comparative statics in m on the equilibrium stable arms level. A surprising feature of equilibrium in the game, however, is that off the equilibrium path there are force levels such that the states play mixed strategies and there is a risk of war if the advantaged state happens to build (and, in cases 3 and 4, if the disadvantaged state does not).

To understand why mixed strategies must arise in equilibrium, consider the case of m = 1.5illustrated in Figure 1, at $a^t = (5, 4)$. State 1 cannot attack successfully at this point, but if it built one more time without a response from state 2, it could. State 2's best reply if it expected this would be to build, which would lead to (6, 5), where state 1 could not attack and win. And if state 1 expects state 2 to build, then its best reply is not to build in period t. Finally, if state 1's Markov strategy is to not build at (5, 4), then state 2 would have no reason to build to prevent attack. So no pair of pure actions in stage games with $a^t = (5, 4)$ are mutual best replies. Equilibrium requires mixing, such that state 1 is indifferent between not building which yields $v/(1 - \delta)$, and building which produces a chance of being able to attack successfully if state 2 happens not to build. Likewise, state 2 must be indifferent between not building, which saves c today but may lead to elimination, and building, which assures a payoff of $v - c + \delta v/(1 - \delta)$. If the discount factor is close to 1 or if the cost of building is small, then the equilibrium mixed strategies have it that the advantaged state almost surely does not build and the disadvantaged state almost certainly builds, thus leading to a high probability of return to (or towards) equal arms levels. Pure and mixed strategy build probabilities are illustrated graphically with vectors in Figure 2.⁸

A different sort of mix arises at certain "attack points" where the balance of forces is not so extreme that state 1 can attack and win even if 2 builds. Consider m = 1.5 and $a^t = (5,3)$, so that 1 could attack and win at these arms levels. To reach this "off path" point state 1 must have built twice without attacking from (3,3), without a reply form state 2.⁹ At (5,3), state 1's best reply if state 2 does not build is to not build and then attack. State 2's best reply to state 1 not building, however, is to build to get to (5,4), which prevents attack. And state 1's best reply to state 2 building is to build and attack, which is possible at (6,4). And to this state 2's best reply would be to not build and let itself be attacked. Again equilibrium is in mixed strategies. In this case, when the discount factor is close to 1 or when the cost of building is small, the advantaged state almost certainly builds and the disadvantaged state probably does not, leading to probable elimination of the weaker state. The case 7 and 9 mixes are closely related, differing because state 2's continuation payoffs involve more expected building costs.

If there are mixed strategies off the equilibrium path where the advantaged state may be able to attack and win, why don't states have an incentive to build to get to such a position? In brief, a state can assure that it won't be attacked by keeping a_i large enough, and it will always want to do this if it is patient enough (large enough discount factor). Thus, in the first sort of mixed

⁸A similar but slightly different mix can occur at points like (4,3) in this case of m = 1.5; the difference is that here, if both build state 2 has to build again to ensure peace and stability, whereas both building leads to a stable point from (5, 4).

⁹Alternatively, we can analyze the game under the assumption that the states begin with initial levels of arms $a^0 = (a_1^0, a_2^0)$, and ask about what would happen in the MPE. For instance, if the game "begins" with state 1 having a 5-to-3 advantage and m = 1.5, then in the first period there is a positive chance of war as given by case 6 in Proposition 2.

strategy subgames described above, the advantaged player's expected payoff cannot be better than $v/(1 - \delta)$, which is also the expected payoff once the equilibrium stable arms level is reached. So there is no incentive to pay a cost of c to get to a point where one's expected payoff is the same as it was before, even if a chance of victory figures into that expected payoff.

As noted, Figures 1 and 2 illustrate the MPE in this simple arming game for a case where m = 1.5 and for $a_1 \ge a_2$ (the results are symmetric for $a_2 > a_1$). In this example, if the game starts at (0,0) then on the equilibrium path both build up to (3,3) and then stop. Off the path, failing to build in any of the first three periods, during the arms race, leads to being attacked (or coerced) for sure by the other player. If the states "start" from an arms level such as (1,0) or (2,1), war (or coercion) is highly likely but not certain due to the mixed strategies. At the higher level (3,2), stability at (3,3) is very likely to follow but there is a small chance that state 1 will build, 2 will not, and war will occur. Small deviations from equality at higher arms levels are likely to be self-correcting, while larger deviations make for conflict.

Proposition 1 required two conditions. The first, $\delta\beta v > c(1 - \delta)$, implies that the states are "greedy" enough about taking new territory relative to the costs of building. If this inequality goes the other way, then (0,0) is a stable point, and the unique efficient outcome in an MPE. So, in a somewhat reductive fashion, we have the additional comparative static that states arm more the more weight they put on conquest, the lower their opportunity costs for arming, and the more they care about future payoffs.

The second condition, $\delta v > c$, plays a more subtle role: When this is the case, the states would prefer to build forever rather than to not build and be attacked in the first period $((v - c)/(1 - \delta) > v)$. This condition ensures that in the first period, when neither has any arms, it is preferable to "race" to the stable point rather than not arm and allow oneself to be attacked. Without this condition, it is possible that the anticipated costs of the arms race will make it not worth running, so that in a pure strategy MPE one state builds and the other does not in the first period, and the non-builder is attacked and absorbed. There will be a symmetric mixed strategy MPE as well, in which both states mix on whether to build. Either way, of course, war occurs with positive probability.¹⁰

 $\delta v > c$ is a sufficient but not necessary condition, since being willing to race forever rather than not build and be attacked is stronger than needed. The necessary and sufficient condition for a MPE to imply that the states arms race to a stable point, with no chance of war on the equilibrium path, is $\delta v > c(1 - \delta^{a^*})$, where a^* is the first stable point, the smallest integer greater than 1/(m-1). More comparative statics follow. In particular, greater offense dominance (smaller m) increases the likelihood that equilibrium will entail a positive risk of war, since greater offense dominance increases the amount states have to build to get to a stable point, making war early on more attractive. This result appears again, in a somewhat more realistic way, in the model considered in the section.

3 A more general version

In this section I consider a more general version of the model, now allowing that if a state attacks, victory or defeat is not certain and the probabilities depend on the balance of forces. Even though we will keep the assumption of symmetry in this section, the analysis turns out to be more involved, as there are more cases off the equilibrium path. The richer set up is worth considering for three main reasons. First, it provides a richer set of comparative statics. Second, one may wonder how much the mixed strategy results above are due to the threshold contest success function. And third, it is a necessary bridge towards the asymmetric case where states can differ in various ways.

Let p(x, y) be the probability that the state with $x \in \mathbb{R}$ units of arms wins if there is a war, with p strictly increasing in x and strictly decreasing in y. Symmetry is implied: p(x, y) = 1 - p(y, x).

It will be useful to let p_{mn} represent state 1's probability of winning when 1 and 2's building choices were m and n, where $m, n \in \{0, 1\}$ and 1 means "build." For example, $p_{10} = p(a_1 + 1, a_2)$ and $p_{11} = p(a_1 + 1, a_2 + 1)$. Note that p_{mn} is always a function of the arms level (a_1, a_2) .

¹⁰This observation also implies that for some parameter values, the game has multiple MPEs.

I will make the additional assumption that when both states add equal amounts to their forces – so moving the force ratio towards equality – the probability that the stronger would win a war moves towards one half. Formally,

A1: For x > y, p(x + a, y + a) is strictly decreasing in a > 0.

This is satisfied for all ratio form contest success functions, which are defined by $p(a_1, a_2) = a_1^n/(a_1^n + a_2^n)$ where n is a parameter greater than zero (Hirshleifer, 1989). Higher values of n imply greater "decisiveness," or "offensive advantage," in military contests, in the sense that a one unit increase from equal arms translates to a larger increase in the odds of winning.¹¹

In the attack subgames, if either or both states attack, a war occurs that state *i* wins with probability $p(a_i, a_j)$, with the losing state eliminated, so state *i*'s time-averaged payoff for war when the balance of forces is (a_i, a_j) is $p(a_i, a_j)\delta v(1 + \beta)$. Once again, Markov strategies are functions $\sigma_i(a) \in \{0, 1\}$ and $\tau_i(a) \in \{0, 1\}$, i = 1, 2, that describe the states' build and attack decisions whenever the arms levels are $a = (a_1, a_2)$.

Taking a similar approach to the last section, consider first whether there are arms levels so imbalanced that, in an attack subgame, the advantaged state (say, state 1) would surely want to attack. Attacking is better for state 1 than subsequent stability when $p(a_1, a_2)\delta v(1 + \beta) > \delta v$, thus at points a such that

$$p(a) > p_l \equiv \frac{1}{1+\beta}.$$
(1)

However, contrary to the case of the threshold contest function in section 2, it is now hypothetically possible that state 1 would want to delay in hopes of attacking with a better chance of winning later on. Intuitively one might think this unlikely, since state 2 would have an incentive to build to prevent being attacked at worse odds. And in fact, for military technologies that satisfy A1, we can show that this is the case and so in an MPE state 1 attacks in attack subgames that satisfy (1). **Proposition 3**: For $a_i > a_j$ and given A1, in any MPE state i attacks in any attack subgame if

¹¹A1 is not satisfied for so-called "difference" contest success functions, for which the probability of winning depends only on the difference between the combatants forces rather than the ratio (Hirshleifer, 1989). Although there may be contexts where this conflict technology makes sense, I find it hard to imagine it applying in the case of fights between armed groups.

arms levels are such that $p(a_i, a_j) > 1/(1 + \beta)$.

Figure 3 illustrates "attack points" that satisfy (1) with red dots. Define an *internal* attack point as an attack point (a_1, a_2) such that $(a_1, a_2 + 1)$ is also an attack point. War must occur following such points in any MPE; the only question is who builds. Under A1, there prove to be three possibilities, which are illustrated in Figure 3.

Proposition 4. Let $C = c(1-\delta)/\delta v(1+\beta)$, which is the cost-benefit ratio such that if war is going to occur, a state prefers to build if its gain in probability of winning is greater than C. For $a_1 > a_2$ and given A1, at an internal attack point build strategies are as follows:

- 1. If $p_{00} p_{01} < C$, neither builds and 1 attacks.
- 2. If $p_{00} p_{01} > C > p_{11} p_{01}$, then 2 builds, 1 does not, and 1 attacks.
- 3. If $p_{11} p_{01} > C$, then both build and 1 attacks.

Al ensures that these are exclusive cases.

As in section 2 above, an arms level is *stable* if neither state builds or attacks at this point in equilibrium. Following the earlier approach, the next step is to conjecture that there is an MPE in which a is stable if it is not an attack point and if neither state would prefer to deviate to build and attack. For $a_1 \ge a_2$, the latter constraint is $v \ge (v - c)(1 - \delta) + \delta p_{10}v(1 + \beta)$, or

$$p(a_1 + 1, a_2) \le p_h \equiv \frac{\delta v + c(1 - \delta)}{\delta v(1 + \beta)} = p_l + C.$$
 (2)

In the case of ratio military technologies, it is easy to show that points with $a_1 \ge a_2$ are stable in this sense if

$$a_2 \ge \left(\frac{1-p_h}{p_h}\right)^{\frac{1}{n}} (a_1+1) \text{ and}$$
 (3)

$$a_2 \ge \left(\frac{1-p_l}{p_l}\right)^{\frac{1}{n}} a_1 = \beta^{\frac{1}{n}} a_1.$$
 (4)

Figure 3 illustrates the set of stable points for a simple ratio form case,¹² with the two lines defined

¹²That is, n = 1 so that $p(a_1, a_2) = a_1/(a_1 + a_2)$. (Let p(0, 0) = 1/2.) The other parameters are c/v = .5, $\beta = .8$, $\delta = .9$.

by (3) and (4) also shown. Since $p_l < p_h$, line (4) has higher slope and eventually becomes the binding constraint. It is evident that the first stable point is the smallest integer a^* such that $p(a^* + 1, a^*) \le p_h$. With a ratio form contest success function, this implies that

$$a^* = \lceil rac{1}{\left(rac{p_h}{1-p_h}
ight)^{1/n}-1}
ceil,$$

where $\lceil x \rceil$ is the smallest integer greater than or equal to x.

With ratio-form contest functions, a first stable point a^* always exists. For the more general case of $p(a_1, a_2)$ that satisfy A1, an additional assumption is needed to guarantee this. Namely, for large enough arms levels, we assume that adding one more "tank" has almost no effect on the chance of winning, which seems reasonable.

A2: p(a + 1, a) approaches one half as a gets large.

To find an MPE with these characteristics it remains to fill in strategies for the arms levels that are neither stable nor internal attack points, and to check that the whole assembly works. We begin with the equilibrium path following (0, 0). In effect, the states have a choice between both building up to the first stable point a^* , which will yield the time-averaged payoff $v - c(1 - \delta a^*)$, or both building and then attacking straight away, which gives $(1 - \delta)(v - c) + \delta v(1 + \beta)/2$. For large enough a^* , getting to stability may be so costly that the states prefer to fight rather than arms race. The condition is given in Proposition 5.

Proposition 5. Suppose A1 and A2 hold and that $\delta v > c$. There is a first stable point a^* . If $c(1 - \delta^{a^*-1}) > v(1 - \beta)/2$, then an MPE with the characteristics given so far must involve both states arming and at least one state attacking at a = (0, 0). Else both states arm but do not attack for points $a_1 = a_2 < a^*$.

Thus, the *prospect* of a burdensome arms race, even if it would lead to a stable situation, may be dangerous by making a fight in the present more attractive. Conflict is still inefficient, of course. If the states could commit not to build and attack at a level of arms less than a^* , they would not want to fight initially and both would be better off.¹³

¹³The mechanism is parallel to that in Powell's (1993) model, where war may be the only MPE when the level of

The comparative statics on the likelihood that (a prospective) MPE involves war at the outset are as follows.

- Greater advantage for offense in the offense-defense balance increases the risk, since this increases the amount of building needed to get to a stable level of arms.
- The more value the states put on controlling the other's territory, or the lower the costs of war (higher β), the greater the war risk. This increases both the temptation to fight now *and* the amount of building needed to get to stability.
- Smaller costs for building arms relative to national income (and for a given "break out" time) have no clear impact, since this increases the building time to stability, but also lowers the cost.
- More patient states may be more or less likely to fight, as they value more the peace they would eventually reach by building, but will have to build more to get there.

Figure 3 shows the build decisions ($\sigma_i(a_1, a_2)$) for the sets of arms levels that have been pinned down so far. A full characterization of an MPE requires that we specify σ for the remaining points (and for any admissable set of parameters). This can be done case by case; results for this particular set of parameters are illustrated in Figure 4. Here, if the states were to start from (0,0), they would build three times and then stop at (3,3). Mixed strategies are still possible off the path (for example, at (5,4)), and with other parameters a variety of other mixed strategy points are possible.

Figure 5 gives (almost all of) the MPE when we increase the offense-defense parameter n in the contest function to n = 2, which increases the marginal advantage gained from building. Now if the states were to "begin" from equal arms levels less than or equal to (3, 3), they prefer fighting at even odds to building up to the first stable point, (6, 6). If the initial levels are (4, 4) or (5, 5), then the building to stability at (6, 6) is preferred.

arms spending required for stability is so high that it's better to just fight.

The complete characterization of MPE for contest functions $p(a_1, a_2)$ that satisfy A1 is rendered particularly tedious by points like (10, 9) in Figure 5. Here, state 2's continuation payoff if both build depends on state 1's mix at (11, 10), which leads to a quadratic for state 1's equilibrium mix at (10, 9). This is tractable, but for other parameter values there can be a long succession of such points, with increasingly complicated recursions for each one.

4 Arms races when states differ in aims or capacity to build

So far the two states have been assumed to have the same value for acquiring new territory, the same arms budgets, the same costs for building and for war, and the same capabilities to build arms. In this section I provide a partial analysis of the more interesting situation where the states differ in either preferences or capabilities. Call a state "status quo" if it has no value at all for acquiring new territory ($\beta_i = 0$), and "greedy" if it has value $\beta_i > 0$ (Glaser, 2010). This case is analyzed next. Mixed strategies now appear on the equilibrium path, with the status quo state typically more likely to build. War may occur if the greedy state happens to build when the status quo state does not, although for almost any reasonable-seeming values of the parameters, this is quite unlikely. The more likely outcome is that there is a brief arms race that ends peacefully when the status quo state builds and the greedy state does not. One implication is that the expected equilibrium arms level of the status quo state is higher than that of the greedy state.

The next subsection considers a different asymmetry: Suppose that one state can build weapons faster, or can build more effective weapons, than the other. The results are broadly similar. If the advantaged state is not so advantaged that it wants to build just once and attack at the outset, then the states arm for sure up to the first level where the disadvantaged state is not tempted to "defect," and then start mixing as in the previous case. The race may end in either conflict or stability, with the militarily advantaged state having more arms.

These analyses are suggestive but partial, in that there are a number of natural ways that two states may differ on the main parameters considered in the model, and these may be related. For instance, the most natural way of interpreting the game's timing empirically is to suppose that building takes place in continuous time, but a state cannot observe (or react to) what the other state has been doing for the last $\Delta > 0$ amount of time. Δ thus corresponds to what is sometimes called the "breakout" capability or time in the informal arms race literature. It depends on, among other things, the technology for monitoring and responding to military acquisitions and deployments. In time Δ , state *i* can build, say, $m_i \Delta$ units of arms at cost $c_i \Delta$. The discount factor per period is $\delta = e^{-r\Delta}$ where r > 0 is the discount rate. A natural assumption would be that m_i is proportional to the size of state *i*'s resource flow, v_i , in which case $v_i > v_j$ implies that *i* has an advantage in terms of building arms but is also a more attractive target for a greedy state *j*. This would roughly correspond to a common (U.S.-based) characterization of the U.S.-Soviet arms race (economically stronger but status quo U.S. versus possibly greedy U.S.S.R) and to the India-Pakistan rivarly (economically stronger India versus territorially dissatisfied Pakistan). [no analysis of such cases yet.]

4.1 A greedy state and a status quo state

Let state 1 be greedy, with $\beta_1 > 0$ value for territory, while state 2 has $\beta_2 = 0$ indicating either zero value for controlling new territory or prohibitive costs for war. Otherwise, use the model from the last section.

The only reason for the security-seeker to arm is to dissuade attack by the greedy state. But if the security seeker would certainly arm to prevent attack, then the greedy state has no reason to build a large and costly army.¹⁴ These observations immediately suggest why equilibrium in this case must involve mixed strategies on the equilibrium path.

Proposition 6. Suppose the military technology is ratio form. In the case of a greedy state with $\beta_1 = \beta > 0$ and a security seeker with $\beta_2 = 0$, if the states are sufficiently patient then on the equilibrium path both states mix in each period until one of the following things happens. (a)

¹⁴In this very "third image" model, at any rate. Empirically, it could be that greedy states are greedy for domestic political reasons that at the same time incline them to build large militaries, irrespective of the possibility of conquest. Militarism may be domestically useful for rulers even if the military is not used (e.g., Snyder (1991))

The status quo state builds while the greedy state does not. This leads to a stable point. (b) The greedy state builds and the status quo state does not. If this occurs then the greedy state attacks. (c) Mutual building leads to the stable point (a^*, a^*) , where a^* is the first arms level such that $p(a^* + 1, a^*) \leq p_h$.

With a greedy state and security seeker, then, the length and results of an arms race are stochastic. It can end in peace or war, and if it ends in peace at a stable arms level, this level might be any of the following: $(0, 1), (1, 2), \ldots, (a^* - 1, a^*), (a^*, a^*)$. An immediate implication is that conditional on stability, the status quo state arms at least as much or more than the greedy state. This is because the greedy state has no need to deter attack or coercion by the status quo state, while the status quo state does need to deter the greedy state.

The probability distribution on the various possible outcomes of course depends on the specific mixed strategies, and unfortunately these are difficult to obtain in closed form for the greedy state except for the cases of $a^* = 1$ and $a^* = 2$. The problem is that the mix on building and not building by the greedy state that makes the status quo state indifferent in a period must be derived by working backwards from the point $(a^* - 1, a^* - 1)$, and the mix in one period depends on the continuation value if both build, creating a complicated recursion. It is, however, straightforward to show that the mix by the status quo state that renders the greedy state indifferent between building and not building at point $(a, a), a = 0, 1, \dots, a^* - 1$, is

$$b_2^a = 1 - \frac{c(1-\delta)}{\delta v[p(a+1,a)(1+\beta)-1]}.$$
¹⁵

Thus, the probability that the status quo builds is highest at the start of the race and decreases thereafter. This is because the greedy state's temptation to build and attack is diminishing as arms increase, which in effect somewhat reduces the pressure on state 2 to build. It is also evident that as δ approaches one, the security seeker becomes virtually certain to build at any level of arms (less than a^*), so that the probability of war must approach zero.

¹⁵This comes from setting what the greedy state can always get by never building, v, equal to its expected payoff for building, which is $(v - c)(1 - \delta) + \delta(1 - b_2^a)p_{10}(a)v(1 + \beta) + \delta b_2^a v$.

As noted, there is no such simple result for the greedy state's equilibrium mixture on building and not building. For the case of $a^* = 2$ (i.e., stability if both build twice), it can be shown that at (1, 1) state 1 builds with probability

$$b_1^1 = \frac{c(1-\delta)^2}{\delta[v - c(1-\delta) - \delta(1-p(2,1))]}$$

while in the first period b_1^0 is the positive root of the quadratic $b_1^2 + Bb_1 - C$, where $B = 1 + v/c\delta(1-\delta)$ and $C = (1-\delta)/\delta^2$.

Table 1 illustrates the likelihood of different arm race outcomes for five different sets of parameters. The first two columns show the probability of four different outcomes in a given period of play at points (0, 0 and (1, 1)). For example, in the first scenario the probability that 2 builds and 1 does not ("peace") is about .879 in a given period at (0, 0), and is .541 conditional on being at (1, 1). Since there is a positive probability that neither will build at each of these points ("sq") and so the situation will be repeated, the "total probability" of a move from, say, (0, 0) to (0, 1) is larger. These "total" probabilities (conditional on being at a point) are shown in columns three and four. The final column shows the probability distribution on outcomes in the game.

The five different scenarios sequentially vary one parameter at a time (bolded), always subject to the constraint that the parameter set implies that $a^* = 2$. Observe that in all cases, at least for these parameters, by far the most likely outcome is peace after the status quo state arms and the greedy state does not. For small discount factors and/or greater greed it is possible to get somewhat lower probabilities of this outcome, but not by much.

A second interesting result is that while mutual escalation to (1, 1) is not very likely, if it happens then war might be quite likely to occur. Note that the total probability of war is always greater from (1, 1) than from (0, 0), so at least in the case of $a^* = 2$ it appears that these greedy/security-seeker arms races become more dangerous as they proceed (albeit not especially dangerous, from an ex ante perspective). The intuition is that, in effect, the security seeker is more likely to err on the side of hoping that deterrence will hold at higher arms levels, because war is less bad if it

happens to underestimate the greedy state's aggression.¹⁶

Some of the other examples show surprising comparative statics. Reducing the greedy state's value for conquest can increase the risk of war; reducing the costs of building can lower it; increasing offensive advantage can also decrease war risk. A partial intuition is that (for instance) greater defensive advantage makes war less attractive for the greedy state, which makes the security seeker more willing to risk attack by not building. Note, however, that at least some of these are "local" relationships that will not hold if a^* changes. For example, for β small enough peace at (0, 0) is certain, so eventually reducing the greedy state's "greed" must reduce war risk.

The example considered here is extreme in that state 2 was assumed to be a pure securityseeker that has zero value for taking territory (or other international issues where it could potentially gain by military coercion). What would happen if both states are at least somewhat greedy, but one is greedier than the other? Say $1 > \beta_1 > \beta_2 > 0$. Then the first arms level at which state 2 would be willing to stop and remain at peace – call this \hat{a}_2 – is (weakly) smaller than the first such arms level for state 1 (\hat{a}_1). And \hat{a}_2 can be greater than zero if state 2 is somewhat greedy. If the costs of arming are not so high that one or both prefer war at the outset, then both will build for sure up to \hat{a}_2 , and then mixed strategies will begin. In such cases, an arms race that starts from a low level becomes more "dangerous" in the sense of having higher risks of conflict farther into the race.¹⁷

4.2 One state builds faster than the other

Now assume two greedy states with values for expansion $\beta_1 = \beta_2 = \beta > 0$, but state 1 produces m > 1 units of arms each period it builds whereas state 2 produces just one unit.¹⁸ Let state 1's probability of winning if war occurs at point (a_1, a_2) be $p(ma_1, a_2)$. With the ratio form contest function, state 1's odds of winning at (a_1, a_2) are $(ma_1/a_2)^n$, which implies that if both have built

¹⁶More on this "purification" interpretation of mixed strategies and conflict in the model in the next section.

¹⁷Although I think it is not necessarily the case that once the mixed strategies start, the risk of war increases with the arms levels.

¹⁸This section is more preliminary than the others ...

Table 1: Outcome prob's for a greedy vs security-seeker arms race

	Period probs $(0, 0)$ $(1, 1)$		Total probs $(0, 0)$ $(1, 1)$		Outcome probs	
				$\frac{(1,1)}{5, \delta = .85}$, n = 1	
peace	.87864	.541	.98605	.94433	peace at 0,1	.98605
war	.00137	.01407	.00154	.02456	peace at 1,2	.01172
esc	.01106	.01782	.01241	.03111	war at 1,0	.00154
sq	.10892	.42711			war at 2,1	3e-04
					peace at 2,2	.00039
	ı	$\gamma = 1, \beta =$	= .8, <i>c</i> = .	5, $\delta = .8$, n = 1	
peace	.82425	.35268	.97273	.8556	peace at 0,1	.97273
war	.00361	.0372	.00426	.09025	peace at 1,2	.01969
esc	.0195	.02232	.02301	.05415	war at 1,0	.00426
sq	.15264	.5878			war at 2,1	.00208
•					peace at 2,2	.00125
		1.0	_		1	
n 00000	.80245	$\frac{y = 1, \beta}{.059}$	=.7, c = .97201	$\frac{.5, \delta = .8}{.51318}$	$\frac{n}{n} = 1$ peace at 0,1	.97201
peace	.00413	.039 .05247	.005	.45639	peace at 0,1 peace at 1,2	.97201
war	.00413	.003247	.003	.03043	war at 1,0	.001
esc	.01898	.88503	.02299	.03045	war at $1,0$ war at $2,1$.005
sq	.1/444	.88505			peace at $2,1$.01049 7e-04
					peace at 2,2	10 01
				. 4, $\delta = .8$		
peace	.84105	.23929	.97816	.8481	peace at 0,1	.97816
war	.00268	.03214	.00312	.11392	peace at 1,2	.01587
esc	.01609	.01071	.01872	.03797	war at 1,0	.00312
sq	.14017	.71786			war at 2,1	.00213
					peace at 2,2	.00071
	v	$= 1, \beta =$.7, c = .4	$, \delta = .8, r$	n = 1.4	
peace	.84105	.54963	.97816	.93755	peace at 0,1	.97816
war	.00268	.01572	.00312	.02682	peace at 1,2	.01755
esc	.01609	.02089	.01872	.03563	war at 1,0	.00312
sq	.14017	.41375			war at 2,1	5e-04
					peace at 2,2	.00067

a times, 1's odds of victory are m^n . (To recall, n is the "decisiveness" parameter in the contest function.)

With the ratio form, because both arming at the same rate preserves the ratio of forces, state 1's probability of winning stays constant as both build. Intuitively, then, if state 1's advantage is

large enough that it prefers attacking at odds m^n to peace and stability – that is, when $m^n > 1/\beta$ – then an MPE will certainly involve both building and state 1 attacking in the first period. Racing is costly and brings no advantage if state 2 would match state 1, which it would if it expects to be attacked.

If, on the other hand, $m^n < 1/\beta$ then there is an arms level \hat{a}_1 large enough that if reached, state 1 would not find it worthwhile to "defect" by building and attacking. There is an analogous level, \hat{a}_2 , for state 2, and it is easy to show that $\hat{a}_2 \leq \hat{a}_1$.¹⁹ So the point (\hat{a}_1, \hat{a}_1) is certainly stable. Now consider $(\hat{a}_1 - 1, \hat{a}_1 - 1)$. Here we have the same kind of "matching pennies" dynamic seen in the case of a greedy state and security seeker. If state 1 builds for sure then state 2 wants to build to prevent attack, but if state 2 builds for sure then state 1 will get stability whether it builds or not and so prefers to not build so as to save on costs. But if state 1 never builds at this point, then state 2 would have no reason to build either, to which building and attacking is a best reply for state 1.

So the results for asymmetric building rates appear to be similar to those for the "greedy state vs. security seeker" case, in that the equilibrium path will involve mixed strategies that could lead to war or coercion if the asymmetry is not so large that war occurs right away. There is a difference, however, in that here, in an arms race, both will build for sure at the outset and mixing will start farther into the race. This is because we have assumed some "greed" on the part of state 2, so that it would want to build and attack at (0, 0) if state 1 was not building.

5 Discussion

5.1 If you want peace, prepare for war?

At least for the case of symmetric states, the model and analysis here are largely consistent with the Roman adage "if you want peace, prepare for war." The reason states build is to deter attack (or coercion) by other states that would like to expand or otherwise change the status quo away from what the state likes. Arms build ups are costly and inefficient. The states could do better if they

¹⁹Specifically, $\hat{a}_1 = \lceil ((o_h)^{1/n}/m - 1)^{-1}) \rceil$, and $\hat{a}_2 = \lceil (m(o_h)^{1/n} - 1)^{-1}) \rceil$, where $o_h = p_h/(1 - p_h)$.

could commit themselves not to arm, or not to use arms for attack or coercion. Either possibility would enable a stable outcome at zero or lower levels of arms.

In the symmetric case, however, even if arms build ups are costly they are not themselves "dangerous." They do not raise the probability of war by any sort of dynamic internal to the build up itself. Indeed, it is *failing* to race that would be dangerous, if states start from or newly find themselves in a position where current arms are inadequate for deterrence.

Contrary to the Roman adage, or going beyond it, the analysis does make clear that while build ups in the symmetric case are not in themselves dangerous, *the anticipation of a costly arms race* could be dangerous, by inclining states to fight in the present rather bear the costs of arming to a stable level. In this complete information model this is "all or nothing" in the sense that either the states fight at the outset or they do not. Note, however, that the mechanism is more general in that the greater the anticipated costs of getting to stability, the more willing the states would be to run risks of armed conflict in crisis bargaining over other issues. This is a potential source of pressure for war that is almost completely missed in standard Realist discussions.²⁰

5.2 Arms control?

Contrary to typical Prisoners' Dilemma models of arms races, there is no role here for arms control to prevent costly arms build ups. Because they do not explicitly incorporate an option to use the arms for attack or coercion, PD models implicitly assume that the players are always around to punish a state that defects from an agreement, by returning to arming or competition on the same terms. When we explicitly model the possibility of using the weapons, it is immediately clear that a state that is the "sucker" may be permanently disadvantaged – here, eliminated in a war – so that lower, more efficient arms levels cannot be sustained by the conditional retaliation mechanism available in a standard repeated PD (such as Tit for Tat).

Put differently, in the repeated PD understanding of the arms problem, "arms control" is

²⁰Section 3 of Fearon (2008) shows that in a model where states choose arms levels and then may "crisis bargain" with private information about costs for fighting, greater arms burdens (and the things that cause them, like offensive advantages) lead the states to run more war risk in bargaining, by making more aggressive offers.

talking to facilitate coordination on relatively efficient equilibria that involve lower levels of arms than do other equilibria. But when we allow for actual use of the weapons that may be acquired, we find that there is a minimum stable level of arms for given parameters and this will be reached if the states start from below this level. No arms levels below this are equilibria, so there is nothing better to coordinate on via arms control negotiations.

In the basic model analyzed above, there are no "carrying costs" for arms. That is, once they are built, there are no continuing costs for maintaining them. An implication is that in this model, there is also no role for arms control talks to coordinate *reductions* in arms from a high level, if an exogenous shock were to lower one or both states' value for taking territory, increase the costs of war, or create new advantages for defense.

One can imagine, however, that if there were carrying costs there might be a role for arms control talks for facilitating coordination on lower levels of arms if, due to a change in (say) one or both states territorial aspirations, a lower level became a stable point. Consider for example arms level (4, 4) in Figure 4, and suppose that this was the first stable point until some time when there was a change in the military technology or both states' values for taking territory that reduced the first stable point to (3, 3). Because (4, 3) and (3, 4) are both attack points, neither state can safely cut its arms level unilaterally – the other side would then want to attack it. If, however, there is a possibility of a coordinated reduction to (3, 3), this is a stable point that both prefer. So in this specific situation, or in any attempt to coordinate on (3, 3) from higher levels, "arms control" in the received sense could play a role.

Note, though, that there are other moves from high to lower levels of arms that do not require coordination – for instance, if there are small carrying costs for arms levels, then *unilateral* disarmament moves would take the states from any stable points above (4, 4) down to (4, 4).²¹ So the results suggest a somewhat nuanced take on arms control, if arms control is understood in the sense of coordination on Pareto-superior equilibra at lower arms levels.

²¹Carrying costs introduce some new issues, however; see the discussion below. The claims in the text here are surely valid, however, for small enough carrying costs.

To the extent that these objections to the standard PD story about arms control as a means to prevent arms build ups, other interpretations of arms control may gain in plausibility. In particular, arms control might have more to do with signaling and screening among types in a incomplete information problem, where states' levels of "greed" β_i are private information (Kydd, 2001). Alternatively, arms control might be more about facilitating coordination on monitoring arrangements that affect underlying parameters in the model studied here, concerning the ability to "break out" without detection.

5.3 Interpreting the model: Races and adjustments.

Almost all of the analysis above started from the assumption that the two states in the model begin with zero arms, so that if one builds and the other does not the former can defeat the latter for sure. Empirically, it is hard to think of any situation in which states simultaneously "start" with no military capabilities at all. Instead, at any given time states have some allocation $(a_1, a_2) > 0$, and then make decisions about whether to build more or less in light of current circumstances.

In an MPE, strategies depend only on the current arms balance and not on how this was reached, so that we can "start" the interaction at any point (a_1, a_2) and ask what would follow.²² The most natural interpretation would then be that the states "start" at some level (a_1, a_2) which is a stable point, and then may experience an exogenous shock that changes some parameter(s). For example, an arms race from the current level of arms could occur if a technological change created some new offensive advantage, or if regime changes made one state more "greedy" than it had been before.

Arms races, in the ordinary language sense of rival states rapidly building up their militaries out of fear or aggressive aims with respect to each other, are not that common. Large-N studies typically identify between 20 and 30 "arms races" involving major powers between 1816 and the

²²I suspect that subgame perfect equilibria of the game are always MPEs, but this is only a conjecture at this point. This is almost certainly true for a finite horizon version of the model, where backward induction would lead to an MPE, so I am guessing it will also be true in the infinite horizon case with diminishing returns in the contest function. If this is right, then the restriction to MPE is not consequential.

end of the Cold War (Sample, 1997). States do, however, quite often *adjust* their military's size in light of perceived changes in other states' military policies and aims. The comparative statics of the models analyzed here can also be used to consider how states would optimally adjust their arms levels in light of such changes.

5.4 When and why are arms races "dangerous"?

As noted above, from the perspective of a state in the model, the biggest danger would be *failing* to arm against a greedy adversary, as this would lead to attack or coercion if they are starting from low enough arms levels. Nonetheless, armed conflict can arise in equilibrium in these models even when the states are well aware of this danger and are trying to avoid it.

It is important to distinguish between two different types of equilibrium armed conflict that emerge in these models. First, war may occur at the outset if the states anticipate that the amount of arming needed to get to stability is large enough that simply fighting is the better option. In this case, armed conflict is due to a commitment problem that is "robust" in the sense even if we were to change the extensive form of the game to allow for bargaining between the states, that wouldn't help. The problem in this case is not getting to a distribution of resources that both prefer to fighting, but that it is too costly to get to a level of arms such that peace is self-enforcing.

Second, war may occur if, at a mixed strategy point or when one state has a large enough building advantage, one state builds while the other does not. The advantaged state can then fight to "lock in," in expected terms, a stream of payoffs that is better than a status quo of v every period, and the disadvantaged state may as well fight because this provides some chance of surviving to enjoy future consumption. In this second case, allowing for bargaining and resource transfers *could* potentially allow the states to avoid a violent armed conflict. Once one state has a military advantage that makes it dissatisfied with the status quo, the two states are in the situation of a "dissatisfied" and "satisfied" state as appear in many bargaining models,²³ and there will be agreements on transfers that both sides would prefer to a costly fight.

²³See especially Powell (1999).

Bargaining may fail, of course, whether due to private information about willingness to fight, or commitment problems related to whether the stronger state can guarantee to uphold a deal in the future (Fearon, 1995, 2007; Powell, 1999). So we should probably interpret this type of armed conflict in these arms race models as arising as follows: Arming can produce an imbalance of power, which can make one state prefer armed conflict to the status quo, which in turn raises the odds of armed conflict if bargaining over redistribution fails for whatever reason.

Alternatively, bargaining may not fail, so that resources are redistributed but there is no destructive fight. If so, then what happens is coercion rather than war. Although this is bad for the coerced state, the outcome is not necessarily inefficient.²⁴

But how is it that arming might produce a dangerous imbalance of power and thus conflict or coercion? As we have seen, in the case of symmetric states imbalances do not arise endogenously if the states start from a position of equality (or rough equality if at higher levels of arms). With symmetric states the only conflict risk from arms races is of the first sort. Possibly the most interesting result is that when states differ in their "greed" or ability to build, arms races may necessarily involve states mixing on whether to build, so that power imbalances and thus war or coercion can arise endogenously.

It is still not clear, however, how to interpret these mixed strategies. Surely state leaders never consciously randomize over decisions to build more weapons, or increase the size of the army.

Harsanyi (1973) showed that for any mixed strategy Nash equilibrium in almost any complete information game there is a "near by" game with incomplete information such that players choose pure strategies but the equilibrium distribution on strategies is the same as in the mixed strategy case. This suggests interpreting mixed strategy equilibria as artifacts of too "coarse" a model: if we allow for some small private variation in preferences, then the players are choosing

²⁴A natural alternative model assumes that the balance of forces determines the disposition of an international issue (or composite of issues) automatically, so there is never any fighting but there can be coercion. [I'm working on this one ...]

pure strategies in light of their particular preferences, while they are uncertain about what the other players are doing.

Applied to the arms race model, the interpretation is that, depending on their idiosyncratic preferences, the status quo state's leaders may or may not want to take a chance that the adversary will keep building, even though they know war or coercion could follow if they don't build. Likewise, depending on idiosyncratic, privately known leadership preferences, the greedy state may or may not plunge ahead with its armaments program, hoping but not certain that the security seeker will not keep up.

The difference between a mixed strategy point and a pure strategy point in the arms race models considered here is that in the former, the strategic situation forces a "tough call" by the leadership, because it has to be unclear what the adversary will do. The mixed strategy result suggests that, in an arms race between asymmetric states, a time may come when it cannot be common knowledge that the adversary will continue to build, or will not build. This is why arms races between asymmetric adversaries may endogenously generate a risk of armed conflict or coercion.

5.5 What about "carrying costs"?

For simplicity and to work through the most relevant limiting case, the models above assumed there were no costs to maintaining arms once built. But of course it costs a lot to maintain aircraft carriers and nuclear forces, not to mention the direct wage costs of soldiers. What would happen if we introduced such carrying costs into the model? Suppose, for example, that the cost of maintaining a_i units of arms is ka_i per period, where k is a parameter greater than zero.

So modified, the model moves a step in the direction of Powell's (1993) guns-butter game, where the states choose how much to spend on arms out of total resources in alternate periods, and there is no accumulation over time (see also Jackson and Morelli 2009). Powell showed that if there is a peaceful MPE with deterrence, then there is a set of peaceful arms allocations that look like the cross-section of a lense in the (a_1, a_2) space, the smallest of which is an efficient MPE. At allocations within the lense, neither state would want to "break out" and attack the other – what it gets from deterrence is better than what it could get by arming up and attacking given the other's allocation. This set is bounded above because for high enough levels of arms, the states would prefer to fight even at even odds of winning to living at the low levels of consumption needed to sustain deterrence.

In Figures 1-5, the set of stable points forms (half of) the lower end of the lense seen in Powell's model, but it does not "close" above. This is because there are no carrying costs. If the states were somehow to reach or start at any level of arms in the stable set, the fact of no carrying costs means they are happy just to stay there, not building or attacking. Adding carrying costs will most likely cause the set of stable allocations, when it exists, to take the shape of a lense as in Powell's model.

I do not have a full analysis for this extension, but it is easy and instructive to write down a necessary condition for a stable point (a, a) in the case of symmetric states:

$$v - ka \ge (v - ka - c)(1 - \delta) + \delta p(a + 1, a)v(1 + \beta).$$

The left-hand side is the (time averaged) value of residing at peace with carrying costs ka each period; the right side is what a state can get by building and attacking. Note that a new and potentially major attraction of war here is that the winning state no longer has to pay the costs to maintain arms for deterrence (in this two state model). Rearranging, we have

$$p(a+1,a) \le \frac{\delta v + c(1-\delta) - \delta ka}{\delta v(1+\beta)} = p_h - \frac{\delta ka}{\delta v(1+\beta)}.$$

For large enough carrying costs and if states would want to attack an unarmed adversary, no level arms will satisfy this expression, since p(a + 1, a) > 1/2 for any a, and $p(1, 0) = 1 > p_h$. We see here essentially the same trade-off, or tension, as in Powell's model. A stable arms allocation must be at a high enough level that neither side is tempted to "break out" and attack, but if maintaining the forces necessary for deterrence is very costly, trying one's hand at war in hopes of getting to a less burdensome peace may be the better option for both states.

With small enough carrying costs, of course, we should expect that the main results to go through – states arming up to the first stable arms level. An interesting question is how increasing

carrying costs would affect this level. Making the inequality above an equality and implicitly differentiating in k leads the conclusion that increased carrying costs imply a longer arms race, if there is still a stable arms level both prefer to fighting.²⁵

5.6 Uncertainty about the other state's aims

An important line of theorizing about international politics is based on what might be called "the Jervis model," after arguments in Jervis (1976) and Jervis (1978). Glaser (2010) provides the fullest informal development of the model, which to date has been analyzed most thoroughly in formal terms by Kydd (1997, 2005).

The Jervis model posits two states, each of which may be one of two types, a "greedy state" or a "security seeker." As above, a greedy state puts positive value on expanding into new territory or changing the resolution of some international issues, while a status quo state does not. The states in the model know their own type but may be unsure about their adversary's type. In informal treatments, the choice variables range over whether to build more arms, what type of arms to build, what type of military doctrine to implement, whether to try to control the politics of "buffer states," and so on. In general the choice is posed as whether to adopt "cooperative" or "competitive" policies. It is commonly argued that the security seeker faces a dilemma, in that arming may be the best course against a greedy state, but could worsen matters if the other side is a security seeker and arming signals greed (or fails to signal lack of greed).

Full analysis of this extension is too much for this paper, but we can get some first-order conclusions by considering a simple special case. Suppose states 1 and 2 may be either "status quo," S, or "greedy," G; that S types put zero value on taking new territory while G types put value $\beta > 0$; and that the prior probability that state i is greedy is $\alpha_i \in (0, 1)$. Consider the game in the case where (2, 2) is the first stable point for two greedy states, and start the interaction at (1,

²⁵This is easier to see with the graphical argument: p(a + 1, a) is decreasing from 1 towards 1/2 as *a* increases from zero, and so if it intersects the line on the right-hand side it does so from above. Thus making the slope of that line more negative (increasing *k*) increases a^* (*a* at the first intersection, if there is one).

 $1).^{26}$

Proposition 7. Suppose beliefs about the other's probability of being greedy are the same, so that $\alpha_1 = \alpha_2 = \alpha$. Then in the game's MPE, at (1, 1) greedy states certainly build, while S types build if $\alpha > c(1 - \delta)/\delta vp(2, 1)$ and not otherwise. Thus, if this condition does not hold, then in equilibrium the types "separate," with peace at (1, 1) following if both are type S; war at (1, 2) or (2, 1) if one is G and one S; and peace at (2, 2) if both are G. If the condition holds, then all types build at (1, 1) and peace prevails (on the path) at (2, 2). In the latter case, the states' beliefs about the other's type are the same after both build as they were before (there is no "spiral of hostility").

The comparative statics are straightforward: Security seekers are more likely to arm, acting like expansionists, the greater the initial belief that the other is greedy, the more the offense-defense balance favors offense (which makes guessing wrong worse), and the smaller the costs of a round of arming relative to total resources. This last depends on "break out" time as well as the state of military technology: better monitoring implies smaller arms costs per "period," so that better monitoring and shorter response times actually incline security seekers to act like greedy states, since there is little cost to doing so (and equilibrium arms levels will be lower for greedy states as well).²⁷

Now suppose that one state is thought much more likely to be greedy than the other – say that α_1 is close to 1 and α_2 is close to zero, so that state 1 is probably greedy and state 2 is probably a security seeker. As in section 4.1, equilibrium has to involve mixed strategies. It can be shown that state 1_S does not build while 1_G mixes, and state 2_G builds for sure while 2_S mixes. If it happens that neither builds, then it becomes known that state 2 is a security seeker, but mixed strategies

²⁶The game is similar to that analyzed by Kydd (2005, chapter 2), differing most of all in that his game does not have attack decisions (it is what might be called a coercion model).

²⁷The comparative statics on beliefs α and the offense-defense balance economically summarize core arguments in Jervis (1976) and Glaser (2010). One difference concerns the role of costs of arming. This are typically put to the side in informal treatments, where the S types sole objective is usually said to be "security." The formal analysis shows that at least in this first-order formalization of the Jervis model, costs of arming are necessary for there to be any dilemma. If c = 0, then security-seekers should just arm irrespective of their beliefs. It requires a more elaborate argument that implies that arming produces a "spiral of hostility" and greater risk of preemptive war to have a hope of making the standard approach coherent if c = 0.

may continue in the next round since 1_G wants to build if 2_S does not build but not otherwise. If both happen to build in the first period, both states correctly increase their beliefs that the other is a greedy type – there is a "spiral of hostility," although it is not consequential since (2, 2) is a stable point even between greedy states. If state 2 builds and 1 does not, then peace prevails at (1, 2), with state 2 increasing its belief that 1 is a security seeker, but not completely sure. If 1 builds and 2 does not there is conflict, which is more likely to be between greedy states than the ex ante probabilities of the two types, but can still involve greedy 1 and security seeking 2.

This brief analysis of a special case suggests that, as with complete information arms races, asymmetries between the states imply behaviors that have not been seen in standard analyses of the "security dilemma" problem. In sum, asymmetries in preferences for expansion, or beliefs about these, imply strategic dynamics that can lead to endogenously created power imbalances, which in turn favor either stable peace if security seekers end up favored, or conflict if greedy states end up favored.

6 Appendix

Proof of Proposition 2. [tedious]

Proof of Proposition 3: Consider $a_1 > a_2$. Note first that if at a point a state 1 builds and then attacks in an MPE, then state 2 also builds. If state 2 does not build in this situation, for state 1 to want to build and attack it is necessary that $(v-c)(1-\delta) + \delta p_{10}v(1+\beta) > v(1-\delta) + \delta p_{00}v(1+\beta)$, which reduces to

$$p_{10} - p_{00} > C \equiv \frac{c(1-\delta)}{\delta v(1+\beta)}$$

By A1, $p_{11} < p_{00}$ and thus $p_{10} - p_{11} > p_{10} - p_{00} > C$. But $p_{10} - p_{11} > C$ proves to be the condition for state 2 to prefer to build if it expects 1 to build and attack.

Now suppose the proposition is false, so that there is an MPE with an *a* such that in an attack subgame, $p(a) > 1/(1 + \beta)$ but state 1 does not attack. The only way for state 1 to do better than $\delta p_{00}v(1 + \beta)/(1 - \delta)$ is by attacking at a subsequent point with better odds of winning. Thus there

is an $a' = (a_1 + x, a_2 + y)$ where x and y are non-negative integers, at which both state 1 and state 2 build and state 1 then attacks at $(a_1 + x + 1, a_2 + y + 1)$. But this could occur only if state 1 did not attack at a' when, by A1, it had a better chance of winning. This is impossible in an MPE (why pay cost c to fight a war at worse odds?).

Proof of Proposition 4. A1 implies that when x > y,

$$\frac{\partial p(x,y)}{\partial x} < -\frac{\partial p(x,y)}{\partial y}$$

or in words, 1's probability gain from building another time is smaller than 2's.

Neither state wishes to build given that 1 will attack when

$$v + \delta p_{00} \delta v (1+\beta) \ge v - c + \delta p_{10} \delta v (1+\beta),$$

and, for state 2,

$$v + \delta(1 - p_{00})\delta v(1 + \beta) \ge v - c + \delta(1 - p_{01})\delta v(1 + \beta).$$

These reduce to $p_{10} - p_{00} < C$ and $p_{00} - p_{01} < C$. By A1, the second implies the first, so this is the condition for (NB, NB) at such build points.

Similarly, we find that 2 prefers to build and 1 not to build when $p_{00} - p_{01} > C > p_{11} - p_{01}$. (This range always exists since $p_{11} < p_{00}$.) And (B, B) are best replies when $p_{11} - p_{01} > C$ and $p_{10} - p_{11} > C$.

The first inequality for (NB, B) contradicts the second inequality for (NB, NB), there is likewise a contradiction that means we cannot have both (NB, B) and (B, B). (NB, NB) and (B, B) cannot both obtain because the first inequality for (B, B) implies $p_{00} - p_{01} > C$, which contradicts the second condition for (NB, NB).

Proof of Proposition 5. First, $\delta v > c$ implies that if state *i* expects that building up to a^* , it will prefer to do so rather than not build at t = 0 and be attacked. The latter yields, time-averaged payoffs, $v(1 - \delta)$, and the former $v - c(1 - \delta^{a^*})$. So $\delta v > c$ is a sufficient condition for the second to be greater than the first.

A2 implies that there is a first stable point a^* : By A2 for large enough x, $p(x + 1, x) < p_h$, since $p_h > 1/2$. So there is a smallest integer a such that $p(a + 1, a) \le p_h$, the condition for a stable point.

The time-averaged payoff for building up to the first stable point is $v - c(1 - \delta^{a^*})$, which is less than the payoff for attacking after both build in t = 0 when

$$v - c(1 - \delta^{a^*}) < v(1 - \delta) + \delta v(1 + \beta)/2$$

This gives the condition in the proposition. By $\delta v > c$, a state does worse to not build and get attacked with no chance of winning. If the inequality is reversed, then a state prefers building to a^* to fighting at (0,0), which ensures that the state also prefers building to a^* from any (a,a), $a < a^*$, since the costs of getting there are smaller.

References

- Acemoglu, Daron, Michael Golosov, Aleh Tsyvinski and Pierre Yared. 2011. "A Dynamic Theory of Resource Wars." Unpublished paper, MIT, Yale, Columbia Universities.
- Axelrod, Robert. 1984. The Evolution of Cooperation. New York: Basic Books.
- Brito, Dagobert L. and Michael D. Intriligator. 1985. "Conflict, War, and Redistribution." *American Political Science Review* 79:943–957.
- Downs, George W. 1991. Arms Races and War. In *Behavior, Society, and Nuclear War*, ed. Philip Tetlock et al. New York: Oxford University Press.
- Downs, George W. and David M. Rocke. 1990. *Tacit Bargaining, Arms Races, and Arms Control*. Ann Arbor, MI: University of Michigan Press.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49:379–414.
- Fearon, James D. 2007. "Fighting rather than Bargaining.". Paper presented at the Annual Meetings of the American Political Science Association.
- Fearon, James D. 2008. "A Simple Political Economy of Relations among Democracies and Autocracies." Paper presented at the Annual Meetings of the American Political Science Association, Boston, MA, August 27-31.
- Glaser, Charles L. 1992. "Political Consequences of Military Strategy: Expanding and Refining the Spiral and Deterrence Models." *World Politics* 44(4):497–538.
- Glaser, Charles L. 2000. "The Causes and Consequences of Arms Races." Annual Review of Political Science 3:251–76.
- Glaser, Charles L. 2010. *Rational Theory of International Politics*. Princeton: Princeton University Press.
- Glaser, Charles L. and Chaim Kaufmann. 1998. "What is the Offense-Defense Balance and Can We Measure It?" *International Security* 22(4):44–82.
- Harsanyi, John. 1973. "Games with randomly disturbed payoffs: A new rationale for mixedstrategy equilibrium points." *International Journal of Game Theory* 2:1–23.
- Hirshleifer, Jack. 1989. "Conflict and Rent-seeking Success Functions: Ratio and Difference Models of Relative Success." *Public Choice* 63:101–112.
- Hirshleifer, Jack. 1995. "Anarchy and Its Breakdown." Journal of Political Economy 103(1):26-52.
- Jackson, Matthew O. and Massimo Morelli. 2009. "Strategic Militarization, Deterrence and Wars." *Quarterly Journal of Political Science* 4(4):279–313.

- Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton: Princeton University Press.
- Jervis, Robert. 1978. "Cooperation Under the Security Dilemma." World Politics 30:167-214.
- Kydd, Andrew. 1997. "Game Theory and the Spiral Model." World Politics 49(3):371-400.
- Kydd, Andrew. 2000. "Arms Races and Arms Control: Modeling the Hawk Perspective." *American Journal of Political Science* 44(2):228–44.
- Kydd, Andrew. 2001. "Trust Building, Trust Breaking: The Dilemma of NATO Enlargement." *International Organization* 55(4):801–828.
- Kydd, Andrew H. 2005. *Trust and Mistrust in International Relations*. Princeton: Princeton University Press.
- Powell, Robert. 1999. In the Shadow of Power. Princeton University Press.
- Powell, Robert M. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1):115–132.
- Richardson, Lewis F. 1919. Mathematical Psychology of War. Oxford: William Hunt.
- Sample, Susan G. 1997. "Arms Races and Dispute Escalation: Resolving the Debate." *Journal of Peace Research* 34(1):7–22.
- Skaperdas, Stergios. 1992. "Conflict, Cooperation, and Power in the Absence of Property Rights." *American Economic Review* 82(4):720–739.
- Slantchev, Branislav L. 2005. "Military Coercion in Interstate Crises." *American Political Science Review* 99(4):533–547.
- Snyder, Glenn H. 1971. "'Prisoner's Dilemma' and 'Chicken' Models in International Politics." International Studies Quarterly 15(1):66–103.
- Snyder, Jack L. 1991. *Myths of Empire: Domestic Politics and International Ambition*. Ithaca: Cornell University Press.

Figure 1. Threshold csf with m = 1.5.

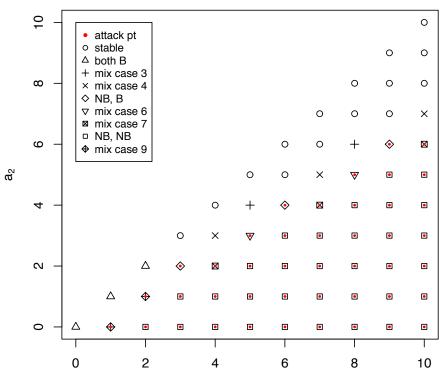


Figure 2. Eqm path and mixed strategies, m = 1.5

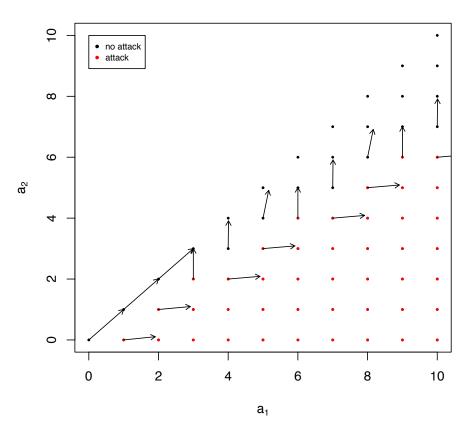


Figure 3. Eqm path, stable, and attack points

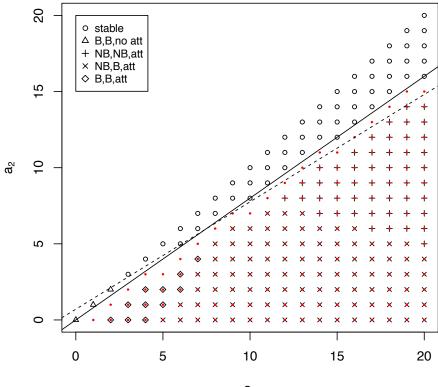


Figure 4. MPE for ratio form with c/v = .5, beta = .8, delta = .9.

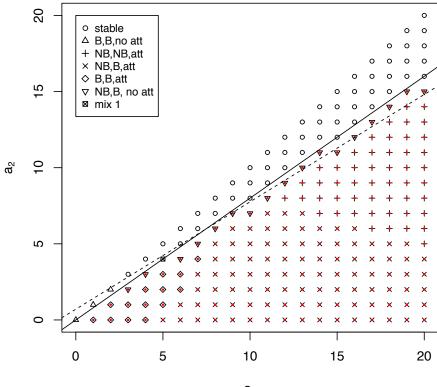


Figure 5. MPE for c/v = .5, beta = .8, delta = .9, n = 2.

